UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ARTES, CIÊNCIAS E HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO

THAIS LUIZA DONEGA E SOUZA

Os efeitos das revisões críticas online sobre o mercado cinematográfico americano

São Paulo

2017

THAÍS LUIZA DONEGA E SOUZA

Os efeitos das revisões críticas online sobre o mercado cinematográfico americano

Versão original

Dissertação apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Sistemas de Informação.

Área de Concentração: Metodologia e Técnicas da Computação

Orientador: Profa. Dra. Marislei Nishijima

São Paulo
2017

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

CATALOGAÇÃO-NA-PUBLICAÇÃO

(Universidade de São Paulo. Escola de Artes, Ciências e Humanidades. Biblioteca)

Souza, Thais Luiza Donega e

Os efeitos das revisões críticas online sobre o mercado cinematográfico americano / Thais Luiza Donega e Souza ; orientadora, Marislei Nishijima. – São Paulo, 2017.

160 f.

Dissertação (Mestrado em Ciências) - Programa de Pós-Graduação em Sistemas de Informação, Escola de Artes, Ciências e Humanidades, Universidade de São Paulo.

Versão original

1. Inteligência artificial. 2. Crítica cinematográfica - Efeitos. 3. Mineração de dados. I. Nishijima, Marislei, orient. II. Título

CDD 22.ed.— 006.3

Dissertação de autoria de Thaís Luiza Donega e Souza, sob o título "**Os efeitos das revisões críticas online sobre o mercado cinematográfico americano**", apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo, para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Sistemas de Informação, na área de concentração Metodologia e Técnicas da Computação, aprovada em _____ de _____ de _____ pela comissão julgadora constituída pelos doutores:

Prof. Dr. _____
Instituição: _____
Presidente

Prof. Dr. _____
Instituição: _____

Prof. Dr. _____
Instituição: _____

Prof. Dr. _____
Instituição: _____

Agradecimentos

Primeiro, aos meus pais. Sem dúvidas os melhores do mundo, mesmo que as vezes não caibamos no mesmo espaço. Acredito que é o amor que transborda. Se um dia eu puder retribuir por todos os esforços imensuráveis que já me deram, prometo-lhes mais do que papéis, escritas e certificados.

Aos meus avós, àquele que ficaria discretamente orgulhoso e “me daria dez reais”, àquela que neste momento (ou em qualquer outro) está pedindo pela minha felicidade e sucesso e ainda cuida de mim e àquela que já esqueceu o que eu disse e que mesmo sem se lembrar me faz grande companhia.

Todos os outros gordinhos do meu coração, de uma maneira ou de outra, me apoiaram o tempo todo com as minhas decisões malucas. Vinícius e Karem, meu muito obrigada pela compreensão, apoio, torcida, marmitas e impressões infinitas.

Bom, é só neste momento que se torna real, certo?! A resposta desta inscrição foi uma transformação de vida. Exageros à parte, é um grande feito! Não o fiz sozinha, claro. Devo desde a intenção de querer a academia até a conclusão desta e das próximas etapas que virão aos meus grandes inspiradores.

A minha admiração “pedestalesca” pelos meus queridos mestres que me acompanharam me trouxe até aqui e a depender dos que estiveram comigo, eu vou continuar. Obrigada aos professores que me ensinaram, me motivaram, me guiaram. Obrigada àqueles que atenderam meus e-mails da madrugada, escreveram cartas de recomendação, tiveram paciência mesmo quando eu nem os compreendia. Obrigada não por serem maravilhosos, mas por serem professores! Obrigada, em especial, a esta minha orientadora. Secreta e carinhosamente chamada de Marismãe pelos seus orientandos. O seu apoio foi fundamental, sorte daqueles que podem contar com você, ouvir seus elogios motivadores e suas críticas em forma de sugestão. Espero que continuemos trabalhando juntas para coisas improváveis, mirabolantes e sensacionais.

Agradeço também aos meus queridos, aos que ganhei com o mestrado (sinto saudades da gente!), obrigada pelo apoio, pelas fofocas, risadas, piadas, discussões, trabalhos, noites varadas, “me passa o seu”, “revisa pra mim”, “o meu não roda”, cafés, cervejas, mais cervejas e pela alegria toda!

Em especial, meu eterno agradecimento. Meus queridos da vida toda, de todos os anos: Robs, Du, Erika, Danilo, Morelli e Aca. Gente, vocês são as melhores pessoas do mundo e somos invencíveis (pode crer, todos somos um e juntos não existe mal nenhum)! Obrigada por

acreditarem em mim, pelo apoio, por me aguentarem mesmo surtando, por me amarem incondicionalmente. Eu simplesmente não tenho palavras suficientes para descrever minha gratidão, mas tudo bem, eu nunca fui boa com palavras e vocês gostam de mim mesmo assim :) , mas segue o clichê “Sem vocês, nada disso seria possível”. Ah, agregados e parças também contam!

Por último e para que eu possa me ler no futuro com orgulho – e uma certa vergonha – estou feliz por ter tentado, arriscado e não desistido: eu arrasei!

“In our culture we always search for closure, but closure is an illusion. Science shows us that the universe is constantly in flux. It is what allows ours friendships and our love to constantly surprise us.”

Temperance Brennan - The Next in the Last.

"A film is simply a series of emotions strung together with a plot... though flippant, this thought is not far from the truth. It is the filmmaker's job to create moods in such a realistic manner that the audience will experience those same emotions enacted on the screen, and thus feel part of the experience"

Ian Maitland. Emmy Award winning director and editor.

Resumo

SOUZA, Thaís Luiza Donega. **Os efeitos das revisões críticas online sobre o mercado cinematográfico americano**. 2017. 160 f. Dissertação (Mestrado em Ciências) – Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, São Paulo, 2017.

O mercado cinematográfico pode ser caracterizado como uma indústria de entretenimento com a produção de bens de informação que são também bens de experiência, cuja qualidade só é conhecida após o consumo. Deste modo, a revisão crítica se torna importante para induzir seu consumo, fornecendo previamente algum grau de informação sobre a qualidade do bem. Segue-se o trabalho de Reinstein e Snyder (2005) para determinar se as revisões críticas conduzidas por consumidores e por críticos profissionais online afetam o tempo de exibição de filmes no mercado americano de cinema, medido em quantidades de semanas, conforme modelos de duração/sobrevivência na literatura. Para esta finalidade foi gerado, a partir de sites de cinemas americanos (Box Office Mojo e Rotten Tomatoes), um banco de dados extremamente rico com informações semanais de todos os filmes disponíveis no cinema americano de 2004 a 2015. Especificamente, investigou-se os efeitos das revisões críticas de críticos profissionais de primeira linha (*Tops*) e de consumidores, conforme a média das notas atribuídas na semana de lançamento de cada filme. No que se refere à avaliação dos consumidores foi aplicada a computação afetiva, que reconhece o sentimento e a emoção em suas resenhas online para captar o efeito boca a boca potencializado pelas mídias sociais e fornecendo, portanto, uma análise mais profunda do boca a boca online. O estudo controla por possíveis problemas de endogeneidade decorrente de simultaneidade, usando as críticas somente antes e durante a semana de lançamento dos filmes. Os resultados sugerem que os críticos profissionais exercem grande influência no tempo de duração dos filmes em cartaz, bem como a positividade dos consumidores em relação ao filme. No entanto, o efeito dos críticos profissionais é em média 3 vezes maior do que dos consumidores. Adicionalmente, pode-se observar que algumas emoções afetam a expectativa de vida dos filmes a depender do gênero do mesmo.

Palavras-chave: Bens de experiência; Cinema; Análise de sentimento; Mineração de emoção; IBM Watson; Boca a boca; Revisões críticas; Computação afetiva; Análise de sobrevivência.

Abstract

SOUZA, Thaís Luiza Donega. **The effects of online critical reviews over the American movie market**. 2017. 160 p. Dissertation (Master of Science) – School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, 2017.

The movie market may be considered as entertainment industry, which produces experience goods that is also information goods, whose quality is only known only after consumption. Thus, critical reviews becomes important to induce consumption, since it provides some level of information about product quality. We follow Reinstein and Snyder (2005) works in order to determine if experts and consumers online critical reviews affect the survival time of movies at the American movie market, measured by number of weeks, according to survival analysis models in the literature. For this purpose, an extremely rich database with weekly information on all the films available in American cinema from 2004 to 2015 was generated from American movie sites (Box Office Mojo and Rotten Tomatoes). Specifically, we investigate the effects of critical reviews from top professionals and from consumers, according to the average ratings assigned in each movie's release week. As far as consumer assessment was concerned, affective computing was applied, which recognizes the sentiment (sentiment analysis) and emotion (emotion mining) in their online reviews to capture the word-of-mouth effect boosted by social media. The study controls for possible problems of endogeneity due to simultaneity, using the criticisms before and during the week of release of the films. The results suggest that the professional critics exert a great influence on the duration of the films in exhibition, as well as the positivity of the consumers in relation to the film. Thus, the effect of professionals are 5 times greater, generally, than the effect of the consumer critics. Additionally, it can be observed that some emotions affect movie life expectancy depending on the its genre.

Keywords: Experience goods; Movie; Motion picture; Sentiment analysis; Emotion mining; IBM Watson; Word of mouth; Critical reviews; Affective computing; Survival analysis.

Lista de figuras

Figura 1 - Roda dos sentimentos proposta por Plutchik	54
Figura 2 – Ferramentas de Compreensão de Linguagem Natural IBM Watson	64
Figura 3 - Extração de relações, entidades, palavras-chave, sentimento e emoção	66
Figura 4 - Diagrama Entidade-Relacionamento do banco de dados criado	90
Figura 5 - Média de emoções por gênero	99
Figura 6 - Média da audiência por gênero	101
Figura 7 - Média do sentimento por gênero	102
Figura 8 - Média das Avaliações dos Críticos	102
Figura 9 - Distribuição dos tempos	105
Figura 10 - Intervalo de confiança para o risco dos filmes saírem das telas de cinema.....	109
Figura 11 - Sobrevivência dos filmes norte-americanos nas salas de cinema.....	109
Figura 12 - Sobrevivência conforme a prevalência de sentimentos	110
Figura 13 - Sobrevivência conforme as avaliações das críticas de profissionais	111
Figura 14 - Estimativa de sobrevivência por emoção predominante.....	111
Figura 15 - Exemplo de crítica retirada do Rotten Tomatoes.....	151
Figura 16 - Requisição de análise do Documento	151
Figura 17 - Análise de Sentimento do Documento.....	152
Figura 18 - Resposta da detecção de Idioma	152
Figura 19 - Resposta da Mineração de Emoção	153
Figura 20 - Taxonomia do Documento	154

Lista de Quadros

Quadro 1 - Pseudo código script da utilização da API.....	21
Quadro 2 - Exemplos de URLs do BoxOfficeMojo.com	88
Quadro 3 - Algoritmo para análise afetiva com API Watson.....	141
Quadro 4 - Script de indexação dos filmes.....	146
Quadro 5 - Script de Coleta Week BoxOffice	147
Quadro 6 - Script de Coleta Revisões Críticas Consumidores RottenTomatoes.....	148
Quadro 7 - Chamada Combinada da API	150
Quadro 8 - Leitura do retorno da chamada da API.....	150
Quadro 9 - Chamada da API AlchemyLanguage para emoção e sentimento	151
Quadro 10 - Resposta via API da Análise de Sentimento	152
Quadro 11 - Resposta via API da detecção automática do Idioma	153
Quadro 12 - Resposta via API da Mineração de Emoção	153
Quadro 13 - Retorno completo da chamada combinada da API	154

Lista de tabelas

Tabela 1 - Modelos Paramétricos de Sobrevivência e Risco.....	26
Tabela 2 - Autores e as emoções básicas propostas	54
Tabela 3 - Métodos disponíveis pela funcionalidade da análise de sentimento da API AlchemyLanguage.....	67
Tabela 4 - Métodos disponíveis pela funcionalidade da análise de emoção da API AlchemyLanguage.....	68
Tabela 5 - Trabalhos revisados	82
Tabela 6 - Acessos do Web Crawler	89
Tabela 7 - Análise descritiva dos dados	91
Tabela 8 - Tabulação avaliação críticos e audiência por filme.....	92
Tabela 9 - Avaliação dos críticos por crítica	92
Tabela 10 - Avaliação dos consumidores por crítica.....	93
Tabela 11 - Variáveis de interesse consumidores.....	103
Tabela 12 - Variáveis de interesse críticos profissionais.....	104
Tabela 13 - Matriz de correlação entre as variáveis contínuas.....	108
Tabela 14 – Estimativa paramétrica modelo (i).....	113
Tabela 15 – Estimativa paramétrica modelo (ii)	114
Tabela 16 - Estimativa modelo paramétrico (i) por gênero.....	115
Tabela 17 - Estimativa modelo paramétrico (ii) por gênero.....	116
Tabela 18 - Descrição das Variáveis	143
Tabela 19 - Exemplos da Análise Afetiva das Revisões Críticas.....	155
Tabela 20 - Exemplos de Revisões Críticas rejeitadas pelo Watson AlchemyLanguage	156
Tabela 21 - Resultado da Análise Afetiva Watson AlchemyAPI.....	156
Tabela 22 - Dados da Análise Afetiva agregados por Filme	156
Tabela 23 - Taxas de risco modelo (i)	157
Tabela 24 - Taxas de risco modelo (ii)	158
Tabela 25 - Taxas de risco modelo (i) por gênero – estimação Weibull	159
Tabela 26 - Taxas de risco modelo (ii) por gênero – estimação Weibull.....	160

Sumário

1	INTRODUÇÃO	14
1.1	CONTEXTUALIZAÇÃO	14
1.2	OBJETIVOS	15
1.3	JUSTIFICATIVA E MOTIVAÇÃO.....	16
1.4	CONTRIBUIÇÕES	18
1.5	MATERIAIS E MÉTODOS	19
1.5.1	<i>Materiais e Métodos em Computação Afetiva</i>	<i>20</i>
1.5.2	<i>Materiais e Métodos em Economia</i>	<i>23</i>
1.6	ORGANIZAÇÃO DO TRABALHO.....	30
2	COMPUTAÇÃO AFETIVA.....	31
2.1	VISÃO GERAL.....	34
2.2	PROCESSAMENTO DE LINGUAGEM NATURAL	37
2.2.1	<i>Definição e História.....</i>	<i>37</i>
2.2.2	<i>Métodos.....</i>	<i>39</i>
2.3	ANÁLISE DE SENTIMENTO	47
2.3.1	<i>Definição.....</i>	<i>47</i>
2.3.2	<i>Revisão Bibliográfica Análise de Sentimentos</i>	<i>49</i>
2.4	MINERAÇÃO DE EMOÇÃO	52
2.4.1	<i>Definição.....</i>	<i>52</i>
2.4.2	<i>Revisão Bibliográfica Mineração de Emoção.....</i>	<i>57</i>
2.5	WATSON.....	59
2.5.1	<i>Computação Cognitiva, Watson e Jeopardy!.....</i>	<i>59</i>
2.5.2	<i>Inteligência Artificial, PLN e Aplicações</i>	<i>62</i>
2.5.3	<i>Alchemy API, Análise de Sentimento e Mineração de Emoção.....</i>	<i>65</i>
2.5.4	<i>Revisão AlchemyAPI.....</i>	<i>68</i>
2.6	CONSIDERAÇÕES FINAIS.....	69
3	LITERATURA DE REVISÃO CRÍTICA EM ECONOMIA.....	71
3.1	VISÃO GERAL.....	71
3.2	REVISÃO BIBLIOGRÁFICA.....	73
3.3	CONSIDERAÇÕES FINAIS.....	83
4	DADOS	85
4.1	VISÃO GERAL.....	85
4.2	A COLETA	87
4.3	ANÁLISE DESCRITIVA DOS DADOS	90
4.4	CONTEXTO: COMPUTAÇÃO AFETIVA.....	94

4.5	CONTEXTO: ECONOMIA	96
4.6	CONSIDERAÇÕES FINAIS	97
5	RESULTADOS	98
5.1	RESULTADOS PRELIMINARES	98
5.2	ESTIMATIVAS	103
5.2.1	<i>Resultados não paramétricos</i>	<i>108</i>
5.2.2	<i>Resultados Paramétricos</i>	<i>112</i>
6	CONCLUSÃO	119
6.1	PRINCIPAIS CONTRIBUIÇÕES	122
6.2	TRABALHOS FUTUROS	122
	REFERÊNCIAS	124
	ANEXO A – SCRIPT DE ANÁLISE AFETIVA COM API WATSON	141
	ANEXO B – DESCRIÇÃO DAS VARIÁVEIS	143
	ANEXO C – SCRIPT DE COLETA DO BANCO DE DADOS	146
	ANEXO D – RESULTADOS DA COMPUTAÇÃO AFETIVA	150
	ANEXO E – TAXAS DE RISCO MODELOS (I) E (II).....	157

1 Introdução

Em um mercado caracterizado por um grande número de consumidores, que de maneira irregular e infrequente ao longo do tempo fazem escolhas a partir de uma enorme gama de possibilidades, informações imparciais ou parciais altamente disponíveis sobre a qualidade e características das opções podem exercer um impacto significativo em sua demanda.

Existem diversas indústrias em que profissionais oferecem suas opiniões sobre a qualidade dos produtos ou marcas, tais como filmes, shows, vinhos, livros, entre outros. Adicionalmente, a quantidade de postagens online dos consumidores sobre esses produtos e marcas tem sido cada vez maior.

Dentro deste contexto de mudanças tecnológicas grandes e rápidas, torna-se quase impossível não ser interdisciplinar em algum aspecto. As revisões críticas de críticos profissionais reconhecidos (*top*) e outros profissionais, além das revisões críticas de consumidores registradas nas páginas da web se tornaram bens de consumo – na forma de bens de informação – que sinalizam a qualidade dos bens de experiência a serem consumidos.

1.1 Contextualização

Bens de experiência são bens cuja qualidade só é conhecida após o seu consumo. Filmes em salas de cinema são bens de experiência, além de serem típicos bens de informação (SHAPIRO e VARIAN, 1999). As revisões críticas são informações sobre a qualidade do filme fornecidas por quem já consumiu o filme por profissão, caso dos críticos profissionais, ou por altruísmo ou outra motivação, caso dos consumidores que desejam compartilhar suas experiências de consumo. Ao mesmo tempo, as revisões críticas de usuários podem ser carregadas de emoções e juízos de valores que tendem a influenciar as pessoas.

Bens de informação são bens que podem ser copiados, transmitidos e armazenados em códigos numéricos (SHAPIRO e VARIAN, 1999). Dito isso, tem-se que os recentes desenvolvimentos tecnológicos que permitiram a criação e viabilidade de consumo em massa de bens de informação associados com o advento da Internet mudaram muitos

aspectos da vida cotidiana dos agentes. Buscando explorar este fenômeno e necessariamente se inserindo numa área interdisciplinar de sistemas de informação e economia, este estudo foca nas mudanças no consumo de filmes em salas de cinema sob o efeito de revisões críticas de consumidores e de profissionais comuns deste bem.

Com a disponibilização em massa de uso da Internet e a consequente proliferação dos sistemas de revisão online dos consumidores, a dinâmica do fenômeno boca a boca tem sofrido impacto por efetivamente conectar os pares (CHEN e XIE 2008; ELIASHBERG, ELBERSE e LEENDERS, 2006; GODES e MAYZLIN, 2004; GODES *et al.* 2005; MAYZLIN 2006; TRUSOV, BUCKLIN e PAUWELS, 2009). Esses fóruns online diminuem o custo de busca de informações sobre um produto, o que motiva novos consumidores a procurá-los (STIGLER, 1961). Afinal, as comunidades de opiniões coletivas dos consumidores podem praticar tanta influência quanto as opiniões de críticos profissionais. Desta forma, consumidores julgam que as avaliações feitas tanto por críticos profissionais quanto pela comunidade amadora, formado por outros consumidores, são úteis no seu processo de escolha.

Por um lado, o boca a boca fornece uma fonte alternativa de informação aos consumidores e portanto, reduz a habilidade das empresas de influenciar tais consumidores pelas maneiras tradicionais de marketing. Por outro lado, o boca a boca talvez seja um novo canal para que as empresas alcancem seus consumidores e estrategicamente os influencie em suas opiniões (DUAN *et al.*, 2008).

1.2 Objetivos

O objetivo geral deste estudo é verificar o efeito das revisões críticas, dos críticos profissionais e dos consumidores, disponíveis na Internet sobre o tempo de duração dos filmes nas salas de cinema do mercado norte americano. Mais especificamente, este trabalho estuda o efeito das avaliações de críticos profissionais conhecidos e não conhecidos (*top* e não *top*) e dos consumidores/usuários sobre o consumo de filmes em salas de cinema no mercado norte americano nas últimas duas décadas.

Para esta finalidade, seguimos a modelagem proposta por Reinstein e Snyder (2005), alterando a variável dependente de faturamento para tempo de duração, medido em semanas, conforme McKenzie (2009). Mas, tal como Reinstein e Snyder (2005) controlamos por viés de omissão de variável relevante utilizando *proxies* do efeito boca

a boca, obtidas a partir das revisões críticas dos consumidores e por endogeneidade usando somente as críticas dos críticos profissionais disponíveis apenas na primeira semana de lançamento, incluindo somente o final de semana em que foi lançado o filme. Adicionalmente, recursos computacionais, como técnicas de processamento de linguagem natural e o uso de ferramentas de computação cognitiva para computação afetiva, foram utilizados conjuntamente para viabilizar a estimação dos modelos. Esses procedimentos explicam os objetivos específicos associados.

Um dos objetivos específicos é avaliar como a revisão dos consumidores pode se desdobrar no sentido de refletir mais adequadamente os efeitos do fenômeno de boca a boca, se afeta e como afeta o tempo de exibição dos filmes em salas de cinema. Isso é relevante uma vez que o boca a boca tende a impactar no consumo dos bens por influência dos pares dos indivíduos, tanto no momento em que tais consumidores se manifestam através de suas opiniões e emoções envolvidas em suas análises, quanto através da forma simplificada de avaliação por ratings/score, tal qual os críticos profissionais.

Ainda sobre o tempo de duração nas salas de cinema do mercado norte-americano, o segundo objetivo específico é avaliar o efeito dos críticos profissionais de primeira linha e dos não classificados nessa rubrica.

1.3 Justificativa e motivação

A escolha do mercado estadunidense está relacionada com a grande disponibilidade de dados, por se tratar do maior produtor de filmes do planeta e de ter uma indústria pioneira em entretenimento. Além disso, o processamento de linguagem natural encontra suas ferramentas desenvolvidas primordialmente em inglês e o uso de revisões de consumidores nesse idioma permite maior acurácia aos objetivos buscados neste estudo.

O setor cinematográfico é um dos maiores em arrecadação nos Estados Unidos e sua importância econômica, ainda que seja um setor de entretenimento (McKENZIE, 2010), constitui objeto de estudo importante pela sua própria dinâmica geradora de riquezas (SILVER, 2007). Assim, estudos econômicos junto ao setor reduzem o grau de desconhecimento de sua lógica de operação e podem gerar indicativos de geração de riquezas para outros países.

De importância econômica global, filmes são principais impulsionadores do mercado de entretenimento, atualmente o número um de exportação dos EUA. A indústria de filmes possui um significado cultural muito forte e atrai atenção. Novamente, nos EUA, os filmes possuem um impacto desproporcional na cultura. Do ponto de vista de pesquisa, o interesse surge pela disponibilidade de grandes e ricos bancos de dados e pelos quebra-cabeças ainda intrigantes que existem sobre o mercado de cinema como um todo (ELIASHBERG *et al.*, 2006).

Ainda sobre o mercado de cinema, a mudança constante da lista de filmes em exibição dá origem a ciclo de vida tipicamente curtos e a performance é fundamentalmente ditada pela informação compartilhada entre os consumidores (boca a boca) durante a vida do filme (McKenzie 2009). Desta forma, o número de semanas que um filme permanece em cartaz talvez forneça informação útil relacionada à dinâmica do boca a boca e agregue às decisões marginais dos exibidores, ao decidir sobre a lucratividade de um filme em relação ao outro ou em permanecer sem exibir nenhum filme (SQUIRE, 1992). Assim, o objetivo proposto de medir a duração ou sobrevivência de um filme em cartaz é uma métrica alternativa para investigar performance e sucesso.

Todo filme é uma inovação que compete pelas telas de cinema e pelas audiências durante seu curto período de vida contra uma mudança no vetor de substitutos imperfeitos (DE VANY e WALLS, 1997). Quando os filmes (de alta qualidade) possuem alta demanda, é possível extrapolar a capacidade dos cinemas, limitada pelos seus assentos, fazendo com que a exibição do filme ocorra em um período maior (DEUCHERT *et al.*, 2005). Portanto, o tempo de vida de um filme é a margem pela qual o fornecedor se ajusta à demanda do mercado devido as admissões de demandas invariantes de preços e a capacidade de assentos (WALLS, 1998).

Por fim, as intuições e conclusões adquiridas sobre o mercado cinematográfico podem ajudar a entender, de uma forma melhor, outras indústrias que compartilham das mesmas características, bem como examinar a interface entre a tecnologia e os bens de experiência na era digital. Utilizando como base o mercado de cinema, este modelo poderá ser usado em qualquer mercado semelhante, ou seja, que seja um mercado de bem de experiência (vinhos, shows, teatros, entre outros) e que se utilize de resenhas críticas como forma de mensurar qualidade prévia do produto, tal como o cinema. Do mesmo modo, a aplicabilidade do modelo pode se estender a diversos países, não se limitando às fronteiras culturais, desde que o mercado tenha o comportamento supracitado. Neste caso, podemos incluir os mercados de bens de experiência brasileiros, contanto que se faça as

adaptações cabíveis nos processos e ferramentas computacionais utilizados, no que se refere ao idioma.

1.4 Contribuições

Este estudo pretende contribuir para a literatura de economia aplicada de efeitos de revisão crítica, por meio de uma nova abordagem sobre as métricas de interesse do filme. Muitos trabalhos focam na tradicional avaliação do efeito da revisão crítica sobre o faturamento ou bilheteria (ELIASHBERG e SHUGAN, 1997; REINSTEIN e SNYDER, 2005; BOATWRIGHT et al., 2007; DUAN et al., 2008; CHINTAGUNTA et al., 2010 e MOON et al., 2010), diferentemente do proposto aqui, com uma literatura muito mais tímida, que é a medida do efeito das revisões críticas de críticos profissionais e de consumidores sobre o tempo de duração do filme em exibição (McKENZIE, 2009; LEGOUX et al., 2016; WALLS, 1998; DEUCHERT et al., 2005).

Ainda com o propósito de complementar a literatura atual, os dados que serão utilizados para análise foram coletados exclusivamente para esta finalidade, constituindo um banco de dados extremamente abrangente em quantidade e variedade do conteúdo de informações. Os dados foram coletados de diferentes sites da Internet e contém informações sobre aproximadamente 13000 filmes lançados no mercado americano. Os detalhes sobre o *dataset* estão no capítulo 4.

Do ponto de vista da literatura de revisão crítica (REINSTEIN e SNYDER, 2005; MOON et al., 2010; LUCA, 2011; KING, 2007; BASUROY et al., 2003) este trabalho oferece também uma importante contribuição na medida que controla pelas avaliações dos pares (*peer to peer*). Avaliações estas mensuradas pela revisão crítica dos consumidores junto aos sites de revisões críticas. Adicionalmente, este controle de avaliação dos pares traz um possível ineditismo a partir do uso da computação afetiva, que proporciona uma melhor investigação sobre os efeitos das críticas nas medidas de sucesso do filme.

Finalmente, entende-se como sendo a maior contribuição deste estudo, verificar o efeito da enorme ampliação da comunicação via Internet, que naturalmente reduz o custo de obtenção da informação e, portanto, reduz também o grau de assimetria de informação nos mercados, sobre as decisões de consumo dos agentes.

Sendo assim, o uso de técnicas computacionais possibilita a obtenção de uma variável proxy para o efeito não observável do boca a boca, antes não disponível. A ideia é que a linguagem dos consumidores seja muito mais próxima à de outros consumidores do que a dos críticos profissionais, o que permite mensurar em certa medida este efeito de influência não observável, além do fato já citado acima de que o próprio fenômeno boca a boca está mudando com a disseminação da Internet. Do ponto de vista metodológico, isso se constituiu no emprego de técnicas computacionais existentes em dados e problemas ainda não explorados na economia.

Nossas hipóteses são: as revisões críticas afetam o tempo de duração do filme em cartaz; as revisões de consumidores podem ser usadas como uma proxy para efeito entre pares decorrente do fenômeno “boca a boca”, que por sua vez também afeta a decisão de consumo dos agentes, além de serem indissociáveis do caráter da avaliação crítica, similar a que um crítico profissional poderia empreender, exceto por não ter acesso prévio ao conteúdo pleno do filme antes do lançamento; as boas avaliações críticas dos críticos profissionais e dos consumidores afetam positivamente o tempo de duração dos filmes.

1.5 Materiais e Métodos

Esta seção discute a metodologia utilizada para a solução do problema proposto. Dividida em duas partes, a primeira discorre sobre as técnicas de análise afetiva que serão utilizadas e a segunda sobre as regressões que serão feitas para a análise da sobrevivência de um filme em cartaz.

Por se tratar de um problema interdisciplinar, a metodologia é composta pela utilização de ferramentas computacionais como forma de geração de variável não observável, as proxies do efeito boca a boca, a serem utilizadas nas análises de dados empregadas na forma de modelos de regressões paramétricos de sobrevivência.

Para conhecer os efeitos objetivados, é necessário controlar por um conjunto de características dos filmes, tais como gênero, diretor, custos, etc., bem como dos consumidores, por exemplo, renda. Estes e os demais dados utilizados no estudo foram minerados primordialmente a partir de dois sites, Rotten Tomatoes e Box-office Mojo, e foram trabalhados, relacionados e ordenados conforme a necessidade de estimações. Isto gerou um banco de dados exclusivo sobre o consumo de cinema nos Estados Unidos de

1980 a 2015, com dados sobre aproximadamente 13 mil filmes e suas informações financeiras semanais dispostas em 117 mil linhas. Soma-se ainda a este banco de dados 835 mil revisões críticas profissionais e mais de 4 milhões de revisões críticas dos consumidores. Os detalhes da extensa coleta e da abrangência e qualidade dos dados adquiridos estão apresentados no capítulo 4.

Desta forma, utilizaram-se técnicas de computação cognitiva em processamento de linguagem natural sobre os textos de revisões críticas dos consumidores, com o uso das ferramentas de análise de sentimentos e mineração de opinião, disponíveis pela versão de teste gratuita da ferramenta IBM Watson. Como resultado, foram obtidos alguns conjuntos de variáveis qualitativas dos sentimentos (positivo, negativo e neutro) descritas de forma quantitativa. Para cada revisão crítica, obteve-se também a presença das expressões de cinco emoções consideradas básicas. Ambas as informações, que quando agregadas, formam o que aqui chamaremos de computação afetiva (capítulo 2), foram usadas nas regressões dos modelos de sobrevivência para medir a influência do efeito do boca a boca sobre os pares.

Por fim, a metodologia empregada para atingir estes objetivos corresponde a um trabalho interdisciplinar, conforme Ventura e Lins (2016), na medida em que oferece técnicas computacionais para resolver questões de economia, aplicadas a bens de informação, cuja natureza é também computacional.

1.5.1 Materiais e Métodos em Computação Afetiva

O objetivo desta pesquisa é analisar o efeito das revisões críticas na duração da exibição dos filmes no cinema americano. Para obter uma medida do fenômeno do boca a boca, variável não observável e de extrema importância para influenciar o consumo, é utilizada a computação afetiva – que alia a mineração de emoção e a análise de sentimentos – aplicada às revisões críticas de consumidores disponíveis na comunidade de filmes RottenTomatoes, de forma a extrair a polaridade e as emoções expressadas nestas postagens.

Esta pesquisa não foca, porém, no desenvolvimento de métodos de processamento de linguagem natural. Para isso, fundamentada na literatura atual das ferramentas disponíveis para a realização da análise afetiva (Capítulo 2), é utilizada a compreensão de linguagem natural disponibilizada pela IBM, o Watson AlchemyLanguage.

De posse do conjunto de revisões críticas dos consumidores, detalhado no capítulo 4, foi criada uma aplicação empregando o sistema de compreensão de linguagem natural do IBM Watson para atender os objetivos da pesquisa. A aplicação desenvolvida utiliza a biblioteca de funções da *AlchemyLanguage API* do IBM Watson, disponibilizada pela IBM em sua plataforma de desenvolvimento, o IBM Bluemix. O serviço de compreensão de linguagem natural do IBM Watson e outras funcionalidades do IBM Bluemix de forma geral, são serviços pagos e o valor cobrado depende da quantidade de acessos e/ou requisições ao sistema. Porém, a IBM possui um canal que favorece pesquisadores, fornecendo contas de teste gratuitas, mas com número de requisições e período limitados. Nesta versão acadêmica é possível registrar usuários membros de universidades e submeter uma proposta de uso para que, se aprovado, o volume de chamadas aumente para 30.000 por dia, limitados a um mês.

Para utilizar a API foi criado um script na linguagem de programação Python que faz as chamadas aos métodos **docSentiment**, para sentimentos, e **docEmotion**, para emoções – o *script* está disponível na íntegra no anexo A. Com os dados preparados (vide capítulo 4), a chamada da API é feita passando o texto da revisão crítica a ser minerado e retorna com as informações de sentimento e emoção do texto passado. De forma muito simplificada, o quadro 1 apresenta o pseudocódigo do *script* implementado. Alguns tratamentos adicionais como teste das chaves (APIKeys), tratamento das respostas da API, tratamento dos dados de entrada para passagem à chamada, entre outros, foram omitidos do pseudocódigo, mas podem ser conferidos em sua forma original no anexo A, conforme dito acima.

Quadro 1 - Pseudo código script da utilização da API

```
\Início
Registra a APIKey

Lê arquivo de entrada (CSV Críticas dos Consumidores)
Cria arquivo de saída com os resultados (CSV Críticas Mineradas)
Cria arquivo de saída de erros (CSV Críticas com Erro)

Para cada linha do arquivo de entrada:
    Tente:
        Resposta = Chamada da API
    Exceção:
```

```

    CSV de erros = Resposta
    Continue
    CSV Criticas Mineradas = Linha de Entrada + Resposta
\Fim

```

Fonte: Thaís Luiza Donega e Souza (2017)

Seguindo a definição de Liu (2012), apresentada no capítulo 2, o objetivo do script é obter os retornos das análises cujas respostas compreendam às tuplas da equação 1 para sentimentos e às tuplas da equação 2 para emoção:

$$\mathbf{o}_j, (\mathbf{oo}_{Sentimento,ijkl}), \mathbf{t}_l \quad (1)$$

$$\mathbf{o}_j, (\mathbf{oo}_{Emoção,ijkl}), \mathbf{t}_l \quad (2)$$

Em que $\mathbf{oo}_{ijkl}$ é a pontuação dada para o objeto $\mathbf{o}_j$ no tempo estipulado $\mathbf{t}_l$ para cada atributo minerado. De forma a adaptar as equações anteriores na ferramenta escolhida e o tipo da estrutura de retorno desejado, tem-se como totalização dos dados o conjunto da equação 3, disposta por revisão crítica:

$$\mathbf{o}_j, (\mathbf{oo}_{SentType,ijkl}, \mathbf{oo}_{SentScore,ijkl}), \quad (3)$$

$$(\mathbf{oo}_{SAD,ijkl}, \mathbf{oo}_{FEAR,ijkl}, \mathbf{oo}_{DIS,ijkl}, \mathbf{oo}_{JOY,ijkl}, \mathbf{oo}_{ANG,ijkl}), \mathbf{t}_l$$

Sendo **SentType** e **SentScore** a pontuação da análise de sentimento e o rótulo do sentimento, respectivamente; **SAD** para *sadness*, **FEAR** para *fear*, **DIS** para *disgust*, **JOY** para *joy* e **ANG** para *angry*.

Tem-se assim as medidas de emoção e sentimento que foram propostas como formação de variável que será utilizada como proxy de boca a boca. A análise afetiva adquirida está, neste momento, registrada por revisão crítica, ou seja, o objeto $\mathbf{o}_j$ neste momento consiste na crítica textual do consumidor e o tempo $\mathbf{t}_l$ compreende a data da publicação da crítica. Para compor tal variável, no entanto, os resultados precisam ser agregados por filme, caracterizando apenas um registro por filme.

A sumarização desses resultados dá-se pela média aritmética das medidas obtidas, contemplando a equação 4:

$$o_j, \frac{\sum_1^n (oo_{SentScore,ijkl})}{n}, f_{oo_{SentType,ijkl}}, \frac{\sum_1^n (oo_{SAD,ijkl})}{n}, \frac{\sum_1^n (oo_{FEAR,ijkl})}{n}, \frac{\sum_1^n (oo_{DIS,ijkl})}{n}, \frac{\sum_1^n (oo_{JOY,ijkl})}{n}, \frac{\sum_1^n (oo_{ANG,ijkl})}{n}, t_l \quad (4)$$

Assim, reescrevendo a equação 4 em termos de dados, tem-se a equação 5, que apresenta as variáveis finais da computação afetiva por filme:

$$o_j, SentScore, SentType, Sad, Fear, Dis, Joy, Ang, t_l \quad (5)$$

Em que o_j é o filme avaliado pelos consumidores, a **SentScore** é a pontuação média do sentimento de maior frequência **SentType** atribuído ao filme e **Sad, Fear, Dis, Joy, Ang** são as médias aritméticas de cada uma das emoções. O tempo t_l foi separado por antes, durante e depois da data de estreia do filme para maior estratificação de dados.

Os materiais usados são o Python 2.7 utilizado na plataforma online 9cloud pela agilidade em troca de informações online com APIs e na plataforma PyCharm Community para tratamento de dados. API AlchemyLanguage IBM Watson; Biblioteca de dados Python Pandas; Banco de Dados Oracle na plataforma SQLDeveloper para demais relacionamentos e tratamentos.

1.5.2 Materiais e Métodos em Economia

Para avaliar o efeito das revisões críticas no mercado de cinema, foram empregadas diferentes especificações do modelo de sobrevivência, seguindo o modelo de regressão proposto por Reinstein e Snyder (2005), porém, tendo como variável de sucesso o tempo do filme em cartaz medido pelo número de semanas, substituindo então a variável de interesse de Reinstein e Snyder (2005), que seria o faturamento da bilheteria. A maioria dos estudos encontrados na literatura utiliza o faturamento da bilheteria – *box office* – como medida de sucesso e variável dependente. Entretanto, autores como McKenzie (2009), que estima um modelo de duração para o mercado de cinema australiano, e Walls (2010), que estuda o mercado de cinema de Hong Kong, utilizam modelos de duração para filmes em sala de cinema. Mais recentemente, Legoux *et al.* (2016) exploram o uso de modelo de sobrevivência para explicar a decisão de tempo de exibição, por parte dos

proprietários de cinemas, dos filmes que entram no mercado canadense. Assim, tendo como base estes autores, usamos modelos de sobrevivência para estimar o efeito das revisões críticas e do efeito boca a boca sobre o tempo de duração do mercado norte-americano entre 2004 e 2015.

O termo análise de sobrevivência ou modelo de duração é frequentemente usado para dados médicos, que utilizam tal modelo estatístico para mensurar a probabilidade da ocorrência de um dado evento em um período t , e, desta forma, estimar a probabilidade de que um paciente sobreviva após determinadas intervenções e introduzir tratamentos da forma que melhor aumente sua chance de sobrevivência. Atualmente, a análise de sobrevivência é amplamente aplicada em diferentes áreas. Alguns exemplos são a identificação do tempo que um indivíduo permanece desempregado, o tempo médio de sobrevivência de pequenas e médias empresas, o tempo em que um equipamento permanece em bom estado até precisar de manutenção, entre outros.

Para estimar o tempo em que um filme fica em exibição, serão utilizados modelos de sobrevivência, ou também chamados modelos de duração, como ferramenta econométrica (HOESMER e LEMESHOW, 1999). Trata-se de um conjunto de técnicas estatísticas que consiste na criação de um modelo que permite prever o tempo transcorrente do início identificado até um evento de interesse. O tempo estimado entre o início da análise até o evento de interesse é denominado tempo de falha. No caso aplicado aos modelos de sobrevivência de um filme em exibição nos cinemas, será considerado o tempo inicial como sendo a semana de lançamento do filme nos cinemas e o tempo de falha a data do encerramento da exibição. Considera-se aqui o início da exibição de um filme sua estreia nacional estadunidense, de acordo com os dados coletados, e o término da exibição indica que o filme não está mais em cartaz em nenhum cinema deste mesmo conjunto nacional.

Comum às características de um modelo de sobrevivência, o conceito de censura se caracteriza pela observação parcial da variável de resposta por qualquer fator que a retire do conjunto de dados em uma análise temporal (RODRÍGUEZ, 2012). Temos, no entanto, a ausência da censura no caso estudado, apesar de representar uma possibilidade teórica de uma proibição de exibição de um certo filme. Aplicando ao mercado de cinema americano, entretanto, podemos considerar início e fim da exibição de um filme em cartaz, tendo uma visão completa da variável de duração. Neste mercado, também inexistente uma política de um número máximo de semanas de um filme em cartaz, sendo o tempo de exibição definido conforme a demanda por ele.

Considerando T como o tempo até a ocorrência do evento de interesse, e sendo T uma variável não-negativa contínua que possui uma função de densidade de probabilidade $f(t)$ e função de densidade acumulada $F(t)$, o modelo de sobrevivência, denotado por $S(t)$, é a probabilidade do tempo de sobrevivência ser maior ou igual a um determinado tempo t , sendo definida dessa forma na versão em tempo contínuo, mas que pode ser avaliada igualmente em tempo discreto.

Conforme Wooldridge (2002), a duração de um evento $f(t)$ é supostamente distribuída com probabilidade contínua. A probabilidade de o tempo de duração ser menor que t é então:

$$F(t) = \text{Prob}(T \leq t) = \int_0^t f(s) ds \quad (6)$$

A função de sobrevivência, $S(t)$, é a probabilidade de a duração ser ao menos t :

$$S(t) = 1 - \text{Prob}(T \leq t) \quad (7)$$

A função Hazard ou taxa de risco representa a probabilidade da distribuição da duração, definida por $F(t) = \text{Prob}(T \leq t)$, que especifica a probabilidade de que a variável randômica T seja menor que algum valor de t (RODRÍGUEZ, 2012). Ao estudar dados de duração, é útil definir a função de sobrevivência como sendo $S(t) = 1 - F(t) = \text{Prob}(T \geq t)$. Assim, $\lambda(t)$ é a taxa na qual o evento ocorrerá no período t , uma vez que durou até t . A taxa de risco é, então, a probabilidade de a duração terminar após o período t , dado que ela tenha sobrevivido até o mesmo período t :

$$\lambda(t) = \frac{f(t)}{S(t)} \quad (8)$$

Essas definições apresentadas constituem-se nas abordagens não paramétricas que serão utilizadas, análises cujas abordagens não possuem variáveis explicativas de controle. Além delas, seguimos Allison (1984), que apresenta algumas dimensões que distinguem as abordagens realizadas pelos analistas dos dados de histórico de eventos. Destaca-se aqui, no entanto, o modelo paramétrico, em que a distribuição do tempo é regida por distribuição específica, como nos casos Exponencial e Weibull. Neste trabalho,

o tempo é uma variável contínua e discreta que consiste no número de semanas que um filme permaneceu em cartaz nos cinemas norte-americanos.

Assim, estimamos modelos não paramétricos, ou seja, sem variáveis de controle, e paramétricos, cujo interesse dá-se pela apresentação das importantes relações entre a variável resposta, representada por uma função dos tempos de sobrevivência, e as variáveis explicativas. Esta relação é demonstrada pelos valores dos coeficientes β 's estimados.

A forma funcional paramétrica associada às diferentes possibilidades de funções de risco $\lambda(t)$ permite a utilização das variáveis explicativas adicionais ao fenômeno, o que nos proporcionará avaliar/controlar o efeito das revisões críticas sobre a duração do filme em cartaz em número de semanas.

Modelos paramétricos podem assumir diferentes formas para a função de risco, conforme tabela abaixo. Dadas as implicações destes modelos sobre as probabilidades de duração de um filme em cartaz nos cinemas, estimamos as especificações da tabela 1.

Tabela 1 - Modelos Paramétricos de Sobrevivência e Risco

Distribuição	Métrica	Função de Sobrevivência	Parametrização	Parâmetros Ancilares
Weibull	AFT	$e^{(-\lambda_i t_i^p)}$	$\lambda_i = e^{-p x_i \beta}$	p
Exponencial	AFT	$e^{(-\lambda_i t_i)}$	$\lambda_i = e^{-x_i \beta}$	
Lognormal	AFT	$1 - \phi\left\{\frac{\ln(t_i) - u_i}{\sigma}\right\}$	$u_i = x_i \beta$	σ
Log-Logistic	AFT	$\{1 + (\lambda_i t_i)^{1/\gamma}\}^{-1}$	$\lambda_i = e^{-x_i \beta}$	γ
Gamma	AFT	$1 - I(\gamma, u)$	$u_i = x_i \beta$	γ, κ

Fonte: Kleinbaum e Mitchel (2012)

Contudo, foram apresentados os modelos de duração, em sua definição matemática, que serão utilizados para a estimação do tempo de sobrevivência de um filme em cartaz, objetivo principal deste trabalho. A seguir, será apresentado o modelo paramétrico proposto, no qual, a exemplo de McKenzie (2009), que também estima modelos de duração paramétricos, o evento de interesse corresponde ao momento em que o filme deixa de ser exibido e o tempo de ocorrência até o evento de interesse é representado pelo tempo do filme em exibição, medido em semanas.

Conforme dito nos objetivos, seguimos a proposta de Reinstein e Snyder (2005). Os autores objetivam mensurar o efeito das críticas de dois críticos profissionais considerados tops sobre o faturamento da bilheteria de filmes no mercado americano de cinema. Eles utilizam a polaridade das revisões críticas dos críticos como sinalização de qualidade do filme para mensurar o seu efeito (*thumbs up* e *thumbs down*) controlada pelo tempo em que a crítica ocorreu, se antes ou durante a semana do lançamento do filme ou depois da primeira semana do lançamento. A proxy é criada a partir da mesma informação, ou seja, da variável de revisão crítica é criada a proxy de efeito de boca a boca, gastos com propaganda e marketing.

O problema inerente em medir a influência de críticos profissionais na demanda é que os produtos que recebem críticas positivas muito provavelmente tendem a ter alta qualidade, e é difícil determinar quando a revisão crítica ou a qualidade em si é responsável pela alta demanda. Em outras palavras, o coeficiente da regressão da demanda sobre as revisões teria viés devido à omissão de variável, no caso, variável indicativa de qualidade. A princípio, o viés pode ser removido se houver o controle pela qualidade, mas a qualidade é difícil de medir para qualquer produto, especialmente para produtos cuja qualidade é incerta antes do seu consumo. Desta forma, Reinstein e Snyder (2005) argumentam poder controlar a endogeneidade potencial de variável omitida. Nota-se, portanto, que eles não dispunham de uma variável que pudesse controlar tais efeitos.

Assim, a exemplo de Reinstein e Snyder (2005), utilizamos a mesma estrutura de variáveis explicativas e, adicionalmente, propomos como proxy para os efeitos de boca a boca as variáveis quantitativas provenientes da computação afetiva realizada nas revisões críticas dos consumidores. Deste modo, temos três proxies a serem usadas em diferentes modelos: a nota do consumidor no momento da postagem da sua revisão crítica (score de 0 a 5 estrelas), a polaridade da análise de sentimento (positivo ou negativo) e a média das presenças das emoções da mineração de emoção (*anger*, *disgust*, *fear*, *joy* e *sadness*). Optou-se por duas especificações, de forma a aplicar a análise afetiva realizada, uma para os resultados da análise de sentimento e outra para os resultados da análise de emoção.

Para o controle de publicidade é utilizada a variável de quantidade de salas de cinema onde é estreado um determinado filme, pois filmes de maior orçamento de produção tendem a ter maiores gastos com publicidade, que por sua vez, investem em uma estreia mais abrangente em número de salas. Outro controle de publicidade que dispomos é o mês de lançamento do filme, utilizadas como dummies. Algumas literaturas

apontam que há uma diferença sazonal de como o orçamento em publicidade é investido, dando ênfase para os meses que antecedem as férias de verão e os meses próximos ao final do ano como meses em que os gastos com as campanhas publicitárias são mais generosos (McKENZIE, 2009).

As demais variáveis de controle foram escolhidas e incluídas a exemplo de Reinstein e Snyder (2005). São elas gênero, classificação indicativa – informação sobre faixa etária recomendada –, distribuidor, época do lançamento, entre outros. Contudo, algumas variáveis não entraram no nosso modelo por possuírem alta correlação com outras variáveis já incluídas.

Por exemplo, a variável *budget* foi desconsiderada pela alta correlação com o número de salas de cinema na primeira semana de exibição. Tal correlação deve-se ao fato de que, na primeira aposta dos distribuidores, dado que o custo de produção é relativamente alto, os distribuidores apostam em estreias de grande porte, refletindo os custos. A variável de revisões críticas dos críticos profissionais não tops também foi excluída do modelo por apresentar alta correlação com as avaliações médias dos críticos tops. Ainda, o volume de avaliações dos críticos tops também foi excluído por possuir alta correlação com o número de salas de cinema na primeira semana do filme, mais uma vez refletindo os gastos com publicidade.

Portanto, seguindo o modelo proposto por Reinstein e Snyder (2005), mas com adaptações, ou seja, a mudança da variável dependente - que no caso deles não se tratava de regressão de modelo de duração, mas sim de regressões lineares com faturamento usado como variável dependente - tem-se que S_i é o tempo em que um filme $i = 1, \dots, I$, fica em cartaz medido em número de semanas durante o período T_i . M_i inclui as variáveis proxies do efeito boca-a-boca de outros consumidores que já tenham visto o filme. Já a variável C_i denota os sinais de qualidade, incluindo as revisões críticas de críticos profissionais antes, e antes e durante a primeira semana de exibição do filme.

A Matriz X_i inclui variáveis *dummies* de mês de lançamento para controlar o efeito de publicidade prévia, marketing, pois dependendo do período do ano, tanto o gasto com publicidade quanto o volume de lançamentos tendem a ser maiores antes das férias de verão, meio do ano, e no final do ano (REINSTEIN e SNYDER, 2005). Incluem-se também outros controles, como o número de teatros da primeira semana em que um filme é lançado, revelando a aposta inicial das empresas de distribuição e exibição e uma medida do custo do filme. Espera-se que, quanto maior o orçamento maior seja o sucesso, fenômeno *blockbuster*. Além disso, essa variável controla pela diferença de filmes com

grande abertura em relação a filmes considerados *cult* e que tendem a ser vistos por uma pequena parcela de pessoas, mas que apesar disso pode ficar em cartaz por várias semanas. Inclui ainda a informação sobre se o distribuidor do filme é grande (medindo seu poder de monopólio), *dummies* de gênero do filme (reduzimos os gêneros disponíveis no Rotten Tomatoes para ação, comédia, drama, romance & família, estrangeiro, horror, aventura & animação, thriller, crime & guerra & velho oeste e demais gêneros. Vide descrições no anexo B.

Supondo que esses sinais de qualidade, na forma de revisões críticas de consumidores e revisões críticas de profissionais, influenciem a demanda dos filmes, medida pelo tempo de duração nas salas de cinema, tem-se a seguinte equação:

$$\ln S_i = \alpha + \beta C_i + \delta M_i + \theta X_i + e_i \quad (9)$$

Sendo α , β , δ e θ vetores de coeficientes, e e_i um termo de erro aleatório para as diferentes distribuições estimadas.

O objetivo do estudo é estimar os coeficiente β e δ , que capturam o efeito das revisões críticas, tanto dos profissionais na performance da bilheteria do filme, quanto dos consumidores como proxies para o efeito boca a boca, respectivamente. Assim, M_i representa um problema resolvido do ponto de vista da proposta de Reinstein e Snyder (2005), que discutem o efeito do boca a boca sobre pares como uma variável não observada, representando um problema econométrico de omissão de variável relevante. Buscou-se, então, com este estudo oferecer uma proxy para esta variável não observada e evitar o viés nas estimativas obtidas.

Note que a variável boca a boca não é diretamente observável e que a ausência de alguma *proxy* que pudesse gerar pelo menos alguma medida indireta deste efeito se constituía numa lacuna da literatura. Com a ferramenta computacional de análise de sentimentos, bem como o advento de comunicação das redes sociais com resenhas críticas de consumidores comuns sobre a qualidade de produtos e de filmes, em nosso caso, pode-se oferecer uma variável proxy para o boca-boca e resolver de maneira mais direta o problema apresentado por estes autores. Além de evitar o problema de viés de estimativas, conforme Reinstein e Snyder (2005), estas variáveis propostas permitem a obtenção da estimativa direta deste efeito, gerando uma medida de seu impacto sobre o tempo de duração dos filmes nas salas de cinema.

Estimamos o modelo (9) no contexto do modelo de sobrevivência-duração descrito acima sob diferentes especificações (tabela 1), incluindo a média das revisões críticas de críticos profissionais antes da estreia do filme nas salas de cinema e durante o final de semana de estreia e as médias pertinentes às revisões críticas dos consumidores ao longo do tempo de exibição até um ano após o lançamento do filme. Especificamente, foi usado como proxy de M_i as notas das revisões críticas dos consumidores disponíveis no site de revisões críticas referente ao primeiro ano de estreia do filme nos cinemas. Além da nota média das revisões incluídas pelos consumidores (scores) – de 0 a 5 – na hora de fazer a resenha, também serão utilizadas as cinco médias de emoções mineradas nestas revisões: *fear*, *angry*, *disgust*, *joy* e *sadness*; e a polaridade dos sentimentos obtidos na análise afetiva: positivo e negativo.

Como material, é utilizado o software de análises estatísticas Stata e suas funcionalidades de modelos de sobrevivência.

1.6 Organização do trabalho

Esta proposta de trabalho está estruturada em 6 capítulos. Capítulo 1, este, considera a introdução, objetivos, justificativa geral e as metodologias propostas para alcance dos objetivos. O capítulo 2 dispõe sobre a computação afetiva, englobando os métodos utilizados de mineração de emoção e análise de sentimentos, bem como a ferramenta computacional escolhida, o IBM Watson, e uma breve descrição das mesmas, também fundamentando as escolhas no estado da arte do tema. O capítulo 3 faz a apresentação do problema formalizado em economia, o estado da arte e o contexto de cinema, fundamentando o objetivo desta pesquisa na literatura correspondente. No capítulo 4, tem-se a apresentação dos dados que serão utilizados, o método de coleta de dados, as organizações tabulares e o banco de dados resultante. O capítulo 5 apresenta os resultados obtidos tanto na computação afetiva quanto os resultados paramétricos e não paramétricos dos modelos econométricos estimados. Por fim, o capítulo 6 conclui e faz considerações finais sobre este trabalho acerca dos resultados obtidos, contribuições e trabalhos futuros.

2 Computação Afetiva

Computação afetiva é o estudo e o desenvolvimento de sistemas e equipamentos que podem reconhecer, interpretar, processar e simular afetos humanos (PICARD, 1995). Trata-se de uma área interdisciplinar que abrange ciência da computação, psicologia e ciência cognitiva (TAO e TAN, 2005). Embora as origens do campo possam ser rastreadas desde as primeiras pesquisas filosóficas sobre a emoção (JAMES, 1884), o ramo mais moderno da ciência da computação se originou com o trabalho de Rosalind Picard em 1995.

Inspirado pela conexão inseparável entre emoções e cognição, a área de Computação Afetiva (*Affective Computing*) pretende preencher a lacuna entre a comunicação humana altamente emocional e a comunicação computacional através de sistemas que reconheçam e respondam aos estados afetivos (CALVO et al, 2010). Afeto, na literatura de psicologia, tem sido pensado como o guarda-chuva do termo; tal termo agrupa as emoções, sentimentos e sentidos (MUNEZERO et al., 2014).

Picard (1995) propõe que a computação afetiva seja capaz de (1) reconhecer a afetividade humana, (2) expressar e/ou reproduzir afetividade humana e (3) realmente sentir sentimentos humanos. Neste trabalho, será utilizada a primeira fase da proposta de Picard (1995) ao reconhecer emoções e sentimentos nas revisões críticas de consumidores.

O reconhecimento da afetividade humana em textos pode ser realizado de diferentes formas, como a análise de sentimento e a mineração de emoção propostos aqui. Trata-se de um problema de classificação, resolvido com Processamento de Linguagem Natural (PLN), com o qual é possível criar classificações categóricas, como as emoções e sentimentos, ou de dimensionalidade, como valência, excitação e dominância. A valência (ou polaridade) é a classificação utilizada na análise de sentimentos, enquanto que as três características juntas, valência, excitação e dominância, podem compor uma classificação que tende ao afeto vivenciado pelo usuário ou a emoção intendida na expressão.

A teoria da computação afetiva sobrepõe-se à área de análise de sentimentos, uma vez que pretende identificar também as emoções expressadas além da sua polaridade.

Propõe-se aqui uma análise afetiva que abrange a análise de sentimento, no sentido de detectar a valência das revisões críticas dos consumidores, bem como a mineração de emoção, que objetiva a classificação das mesmas revisões críticas quanto às emoções

nelas expressadas. Para elucidar melhor as diferenças de ambos os termos, uma breve explicação será feita com base no trabalho de Munezero et al (2014).

Afeto, sentir-se, emoção, sentimento e opinião (*Affect, feeling, emotion, sentiment e opinion*) são termos relacionados à subjetividade humana. Atualmente, os esforços de detecção em textos subjetivos são limitados em suas habilidades. Por exemplo, muito dos estudos de análise de sentimento e de mineração de opinião tem focado em detectar a polaridade do texto (positivo, negativo ou neutro). No entanto, sentimentos e opiniões são mais complexos que apenas suas polaridades. Além disso, termos subjetivos indicando afeto, sentimentos, emoções e opiniões em texto são frequentemente entendidos como similares e usados com permutabilidade em pesquisas de Processamento de Linguagem Natural sem uma diferenciação suficiente entre eles.

Emoções, segundo Kleinginna e Kleinginna (1981), são um complexo conjunto de interações entre fatores subjetivos e objetivos, mediados por sistemas neurais e hormonais. Fehr e Russell (1984) afirmam que, embora ainda não exista um consenso sobre sua definição, pode-se dizer que emoções são estados mentais e psicológicos associados a uma grande variedade de sentimentos, pensamentos e comportamentos. Uma descrição mais detalhada sobre emoção será apresentada na seção 2.3.

Sentimentos, por outro lado, são, segundo Cattell (2006), uma disposição neuropsíquica importante adquirida e relativamente permanente para reagir emocional, cognitiva e conativamente acerca de determinado objeto ou situação de certa forma estável, com consciência do objeto e da maneira de reagir. Ainda, Cattell (2006) dispõe que a função dos sentimentos é acompanhada no indivíduo pelo interesse e sentido, por sua vez, dos 'valores' no objeto. Novamente, este tópico será melhor discutido na seção 2.4 deste trabalho.

A partir da teoria de Picard (1995), muitas pesquisas surgiram no âmbito do reconhecimento de afeto e emoções humanas através de expressões faciais, postura corporal, gestos, comportamento humano, discurso e texto. Inicialmente com a intenção de replicar o comportamento humano no sentido afetivo, a área foi ganhando abrangência em suas aplicações.

O avanço tecnológico e a Web 2.0 trouxeram uma infinidade de dados disponíveis online para que esse reconhecimento afetivo seja aplicado. Qualquer pessoa pode expressar suas opiniões sobre qualquer assunto na web, o que transforma essa informação em uma fonte de dado útil para as diversas aplicações da computação afetiva como satisfação do consumidor e sinalização de qualidade do produto. Essa disponibilidade da

sinalização da qualidade de um bem diminui a assimetria de informação sobre ele e gera efeito de boca-a-boca, por conectar os pares (Capítulo 3).

Na linguística computacional, a detecção automática do afeto em texto tem se tornado cada vez mais importante pela perspectiva da aplicabilidade (MOHAMMAD, 2015). As áreas de aplicação são diversas, tais como aprendizado, interação humano-computador, recuperação de informação perceptiva, artes criativas e entretenimento, saúde humana e inteligência de máquina (PICARD, 1997). Outras áreas de aplicação seriam (MOHAMMAD, 2015):

- Saúde Pública: identificar depressão, casos de *cyberbullying*, atributos de saúde; rastrear bem-estar.
- Política: sentimentos públicos que envolvam políticos; eventos eleitorais nacionais e internacionais; questões controversas, como referendos; detectar polarização de votos, eleição de pré-candidatos e de candidatos.
- Gerenciamento de marca, de relacionamento com cliente e de mercado de ações: rastrear a imagem da marca e de seus produtos; respostas dos consumidores; sistemas automáticos de interação empresa x consumidor.
- Educação: desenvolvimento de sistemas automáticos para tutoria e avaliação de estudantes que considere as emoções dos alunos.
- Detectar traços de personalidade: detectar comportamentos extrovertidos, introvertidos, narcisistas, entre outros, e auxiliar estratégias de apoio.
- Entender diferenças entre sexos.
- Análise literária: analisar o fluxo de emoções em romances, peças, roteiros de filmes; detector padrões de sentimentos comuns em uma vasta coleção de texto e rastrear as emoções expressadas por seus personagens; sistemas de composição musical baseadas em emoções de um texto;
- Visualização de emoções: promover a interatividade entre os resultados emocionais de uma análise e o usuário, facilitando novos insights e gerando novas ideias para estratégias e tomadas de decisão em diferentes áreas e mercados.
- Sistemas de recomendação e recuperação de informação: utilizar queries emocionais para recuperação de informação em buscas diversas; classificação e personalização de conteúdo, como música, filmes e notícias para recomendação ao usuário.

2.1 Visão Geral

A opinião de terceiros consiste em uma fonte importante da informação e de influência para a maioria das pessoas durante o processo de tomada de decisão (PANG e LEE, 2008). Muito antes da *World Wide Web* se disseminar, tal fenômeno já era documentado e reportado, sendo evidenciado por guias de consumo, revistas e etc. A Internet tem sido agora o veículo de disseminação destas informações e influências, entre outras coisas, tornando possível encontrar opiniões e experiências de amigos, especialistas e de outros antes não disponíveis, os anônimos.

O processo de tomada de decisão das pessoas tende a ser afetado pelas opiniões formadas por líderes e pessoas comuns. Quando um indivíduo deseja comprar um produto, ele tipicamente começa procurando por revisões e opiniões expressadas por outras pessoas em diversos canais (RAVI e RAVI, 2015). Da mesma forma, emoções possuem um papel crucial na tomada de decisão, na percepção, na interação e inteligência humana. Este poder e importância de influenciar a emoção através da linguagem foi um princípio básico da Retórica de Aristóteles (PICARD, 1997).

Deve-se notar que tais informações/influências provêm de um conjunto grande de indivíduos e a sua mensuração ou contabilização do efeito sobre a decisão de consumo tende a ser bastante complexa e de difícil análise. Entretanto, o desenvolvimento de métodos computacionais recentes tem permitido que fontes de informações na forma de textos escritos possam ser trabalhadas e organizadas para obtenção de informações, entre elas, as técnicas de computação afetiva.

A análise de sentimento é uma das áreas mais movimentadas na ciência da computação. Mais recentemente, a mineração de emoção tem também ganhado espaço no mesmo âmbito de aplicação da análise de sentimentos. Atualmente, existe uma grande explosão de “sentimentos” em geral disponíveis nas mídias sociais (FELDMAN, 2013). Estes recortes de texto são uma mina de ouro para as companhias e para os indivíduos que querem monitorar suas reputações e conseguir *feedbacks* sobre seus produtos e ações, além da exploração das vendas destas informações organizadas.

No passado, para uma companhia conhecer as opiniões ou sentimentos do seu público consumidor sobre seus produtos e serviços, ela conduziria pesquisas e questionários de campo e focavam em seus grupos. No entanto, com a Web, especialmente a partir do crescimento explosivo do conteúdo gerado por usuários (*User*

Generated Content) na Internet nos últimos anos, o mundo tem sido transformado. A Web permite que todas as pessoas possam postar revisões de produtos e expressem seus pareceres sobre praticamente qualquer coisa em fóruns da internet, grupos de discussão, blogs e redes sociais. Este comportamento de boca-boca online representa uma nova e imensurável fonte de informações com muitas aplicações práticas. Atualmente, se alguém quer comprar um produto, não está mais limitado a pedir opinião às pessoas em sua volta, pois existem muitas fontes de avaliação e/ou opinião online para serem utilizadas. No entanto, encontrá-las e monitorá-las na Web ainda é uma tarefa complexa, pois existem um grande número de diferentes fontes, e cada uma delas pode conter imensos volumes de textos opinativos (LIU, 2010).

Em geral, opiniões podem ser expressas sobre qualquer coisa, como um produto ou um serviço, uma pessoa, uma empresa ou marca, um evento ou um tópico. Inúmeros fóruns, blogs, redes sociais, comércio eletrônico, web sites, jornais e outros recursos Web servem de plataformas para manifestar opiniões, que podem ser utilizadas para entendimento das percepções do público em geral e consumidores em eventos sociais, movimentos políticos, estratégias de companhias, campanhas de marketing, referências de produtos e monitoramento de reputação de empresas e marcas (SALEH et al., 2011).

Jurafsky e James (2000, p. 34) argumentam que “este é um momento excitante para o campo do PLN (...) e o crescimento da World-Wide-Web colocaram a aplicação do PLN nos holofotes e têm apontado uma infinidade de aplicações possíveis”. Pode-se então observar que há 16 anos o surgimento da Word Wide Web já contribuía para o interesse e o avanço na área, uma vez que o conteúdo online facilitou o uso das técnicas e ampliou suas aplicações.

A imensa quantidade de textos disponíveis online exige novos métodos que possam melhorar o acesso a este tipo de informação. Surgem então as subáreas da computação afetiva, como dito no início deste capítulo, que consiste no processamento automático de documentos para detectar a afetividade (opinião, sentimento, emoção) expressada dentro deles, como um único corpo de pesquisa. Como também explicado nesta introdução, os termos que denotam a extração e classificação de afetividade em textos são usados com certa displicência e permutabilidade.

Algumas pesquisas relatam que a opinião difere do sentimento no sentido de que este pode ser expresso sem que haja necessariamente um alvo a ser atingido (MUNEZERO et. al, 2014), enquanto a opinião é a expressão de um sentimento sobre um tópico ou uma característica, exigindo assim a racionalização do usuário que a expressa

(LIU, 2010). Neste trabalho, os termos mineração de opinião e análise de sentimento possuem exatamente o mesmo significado, uma vez que os sentimentos a serem extraídos das resenhas críticas são necessariamente sobre os filmes que estão sendo avaliados. Desta forma, a permutabilidade entre os termos referidos é possível e adequada.

Dada as informações apresentadas, pode-se inferir que a computação afetiva é uma área de pesquisa em expansão que abrange tanto a análise de sentimentos quanto a mineração de emoção, ambas técnicas a serem utilizadas neste trabalho. A aplicação aqui proposta é a de medir a correlação entre as emoções e os sentimentos dos consumidores de filmes em cinemas nos Estados Unidos, obtidas a partir das críticas de filmes feitas por consumidores na web – especificamente do site RottenTomatoes – e o tempo de duração destes filmes em cartaz, gerando assim uma estimativa de sucesso ou fracasso para os filmes estudados a fim de mensurar o efeito do boca a boca sobre o consumo de filmes no cinema americano.

Ambas as classificações, análise de sentimento e mineração de emoção, dar-se-ão em um banco de dados exclusivo, conforme apresentado no capítulo 4. Trata-se de um conjunto de opiniões sobre filmes proferidas por seus consumidores. Pretende-se, então, extrair a polaridade presente nas revisões críticas de forma a categorizá-las quanto à valência – positivo e negativo – e quanto às emoções – *joy, fear, anger, sadness e disgust*.

Para estas tarefas, muitas abordagens de processamento de linguagem natural podem ser aplicadas, sendo mais presentes na literatura as abordagens por dicionários léxicos de palavras e expressões e via aprendizado de máquina. Ambas as análises serão realizadas utilizando a ferramenta de computação cognitiva da IBM, o IBM Watson, e sua *feature* de processamento de linguagem natural, a AlchemyLanguage. A escolha e a forma de funcionamento estão detalhadas na seção 2.5.

Este capítulo disserta sobre a análise de sentimento e a mineração de opinião aplicada à revisão crítica dos consumidores sobre filmes. A seção 2.1 discorre rapidamente sobre a abrangência de informações na era da Web 2.0 e a importância desse avanço tecnológico para o processamento de linguagem natural (PLN). Na seção 2.2 um breve *overview* é feito sobre o PLN. Já a seção 2.3 revisa sobre a análise de sentimento, considerando sua origem, aplicações e estado da arte; a seção 2.4 relata a mineração de emoção, também partindo das definições, aplicações e estado da arte do tema. A seção 2.5 apresenta as ferramentas existentes para a realização de ambas as tarefas, bem como a ferramenta escolhida, o IBM Watson, juntamente com o detalhamento da metodologia utilizada. Finalmente, a seção 2.6 apresenta as considerações finais deste capítulo.

2.2 Processamento de Linguagem Natural

Apesar de mineração de emoção e análise de sentimento diferirem quanto às definições e objetivos, as técnicas utilizadas para suas aplicações são as mesmas: processamento de linguagem natural. Trata-se de tarefas de classificação textual, ou seja, classificar um documento de texto em suas categorias desejadas. Neste caso, sentimentos (positivo e negativo) e emoções (*anger*, *joy*, *sadness*, *disgust* e *fear*).

Esta seção discorre sobre a definição do processamento de linguagem natural, sua origem e aplicações além de apresentar alguns dos métodos mais comuns da subtarefa de Inteligência Artificial e da Mineração de Texto.

2.2.1 Definição e História

Processamento de Linguagem Natural (PLN) pode ser definido como o processamento automático ou semiautomático da linguagem humana. PLN é essencialmente multidisciplinar: está intimamente relacionado à linguística, ciência cognitiva, psicologia, filosofia e matemática. Relaciona-se com a teoria formal da linguagem, técnicas de compilação, aprendizado de máquina e interação humano-computador (JURAFSKY e JAMES, 2000).

O PLN consiste no desenvolvimento de modelos computacionais para a realização de tarefas que dependem de informações expressas em alguma língua natural, por exemplo tradução e interpretação de textos, busca de informações em documentos e interface homem-máquina (COVINGTON, 1994; RUSSELL e NORVIG, 1995). Conforme Covington *et. al* (1997) a pesquisa em PLN está voltada, essencialmente, a três aspectos da comunicação em língua natural:

- Som: fonologia
- Estrutura: morfologia e sintaxe
- Significado: semântica e pragmática

A fonologia está relacionada ao reconhecimento dos sons que compõem as palavras de uma língua. A morfologia reconhece as palavras em termos das unidades primitivas que a compõem. A sintaxe define a estrutura de uma frase, com base na forma como as

palavras se relacionam nela. A semântica associa significado a uma estrutura sintática, em termos dos significados das palavras que a compõem. Finalmente, a pragmática verifica se o significado associado a uma estrutura sintática é realmente o significado mais apropriado no contexto considerado (CONVIGTON et al, 1997). Tanto a análise de sentimentos quanto a mineração de emoção utilizam-se do recurso de significado do PLN.

Os modelos pioneiros desenvolvidos por Shannon (1948) e Chomsky (1956) levaram à área de teoria da linguagem formal, que usava álgebra e teoria de conjuntos para definir linguagem formal como uma sequência de símbolos. Isso incluía a gramática livre de contexto, primeiramente definida por Chomsky (1956) para as linguagens naturais.

No final da década de 1950 e início da década de 1960, processamento de linguagem e fala se dividiu claramente em dois paradigmas: simbólico e estocástico. O paradigma simbólico, por sua vez, foi dividido entre duas linhas de pesquisa, uma liderada por Chomsky sobre a teoria da linguagem formal e a outra seria a área de inteligência artificial. No verão de 1956 houve um *workshop* de dois meses, comandado por pesquisadores da área, chamada por eles de inteligência artificial, a qual já existia como teoria em outras aplicações. Nesse ponto, foram construídos os primeiros sistemas de compreensão da linguagem natural, sistemas simples que funcionavam em domínios únicos, principalmente por uma combinação de correspondência de padrões e busca de palavras-chave com heurísticas básicas para raciocínio e resposta a perguntas. No final dos anos 60, sistemas lógicos mais formais foram desenvolvidos (JURAFSKY E JAMES, 2000).

Jurafsky e James (2000) ainda alegam que os anos 1960 também presenciaram o desenvolvimento dos primeiros modelos psicológicos sérios e testáveis de processamento de linguagem humana baseada na gramática transformacional, bem como o primeiro dicionário online: *The Brown corpus of American English*. Durante a década de 1970, muitos programadores começaram a escrever ontologias conceituais, que estruturaram a informação do mundo real em dados compreensíveis por computadores.

O próximo período viu uma explosão na pesquisa de processamento de fala e língua e o desenvolvimento de um número de paradigmas de pesquisa que ainda dominam a área. A próxima década vivenciou o retorno de duas classes de modelos que haviam perdido popularidade no final dos anos 50 e início dos anos 60, parcialmente devido a argumentos teóricos contra eles, como a revisão influente de Chomsky sobre o comportamento verbal proposto por Skinner (Chomsky, 1959b).

Outra tendência foi, mais notavelmente, o surgimento de modelos probabilísticos em todo o processamento de fala e linguagem, influenciado fortemente pelo trabalho sobre o assunto desenvolvido no centro de pesquisa Thomas J. Watson, na IBM. Estes métodos probabilísticos e outras abordagens baseadas em dados se espalharam para tagueamento de partes do discurso, análise e anexação de ambiguidades, e abordagens connexionistas do reconhecimento de fala à semântica.

Este período também viu trabalhos consideráveis em geração de linguagem natural. Pelos últimos cinco anos do milênio, ficou claro que o campo havia vastamente mudado. Primeiro, modelos probabilísticos e baseados em dados tornaram-se um padrão em todos os processamentos de linguagem natural. Algoritmos para análise, tagueamento de partes do discurso, resoluções referentes e processamento do discurso começaram a incorporar probabilidades e empregaram metodologias de avaliação emprestadas do reconhecimento de fala e da recuperação de informação. Segundo, o aumento na velocidade e memória dos computadores permitiu exploração comercial de uma quantidade de subáreas do processamento de linguagem e fala, em particular, reconhecimento de fala e a verificação de ortografia e gramática. Finalmente, o surgimento da Web enfatizou a necessidade de sistemas de recuperação e extração de informação baseada em linguagem (JURAFSKY E JAMES, 2000).

2.2.2 Métodos

Coleta e Granularidade

Devido à ampla disponibilidade de diversos recursos online, a aquisição de dados é altamente subjetiva ao tipo da mídia, formato da informação suportado pela mídia e aos tipos das análises a serem realizadas. Alguns microblogs como Twitter¹, Facebook² e Sina-Weibo³ possuem uma API (*Application Programming Interface*) disponível para coletar os dados públicos dos seus sites. Outra maneira de coletar dados da Web é através de *Crawlers* e *Scraps*, mais trabalhoso do que com o uso da API, porém permite mais liberdade de escolha, formatação e volume de dados.

¹ <https://dev.twitter.com/overview/api>

² <https://developers.facebook.com/docs/apis-and-sdks>

³ http://open.weibo.com/wiki/API文档_V2/en

Para este trabalho, os dados foram coletados utilizando *Crawlers*, uma vez que as fontes destes dados não possuem aplicações do tipo API para auxiliar na coleta e que o método escolhido fornece maior flexibilidade para conteúdo, volume e formatação. A fase da coleta está documentada no capítulo 4 – Dados, bem como a escolha da formatação dos dados e, consequentemente, da sua granularidade.

O PLN pode ser aplicado em diferentes níveis de documentos a fim de se obter resultados distintos de classificação, conforme descrito por Liu (2012) – nível do documento, da sentença, entidade ou aspecto.

A Mineração de texto ao nível do documento objetiva-se em classificar um documento inteiro entre as possíveis categorias desejadas. A aplicabilidade deste nível é frequentemente focada e comumente encontrada no contexto da análise de revisões de produtos. Ao nível de sentença torna-se análogo ao anterior porque uma sentença pode ser considerada como um documento mais curto. No entanto, apresenta um pré-processamento adicional que consiste em quebrar o documento em sentenças separáveis. A análise a nível de entidade ou aspecto representa o maior grau de granularidade possível. Neste ponto a tarefa não é somente encontrar a polaridade da opinião, mas também seu alvo – entidade, aspecto ou ambos (LIU, 2012).

Tanto o nível de documento quanto o de sentença funcionam bem quando o texto examinado contém apenas uma entidade e/ou um aspecto, mas pode ocorrer falhas quando mais de um se faz presente (FELDMAN, 2013). A mineração no nível descrito objetiva solucionar este problema através da detecção de cada aspecto mencionado no texto e associá-lo a um afeto. Contudo, nota-se que as análises com maior nível de granularidade podem ser agregadas com a intenção de se obter maiores níveis. Por exemplo, uma mineração aplicada ao nível do aspecto pode ser simplesmente calculada como a somatória média para formar a classificação afetiva ao nível da sentença (WU, KITA e MATSUMOTO, 2014; FUGI e QUAN, 2012).

Generalizando, para documentos realmente grandes, o problema de indexação granular torna-se evidente. Para uma coleção de livros, usualmente seria uma má ideia usar o livro todo como um único documento. Podem-se separar os capítulos dos livros como sendo um documento, aumentando o nível de granularidade da análise. Também é possível que cada sentença dentro deste capítulo de um livro pertencente a uma coleção seja tratada como um único documento. Se a unidade ficar muito pequena, outro problema ocorre, é suposto que se percam passagens importantes dada a distribuição dos termos através dos pequenos documentos, enquanto que se as unidades forem muito grandes, a

tendência é obter combinações espúrias e tornar a informação relevante difícil de ser analisada (MANNING, RAGHAVAN e SCHÜTZE, 2008).

Para escolher a granularidade, é essencial um bom domínio da coleção de documentos que será usada, considerando suas características de escrita, tamanho e o objetivo do resultado esperado. Dito isso, a escolha neste trabalho é a de usar uma revisão crítica completa como sendo um único documento. Maiores detalhes sobre a estrutura das revisões críticas a serem utilizadas podem ser compreendidos na seção dados.

Pré - Tratamento

Após a coleta, os dados passam por uma etapa de pré-processamento com o objetivo de torná-los compreensíveis por máquinas. Dados não tratados adquiridos de diversas fontes frequentemente precisam ser pré-processados antes de serem lançados para análise completa. Como mostram Ravi e Ravi (2015), alguns dos passos de pré-processamentos mais utilizados são: tokenização, remoção de *stop words*, normalização, derivação e lematização (*stemming* e *lemmanization*), tagueamento de partes do discurso e representação e extração de características. A tokenização é usada para quebrar a sentença em palavras, frases, símbolos ou outros *tokens* importantes através da remoção da pontuação. *Stop words* são palavras que não contribuem, necessariamente, nas análises, portanto são cortadas nesta fase. Derivação e lematização são processos de transformação da palavra de forma a retorná-las para seu formato raiz. O tagueamento de partes do discurso é feito para reconhecer diferentes partes do discurso no texto, o que é essencial para o processamento da linguagem natural.

Devido ao espaçamento e o alto ruído nos textos, frequentemente faz-se necessário um nível extremo de extração de características, que é também um importante passo de pré-processamento (FIGUEIREDO et al, 2011). Além da extração de características, tem-se a seleção de características, outro passo crítico do pré-processamento para o sucesso do PLN (MEDHAT, 2014). As etapas serão apresentadas com mais detalhes a seguir, conforme explica Manning, Raghavan e Schütze (2008).

Tokenização: Dada uma sequência de caracteres e uma unidade de documento definida, a tokenização é a tarefa de cortá-las em partes, chamadas de *tokens*, talvez ao mesmo tempo em que elimina certos caracteres, como pontuação.

Entrada: Amigos, Romanos, Cowboys, me emprestem vossas orelhas

Saída: Amigos | Romanos | Cowboys | me | emprestem | vossas | orelhas

A tokenização envolve diversos tratamentos além da exclusão da pontuação e da separação de palavras. Por exemplo, e-mails, URLs, endereços de IP, códigos de rastreamento de pacotes, palavras hifenadas, contrações e espaços entre palavras devem ser tratados de modo a não perderem o sentido ou não se descaracterizarem. Esta tarefa, contudo, é altamente dependente da língua na qual é aplicada.

Remoção de Stop Words: *Stop Words* são palavras extremamente comuns na construção de frases e não são, necessariamente, importantes à mineração. A estratégia geral para determinar uma lista de *stop words* é ordená-las pelos termos através da sua frequência total, filtrá-las pelo seu conteúdo semântico relacionado ao domínio e excluí-las durante a indexação.

Exemplos: a | o | em | no | na | por | para | de | isto | aquilo | com | onde

Normalização: é o processo de transformar as variações dos *tokens* em uma representação única, ignorando possíveis diferenças superficiais dos caracteres de um *token*. A maneira mais comum de executar a normalização é criar uma classe de equivalências, que são geralmente nomeadas após um membro do conjunto de *tokens*. A normalização pode ocorrer na remoção da acentuação das palavras, por exemplo, *clichê* e *cliche* e *cliché*. Outro uso comum é transformar *tokens* que não são substantivos próprios em letras minúsculas, como acontece no início das frases.

Assim como a remoção de *stop words*, a técnica de normalização é também dependente da língua e varia nas suas regras algorítmicas, de forma a concordar com as regras linguísticas do idioma em que está sendo aplicada.

Derivação e Lematização (*Stemming* e *Lemmanization*): Por razões gramaticais, documentos usam diferentes formas de uma palavra, como “organizar”, “organiza” e “organizando”. Adicionalmente, há toda uma família de palavras derivadas relacionadas com significados similares, como “democracia”, “democrático” e “democratização”. O objetivo de ambas as técnicas é reduzir formas flexionais e formas derivadas relacionadas de uma palavra para uma forma comum: *sou | é | és | somos | sois | são → Ser*.

Entrada: Os carros dos garotos são de cores diferentes.

Input: The boy's cars are different colors.

Saída: O carro do garoto ser cor diferente.

Output: The boy car be differ color.

Tagueamento de Partes do Discurso: O tagueamento de partes do discurso considera que os adjetivos são indicativos importantes da subjetividade e opiniões. Desta forma, os adjetivos têm sido tratados como características especiais. Partes do Discurso (POS –

Parts of Speech) são as classes morfológicas das palavras que também podem ser usadas em complementação às palavras ou termos de opinião. Também se pode aplicar a dependência sintática entre as palavras, com o intuito de auxiliar na definição do alvo e a fonte do sentimento/emoção.

Palavras e frases opinativas são geralmente utilizadas para expressar sentimentos positivos ou negativos. Por exemplo, “bonito”, “maravilhoso”, “bom” e “surpreendente” são palavras de opiniões positivas, enquanto “ruim”, “pobre” e “terrível” são palavras de opiniões negativas. Embora muitas palavras que expressem opinião sejam adjetivos e advérbios, substantivos como “lixo” e “porcaria” e verbos “odiar” e “gostar” podem também indicar opiniões. Além das palavras individuais, há opiniões expressadas por frases e expressões idiomáticas.

Considera-se também a negação. Palavras que indiquem negação são importantes porque sua existência pode mudar a orientação da opinião. Por exemplo, a existência do verbo gostar indica uma orientação positiva, porém, se seguido de negação, como em “não gostei”, a orientação é alterada.

Extração de Características

De acordo com Pang e Lee (2002) e Liu (2012), algumas das técnicas de extração de características mais utilizadas são a extração das palavras que expressam afetividade como sendo presentes ou ausentes no documento, a frequência destes termos, partes do discurso e dependência sintática. Novas abordagens vêm sendo propostas, como a adição de pesos às frequências dos termos, a representação por matrizes e a mistura de técnicas de frequência de termos.

Termos e suas frequências são palavras individuais ou n-gramas de palavras e suas contagens de frequências. Em alguns casos, as posições das palavras são também consideradas. O esquema TF-IDF (termo frequente e frequência inversa do documento) da recuperação de informação podem também serem aplicados. Utilizados na tradicional mineração de texto, esses procedimentos se mostram eficazes também na classificação afetiva de documentos (CAMBRIA et al, 2013).

De uma maneira geral, a extração de características considera todo e cada par de termos ou *tokens* consecutivos como sendo um único termo de vocabulário. A união destes tokens é chamada de bigramas. Da mesma forma, um unigrama consiste em considerar apenas um *token* em uma posição do vetor do documento e n-gramas é o uso

de *n tokens* ou termos em uma mesma posição do vetor documento, considerando este agrupamento um único termo de vocabulário (MANNING, RAGHAVAN e SCHÜTZE, 2008).

Entrada: Amigos, Romanos, Cowboys

Bigrama: Amigos Romanos | Romanos Cowboys

Contudo, há vários métodos para representar o documento em forma de vetor. Além da representação por *n*-gramas, a representação por posição de index também é amplamente utilizada. Neste caso, para cada termo do vocabulário, suas posições são anotadas, sendo que cada posição é um índice de um *token* no documento. Cada posição deve conter também a frequência deste termo.

Ser, 178239 → [1, 6: (7, 18, 33, 72, 86, 231)]

Neste exemplo, o *token* “ser” aparece 178239 vezes na coleção de documentos, sendo esta sua frequência total. No documento 1 o *token* aparece 6 vezes – posições 7, 18, 33, 72, 86 e 231 – sendo 6 sua frequência no documento. A representação por frequência do termo pode ocorrer de forma absoluta ou de forma relativa, formando *n*-gramas.

É possível adicionar outras características baseadas em termos ao vetor de características. Posição refere-se a como uma posição de um *token* em uma unidade de texto pode afetar o próximo sentimento. Adicionalmente, pode-se considerar a presença de *n*-gramas (bigramas e trigramas) como características úteis. Alguns métodos também utilizam a distância entre termos. Análises textuais, de forma geral, utilizam informações de partes do discurso (*Parts of Speech* – POS), como substantivos, adjetivos, advérbios e verbos como uma forma básica de desambiguação. Certos adjetivos são bons indicadores de sentimentos e guiam a seleção de características para classificação (CAMBRIA et al., 2013).

Abordagem Lexical e Aprendizado de Máquina

Todas as definições descritas são tarefas básicas do Processamento de Linguagem Natural. São diversas as aplicações do PLN: verificação ortográfica, pesquisa em documentos de texto, reconhecimento de voz, processamento de páginas web, marcação de partes do discurso, tradução automática, agentes de diálogo de língua falada e escrita, recuperação de informação, análise cognitiva e, recentemente, computação afetiva (JURAFSKY e JAMES, 2000). Portanto, estas tarefas de pré-processamento são comuns

a todas as áreas citadas, incluindo as subáreas da computação afetiva utilizadas neste trabalho, análise de sentimentos e mineração de emoção.

A tarefa mais básica da análise de sentimento é a classificação da polaridade. Ela ocorre quando um documento de texto que expressa uma opinião em um único problema é classificado como um dos dois sentimentos opostos possíveis. Resenhas como “gostei” e “não gostei” são exemplos de classificação de polaridade (CAMBRIA et al, 2013). Da mesma forma, a mineração de emoção consiste em classificar documentos de textos, em algum nível, de acordo com as emoções expressadas neles (MOHAMMAD, 2015). A escolha das categorias emocionais é, de certa forma, arbitrária.

Portanto, as técnicas utilizadas para executar, extrair e classificar emoções ou sentimentos em texto são as mesmas. Ainda, como parte do conjunto de tarefas que compõem o processamento de linguagem natural, as etapas de pré-processamento, conforme já descrito neste capítulo, também são semelhantes entre a análise de sentimentos e a mineração de emoção.

De acordo com Kim *et al* (2015) classificar sentimentos de um conteúdo textual de forma automática é usualmente conduzido de duas formas, abordagem lexical e abordagem de aprendizado de máquina. Dadas estas duas abordagens bem estabelecidas na literatura para a realização do processo central de analisar afeto em texto, define-se que as abordagens lexicais usam os recursos linguísticos, chamados de dicionários de sentimento ou léxicos de sentimentos, como *SentiWordNet*, *SenticNet* e *OpinionFinder*, enquanto as abordagens de aprendizado de máquina geram um algoritmo de classificação através do aprendizado com um conjunto de características linguísticas e os aplica (LIM e KIM, 2014).

De forma detalhada, os trabalhos baseados em dicionários, método também conhecido por abordagem baseada em semântica, pretendem classificar o texto usando um conjunto de regras e heurísticas obtidas a partir do conhecimento da linguagem (VILARES *et al*, 2015; TURNEY, 2002; LU *et al*, 2010; EIRINAKI, 2012; MORRAMED, 2013; SAINI *et al*, 2014; DENG *et al*, 2015; AL-HAJJAR e SYED, 2015; CHUNG *et al*, 2015). Os passos comuns começam por marcar cada palavra e frase com a sua polaridade ou a emoção correspondente com a ajuda de um léxico e então, incorporam-se os modificadores da análise de sentimento, emoção e seus escopos (intensificadores e negadores). Finalmente, gerencia-se as cláusulas adversativas pelo entendimento de como elas afetam a classificação e como é refletida na pontuação final da avaliação afetiva. Alguns passos posteriores podem incluir visualização e sumarização

das classificações obtidas. A abordagem baseada em dicionário pode ser considerada mais simples, pois não requer tratamento de rotulação dos dados e nem sua separação em conjunto de treinamento e teste.

Contudo, os dicionários léxicos de emoção e de sentimento diferem quanto à rotulação. Os dicionários utilizados na mineração de emoção possuem suas classificações em emoções homologadas como humanas na teoria das emoções, enquanto que os dicionários de sentimentos têm anotações de palavras – adjetivos, advérbios, etc. – apenas como positivas ou negativas.

A abordagem baseada em aprendizado de máquina, também conhecida como métodos estatísticos para problemas de classificação em texto, consiste em algoritmos que podem aprender padrões subjacentes de um conjunto de dados de exemplo (REBOLLEDO *et al*, 2010), dados cuja classe ou rótulo são conhecidos, para então classificar novos dados não rotulados (MITCHELL, 1997). Usualmente, os passos na abordagem de aprendizado de máquina consistem na engenharia das características para representar o objeto cuja classe será analisada e utilizar esta representação como entrada para o algoritmo. Métodos existentes de aprendizado supervisionado podem ser aplicados à análise de sentimentos e à mineração de opinião, como os populares Naïve Bayes, Suport Vector Machine e Max Entropy (PANG e LEE, 2002, 2004 e 2008; SALEH *et al*, 2011; DANG *et al*, 2010; BAI, 2011; ZHANG *et al*, 2011; KIM *et al*, 2015; CHAFFAR e INKPEN, 2011; CALVO e SUGHWAN, 2013; BELLAGARDA, 2011; PORIA *et al*, 2014).

Há, também, as abordagens híbridas, que combinam ambas as técnicas de forma eficiente (XIA *et al*, 2011; BALAHUR *et al*, 2012; ABBASI, 2011; PRABOWO e THELWALL, 2009; LI e XU, 2013). Ainda, existe uma tendência de crescimento que usa ontologias para resolver o problema tanto da análise de sentimento quanto da mineração de emoção, a técnica é baseada em conceitos.

O método baseado em conceitos é relativamente novo e consiste no uso de ontologias para dar suporte às tarefas de mineração em texto. Uma ontologia é definida como um modelo que conceitua o conhecimento de um dado domínio de um modo que é compreendido tanto por humanos quanto por máquinas. Ontologias são geralmente apresentadas como grafos nem que os conceitos são mapeados nos nós ligados pelas relações (BALAZS e VELÁSQUEZ, 2016; MOHHAMAD e TURNEY, 2013; CAMBRIA e HUSSAIN, 2012).

2.3 Análise de Sentimento

Esta seção discorre sobre a definição da análise de sentimentos, o estado da arte da área e a metodologia adotada.

2.3.1 Definição

De acordo com Munezero et al (2014), análise de sentimento é a categorização do texto de acordo com sua relevância afetiva. Um exemplo de aplicação é a exploração da opinião para análise de mercado.

A análise de sentimento é uma área de pesquisa recente em mineração de textos que tem sido designada por diferentes termos, como *subjectivity analysis*, *sentiment analysis* or *sentiment orientation*. Existem muitas definições sobre cada uma. Pang and Lee (2008) capturaram diferentes definições sobre esses termos baseados em aplicações realizadas nesta área. Por exemplo, *subjectivity analysis* é descrita como o reconhecimento de linguagem orientada à opinião em ordem de distingui-la da linguagem objetiva, enquanto *sentiment analysis* classifica as revisões de acordo com sua polaridade (positiva ou negativa). De qualquer modo, todos estes termos referem-se ao mesmo campo de pesquisa. Ainda, a mineração de opinião pode ser elucidada como sendo basicamente a tarefa de classificar sentimentos de um texto quanto a sua polaridade (CAMBRIA *et al*, 2013; CHEN e ZIMBRA, 2010; PANG e LEE, 2008).

Segundo Chen e Zimbra (2010, p.74) “Mineração de opinião é uma subdisciplina da mineração de dados e da linguística computacional, e é referido como técnica computacional que extrai, classifica, interpreta e avalia as opiniões expressadas em diversas fontes online”. Análise de sentimento é frequentemente utilizada na mineração de opinião para identificar sentimentos, afetos, subjetividade e outros estados emocionais no conteúdo textual (PANG e LEE, 2008).

As comunidades acadêmicas e de pesquisa estão trabalhando rigorosamente na análise de sentimentos há pelo menos uma década e meia com diferentes abordagens e inúmeras aplicações práticas. Essa ferramenta ajuda a alcançar diferentes objetivos, como observar a opinião do público considerando um movimento político ou de inteligência de mercado (LI e LI, 2013), medir a satisfação do consumidor (KANG e PARK, 2013), prever vendas de cinema (RUI et al., 2013), analisar os sentimentos do boca-boca online

(*online word of mouth*) (CAMBRIA et al, 2013), entre outros. Este interesse se dá principalmente pelo fato de que as opiniões existentes em sites de e-commerce como Amazon, IMDb, epinions.com, etc. podem influenciar a decisão dos consumidores na compra de produtos ou serviços (RAVI e RAVI, 2015).

Segundo Liu (2010), existindo um documento opinativo d , que pode ser uma resenha de um produto, uma postagem em um fórum ou de um blog que avalia um conjunto de objetos, d consiste em uma sequência de sentenças $d = (s_1, s_2, \dots, s_m)$. Um objeto é uma entidade que pode ser um produto, pessoa, evento, organização ou tópico. Ele está associado ao par $o: (T, A)$, onde T é a hierarquia dos componentes e subcomponentes e A é o conjunto de atributos de o ; A opinião sobre uma característica f de um objeto o avaliado em d é um grupo de sentenças consecutivas em d que expressam uma opinião positiva, negativa ou neutra em f . É possível que uma sequência de sentenças em um documento opinativo expresse, coletivamente, uma opinião sobre um objeto ou sobre uma característica de um objeto. É possível também que uma única sentença expresse opiniões sobre mais de uma característica, sendo necessário reconhecê-las; O detentor da opinião h é a pessoa que emite a opinião, sendo a opinião sobre uma característica f uma visão, atitude, emoção ou avaliação orientada do emissor da opinião k sobre f ; A orientação de uma opinião é a sua qualificação quanto à positividade, negatividade ou neutralidade sobre uma característica f .

Deste modo, Liu (2010) define a opinião como 5 tuplas contendo o alvo da opinião (ou entidade) o , o atributo do alvo em que a avaliação é dirigida f , o sentimento (ou polaridade) contendo a opinião oo que pode ser negativa, positiva ou neutra, o detentor da opinião h e a data que a opinião foi emitida t . Tendo assim, formalmente, a opinião definida como:

$$o_j, f_{jk}, oo_{ijkl}, h_i, t_l \quad (10)$$

Tal que o_j seja a opinião no documento d , f_{jk} é o atributo de o_j ; h_i é o detentor da opinião, t_l é a data em que a opinião foi emitida e oo_{ijkl} é a polaridade da opinião sobre o atributo f_{jk} da entidade o_j dado pelo detentor da opinião h_i no tempo t_l em d .

Note que oo_{ijkl} foi descrita como positiva, negativa ou neutra, porém, poderia ser uma representação numérica, por exemplo, com -5 para muito negativo e 5 para muito positivo. Também, em casos de análise não requerer muitos níveis de detalhe, os atributos da entidade podem ser omitidos e denotados como Geral, ao invés de f_{jk} .

Deve ser evidenciado que as cinco peças da informação na quintupla precisam se corresponder, ou seja, a opinião oo_{ijkl} deve ser dada pelo detentor da opinião h_i sobre a característica f_{jk} do objeto o_j no tempo t_l . Esta exigência nos dá alguma dica do porquê de a análise de sentimentos ser um problema tão desafiador, tendo em vista que identificar cada pedaço dessa informação sozinha por si só já é bastante difícil, encontrar todas elas e combiná-las, então, é ainda mais difícil. Para dificultar ainda mais, uma sentença pode não conter uma menção explícita sobre alguns pedaços, mas de forma implícita através de pronomes, convenções da linguagem e contexto.

Portanto, o principal objetivo da mineração de opinião é encontrar as tuplas de opinião em um documento, uma coleção de documentos ou através de várias coleções de documentos (BALAZS e VELÁSQUEZ, 2016). Outros trabalhos definem mineração de opinião de outras formas, apresentando diferentes escopos e níveis de granularidade, no entanto, todas elas podem ser adaptadas ao modelo de opinião proposto por Liu (2010), conforme apresentado por Balazs e Velásquez (2016).

Adaptando a equação 10 ao contexto utilizado neste trabalho, tem-se a equação 11:

$$o_j, oo_{ijkl}, t_l \quad (11)$$

Na qual, o detentor da opinião h_i e a característica do objeto f_{jk} não participam, uma vez que estas informações não são pertinentes à pesquisa. O tempo t_l é dado de acordo com a data da publicação da revisão e pode ser classificado em anterior, durante ou pós estreia do filme em cartaz. O objeto o_j refere-se ao filme ao qual a revisão crítica pertence e a opinião oo_{ijkl} é a tupla de interesse a ser descoberta pela mineração de texto, especificamente, pela análise de sentimento aqui explicada. Maiores detalhes sobre a construção da tupla estão no capítulo 4.

2.3.2 Revisão Bibliográfica Análise de Sentimentos

Ravi e Ravi (2015), em seu *survey* de mineração de opinião e análise de sentimentos, analisaram 219 trabalhos da área e os classificaram em uma série de maneiras, como foco da aplicação, nível de análise e abordagem. Como mostram os autores, a aplicação da Análise de Sentimento (AS) se dá, em grande maioria, no nível do documento. Isso pode ser explicado primeiramente pelo contexto histórico do estudo do

tópico. Inicialmente, o interesse surgiu a esta granularidade de informação e, conseqüentemente, as aplicações eram estudadas apenas neste nível. Com o aprofundamento das pesquisas, diferentes níveis de análise começaram a surgir. Apesar de detalhadas e, talvez mais precisas, a mineração de opinião aplicada nos níveis de sentença e de característica tendem a serem mais trabalhosos, difíceis e de aplicação específica.

No mesmo estudo, Ravi e Ravi (2015) relatam ainda que do total de artigos revisados, 55 utilizam SVM, sendo a técnica de abordagem supervisionada mais popular na literatura, seguida por 28 trabalhos que utilizam o classificador Naïve Bayes, também uma técnica de aprendizado de máquina supervisionado. Nota-se ainda o aumento do volume dos trabalhos por período. O trabalho de Ravi e Ravi (2015, pg. 37) mostra que, de acordo com os trabalhos revisados, a década passada mantém 17% das publicações sobre o tópico, enquanto que a década atual apresenta 82,5% do total do volume analisado, demonstrando um aumento de produção acadêmica no assunto de mais de 400%.

Ainda sobre o trabalho de Ravi e Ravi (2015), as mais variadas aplicações de análise de sentimento incluem inteligência de negócio, decisões de compra, gerenciamento de opiniões de público, Web Marketing além de 30 trabalhos que são aplicados em *Movie Reviews*. Dos trabalhos de mineração de opinião e/ou análise de sentimentos em revisão de filmes, 14 utilizam de abordagens baseadas em dicionários e ontologias, 14 utilizam aprendizado de máquina como supervisionado e 2 apresentam abordagem híbridas, combinando o uso de dicionários ou ontologias e *machine learning*. Dos que usam aprendizado de máquina, 7 utilizam o classificador Naive Bayes, os demais variam entre outras técnicas como SVM, *Max Entropy*, *Decision Trees* e outros.

Pang e Lee (2002) são os pioneiros a aplicar a abordagem de aprendizado de máquina para classificação binária de sentimentos em revisões de filmes. Eles utilizaram *Naïve Bayes*, *Maximum Entropy* e *Support Vector Machine (SVM)*. Em seus muitos experimentos, com muitas variações de extração de características, obtiveram 78% de acurácia utilizando Naïve Bayes e unigramas de frequência. Boiy e Moens (2009) realizam a classificação de sentimentos multilíngue e multidomínio. Eles utilizam unigramas, negação, partes do discurso, componentes de palavras e verbos, entre outras técnicas para seus testes com os classificadores Multinomial Naïve Bayes, Maximum Entropy e SVM. O melhor resultado obtido foi de 87.40% de acurácia utilizando MNB em corpus (Pang e Lee, 2004) de língua inglesa.

Moraes et al. (2013) comparam SVM e NB com a abordagem baseada em Redes Neurais (NN) para classificação de sentimentos. Os experimentos foram executados em conjuntos balanceados e não balanceados. Quatro composições de dados foram selecionadas para os experimentos, incluindo o *dataset* de filmes de Pang e Lee (2004) e revisões de produtos. Para o conjunto de dados desbalanceado, a performance do NB não foi prejudicada, enquanto que a do SVM e a de NN foram. Os autores alegam que estes classificadores podem ser testados com outros métodos de extração de características. Xia et al. (2011) propõem conjunto de características ensemble e aprendizado de máquina para classificação de sentimentos. Dois tipos de características (partes do discurso e relação entre palavras) foram utilizados juntamente com NB, ME e SVM. Os resultados mostram que a característica baseada na relação ponderada entre as palavras atingiu acurácia de 87.7% para o *dataset* de Pang e Lee (2004).

Weichselbraun et al. (2013) criam dicionários multidomínios contextualizados, que podem ser empregados em uma grande variedade de aplicações de mineração de opinião e suporte às decisões. Estes dicionários ontológicos foram criados em três etapas: detecção de termos ambíguos a partir de um corpus existente; cálculo da pontuação de sentimento baseado nas probabilidades da co-ocorrência de termos contextuais e; análise de sentimento através da combinação das polaridades de termos ambíguos e não ambíguos. A avaliação foi executada em diversos conjuntos de dados, incluindo o *dataset* de filmes de Pang e Lee (2004) e revisões de produtos. Utilizando Naïve Bayes como classificador, a abordagem alcançou a acurácia de 79% para os dados de filmes. Yu *et al.* (2013) investigam o efeito da mídia social e convencional, sua importância relativa e sua inter-relação com as performances das firmas. Para os experimentos foram utilizados dados do mercado financeiro. O classificador Naïve Bayes foi treinado no *dataset* de Pang e Lee (2004) para classificação binária de sentimentos e atingiu 86% de F1. Eles concluem que, em média, mídias sociais têm maior impacto sobre as ações de uma companhia do que a mídia convencional, porém ambas as mídias estão inter-relacionadas.

Pang e Lee (2004) utilizam proximidade física entre itens a serem classificados, em que a classificação da base de corte foi aplicada para classificação da subjetividade. O detector de subjetividade baseado em corte determina o status da subjetividade para todas as sentenças do documento utilizando informação de pertinência por item e por par. O banco de dados preparado para estes experimentos foi coletado do rottentomatoes.com para revisões de filmes (5000 revisões) para a análise subjetiva e os resumos (5000)

coletados do site imdb.com para as sentenças objetivas. Os testes com NB e SVM atingiram a acurácia de 92% e 90%, respectivamente, para classificação da subjetividade.

Conforme o objetivo desta revisão, todos os trabalhos aplicam a mineração de opinião de forma a classificar a polaridade em algum *dataset* específico. No entanto, os objetivos destas aplicações, apesar de variados, mostram semelhança quanto à avaliação dos classificadores e de múltiplos métodos combinatórios de extração de características, pré-processamento, técnicas híbridas e ensembles. Também é possível identificar algumas aplicações, sendo a mineração de opinião a forma de obter alguma outra informação, como o sentimento sobre as opiniões do mercado de café, predição de eleições, saúde e sistemas de avaliação de revisões. Nota-se, porém, que não há nenhum trabalho semelhante a este proposto, que tem por objetivo utilizar a polaridade obtida através da análise de sentimento para identificar possíveis efeitos no tempo de exibição dos filmes do mercado de cinema americano, representando uma lacuna no espaço e tempo analisado pela revisão da literatura.

2.4 Mineração de Emoção

Esta seção apresenta a definição da mineração de opinião, suas aplicações e estado da arte e a metodologia escolhida para esta tarefa.

2.4.1 Definição

A teoria da emoção é confusa e não está bem definida. As pesquisas se iniciam de longa data e não há um consenso geral a respeito, resultando em mais de 180 teorias sobre o que é emoção (BAGOZZI et. al, 1999). As teorias de emoção não compreendem no foco deste trabalho, portanto, apenas uma breve elucidação sobre o tema será feita, com a intenção de embasamento teórico nas escolhas realizadas para cumprir o objetivo proposto.

Pioneiros, Charles Darwin (DARWIN, 2002) e Willian James (JAMES, 1884) propuseram as primeiras teorias da emoção, que continuam a influenciar os pensadores atuais. As pesquisas em emoção eram conduzidas pelos filósofos e psicologistas cujos trabalhos eram baseados num pequeno conjunto de ‘teorias da emoção’.

Pesquisas mais recentes, da década de 80, como a de Turkle (1984), começaram a especular sobre como os computadores poderiam ser usados para estudar emoções. Programas de pesquisas sistemáticos começaram a surgir no início dos anos 90. Poucos anos depois, Rosalin Picard lançou seu livro *Affective Computing*, provocando uma onda de interesse entre cientistas da computação e engenheiros, que buscavam maneiras de melhorar a interação humano-computador por coordenar emoção e cognição (CALVO e D'MELLO, 2010).

Detectar afeto é, entretanto, um problema desafiador, pois emoções são construções com fronteiras *fuzzy* e com diferenças individuais substanciais pela expressão e experiência. Em 2009, duas conferências foram dedicadas à computação afetiva: a International Conference on Affective Computing and Intelligent Interaction e a Artificial Intelligence in Education, sendo consideradas como início das pesquisas na área (CALVO e D'MELLO, 2010).

Segundo Calvo e D'Mello (2010), existe um número pequeno de emoções básicas ou primárias – todas as outras emoções são misturas ou estados derivados, ou seja, elas ocorrem através de misturas ou composições das emoções primárias. Emoções primárias são construções hipotéticas ou estados idealizados cujas propriedades e características podem somente serem inferidas de vários tipos de evidências e podem ser conceituadas em termos de pares de polaridades opostas. Segundo Plutchik (1962), emoções básicas são caracterizadas por sinais distintivos universais, presença em outros primatas, psicologia distinta, distinção universal em eventos antecedentes, coerência entre resposta emocional, início rápido, curta duração e avaliação automática e espontânea.

Ekman (1992) propõe como conjunto de emoções básicas *anger, fear, enjoyment, sadness, disgust* e *surprise*. O autor se baseia em evidência robusta, consistente de expressões faciais universais para as 6 emoções. No entanto, as definições sugeridas por Ekman (1992) se aplicam quanto à universalidade de expressões faciais, o que faz sentido quando a computação afetiva e a mineração de emoção eram reduzidas às pesquisas nestas mídias, mas não se aplicam, necessariamente, para a universalidade em termos de texto.

Plutchik (1980, p. 7) baseado nos estudos de Darwin elege como conjunto de emoções básicas o que considera como “comportamentos que promovem a sobrevivência da espécie”: *Fear, anger, joy, sadness, acceptance, disgust, expectation* e *surprise*. Segundo o autor é possível identificar outras emoções através da mistura das emoções básicas, ilustradas na figura 1.

Figura 1 - Roda dos sentimentos proposta por Plutchik

Fonte: Plutchik (2001, p. 349)

Muitos outros conjuntos de emoções consideradas básicas foram propostos na literatura, conforme mostra a tabela 2.

Tabela 2 - Autores e as emoções básicas propostas

Autor	Emoção
Plutchik (1980)	Acceptance, anger , anticipation, disgust , joy , fear , sadness , surprise
Ekman, Friesen e Ellsworth (1982)	Anger , disgust , fear , joy , sadness , surprise
Izard (1971)	Anger , contempt, disgust , distress, fear , guilt, interest, joy , shame, surprise
James (1884)	Fear , grief, love, rage
Tomkins (1984)	Anger , disgust , anxiety, happiness, sadness

Fonte: Ortony e Turner (1990, p. 2)

Muita atenção foi dada à mineração de emoção em faces e gestos. Com o avanço tecnológico, as técnicas de processamento de linguagem natural e a imensa quantidade de textos na web, a mineração em texto foi ganhando espaço. Muitas técnicas foram criadas e aprimoradas para esta tarefa. Por também se tratar de um problema de classificação e PLN, os primeiros métodos – e possivelmente os mais populares – foram os dicionários

léxicos de palavras afetivas. Da mesma forma que na análise de sentimento, outras técnicas são as baseadas em aprendizado de máquina.

Para a abordagem lexical, muitos dicionários e ontologias foram criadas para auxiliar a tarefa do processamento de linguagem natural e melhorar os resultados. Alguns são: Linguistic Inquiry and Word Count (LIWC); The Affective Norm for English Words (ANEWs); Affective Norms for English Text (ANET); e suas evoluções: SentiWord-net, Hourglass of emotions, entre outros.

Contudo, ainda há uma discussão por parte da comunidade acadêmica de que estes dicionários seriam poucos e com muitas limitações, como língua e cultura. Algumas teorias julgam que a expressão de emoção é diferente em distintos lugares, apesar de existirem em qualquer contexto humano. A divergência cultural e linguística molda a forma de expressão da emoção, como a palavra “saúde”, que expressa um sentimento só conhecido na língua portuguesa e a palavra “*blues*” que pode significar triste para algumas culturas e países, mas que tem conotação positiva em outros (MUNEZERO et al, 2014).

Adicionalmente, do ponto de vista construtivista social as emoções são construções culturais, dependentes de fatores como regras sociais, práticas e valores. Em outras palavras, análises sociais e consideração e interpretação de funções culturais podem requerer entendimento emocional. Isso não necessariamente contradiz, tampouco complementa a perspectiva da universalidade proposta pelos conjuntos de emoções básicas (ALM, 2008).

Dada as características culturais, linguísticas e contextuais apresentadas, a escolha do método do PLN se torna ainda mais importante. Essa escolha, o IBM Watson como técnica de mineração de emoção, é reforçada baseada nessas características, uma vez que o processamento *Deep Learning* (IBM, 2015a), o ensemble de algoritmos de aprendizado de máquina e a inteligência artificial cognitiva da compreensão de linguagem natural de Watson (IBM Watson, 2015a) são capazes de considerar tais fatores. Além disso, a mineração de emoção considera somente os textos escritos em inglês, eliminando incompreensões naturais de significados diferentes entre idiomas.

Na maioria dos sistemas de avaliação online, como no caso das revisões críticas de filmes, um mecanismo de pontuação é agregado permitindo que o usuário classifique, dentro de um range de possibilidades, sua positividade ou negatividade sobre determinado assunto. Este esquema de pontuação pode ser feito por nota, estrelas, *likes*, etc. Essa informação agregada facilita – como rótulo – ou até mesmo inutiliza o trabalho da análise

de sentimento, uma vez que a polaridade pode ser extraída diretamente deste mecanismo e não necessariamente utilizar-se de processamento de linguagem natural.

A mineração de emoção, por outro lado, traz uma informação até então escondida nas revisões críticas, que é o estado afetivo do usuário no momento em que ele profere a crítica, não havendo rótulos prontos para aquela opinião ou expressão realizada. Analisar emoções de usuários é útil para entender melhor as perspectivas e preferências dos consumidores individualmente e, portanto, facilitar para que as empresas forneçam recomendações mais personalizadas. A mineração de emoção é aqui considerada como característica enriquecedora desta pesquisa.

Essa informação é tão importante que alguns sistemas, como as mídias sociais, já têm seu rótulo de emoções agregado ao espaço das publicações, como forma de obter dados mais consistentes e ricos sobre o comportamento dos seus usuários. A atualização do Facebook⁴ em 2016 trouxe botões de reações afetivas agregados ao já existente *like*, sendo então possível que os usuários agora expressem *love*, *joy*, *surprise*, *sad* e *angry* sobre uma determinada publicação. Ainda, a mesma rede, permite que o usuário marque *tags* de sentimentos durante suas publicações, como “*feeling thankful*”⁵, outra forma de gerar e armazenar dados rotulados.

Da mesma forma que na análise de sentimento, o que se pretende aqui é extrair a afetividade das revisões críticas de forma a se obter suas emoções. A estrutura de mineração de emoção pode ser representada pela equação 12, adaptada da definição da análise de sentimento proposta por Liu (2010).

$$o_j, oo_{ijkl}, t_l \quad (12)$$

Uma vez que as emoções a serem classificadas – *sadness*, *fear*, *disgust*, *joy* e *angry* – são consideradas básicas por conforme mostrado na tabela 2, tem-se a equação 13:

$$o_j, (oo_{SAD,ijkl}, oo_{FEAR,ijkl}, oo_{DIS,ijkl}, oo_{JOY,ijkl}, oo_{ANG,ijkl}), t_l \quad (13)$$

⁴ <http://www.usatoday.com/story/tech/news/2016/02/24/facebook-reactions-launch/80803468/>

⁵ <http://www.telegraph.co.uk/technology/facebook/9981765/Facebook-wants-to-know-how-youre-feeling.html>

Em que oo_{ijkl} é a pontuação da emoção dada para o objeto o_j no tempo estipulado t_l para cada emoção minerada: **SAD** para sadness, **FEAR** para fear, **DIS** para disgust, **JOY** para joy e **ANG** para angry.

2.4.2 Revisão Bibliográfica Mineração de Emoção

Entre os trabalhos analisados foram identificados diferentes objetivos e aplicações. Focando no método de mineração, SAINI *et al.* (2014) propõem um modelo de mineração de emoção independente de domínio, o EmoXtract, que extrai e classifica emoções a partir de dados não estruturados. Eles usam o dicionário léxico WordNet para criação de sinônimos e para palavras derivadas, o dicionário léxico de emoção WordNet-Affect e regras heurísticas para estabelecer uma relação ponderada entre cada palavra a uma determinada emoção primária. Utilizam manchetes, contos de fada, romances e peças de teatro e os classificam a nível de sentença, utilizando para isso as 6 emoções básicas propostas por Ekman *et al* (1982). Com os mesmos domínios como conjunto de dados e aprendizado de máquina SVM, Chaffar e Inkpen (2011) mineram emoção de cinco diferentes bancos de dados rotulados entre cinco e oito emoções. O objetivo é utilizar aprendizado de máquina supervisionado em corpus heterogêneo de forma a analisar a capacidade de generalização do método para os diferentes domínios.

Calvo e Sughran (2013) classificam emoção de forma categórica e tridimensional em quatro bancos de dados diferentes. Ambas as classificações utilizam LSA, PLSA e NMF para classificar emoções básicas (*anger-disgust, fear, sadness e joy*) e dimensionais (valência, excitação e dominância). Com objetivo de melhorar o método de LSA, Bellegarda (2011) propõe uma abordagem *Latent Affective Embedding* que adiciona camadas supervisionadas a uma metodologia de aprendizado de máquina não supervisionado. Os experimentos foram feitos em manchetes de notícias e classificados em seis emoções básicas.

Com dados multimodais, PORIA *et al* (2014) classificam áudio, vídeo e texto de postagens do YouTube e do Facebook. Utilizam abordagem ensemble de extração de características e aprendizado de máquina supervisionado para classificar os documentos em *anger, disgust, fear, guilty, joy, sadness e shame*.

Utilizando redes neurais, o trabalho de Ren e Kang (2013) minera emoções básicas (*Joy, love, expectation, surprise, anxiety, sorrow, anger e hate*) e suas variações

complexas em tópicos de texto de blogs. Os autores utilizam o modelo de rede bayesiana hierárquica com dependência entre as palavras e seus tópicos. A classificação é feita em duas formas, emoções básicas ou complexas. Desta forma, um documento pode ser classificado com uma das oito emoções básicas majoritárias ou por classificação multirrotulada com intensidades para cada variação de emoções complexas.

Com uma abordagem em lógica *fuzzy*, Subasic e Huettner (2001) realizam análise afetiva em manchetes de notícias e revisões críticas de filmes. A análise afetiva aplicada abrange estados de humor, reações, sentimentos e emoções. A escolha da lógica *fuzzy* se dá pela capacidade de se manter a ambiguidade e a imprecisão dos termos afetivos no texto e à proximidade com a interpretação humana, podendo ser uma melhor ferramenta para representações visuais, por exemplo.

Outros trabalhos vão além da técnica e aplicam a mineração de emoção com diferentes objetivos e diferentes áreas. Com o objetivo de criar um modelo de recomendação de música baseado em filtros de emoção, Deng *et al.* (2015) extraem as preferências emocionais expressadas por usuários do Twitter durante um período de tempo e em diferentes níveis de granularidade para associá-los às músicas postadas por este mesmo usuário. O método utilizado é baseado em léxico próprio, montado a partir de outros três dicionários. A classificação da emoção abrange a valência geral do usuário quanto à música postada (positivo/negativo), utilizando sete emoções básicas de acordo com Izard (1971) e outras 21 emoções complexas.

Al-Hajjar e Syed (2015) propõem mineração de emoção em tweets sobre companhias de tecnologia como ferramenta para monitoramento de marca e estratégia de marketing. Em relação à classificação da emoção, o dicionário NRC Hashtag Emotion Lexicon foi criado para esse propósito. Também para monitoramento de marca e satisfação do consumidor, Ren e Changqin (2012) apresentam um modelo de computação afetiva em textos de blogs e microblogs. Os autores extraem as palavras emocionais e as classifica, de 0 a 1, em oito emoções. A polaridade geral é definida conforme a pontuação da emoção de maior valor. O corpus utilizado é previamente rotulado e a classificação ocorre através das técnicas de *machine learning*: MaxEntropy, NBM, Random Forest e Feature Intervals.

Com foco em segurança e políticas públicas, Chung et al. (2015) exploram a emoção na segurança de fronteiras nos Estados Unidos. Baseados no fato de que emoções possuem um papel importante na modelagem de políticas públicas e tomadas de decisões, os autores mineraram, utilizando dicionário léxico, 189 mil tweets de 105 mil usuários

que postaram sobre eventos de imigração e problemas na segurança de fronteira em novembro de 2014 e identificam agentes influenciadores importantes e seus impactos sobre o evento pesquisado. Dodds e Danforth (2010) mineram emoção em música e blogs para medir os níveis de felicidade em diferentes contextos utilizando métodos estatísticos. Os resultados mostram que é possível utilizar análise afetiva para direcionar políticas de incentivo e prevenção de origem pública e privada.

Outras aplicações foram identificadas, como diferenças de gênero e comportamento (MOHAMMAD, 2012); relacionamento empresa e consumidor (GARCIA-CRESPO *et al*, 2010); recuperação de informação (LEI *et al.*, 2015); educação (ALM *et al.*, 2005); visualização de emoções e resultados (SARAVIA *et al*, 2015; NIE *et. al*, 2015)

Novamente, a revisão bibliográfica proporcionou a identificação do estado da arte atual, o qual revelou-se variado e abrangente. Contudo, não foi possível identificar estudos que utilizassem da mineração de emoção aplicada às revisões críticas, especificamente, que empregasse a mineração de emoção como forma de mensurar boca a boca no mercado de cinema.

2.5 Watson

Este capítulo apresenta o IBM Watson e descreve seu histórico, desenvolvimento, funcionalidades e aplicações.

2.5.1 Computação Cognitiva, Watson e Jeopardy!

Sistemas cognitivos estão prestes a se tornarem um marco, uma vez que a linguagem natural é usada ao interagir com humanos. Esta nova tecnologia permite que os profissionais, em uma ampla gama de áreas, tomem decisões que envolvam Big Data melhores. Além disso, sistemas cognitivos são capazes de aprender através da educação e podem entender a linguagem natural (KELLY III e HAMM, 2013)

O IBM Watson está à frente dessa nova era da computação. A Computação cognitiva é um novo tipo de computação, muito diferente dos sistemas programáveis que precederam e diferente dos sistemas de máquinas de tabulação de um século atrás. Com a riqueza atual de grandes volumes de dados e a necessidade de decisões mais complexas apoiadas em evidências, esses sistemas baseados em princípios matemáticos e regras de

lógica muitas vezes falham ou não conseguem se manter atualizados com as informações disponíveis (IBM, 2014a).

A computação cognitiva permite que as pessoas criem um tipo de valor profundamente novo em encontrar respostas e ideias escondidas em grandes volumes de dados. Desde um médico fazendo um diagnóstico a um paciente, um analista financeiro aconselhando um cliente sobre sua carteira de investimentos até um chef criando uma nova receita, todos precisam de novas abordagens para contextualizar o volume de informações com que lidam diariamente com o intuito de extrair o valor apropriado desses dados (IBM, 2014a).

O Watson é um supercomputador da IBM que combina inteligência artificial e softwares analíticos sofisticados para otimização de performances como uma máquina de “respostas e perguntas”. Trata-se de um Sistema cognitivo que utiliza linguagem natural para receber e responder perguntas. Ele trabalha com dados não estruturados, que compõem 80% dos dados disponíveis online (IBM, 2014a).

O Watson é um sistema de perguntas e respostas capaz de responder questões feitas em língua natural (IBM Research, 2012), desenvolvido no projeto DeepQA da IBM por um time de pesquisadores liderados por David Ferrucci (IBM, 2011b; FERRUCCI et al, 2013). Watson recebeu o nome do primeiro CEO da IBM, o industrial Thomas J. Watson (FERRUCCI, 2012) O supercomputador foi desenvolvido especificamente para responder a perguntas no jogo de teste Jeopardy! (FERRUCCI, 2011).

O Watson processa a uma taxa de 80 teraflops (um trilhão de operações de pontos flutuantes por segundo). Para replicar – ou ultrapassar – a alta habilidade humana de responder perguntas, o sistema é otimizado para a carga de trabalho, integrando processadores POWER7 massivamente paralelos e construído sobre a tecnologia DeepQA da IBM, usada para gerar hipóteses, coletar evidências maciças e analisar dados. O Watson emprega um cluster de 90 servidores IBM Power 750, cada um usa um processador de oito núcleos POWER7 de 3,5 GHz, com quatro threads por núcleo. No total, o sistema possui 2.880 threads de processador POWER7, 16 terabytes de RAM e 500 GB de informações pré-processadas (IBM Research, 2011; IBM, 2011b). O equipamento e seus dados estão armazenados em um espaço que poderia acomodar 10 refrigeradores (IBM, 2011a).

O Watson usa o software DeepQA e o framework Apache UIMA, que foram desenhados para recuperação de informação incorporando processamento de linguagem natural e aprendizado de máquina. O Sistema foi escrito em diversas linguagens de

programação, incluindo Java, C++ e Prolog e roda no sistema operacional SUSE Linux Enterprise Server 11 usando Apache Hadoop para prover o processamento distribuído (FERRUCCI, 2012).

O Deep Blue, outro supercomputador da IBM, é o predecessor do Watson e foi a primeira máquina construída pela empresa capaz de vencer um campeão do mundo em xadrez usando cálculos matemáticos complexos. Consequentemente, devido às suas capacidades matemáticas, a máquina foi usada de forma a executar os cálculos que eram necessários em muitas outras áreas científicas (CAMPBELL et al, 2002). Seguindo essa ideia, a IBM quis ir além e construir uma máquina capaz de entender questões em língua natural, subárea de inteligência artificial que tem interessado muitos pesquisadores ao longo do tempo. Devido a isso, o time de DeepQA quis construir uma máquina capaz de utilizar Linguagem Natural de forma a ajudar em áreas como finanças, saúde e advocacia. Surge então, em 2005, o Watson (FERRUCCI, 2011).

Para mostrar suas habilidades, Watson desafiou os dois melhores jogadores do programa de TV Jeopardy!, Ken Jennings e Brad Rutter, em 2011 (JENNINGS, 2013). O supercomputador ganhou o programa utilizando seu sistema de diálogo, que foi sua primeira função. O avatar do Watson “sentou” entre os competidores, como um competidor humano faria, enquanto sua “massa cinzenta” estava alocada em outro lugar do prédio. Assim como os outros dois competidores, o Watson também não teve acesso à Internet (IBM Research, 2011b).

Imediatamente após vencer o jogo em 2011, o Watson anunciou que suas primeiras aplicações para o mundo real seriam na área médica. Assim a IBM começou a trabalhar com diversas companhias e universidades para desenvolver as tecnologias do Watson e auxiliar médicos e áreas relacionadas à medicina. Em 2013, o Watson se tornou comercial, iniciando sua carreira na advocacia. Em 2014 ele foi expandido para novas funcionalidades, como extração de conhecimento e visualização gráfica, o Watson Explorer e o Watson Ecosystem. Um ano depois, a plataforma Bluemix foi desenvolvida e o Watson developer cloud foi lançado. Em 2016, o Watson foi expandido pela melhora no engajamento humano focando em detecção e expressão emocional, incluindo robótica (HIGH, 2016). O sistema cresceu em cima de suas respostas curtas no Jeopardy! e passou para aplicações realistas.

Atualmente o Watson está revolucionando a forma como tomamos decisões, como nos tornamos especialistas e como compartilhamos conhecimento em campos tão diversos como medicina, advocacia e até mesmo cozinhar. Além do mais, Watson está

descobrimos e oferecendo respostas e padrões que não sabíamos que existiam mais rápido do que qualquer pessoa ou grupo de pessoas poderia, em formas que fazem diferença material todos os dias. O mais importante de tudo é que o Watson aprende, se adapta e se mantém inteligente. Ele realmente ganha valor com a idade, aprendendo a partir de suas interações com humanos e de seus próprios sucessos e fracassos, assim como nós (IBM Watson, 2014).

2.5.2 Inteligência Artificial, PLN e Aplicações

Quando o Watson trabalha em um campo particular, ele aprende a linguagem, o jargão e o modo de pensar desse domínio. Com a orientação de especialistas, Watson coleta o conhecimento exigido para obter a instrução em um particular domínio, isto é chamado de corpus do conhecimento. Coletar um corpus começa por carregar, dentro do Watson, a literatura relevante do assunto. Construir o corpus também requer alguma intervenção humana para chamar a informação correta e descartar tudo que está desatualizado ou seja, considerado irrelevante para o domínio do problema, como uma curadoria. Em seguida, os dados são pré-processados pelo Watson de forma a se construir índices e metadados que tornam o trabalho com este novo conteúdo mais eficiente, o que é conhecido como ingestão. Neste momento, o Watson também pode criar um gráfico de conhecimento para ajudar a responder perguntas mais precisas (IBM Watson, 2014).

Agora que o Watson ingeriu o corpus, ele precisa ser treinado por um especialista humano para aprender a como interpretar a informação, as melhores respostas possíveis e adquirir a habilidade de encontrar padrões. Os parceiros do Watson e os especialistas que o treinam usam a abordagem de aprendizado de máquina. Um especialista envia os dados treinados a Watson na forma de pares pergunta-resposta que são considerados verdadeiros. Isso não dá ao Watson respostas explícitas para todas as questões, mas sim ensina os padrões linguísticos do significado no domínio (IBM Watson, 2014).

Uma vez que o Watson foi treinado no formato de pares pergunta-resposta, ele continua a aprender através da interação contínua. Interações entre usuários e o Watson são periodicamente revisadas por especialistas e alimentadas novamente no sistema para ajudá-lo a entender e interpretar melhor as informações. Da mesma forma, conforme novas informações vão sendo publicadas, o Watson é atualizado, de forma que ele está constantemente se adaptando às mudanças, conhecimentos e interpretações linguísticas de uma determinada área. As fontes de informação do Watson incluem enciclopédias,

dicionários, artigos e obras literárias. Watson também usou bancos de dados, taxonomias e ontologias. Especificamente, DBPedia, WordNet, e Yago foram usados (FERRUCCI, 2010).

Quando uma pergunta é feita, o Watson considera os padrões de linguagem e cria hipóteses, então, procura por evidências para embasar ou refutar a questão. Finalmente, Watson estima um grau de confiança na resposta antes de informá-la. Watson está agora pronto para responder questões sobre situações altamente complexas e rapidamente fornecer um range de potenciais respostas e recomendações que são suportadas por evidências. Ele está também preparado para identificar novas ideias ou padrões escondidos nas informações (IBM Watson, 2014).

Tratando-se de interação humano-computador, grande parte da inteligência do Watson se dá pelo processamento da linguagem natural, que permite a Watson a compreensão de dados não estruturados, ou seja, 80% dos dados disponíveis atualmente, incluindo desde artigos, literatura e relatórios de pesquisa a postagens, blogs e tweets (IBM Watson, 2014). De acordo com a IBM (2011c, p. 3) “mais de 100 técnicas diferentes são usadas para analisar a língua natural, identificar as fontes, encontrar e gerar hipóteses, encontrar uma pontuação de evidência e agregar e ordenar as hipóteses”.

Quando se trata de texto, Watson não procura apenas por combinações de palavras ou seus sinônimos, como uma máquina de busca. Ele interpreta o texto como uma pessoa, quebrando uma sentença gramaticalmente, relacional e estruturalmente, discernindo o significado da semântica do material escrito. Watson compreende o contexto, o que é muito diferente de simples reconhecimento de fala, modo como um computador traduz a fala humana em um conjunto de palavras. Watson tenta compreender a real intenção da linguagem dos usuários e utiliza esse entendimento para possivelmente extrair respostas lógicas e desenhar inferências a potenciais respostas através de uma ampla gama de modelos linguísticos e algoritmos (IBM Watson, 2014).

Como ferramenta potente de mineração de texto através do processamento de linguagem natural, as aplicações para a tecnologia cognitiva subjacente do Watson são quase infinitas. Com a Web 2.0 e a explosão de dados, mídias sociais, blogs, relatórios de pesquisa e e-mails inspiradores, a riqueza dos *insights* e das perguntas que podem ser extraídas a partir da união dos dados é imensurável, como detalhado na seção Web 2.0.

Por exemplo, monitoramento de marca e de produtos. A Compreensão de linguagem natural do Watson oferece um conjunto de ferramentas de análise de texto, com as quais os usuários podem responder questões como “como posso usar inteligência

de negócios para criar painéis gráficos e relatórios?”, “como posso monitorar as tendências das mídias sociais e as mudanças das opiniões públicas?”, “como posso recomendar um conteúdo que a minha audiência irá gostar?” e “como posso saber o que os consumidores estão dizendo sobre a minha marca?” (figura 2). O serviço de compreensão de linguagem natural do Watson permite que os usuários compreendam os tópicos presentes no texto ao analisar palavras-chave, conceitos e categorias. Com as entidades, é possível extrair as pessoas, lugares e organizações mencionadas no texto. Os usuários podem também entender o sentimento expressado, incluindo os sentimentos expressados acerca de uma palavra, entidade ou frase específica. Adicionalmente, eles podem se aprofundar na análise de sentimentos e obter as emoções distintas por trás da escrita das pessoas, quando determinar a polaridade do texto não for suficiente (IBM Watson, 2017).

Figura 2 – Ferramentas de Compreensão de Linguagem Natural IBM Watson

Fonte: IBM Watson (2017)

Watson agora está incorporado dentro de cerca de outras 17 indústrias diferentes e a IBM está trabalhando com mais de 500 parceiros (IBM, 2016) para construir as aplicações cognitivas, incluindo varejo, direito, música, reconhecimento de imagem, hotelaria, moda e cozinha (IBM, 2017).

2.5.3 Alchemy API, Análise de Sentimento e Mineração de Emoção

O serviço do IBM Watson AlchemyLanguage⁶ é uma coleção de funções de análise de texto, por meio de processamento de linguagem natural, que extrai informações semânticas, o que permite aos usuários construir aplicações cognitivas e entender conteúdos e contextos. Ao introduzir um texto, HTML ou um link URL público, o sistema executa técnicas sofisticadas de processamento de linguagem natural para obter uma rápida compreensão de alto nível do conteúdo e extrair insights detalhados, como sentimento para entidades e palavras-chave detectadas (IBM Watson, 2016).

O IBM Watson AlchemyLanguage API⁷ pode extrair diversas informações de conteúdos textuais, como por exemplo, conceitos – identifica conceitos que estão associados ao texto, baseado em outros conceitos e entidades também presentes no mesmo documento de texto; entidades – retorna informações como pessoas, lugares e organizações que estão presentes no texto; palavras-chave – tópicos importantes no conteúdo que são tipicamente usados quando se indexa dados, geração de *tags* ou busca; língua – reconhecimento automático do idioma; relações – identifica relações entre pessoas, ações e objetos nas sentenças do texto; sentimento – identifica atitude, opiniões ou sentimentos no conteúdo analisado (figura 3). A classificação da polaridade pode ser dada sobre um documento inteiro ou em níveis menores como entidade, palavras-chave e citações; e análise de emoção – detecta raiva, desgosto, medo, alegria e tristeza implícitos e textos. A análise de emoção pode detectar emoções que estão associadas às frases, entidades, palavras-chave ou analisar a emoção geral do conteúdo (IBM Watson, 2016).

⁶ <https://www.ibm.com/watson/developercloud/alchemy-language/api/v1/#introduction>;
<https://www.ibm.com/watson/developercloud/doc/alchemylanguage/index.shtml>

⁷ A API de PLN Alchemy foi comprada da Alchemy Company pela IBM em 2015 para integrar o conjunto de ferramentas do Watson. Disponível em: <https://www.ibm.com/watson/alchemy-api.html>.

Figura 3 - Extração de relações, entidades, palavras-chave, sentimento e emoção

Fonte: IBM Watson (2017b)

O Watson já está treinado com dados de diversos domínios, havendo assim a detecção automática do contexto e tornando desnecessário um treinamento adicional de dados. Isso faz com que sua utilização, apesar de se tratar de algoritmos de aprendizado de máquina supervisionado, seja fácil e menos onerosa, uma vez que a fase de treinamento está pronta e basta enviar os dados de teste.

Sentimento é a atitude, opinião ou o afeto sobre alguma coisa, como uma pessoa, organização, produto ou local (vide seção 2.3). A análise de sentimento do Alchemy API⁸ oferece um mecanismo de fácil uso para identificar a polaridade negativa, positiva ou neutra, de qualquer documento de texto ou página da web (IBM Watson, 2016b).

O algoritmo de análise de sentimento da API Alchemy busca por palavras que carreguem conotações positivas ou negativas e então descobre sobre qual pessoa, lugar ou coisa elas se referem. O algoritmo também lida com negações (“good” vs. “not good”) e modificadores (“good” e “really good”). O Watson está treinado com dados da web rotulados, como Twitter e Wikipedia, incluindo *emoticons*, *emojis*, *hashtags* e *likes*, que são avaliados com o intuito de se obter a polaridade. A API de análise de sentimento

⁸ <http://www.alchemyapi.com/products/alchemylanguage/sentiment-analysis>

funciona em documentos grandes e pequenos, incluindo manchetes de notícias, postagens em blogs, revisão de produtos, comentários e Tweets (IBM Watson, 2016b).

Tabela 3 - Métodos disponíveis pela funcionalidade da análise de sentimento da API AlchemyLanguage

Método	Descrição
Status	OK / ERROR – Indica quando a requisição obteve sucesso
URL	HTTP URL passada na entrada
Language	Detecção automática do idioma do texto
docSentiment	Objeto contendo a informação de sentimento a nível de documento
Mixed	Indica quando o sentimento é positivo e negativo (1)
Score	Força do sentimento: [-1, 1]
Type	Polaridade do sentimento: Negativo / Neutro / Positivo

Fonte: https://www.ibm.com/watson/developercloud/alchemy-language/api/v1/#targeted_sentiment

Com a análise de emoção da API AlchemyLanguage Emotion Analysis^{9,10}, pode-se agregar sensibilidade às análises. Manter-se sintonizado com os sentimentos dos clientes durante as sessões de bate-papo, acompanhar as reações das mídias sociais aos comunicados de imprensa ou avaliar a visão pública sobre notícias financeiras. A análise emocional pode inferir emoções de amostras de texto de qualquer tamanho até 50KB.

O texto a ser minerado pode ser um texto publicado, um texto HTML, ou um link para uma página da web e a análise de emoção iniciará um sofisticado aprendizado de máquina para retornar escores de confiança para raiva (*angry*), desgosto (*disgust*), medo (*fear*), alegria (*joy*) e tristeza (*sadness*). Cada score indica a probabilidade de que a emoção correspondente esteja implícita no texto em questão (IBM Watson, 2017b).

A análise de emoção utiliza uma estrutura ensemble baseada em generalizações para extrair as pontuações finais das emoções de resultados iniciais de vários modelos de aprendizado de máquina. Cada nível inferior usa uma combinação única de algoritmos de aprendizado de máquina e características de linguagens, como palavras, frases, pontuações e sentimentos gerais (GUNDECHA, 2016).

⁹ <http://www.alchemyapi.com/products/alchemylanguage/emotion-analysis>

¹⁰ https://www.ibm.com/support/knowledgecenter/en/SSWTQQ_1.2.0/com.ibm.swg.ba.cognos.sifs_solution_guide.1.2.0.doc/c_trd_emotion_detection_library.html

Tabela 4 - Métodos disponíveis pela funcionalidade da análise de emoção da API AlchemyLanguage

Método	Descrição
Status	OK / ERROR – Indica quando a requisição obteve sucesso
URL	HTTP URL passada na entrada
Language	Detecção automática do idioma do texto
docEmotions	Objeto contendo as emoções (<i>sadness, joy, fear, disgust e anger</i>) e seus escores [0, 1], sendo 0 não contém a emoção e 1 definitivamente contém e suas graduações graus de força da emoção.

Fonte: https://www.ibm.com/watson/developercloud/alchemy-language/api/v1/#emotion_analysis

2.5.4 Revisão AlchemyAPI

A API AlchemyLanguage tem sido usada de forma consistente na literatura atual. Por se tratar de uma ferramenta recente, o seu uso em pesquisas se inicia em 2012, com as tarefas mais simples de extração de características do texto analisado e evolui, a partir de 2013, com a análise de sentimento. A mineração de emoção só foi implementada à suíte de funcionalidades do IBM Watson em julho de 2016¹¹, o que é ainda muito recente para um reflexo visível na área acadêmica.

Saif, Yulan e Harith (2012) e Saif *et al* (2012) utilizam a funcionalidade de extração de características da API AlchemyLanguage e a compara com outras ferramentas disponíveis no mercado: Zemantra e OpenCalais. Em ambos os trabalhos os resultados mostram que a Alchemy executa melhor que em termos de qualidade, acurácia e de volume de entidades extraídas. Kulshrestha *et al* (2015) utilizam a ferramenta como *baseline* de qualidade a ser atingida pelo método PLN que apresentam.

Muitos outros trabalhos utilizam a análise de sentimento da AlchemyLanguage para detectar a polaridade em seus textos do Twitter (BATOOL *et al*, 2013; MEEHAN *et al*, 2013), do Facebook e outras mídias sociais (RODRIGUES *et al*, 2016; CIULLO *et al*, 2016; HOFMANN *et al*, 2013) e postagens e revisões críticas de filmes (SINGH *et al*, 2013). Singh *et al* (2013) e Loia e Senatore (2014) utilizam as métricas de resultados da classificação da polaridade da API Alchemy método avaliativo das suas análises de sentimento.

¹¹ Notas de Releases, disponível em: <https://www.ibm.com/blogs/bluemix/2016/08/alchemylanguage-emotion-analysis-api-even-better>

Ma *et al* (2013) comparam a análise de sentimentos da API AlchemyLanguage com outras três ferramentas que realizam a mesma tarefa: SentiStrenght, Stanford NLP sentiment analyser e NL ToolKit. Os resultados da AlchemyLanguage se mostraram mais robustos, uma vez que a ferramenta é mais completa em termos gerais de processamento e funcionalidade. Com o mesmo foco comparativo, Serrano-Gurrero *et al* (2015) realizam a comparação da Alchemy com outras ferramentas disponíveis no mercado para monitoramento de mídias sociais (Lymbix, Musicmetric, Openamplify, Opinion Crawl, Opendover, Repustate, Semantria, Sentiment140, SentimentAnalyzer, SentiRate, Sentimetrix, Uclassify, ViralHeat e Wingify). Na comparação, a API AlchemyLanguage é a mais completa, cobrindo as funcionalidades requisitadas como diferentes níveis de granularidade nas análises, polaridade e emoção. Apesar de apresentar somente a quarta melhor acurácia do conjunto de ferramentas testadas para o conjunto de dados de revisões críticas de filme, a AlchemyLanguage obteve a maior acurácia do conjunto para análises de Tweets e apresentou a menor taxa de erro e a maior assertividade para verdadeiros positivos na média geral.

Por último, Wu, Bartram e Shaw (2017) e Swati e Sureka (2016) utilizam mineração de emoção da AlchemyLanguage API em dados do Twitter e do Tumblr, respectivamente.

Com base nas avaliações apresentadas em pesquisas anteriores, acredita-se que a API AlchemyLanguage seja uma escolha adequada para a atividade pretendida. A AlchemyLanguage aparece na literatura como melhor em termos de facilidade de uso, acurácia, volume de processamento e abertura para acadêmicos. Adicionalmente, a IBM, empresa que detém a tecnologia do Watson, é pioneira em PLN, conforme citado na seção 2.2.

2.6 Considerações Finais

Análise de sentimento e mineração de emoção são áreas interligadas da pesquisa que utilizam de técnicas de Processamento de Linguagem Natural, Recuperação de Informação e Mineração de Dados. A junção das técnicas de forma a se obter um melhor resultado foi aqui fundamentada em computação afetiva.

A maior parte dos dados disponíveis são provenientes de fontes não estruturadas, o que configura em um importante desafio. Vários passos são necessários para realizar a mineração dos textos, uma vez que estes textos vêm de diferentes fontes e em diferentes

formatos. A aquisição dos dados e o pré-processamento dos mesmos são as subtarefas mais comuns requeridas para as análises.

Apesar da computação afetiva poder parecer uma tarefa de simples categorização de texto, ela inclui vários desafios, fazendo com que pesquisadores foquem neste tópico estimulante. A interpretação semântica do documento como em “Como alguém pôde se sentar em frente a este filme?” e as diferentes interpretações que a mesma palavra/expressão pode ter em diferentes contextos são exemplos claros de desafios que tornam a tarefa de extrair sentimento e emoção ainda mais complexas. Outros desafios ocorrem devido à ambiguidade das palavras no texto, à complexidade de significado e aos papéis de vários fatores como ironia, educação, estilo da escrita, bem como a variedade da linguagem de pessoa para pessoa e de cultura para cultura. Todos estes fatores fazem da análise afetiva uma tarefa muito interessante e desafiante (SALEH *et al.*, 2011).

O conteúdo gerado por usuário se apresenta de forma crescente diariamente, especialmente nas redes sociais e em *e-commerces*, o que motiva o crescimento dos estudos da área com o intuito de criar novas técnicas e combinações capazes de processar de forma automática e robusta os dados opinativos existentes. Muitas são as aplicações centradas na sumarização e visualização das afetividades em geral, ou na predição de comportamentos com base no sentimento/emoção existente. Tanto o consumidor quanto as empresas estão interessadas nas múltiplas aplicações que a análise afetiva pode proporcionar, variando entre a monitoração de mercados, apoio à tomada de decisão e predições de comportamentos. Portanto, ainda há muito que se explorar dentro desta área relativamente recente, permitindo a evolução das técnicas de classificação entre línguas, domínios e escalabilidade.

3 Literatura de revisão crítica em Economia

Considerada como bem de experiência, os filmes são de difícil avaliação, sendo necessário primeiro o consumo para somente depois avaliar sua qualidade. Para que o consumidor tome sua decisão de compra sem dispor do conhecimento sobre a qualidade do filme, os consumidores procuram por informações que possam fornecer algum parecer desta, como críticas de críticos profissionais e de seus pares, outros consumidores. Tal característica faz com que investigar o que torna um filme de sucesso ou não seja um desafio interessante que vem sendo estudado a longo prazo.

Este capítulo disserta sobre a contextualização do problema no âmbito da economia e da literatura de revisão crítica. Apresenta uma revisão bibliográfica relacionado com o intuito de identificar o estado da arte atual, suas lacunas e limitações. Em seguida, conclui-se com as evidências encontrada e atribui-se então, as contribuições pertinentes.

3.1 Visão Geral

É comum ver livros, concertos, filmes, peças de teatro, restaurantes, programas na TV e outros produtos da indústria do entretenimento serem revisados por críticos profissionais. Há, inclusive, publicações específicas para esta finalidade, como revistas, jornais e alguns boletins televisionados, apesar de que tais formatos estão sendo migrados para meios online. Críticas profissionais têm existido ao longo do tempo em diversas categorias de produtos, especialmente para produtos considerados bens de experiência (*experience goods*), onde é preciso primeiro consumir para então avaliar a sua qualidade (SHAPIRO e VARIAN, 1997). Portanto, a informação de especialistas é de extrema relevância para induzir consumo de bens de experiência. Como consequência, para tais produtos a opinião de *experts* é frequentemente solicitada e, até um certo momento no ciclo de vida do produto, tal crítica profissional é a única informação que se tem disponível a respeito da qualidade do mesmo.

Espera-se que os consumidores sejam influenciados pelos julgamentos de críticos profissionais, pelas comunidades amadoras (Amazon.com, Netflix.com, vídeo games em GameSpot.com, restaurantes em Citysearch.com, etc.), outros consumidores, e por seus pares (REINSTEIN e SNYDER, 2005; LUCA, 2011; DUAN et al., 2008;

CHINTAGUNTA et al., 2010 e BOATWRIGHT et al., 2007). Isso porque o grande número de novos bens de experiência e a alta frequência com que são lançados nos mercados pode compor informação em quantidade exagerada para os processos de escolha dos consumidores (MOON et al., 2010).

Ilustrando esta influência por parte de outros consumidores, o relatório de decisões dos consumidores *The Internet and Consumer Choice* do órgão de pesquisa americano PEW (2008) mostra que informações online geralmente são modestas em seu impacto sobre as decisões dos consumidores, porém existem e a importância destas informações cresce à medida da importância da compra e sobre produtos considerados bens de experiência. Em outro de seus relatórios de pesquisa, o *online shopping*, revela-se que aproximadamente entre 73% e 87% dos leitores de críticas online de restaurantes, hotéis e outros serviços como agências de viagem ou médicos, relatam que as críticas têm influência significativa em suas compras. Outro fato apontado no mesmo relatório é que os consumidores afirmam estarem dispostos a pagar entre 20 e 99% a mais em um produto com pontuação 5 estrelas – máxima – em comparação a um produto de 4 estrelas (PEW, 2008b).

Com a explosão da Web 2.0, aproximadamente em 2004, em conjunto com as plataformas como blogs, fóruns de discussão, redes *peer-to-peer* e vários outros tipos de mídias sociais, consumidores tem hoje a sua disposição um leque de alcance sem precedentes e poder pelo qual partilham suas experiências de marca e opiniões, positivas ou negativas, em relação a qualquer produto ou serviço. Como as grandes empresas estão cada vez mais a notar, essas vozes podem exercer enorme influência na formação das opiniões de outros consumidores – e, finalmente, à sua lealdade de marca e suas decisões de compra (PANG e LEE, 2008).

Para embasamento teórico e levantamento do estado da arte sobre o tema, este capítulo apresenta uma revisão bibliográfica na área com o objetivo de determinar se as revisões críticas disponibilizadas por críticos profissionais de primeira e segunda linha (tops e não tops) e por parte dos consumidores afetam o tempo de duração de exibição de um filme no mercado americano de cinema. Assim, a investigação na literatura de economia busca levantar estudos empíricos que relacionem estes efeitos, uma vez que não foram encontradas literaturas que estimassem o efeito do boca a boca de maneira quantitativa tal como propomos aqui.

3.2 Revisão Bibliográfica

Alguns estudos nas áreas de economia, marketing e computação tem investigado empiricamente o efeito das revisões críticas sobre o sucesso ou fracasso de produtos, medido em termos de receita de vendas. Estes estudos, entretanto, embora reconheçam o efeito das críticas em geral, não consentem a respeito de como a revisão crítica afeta o faturamento (DUAN et al., 2008; KING, 2007). Mais especificamente, não há consenso a respeito de como a influência dos críticos profissionais afeta a arrecadação da bilheteria, havendo divergências conforme a categoria dos filmes, reputação do crítico, orçamento da produção e no enfoque da influência das críticas (se de influência ou de previsão).

Investigando revisões críticas de críticos profissionais, Eliashberg e Shugan (1997) sustentam que os críticos exercem com maior frequência o papel de preditores, sendo capazes apenas de antecipar o fracasso ou sucesso do filme por terem relação com a receita total da bilheteria, e não com o faturamento da primeira semana de exibição, onde tal correlação implicaria em efeito influenciador.

Outros trabalhos, no entanto, analisam exclusivamente a relação entre o faturamento e as críticas dos revisores profissionais de primeira linha. Reinstein e Snyder (2005) analisam o impacto das críticas de dois críticos profissionais de primeira linha: Siskel e Ebert. Famosos, Siskel e Ebert têm suas críticas transmitidas todos os finais de semana antes da estreia dos filmes. Utilizando modelos de regressões os autores eliminam o viés de qualidade do produto, que tradicionalmente, são proporcionais às notas críticas ao incluírem as notas das avaliações críticas como proxy. Os resultados mostram alguma fraca evidência de um efeito de influência na amostra contendo todos os filmes. A influência de um *thumb up* é 11% da receita do final de semana de estreia, mas é estatisticamente insignificante. A influência de dois *thumbs up* é maior, 25%, mas somente marginalmente estatisticamente significativa. Foi encontrado que o efeito de influência difere através de diferentes categorias de filmes, sendo maior para filmes de pequena distribuição (filmes estrangeiros, menos populares, documentários, filmes de arte) e para dramas, virtualmente insignificante para os filmes mais populares e para filmes de comédia e ação. Uma crítica positiva realizada próxima à semana de estreia do filme aumenta o número de consumidores no cinema como um todo, ao invés de apenas mudar a data da ida ao cinema para mais tarde. Este aumento na receita pode vir da disputa

entre os filmes competidores no mesmo final de semana, quando a crítica ajuda na escolha entre qual filme em cartaz escolher, uma vez que já se decidiu ir ao cinema.

Boatwright et al. (2007) mostram, para o mercado americano, que alguns revisores críticos de primeira linha tendem a influenciar mais o público do que outros revisores também do mesmo grupo. Esta posição destacada de alguns revisores críticos, entretanto, pode ter influências contraditórias, alguns podem influenciar o público positivamente e outros negativamente. Também exibem evidências de que o efeito influenciador dos revisores críticos de primeira linha é maior para filmes *narrow-release*. Ainda, considerando o boca a boca como forma de publicidade inicial do produto, os autores determinam como os consumidores descobrem um novo produto, considerando a publicidade como fonte externa e o boca a boca como fonte interna de conhecimento. Os resultados mostram que a publicidade está positivamente relacionada com o coeficiente de inovação para filmes de baixa abertura enquanto inibe a queda das vendas, tipicamente rápidas, para filmes de alta abertura. Destacam também que se os críticos são primeiramente líderes de opinião então suas revisões terão um impacto significativo sobre os valores iniciais da receita da bilheteria, mas não necessariamente sobre o resultado final.

Utilizando as críticas de apenas críticos profissionais King (2007) pesquisa a relação entre resenhas críticas de críticos profissionais e a arrecadação da bilheteria e encontra que os críticos são mais propensos a gostarem mais de certos tipos de filmes, especialmente estrangeiros e documentários, mais do que o público como um todo. Nesta categoria, incluindo filmes de baixo lançamento, classificações críticas tem certa influência no faturamento da bilheteria, mas existe pouca probabilidade que mesmo excelentes revisões vão impulsionar esta categoria de filme a uma distribuição ampla. Para filmes de alto orçamento não existe diferença sistemática de críticas entre críticos e o público quanto à qualidade relativa de diferentes filmes. Entretanto, sinaliza uma relação positiva entre pontos de classificação média e bilheteria. Seus resultados apontam que um aumento de 20 pontos na classificação média do filme está associado a um aumento de aproximadamente USD26 milhões no faturamento.

De forma a descobrir outras possíveis influências para a arrecadação da bilheteria, além das resenhas críticas, Basuroy et al. (2003) investigam de que forma as críticas afetam a performance de arrecadação da bilheteria de filmes e como os efeitos podem ser moderados pelo elenco e orçamento. Eles mostram que ambas resenhas críticas, positivas e negativas, possuem papel duplo de influenciador e de preditor sobre

o faturamento, no entanto, o impacto das revisões negativas tendem a diminuir com o tempo, um padrão que é mais consistente com o papel de influenciador do que de preditor. Apontam também que durante a primeira semana do filme as críticas negativas atrapalham mais a performance financeira do que críticas positivas ajudam. Adicionalmente, os resultados indicam que o peso do elenco e o orçamento da produção melhoram a arrecadação da bilheteria para filmes que possuem mais críticas negativas do que positivas e tem mínimo impacto sobre aqueles cujas críticas são majoritariamente positivas.

De um modo geral existem fatores na indústria cinematográfica que podem determinar o fracasso ou sucesso de um filme (McKENZIE, 2010). Esses fatores são de difícil identificação, principalmente considerando que múltiplos fatores podem ser importantes para identificar um filme como sendo um sucesso. Deve-se observar que a revisão crítica é apenas um deles, mas necessário que se controle por todas as influências para se evitar o viés de omissão de variável relevante. Ainda, nota-se que o próprio conceito de sucesso precisa ser definido, Simonton (2009) é um autor que oferece uma medida ao fenômeno.

Com referência aos estudos mencionados acima, pode-se dizer que a maioria deles coletou informações de revisões críticas em jornais e revistas, ou seja, em mídias impressas que hoje estão simultaneamente disponíveis na Internet. Desta forma, podemos considerar que a Internet engloba todos os meios de revisões críticas existentes no mercado, sendo, portanto, mais abrangente. Uma vez que o conteúdo impresso nestes veículos está inteiramente disponível também online para o público em geral. Sendo assim, as revisões críticas disponíveis na Internet são uma boa medida para todo o conjunto de informação de revisão crítica disponível para os consumidores, ainda que estes não tenham acesso à Internet.

Com o advento da Internet e mais recentemente da generalização da banda larga, iniciada em 1996 nos Estados Unidos, houve aumento da disseminação dessas informações através de postagens, publicações, comentários e compartilhamentos, de modo que todas as publicações de revisão crítica foram incorporadas à rede e extremamente ampliadas. Assim, a Internet se tornou o principal meio de divulgação de revisões críticas. Neste contexto, a ideia da classificação e avaliação online sobre os mais variados tipos de bens e serviços tem se popularizado e ampliado a participação de todos os agentes: dos revisores críticos, de usuários que fazem revisão crítica e consumidores de resenhas críticas. Esse fenômeno tem causado um enorme impacto na dinâmica do

efeito boca a boca (*word of mouth*) ao efetivamente conectar os consumidores com seus pares (MOON et al., 2010). Todos indivíduos podem compartilhar seus pensamentos com milhões de outros usuários na Internet e influenciar outras decisões através do boca a boca (DUAN et al., 2008).

O fenômeno boca a boca tem sido reconhecido como um dos recursos mais influenciadores da transmissão da informação nas sociedades, especialmente para bens de experiência. Sistemas de comunicação online são uns dos mais poderosos canais para potencializar este fenômeno, pois garantem agilidade, flexibilidade, maior alcance e não tem limites de tempo ou fronteira, como ocorria na era sem Internet (DUAN et al., 2008).

Essa disseminação e popularização das revisões críticas via web acontece principalmente por que não só o filme, mas a própria revisão crítica, é também um bem de experiência (REINSTEIN e SNYDER, 2005) e um bem de informação. Deve-se notar que bem de informação possui característica de bem público, ou seja, é não excludente e não rival, justificando Duan et al. (2008) sobre o alcance a falta de fronteiras sobre os sistemas de comunicação.

Neste âmbito, da influência do boca a boca, online ou não, sobre o mercado de filmes, se destacam algumas literaturas. Dellarocas e Ritu (2006) introduzem uma nova métrica de medida para proxy de boca a boca, a densidade, além das métricas comuns de volume e valência. A densidade é definida pelo total de pessoas que postaram revisões online para um produto num determinado período de tempo em razão do número de pessoas que compraram o produto no mesmo período. A densidade das avaliações online pode ser útil como um proxy de consumidores propensos a engajar boca a boca pós compra. Gastos com marketing tem relação positiva com a densidade das revisões, e a valência e a densidade das críticas têm relação em forma de U.

Chintagunta et al (2010) propõem medir o impacto (valência, volume e variância) das revisões online de consumidores na arrecadação da bilheteria do mercado cinematográfico estadunidense separado por áreas. Os autores mostram que no nível desagregado local pesquisado a valência é mais importante que o volume e que, no nível agregado nacional, o volume é mais influente na arrecadação da bilheteria. De forma semelhante Duan et al. (2008) examinam o efeito de persuasão e de conscientização das críticas online de usuários na performance diária das bilheterias do cinema americano. Os resultados mostram que a classificação das críticas online dos usuários não tem impacto significativo nas receitas da bilheteria se considerada a possível endogeneidade, indicando, portanto, que as críticas online têm um efeito de persuasão pequeno nas decisões de

compra do consumidor. Porém, as vendas de bilheteria são fortemente influenciadas pelo volume das postagens online, sugerindo a importância da propagação da informação provocada por essas revisões críticas. Apresentam ainda que o efeito do boca a boca é melhor após a primeira semana, mas juntamente com seu ‘pico’ tal fenômeno começa a perder força.

Liu (2006) explora, entre outras coisas, o mecanismo por trás dos efeitos do boca a boca e qual deles ajuda, de fato, a explicar as vendas da bilheteria nos cinemas americanos. Os resultados mostram que as atividades de boca a boca são mais ativas antes e durante a estreia e que os consumidores tendem a elevar suas expectativas no período pré-estreia e ficam mais críticos após ela, justamente por poder causar desapontamentos. Encontra também que o efeito do boca a boca oferece um poder explanatório significativo para a receita agregada e semanal da bilheteria, especialmente nas semanas após o lançamento do filme, no entanto, esse poder explanatório é mais significativo no volume do boca a boca do que de sua valência.

Moon et al (2010) investigam como as notas dadas por críticos profissionais, comunidades amadores e consumidores aos filmes influenciam nas medidas de performance do filme. Apresentam que as arrecadações mais próximas à estreia melhoram as críticas posteriores do filme. Mostram que as críticas profissionais altamente positivas da primeira semana tendem a liderar altas avaliações também por parte das comunidades amadoras e, conseqüentemente, aumentar a arrecadação da bilheteria, que por sua vez, pode levar a um aumento das avaliações nas semanas subsequentes. Os autores afirmam ainda que para as semanas posteriores ao lançamento, apenas as avaliações, sem suporte simultâneo da publicidade, não são suficientes para manter a atenção dos cinéfilos.

Rui et al (2013) investigam como e quando as revisões críticas da rede social Twitter afetam as vendas da bilheteria dos cinemas. Através dos tweets, sua classificação de polaridade e a influência do usuário que posta a revisão crítica (baseada na sua rede de relacionamento), os autores encontram que a valência de fato exerce influência sobre a bilheteria de acordo com sua polaridade e que usuários que possuem um maior número de seguidores exercem também maior influência do que os usuários com poucos seguidores. Kim et al (2013) examinam o impacto do boca a boca online e das revisões críticas de profissionais na receita de bilheteria do cinema Americano e mundial. Eles concluem que para os críticos profissionais a valência importa no faturamento para o mercado nacional, mas não tem efeito sobre o mercado internacional. Para as críticas dos

consumidores encontram que o volume de revisões importa em ambos mercados, mas a valência é estatisticamente insignificante.

Este efeito de influência do volume do boca a boca pode ocorrer, em partes, pela característica endógena do fenômeno, conforme afirmam Duan et al (2008b): o boca a boca leva a maiores vendas, que em troca, gera mais boca a boca e então, mais vendas. Os autores desenvolvem um sistema dinâmico para capturar a inter-relação entre o boca a boca online e as vendas do cinema. O modelo especifica a relação dupla de causalidade e revela o retorno positivo entre o boca a boca online e as vendas. Os resultados mostram que é preciso controlar a endogeneidade neste tipo avaliação, mas suporta os resultados de trabalhos anteriores, que o volume do boca a boca online possui uma relação positiva com as vendas da bilheteria.

Alguns trabalhos focam na qualidade da revisão, na quantidade de revisão crítica disponível sua valência e densidade. Yuan-yuan et al (2008) exploram as diferenças de relacionamento entre revisões críticas online e a receita da bilheteria em diferentes estágios do ciclo de vida do filme (semanas). Os resultados mostram que tanto o volume quanto a valência das revisões online são insignificantes para a bilheteria na primeira semana de estreia, enquanto as revisões críticas e o orçamento de produção do filme mostram influências positivas, principalmente as revisões críticas dos críticos profissionais que exercem maior impacto. A partir da segunda semana, o impacto positivo do volume de revisões online se torna significativo, mas vai perdendo ao longo das semanas de exibição. Sobre a qualidade da revisão, Hu et al (2008) consideram ambos aspectos, quantitativo e qualitativo das revisões online como qualidade do revisor, exposição do revisor, cobertura do produto e efeitos temporais. Os autores mostram que as mudanças nas revisões online estão associadas às mudanças nas vendas. Além da medida quantitativa das revisões (notas), os consumidores também se atentam para os aspectos qualitativos da revisão, como a qualidade do revisor e a exposição do mesmo. As reações dos consumidores são mais fortes para os itens com menos cobertura, ou seja, quando há poucas revisões existentes. O impacto das revisões online nas vendas decresce de acordo com a idade do produto.

Focando no comportamento do consumidor Moretti (2011) utiliza modelos de aprendizagem social cuja decisão de compra de um indivíduo depende da informação que ele recebe de seus pares. O modelo conta com uma dinâmica que varia de acordo com a expectativa de venda, medida pelo número de cinemas na data de estreia. Se um filme é melhor do que se espera, ou seja, surpreende positivamente às expectativas do público, a

bilheteria aumenta mais do que o esperado e ocorre um tipo de atualização sobre a qualidade do filme, que eleva ainda mais a bilheteria da segunda semana. A diferença na qualidade entre o esperado e o real é mais impactante para filmes com baixas expectativas antecedentes. Esta surpresa positiva é maior entre os consumidores que possuem grandes redes sociais. Isso implica a existência de um grande "multiplicador social" de tal forma que a elasticidade da demanda agregada à qualidade do filme é maior do que a elasticidade da procura individual de qualidade do filme.

No âmbito da análise de sobrevivência no mercado de cinema em geral as pesquisas são bem menos numerosas. De Vany e Walls (1997) examinam a relação entre práticas contratuais, o risco e a sobrevivência de distribuição dos filmes no mercado americano de cinema. Há apenas informação sobre a quantidade de salas de cinema, o número de semanas exibidas, o ranque e o faturamento. Os resultados indicam que a taxa de falha de um filme é dependente do tempo e o tempo de sobrevivência está fortemente relacionado ao número inicial de tickets da primeira semana. No ano seguinte, ainda Walls (1998) investiga as influências na sobrevivência de um filme em cartaz no mercado de cinema em Hong Kong. Os resultados obtidos apontam que o retorno financeiro das primeiras semanas é estatisticamente e economicamente significativo para determinar a expectativa de vida do filme em cartaz.

Outro trabalho de De Vany e Walls (1999) investiga estratégias que diminuiriam o risco de um filme 'falhar', focando no elenco de atores famosos, através da análise de sobrevivência. Os autores reforçam que não existe um filme típico porque o faturamento não converge em uma média, eles divergem em todas as escalas. Por fim, concluem que "é a audiência que faz do filme um hit e nenhuma quantidade de 'estrelões' ou publicidade pode alterar isso" (DE VANY e WALLS, 1999, pg. 285).

Com foco nas nomeações e premiações, Deuchert e Pauly (2005) pesquisam se a nomeação ou o recebimento de um Oscar realmente aumentam os lucros com as bilheterias e se afetam o tempo de exibição do filme nos cinemas. Encontram, porém, que receber o Oscar não é significativo para o tempo de exibição, mas a nomeação em melhor filme e melhores atores/atrizes no papel principal é. Notam também que filmes que ganham o Oscar são consideravelmente exibidos por mais tempo do que a média geral dos demais filmes. McKenzie (2009) verifica o efeito de algumas variáveis na duração da exibição do filme em cinemas Australianos de 2000 a 2005. O autor encontra que *previewing*, propaganda, revisões críticas e a arrecadação americana possuem influência positiva no tempo de duração enquanto orçamento, elenco e número de telas na primeira

semana não exerce influência. Das 28 variáveis utilizadas, 17 mostraram ter alguma influência na performance de sobrevivência.

Fu (2009) pesquisa os efeitos de integração vertical e controle horizontal nas exhibições de cinemas concorrentes considerando diferentes formações de cinema em Cingapura. O autor encontra que os filmes possuem maior duração em cinemas independentes e cadeias pequenas do que em grandes cadeias ou cinemas que são integrados às distribuidoras. Também focando nas decisões do distribuidor, Chisolm e Norman (2006) analisam a estratégia de término de exibição de um filme com dados de escolhas dos exibidores para estimar a função de sobrevivência dos filmes nos cinemas de Boston. Os resultados mostram que as escolhas de continuar ou terminar a exibição de um filme em cartaz sofre influência da performance interna da empresa e da pressão da competição externa à empresa. Observam também que os cinemas aumentam o tempo de exibição dos filmes em uma possível competição com outros cinemas das vizinhanças. Adicionalmente, dos modelos estimados 5 das 5 variáveis utilizadas exercem alguma influência, destacando o ranque e o elenco com influência negativa e a quantidade de telas na primeira semana e a bilheteria do primeiro final de semana com influência positiva.

Mais recentemente outro artigo relacionado ao tema foi encontrado. Legoux, Larocque, Laporte, Belmati e Boquet (2016) propõem demonstrar que os exibidores, os cinemas, são influenciados pelas revisões críticas em suas decisões de programação após a estreia de um filme. Mais especificamente, testam o efeito das revisões críticas na decisão dos exibidores para manter ou retirar um filme de cartaz dos seus cinemas semana após semana para o mercado cinematográfico canadense, exclusivamente Quebec. As revisões críticas de críticos profissionais foram extraídas do RottenTomatoes (aprovação geral dos críticos em %) e as revisões críticas dos consumidores do IMDb. Os resultados obtidos apontam que um filme com revisões críticas excelentes tem maiores chances de permanecer em cartaz por mais tempo em comparação a filmes com revisões críticas muito ruins, ruins ou até mesmo boas.

Nota-se, portanto, que os artigos citados sobre análise de sobrevivência em cinema diferem em algumas ou muitas partes do que é proposto neste trabalho. As pesquisas no mercado cinematográfico americano (DE VANY e WALLS, 1997 e 1999; CHISOLM e NORMAN, 2006; DEUCHERT e PAULY, 2005) divergem principalmente por não utilizarem das variáveis de qualidade propostas por este trabalho, como as resenhas críticas de profissionais e consumidores, ou seja, não investigam o boca a boca. Por outro lado, McKenzie (2009) e Legoux et al (2016) utilizam revisões críticas de

especialistas. McKenzie (2009) não controla a possível relação de viés por omissão de variável relevante e tem seu mercado alvo na Austrália. Legoux et al (2016) investiga as influências do ponto de vista do cinema, com dados agregados neste nível sobre o mercado Canadense, apesar de ser o mais próximo desta pesquisa, em termos de metodologia.

A tabela 5 apresenta alguns dos dados sumários levantados dos trabalhos revisados. A partir desta, é possível identificar que o artigo de Eliashberg and Shugan, (1997) é o mais antigo do conjunto e representa um precursor sobre o tema, sendo citado por quase todos os demais artigos. Os estudos de Kim *et al.* (2013) e Legoux *et al.* (2016) são os mais recentes, evidenciando a relevância do tema e do mercado de cinema e seu potencial de exploração, dado que o assunto está sendo pesquisado há pelo menos 17 anos e se mostra ativo.

Também é possível observar a pluralidade dos trabalhos relacionados, que apesar de possuírem o mesmo tema, divergem bastante entre foco, aplicabilidade e conclusões. Desta forma, constatou-se que não há um consenso empírico quanto aos resultados apresentados: alguns trabalhos mostram correlações positivas entre críticas e bilheteria dos filmes mas divergem entre especificidades como gênero, classificação MPAA, e orçamento. Outros trabalhos notificam a ausência de correlação estatística entre as críticas e a bilheteria, ou o peso das críticas online variando conforme a origem (críticos profissionais de renome, críticos profissionais e revisões de consumidores). Ainda, foram encontrados trabalhos que focam no volume, valência, densidade e qualidade das críticas sobre o produto, considerando que cada uma dessas características mudam o peso de como a crítica se relaciona com a variável de interesse.

A revisão possibilitou identificar também a lacuna de interdisciplinaridade que é proposta com o objetivo geral deste trabalho. Diversos trabalhos usam as classificações (*ratings* / *scores*) dos críticos profissionais e de consumidores como variável de controle para correlação estatística, cuja variável dependente é o faturamento, seja ele semanal ou doméstico total. Da mesma forma, outros trabalhos utilizam tais classificações para determinar influência sobre o tempo de duração de um filme nos cinemas. A lacuna parte do princípio que, dentre os trabalhos revisados, não se tem esse cruzamento de dados, utilizando a análise afetiva de texto para construir a variável a ser utilizada nos modelos correlacionais, independente do seu método, constituindo desta uma contribuição.

Tabela 5 - Trabalhos revisados

Trabalho	Revisões Críticas		Variável Dependente	Mercado	DataSet	Computação
	Profissionais	Consumidores				
Eliashberg & Shugan (1997)	Sim	Não	Faturamento	Americano	172 Filmes	Não
King (2007)	Sim	Não	Faturamento	Americano	152 Filmes; 2003	Não
Reinstein & Snyder (2005)	Sim	Não	Faturamento	Americano	609 Filmes	Não
Basuroy et al (2003)	Sim	Não	Faturamento	Americano	200 Filmes	Não
Boatwright et al (2007)	Sim	Não	Faturamento	Americano	466 Filmes	Não
Dellarocas e Ritu (2006)	Não	Sim	Faturamento	Americano		Não
Chintagunta et al (2010)	Não	Sim	Faturamento	Americano	148 Filmes	Não
Duan et al (2008)	Não	Sim	Faturamento	Americano	71 Filmes	Não
Liu (2006)	Não	Sim	Faturamento	Americano		Não
Rui et al (2013)	Não	Sim	Faturamento	Americano	63 filmes; 2009/2010	Análise Sent
Duan et al (2008b)	Não	Sim	Faturamento	Americano		Não
Yuan-Yuan et al (2008)	Não	Sim	Faturamento	Americano	200 Filmes	Não
Hu et al (2008)	Não	Sim	Faturamento	Americano		Não
Moretti (2011)	Não	Sim	Faturamento	Americano	4992 Filmes	Não
Moon et al (2010)	Sim	Sim	Faturamento	Americano	246 Filmes	Não
Kim et al (2013)	Sim	Sim	Faturamento	Americano	169 filmes; 2008	Não
Vany & Walls (1997)	Não	Não	Tempo	Americano	350 filmes; 1985/1986	Não
Walls (1998)	Não	Não	Tempo	Hong Kong	493 filmes; 1994/1996	Não
Vany & Walls (1999)	Não	Não	Tempo	Americano		Não
Deuchert e Pauly (2005)	Não	Não	Tempo	Americano	2.244 filmes; 1990/2000	Não
Fu (2009)	Não	Não	Tempo	Cingapura	360 filmes; 2002/2003	Não
Chisolm e Norman (2006)	Não	Não	Tempo	Boston		Não
McKenzie (2009)	Sim	Não	Tempo	Austrália	360 filmes; 2000/2005	Não
Legoux et al (2016)	Sim	Sim	Tempo	Quebec	788 filmes; 2002/2011	Não

Fonte: Thaís Luiza Donega e Souza, 2017

3.3 Considerações Finais

O mercado de cinema americano movimenta bilhões de dólares por ano e milhões de empregos diretos e indiretos. Considerada como bem de experiência, a indústria cinematográfica é de difícil avaliação, sendo necessário primeiro o consumo para somente depois julgar a qualidade. Para que o consumidor tome sua decisão de compra sem dispor do conhecimento sobre a qualidade do filme, os cinéfilos procuram por informações que possam fornecer algum parecer desta, como críticas de críticos profissionais e de seus pares, outros consumidores. Tal característica faz com que investigar o que torna um filme um sucesso ou um fracasso seja um desafio interessante que vem sendo estudado a longo prazo.

Outro problema em potencial com o uso de revisões agregadas na predição de sucesso para um novo produto é a potencial endogeneidade nas críticas por elas mesmas. Por exemplo, um filme baseado num roteiro excelente por um autor famoso e utilizando de atores populares liderados por diretores estrelados é mais propenso a ser revisado devido ao seu alto apelo. Consequentemente, talvez possa-se notar uma forte correlação entre os resultados da bilheteria e o volume e a valência média das críticas, não pelo poder preditivo dos críticos e sua influência, mas por causa da correlação de um fator comum – o apelo global do filme. Filmes melhores vendem mais tíquetes cumulativamente e tem vendas maiores em ambos períodos, mais cedo e mais tarde no ciclo de vida de exibição. Um modelo que ignora esta correlação talvez não tenha o poder para descobrir o papel real dos críticos (BOATWRIGHT, 2007). Deste modo, identificou-se que a maioria das metodologias revisadas não conseguem tratar o viés causado pelo efeito de predição, enquanto a proposta neste trabalho é plausível para dissolver este problema, uma vez que se controla o efeito de viés de omissão de variável relevante.

Através da revisão bibliográfica realizada foi possível identificar as diferentes técnicas, métodos, limitações e lacunas atuais. A maioria dos artigos utilizam as críticas online de críticos profissionais e de comunidades amadoras como variável de controle em seus modelos de regressão (personalizados ou não) para identificar correlação estatística. Outros trabalhos investigam o efeito do boca a boca, online ou não, sobre o faturamento da bilheteria dos cinemas. Dos estudos revisados, apenas três utilizam críticas tanto de especialistas quanto de consumidores e apenas um destes o faz para identificar efeito sobre o tempo de sobrevivência do filme em cartaz, qual tal propomos aqui. Ainda como

resultado da análise, mostramos que mesmo se tratando de um tema relativamente popular e com uma boa abrangência de estudos, se mostra ativo e com potencial de exploração nas mais diversas aplicações, peculiaridades e interdisciplinaridades.

Contudo, identificou-se também que nenhum destes trabalhos utiliza técnicas computacionais para criação da variável de boca a boca conforme proposto aqui, através da computação afetiva e com maior nível de informação para as análises (Tabela 5).

Com isso, atingimos os objetivos iniciais sobre identificar e obter um panorama geral sobre o papel dos revisores críticos no mercado de bens de experiência, sobre os métodos e técnicas existentes para mensurar, qualificar e correlacionar o sucesso de produções cinematográficas e sobre a utilização da computação afetiva como ferramenta auxiliar na identificação de sucessos e fracassos de produções de cinema.

4 Dados

Com o propósito de complementar a literatura atual, os dados utilizados para análises foram coletados exclusivamente para esta finalidade, constituindo um banco de dados extremamente abrangente em quantidade e variedade do conteúdo de informações. Os dados foram coletados de diferentes sites da Internet, conforme descrito abaixo, e contém informações financeiras sobre aproximadamente 13.000 filmes lançados no mercado americano entre 1980 e 2015. As revisões críticas de consumidores e críticos profissionais somam 4.150.587 e 835.578, respectivamente, para aproximadamente 12 mil filmes dos quase 13 mil da amostra.

Tais dados foram coletados dos sites BoxOfficeMojo.org (Box Office Mojo) de propriedade da Internet Movie Database, e RottenTomatoes.com (Rotten Tomatoes), popular agregador de críticas de cinema, centralizando críticas de diversos autores certificados em múltiplas associações e publicações de consumidores. Ambos sites são líderes em conteúdo do mercado cinematográfico americano e são regularmente utilizados em outras literaturas (MOON et al., 2010; DUAN et al., 2008; PANG e LEE, 2004). Do site BoxOfficeMojo.org foram colhidas as informações técnicas de produção e exibição dos filmes e dados econômicos e financeiros. As informações de críticas de críticos profissionais (tops e não tops) e críticas não profissionais (de consumidores registrados no site) foram coletadas do site RottenTomatoes.com que, por sua vez, reúne tais críticas de uma coleção de outros sites.

Buscou-se coletar todas as informações com maior nível de estratificação possível, o que permite maior variabilidade a ser explorada nas estimativas. Ou seja, a coleta se fundamentou na maior granularidade disponível, permitindo agregações conforme as necessidades de cada especificação.

Este capítulo descreve sobre o banco de dados coletado, como esta coleta foi realizada, a escolha dos dados, a montagem do conjunto de dados para cada contexto, as especificações do conjunto final e a contribuição deste na literatura atual.

4.1 Visão Geral

Segundo o relatório anual de estatística do mercado cinematográfico americano da MPAA (Motion Picture Association of America, 2013), a arrecadação total de

bilheteria por todos os filmes em cada país ao redor do mundo foi de aproximadamente US\$36.4 bilhões.

A economia movimentada nos Estados Unidos e Canadá na bilheteria dos cinemas foi de aproximadamente US\$10.4 bilhões, representando 28,58% da arrecadação global. Além de deter a maior parte do faturamento mundial, os Estados Unidos e sua fábrica de filmes Hollywood, compreendem cerca de 85% da produção cinematográfica de amplo lançamento no mundo. A economia de toda produção cinematográfica estadunidense vai além das bilheterias e movimenta aproximadamente US\$113 bilhões e 1.9 milhões de postos de trabalhos diretos e indiretos (MPAA 2013).

Além do impacto financeiro do cinema no mercado americano, optou-se por utilizar os dados estrangeiros devido alguns fatores. Poucas produções cinematográficas conseguem realizar estreias mundiais, ou seja, produções que não são lançados simultaneamente através do mundo. O site IMDB (www.imdb.com) mostra, em sua seção de estreias, as datas de lançamentos por país. É possível observar que as estreias acontecem primeiro majoritariamente nos Estados Unidos para então estreiar em outros países. Este atraso no lançamento acaba por ser mais lento que a velocidade da informação que a Internet proporciona. Desta forma, as críticas de profissionais e consumidores chegam primeiro aos mercados mundiais que o próprio filme. Esta assimetria de informação pode gerar um mecanismo influenciador no comportamento dos críticos e consumidores de outros países, tornando a mensuração de influência sobre o consumo mais complexo de ser avaliado.

Adicionalmente, o mercado americano de cinema possui uma extensa quantidade de dados colecionáveis e rotulados, como os obtidos nos sites BoxOffice e RottenTomatoes. No site fonte de dados financeiros, BoxOffice, é possível obter informações da arrecadação da bilheteria internacional, porém as informações agregadas não são tão detalhadas e a informação da arrecadação da semana de estreia para o Brasil, por exemplo, frequentemente não corresponde à data de lançamento do filme fixado em sua ficha técnica.

Note que os sites são norte-americanos e possuem seus dados dispostos de forma a refletir a economia cinematográfica americana. Para o caso de filmes estrangeiros estreados primeiro fora dos Estados Unidos, a contabilidade das exhibições e da data de estreia é feita dentro do mercado estadunidense. No entanto, filmes estrangeiros, no geral, compõem uma pequena parcela da amostra até porque, como dito anteriormente, os Estados Unidos são responsáveis por 85% da produção cinematográfica mundial.

Outro fator considerado foi para a escolha dos dados foi a computação afetiva proposta. Trata-se, basicamente, de problemas computacionais de mineração de texto que, em geral, utilizam de dicionários auxiliares à categorização das palavras mineradas, bem como dependência de domínio e idioma. Tanto a análise de sentimento quanto a mineração de emoção, discutidos no capítulo 2, são dependentes do idioma e a ferramenta escolhida para realizar estas tarefas, o serviço de compreensão de linguagem natural do IBM Watson, suporta apenas conteúdo em inglês para mineração de emoção, tornando qualquer outra língua um desafio adicional e um esforço fora do escopo deste trabalho.

4.2 A Coleta

A coleta foi feita através de *scripts* desenvolvidos em linguagem Python com o auxílio da biblioteca *Scrapy*. A biblioteca em questão permite que seja construído um *spider* com *crawler* de modo a coletar conteúdo HTML delimitados por *xPath* e gerenciados por regras. A documentação da biblioteca auxiliar está disponível em <http://scrapy.org/> e os scripts de coleta estão disponíveis na íntegra no anexo C.

Web Crawling é o processo pelo qual se coleta as páginas da Web com o objetivo de indexá-las. O processo inteiro, que pode ser visto como um grafo transversal da Web, tem o objetivo de “crawlear”, ou seja, reunir de uma maneira rápida e eficiente o maior número de páginas possíveis de acordo com a estrutura de links que as interconecta (MANNING et al, 2008).

A operação básica de qualquer crawler é: inicia-se com uma ou mais URLs que constituem um conjunto de sementes (*seeds*); o crawler então, escolhe uma URL deste conjunto e busca a página Web correspondente à URL; a página buscada é analisada sintaticamente para extração de ambos os textos, visual e HTML, e outros links da página, que por sua vez apontarão para outras URLs. O texto extraído é alimentado em um indexador e os links extraídos são adicionados à URL *frontier*, que consiste nas URLs cujas páginas correspondentes já foram buscadas pelo crawler. Inicialmente, a URL *frontier* contém o conjunto de URLs iniciais – as sementes. Conforme as páginas vão sendo buscadas, as URLs correspondentes são deletadas da lista de URLs *frontier* (MANNING et al, 2008).

Quadro 2 - Exemplos de URLs do BoxOfficeMojo.com

Exemplo: `www.boxofficemojo.com/*`

1: `www.boxofficemojo.com/yearly/`

2: `www.boxofficemojo.com/yearly/chart/`

3: `www.boxofficemojo.com/alltime/`

4: `www.boxofficemojo.com/alltime/domestic/`

5: `www.boxofficemojo.com/alltime/world/`

6: `www.boxofficemojo.com/movies/`

A partir da indexação de todos os links a serem seguidos o *crawler* começa a acessar e varrer as páginas elencadas. Desta forma, torna-se um *spider* e faz o download das informações especificadas. Esse “download”, ou seja, a ação de coletar as informações não estruturadas das páginas acessadas, é chamada de *scrap*.

Para a coleta no site Box Office Mojo, inicialmente foram utilizadas as tabelas de lançamentos dos filmes por ano (`http://www.boxofficemojo.com/yearly/`). A partir dos links coletados nas tabelas dos anos 1980 a 2015, montou-se um *spider* inicial, capaz de visitar cada um dos links e realizar o *scrapping* das informações solicitadas. O link, que possui expressão padrão, identificado por “*movie_id*”, (`http://www.boxofficemojo.com/movies/?id=movie_id.htm`), direciona o *spider* para a página de dados sumários do filme (apresentados na tabela 7), montando assim a tabela BOXOFFICE_SUMMARY. Com os mesmos links, o *spider* foi redirecionado pelo *script* para seguir a página com terminação *weekend* – `http://www.boxofficemojo.com/movies/?page=weekend&id=movie_id.htm`. Destas páginas foram coletadas as informações da bilheteria por semana, variáveis também representadas na tabela 7, constituindo a tabela de dados BOXOFFICE_WEEKEND. Ainda, de forma semelhante, o *spider* acessa e coleta os dados do link `http://www.boxofficemojo.com/movies/?page=week&id=movie_id.htm`, de terminação *week*, montando assim a tabela BOXOFFICE_WEEK.

A tabela 6 resume as informações dadas acima, apresentando a página a ser acessada, o link *crawlado* pelo *spider* e o elemento indexado pelo *crawler* a partir da coleta inicial dos filmes lançados. Neste caso, os índices consistem da informação de identificação do filme, chamada pelo *xPath* do BoxOffice como *movie_id*.

Tabela 6 - Acessos do Web Crawler

Página	Link Crawlado	Indexado
Summary	http://www.boxofficemojo.com/movies/?id=	<i>movie_id</i>
Weekend	http://www.boxofficemojo.com/movies/?page=weekend&id=	<i>movie_id</i>
Week	http://www.boxofficemojo.com/movies/?page=week&id=	<i>movie_id</i>

Fonte: Thaís Luiza Donega e Souza, 2016

Com base nos títulos dos filmes e o controle do ano de lançamento, visto que existem filmes com nomes iguais, montou-se uma listagem de links a serem visitados pelo *web spider* no site de críticas RottenTomatoes. De forma semelhante à coleta realizada no site BoxOffice, primeiro coletou-se as informações gerais de avaliação do filme, como porcentagens, ícones de qualidade, volume de resenhas críticas, entre outros, formando a tabela ROTTEN_MOVIES. Com o link do filme já coletado (https://www.rottentomatoes.com/m/movie_id/) o *spider* foi direcionado primeiramente ao link da página de revisões críticas de críticos profissionais – https://www.rottentomatoes.com/m/movie_id/reviews/. Desta página obteve-se as resenhas críticas dos críticos profissionais e os dados foram armazenados na tabela ROTTEN_REVIEWS. Analogamente, a composição de dados da tabela ROTTEN_AUDIENCE iniciou-se do link original do filme e se seguiu pela terminação padrão “*?type=user*”, adquirindo assim as críticas dos consumidores registrados no site.

Tanto para as críticas profissionais quanto para as críticas dos consumidores o *spider* foi programado para acompanhar todas as páginas existentes dos links perseguidos. Assim, todas as páginas de resenhas existentes no site foram acessadas e coletadas. Contudo, houve uma limitação na quantidade de informações disponíveis, pois o RottenTomatoes mantém online a quantidade máxima de 50 páginas de revisões críticas de consumidores, mesmo que existam mais críticas.

De posse destas 6 tabelas foi montado um banco de dados de forma a auxiliar o cruzamento entre as informações dos filmes correspondente entre as fontes. A figura 4 mostra o diagrama entidade-relacionamento entre estas tabelas, que serviu como base para criação do dataset final, conforme necessidade e especificação. A seção 4.3 e 4.4 descrevem, em mais detalhes, o conjunto de dados montado para cada cenário, computação afetiva e economia, respectivamente.

Figura 4 - Diagrama Entidade-Relacionamento do banco de dados criado

Fonte: Thaís Luiza Donega e Souza, 2016

4.3 Análise descritiva dos dados

A tabela 7 traz a análise descritiva dos dados quantitativos com as variáveis, separadas pelas fontes de obtenção, suas descrições, número de observações, média, desvio padrão, mínimo e máximo.

A primeira parte da tabela 8 sumariza as informações de positividade e negatividade das críticas dos críticos profissionais. Os dados apresentados como “Sem Avaliações” são de filmes que não obtiveram nenhuma crítica para que fosse feita a classificação. O RottenTomatoes considera a classificação de um filme, em relação às críticas profissionais, como sendo Muito Bom (*Certified Fresh*), Bom (*Fresh*) e Ruim (*Rotten*). A classificação “Muito Bom” ocorre quando a porcentagem de revisões positivas de críticos profissionais ultrapassa 75%, considerando uma quantidade mínima de críticas de 80 para críticos não *tops* e com ao menos 5 críticas de críticos profissionais credenciados como *top*. Para o rótulo “Bom” o filme deve que receber 59% ou mais de críticas positivas e o rótulo Ruim 59% de críticas positivas ou menos. Toda a contagem é feita sob as revisões positivas, representadas por um ícone de tomate vermelho. A classificação sumarizada do RottenTomatoes é dinâmica no tempo, logo, se as críticas melhoram ou pioram com o passar do tempo a classificação geral do filme acompanha a nova métrica.

Tabela 7 - Análise descritiva dos dados

Descrição da Variável	Obs	Média	Desvio	Mín	Máx
Box Office¹ Summary					
Filmes	13.698				
Escritor(es)	4.827				
Ator(es)	4.885				
Diretor(es)	7.208				
Produtor(es)	2.778				
Distribuidor	13.698				
Gênero	13.698				
Classificação MPAA	13.698				
Data de Lançamento	13.698				
Orçamento (USD em milhões)	2.881	43,489	43,868	0,001	300
Bilheteria (USD em milhões)	13.698	17,6	43	0,00003	879,3
Box Office¹ Weekends					
Número de Finais de Semana	111.247	19,528	57,571	-1	790
Faturamento por Final de Semana (USD em milhares)	111.247	1.449	5.285	0,006	248
Rank de exibição nacional por Final de Semana	104.537	50,689	36,874	1	162
Telas de cinemas em exibição do filme por Final de Semana	111.229	457,81	823,19	1	4468
Número de Tickets por Final de Semana (em milhares)	111.247	267,62	857,33	0,1	28.501
Box Office¹ Weeks					
Número de Semanas	100791	17,377	51,886	-1	789
Faturamento por Semana (USD em milhares)	100791	2.210	7.737	0,001	390.856
Rank de exibição nacional por Semana	104.537	50,689	36,874	1	162
Telas de cinemas em exibição do filme por semana	111.229	457,81	823,19	1	4468
Número de Tickets por Semana (em milhares)	100.791	361,90	1.163,18	0,1	44.926
Rotten² Tomatoes Movies					
Filmes	12813				
Aprovação Geral dos Críticos (%)	9864	59,449	27,054	1	100
Aprovação Geral da Audiência (%)	12151	61,691	20,723	1	100
Avaliações negativas de críticos (Rotten)	10596	22,133	28,357	1	195
Avaliações positivas de críticos profissionais (Fresh)	10684	34,669	43,233	1	316
Total de Críticas dos críticos profissionais (Volume)	11252	53,762	57,470	1	341

RottenTomatoes: rottentomatoes.com;² BoxOffice: boxoffice.com

Fonte: Thaís Luiza Donega e Souza, 2017

A segunda parte da tabela 8 apresenta a frequência da avaliação dos consumidores, denotados como audiência pelo RottenTomatoes. As classificações de Bom e Ruim são menos severas para a audiência do que para os Críticos. Para conseguir um rótulo de “Bom”, o filme deve receber avaliações que ultrapassem 3.5 estrelas na média e, para “Ruim”, a média das avaliações dos consumidores deve ser menor que 3.5. Para a avaliação geral dos filmes quanto a audiência, o RottenTomatoes une as resenhas recebidas por eles e por outro site, também agregador de revisões críticas de consumidores, o Flixter (flixter.com). Do mesmo modo, a classificação apresentada como “Sem Avaliações” significa que o filme não possui nenhuma resenha para que seja feita a média das notas.

A tabela 9 mostra a sumarização de revisões de críticos profissionais separadas pela rotulação dada pelo crítico no momento da resenha e pela posição profissional do crítico profissional junto ao site, se top ou não. A tabela 10 apresenta a sumarização das resenhas dos usuários, divididos conforme sua variável de avaliação numérica.

Tabela 8 - Tabulação avaliação críticos e audiência por filme

Descrição da Variável	Frequência	Percentual	Acumulado
Avaliação Críticos¹			
Muito Bom	2.012	15,70	15,70
Bom	3.371	26,31	42,01
Ruim	4.643	36,24	78,25
Sem Avaliações	2.787	21,75	100,00
Total	12.813	100,00	
Avaliação Audiência¹			
Bom	6.504	50,76	50,76
Ruim	5.442	42,47	93,23
Sem Avaliações	867	6,77	100,00
Total	12.813	100,00	

RottenTomatoes: rottentomatoes.com

Fonte: Thaís Luiza Donega e Souza, 2017

Tabela 9 - Avaliação dos críticos por crítica

Descrição da Variável		Frequência	Percentual	Acumulado
Avaliação Críticos¹				
Bom	<i>Top</i>	121.045	15,21	15,21
	<i>Não Top</i>	366.311	46,02	61,23
Ruim	<i>Top</i>	87.126	10,95	72,18
	<i>Não Top</i>	221.486	27,82	100,00
		795.968	100,00	

RottenTomatoes: rottentomatoes.com

Fonte: Thaís Luiza Donega e Souza, 2017

Tabela 10 - Avaliação dos consumidores por crítica

Descrição da Variável	Frequência	Percentual	Acumulado
Avaliação Consumidores¹			
0	3.472	0,08	0,08
0.5	408.597	9,84	9,93
1	135.455	3,26	13,19
1.5	122.806	2,96	16,15
2	237.109	5,71	21,86
2.5	271.744	6,55	28,41
3	545.105	13,13	41,54
3.5	566.687	13,65	55,20
4	738.849	17,80	73,00
4.5	392.509	9,46	82,45
5	728.254	17,55	100,00
	4.150.587	100,00	

RottenTomatoes: rottentomatoes.com

Fonte: Thaís Luiza Donega e Souza, 2017

Com esses dados primários reunidos, alguns tratamentos foram realizados de forma a viabilizar as análises a serem feitas, como o cruzamento entre os dados financeiros, oriundos do BoxOffice e as revisões críticas, provenientes do site RottenTomatoes. Portanto, todas as informações de cada filme precisam estar disponíveis em ambos os sites de coleta. Este cruzamento de informações, no entanto, reduziu o número de observações, uma vez que o conjunto de filmes avaliados por críticos e consumidores é um subconjunto dos filmes que foram estreados nas salas de cinema estadunidenses.

Ainda, apesar dos mais de 13 mil filmes coletados, como mostra a tabela 7, nem todos estão completos. Ou seja, para que o filme seja considerado completo seus finais de semana e suas semanas de exibição precisam ser contínuas, com visibilidade de começo (primeira semana) e fim (última semana). Filmes que possuem lacunas entre as semanas de exibição foram excluídos da amostra.

Para cada modelo especificado na metodologia (Capítulo 1) considerou-se a disponibilidade de informações financeiras, técnicas e de avaliação crítica de críticos profissionais e de consumidores e, desta forma, o volume de dados da amostra foi reduzido substancialmente. A quantidade final de observações utilizada nas estimativas, bem como as variáveis empregadas, estão detalhadas no anexo B.

4.4 Contexto: Computação Afetiva

Na mineração de texto e, por consequência, no processamento de linguagem natural, os conjuntos de dados são, em geral, conhecidos como corpus. Um dos recursos mais importantes para diversas tarefas de Processamento de Linguagem Natural é o corpus: um conjunto de textos que podem estar pré-processados e, geralmente, rotulados, dependendo do fim para o qual o corpus foi compilado. O método escolhido, já discutido no capítulo 3, permite as análises afetivas sem a necessidade de rótulos, uma vez que o IBM Watson já é treinado por *deep learning* e é capaz de reconhecer o domínio do corpus e o idioma.

Apesar de ser um agregador de revisões de consumidores, o RottenTomatoes aceita apenas comentários de usuários devidamente registrados, o que favorece um certo controle contra *spams*. Ainda, a estrutura e a proposta do site auxiliam para que as resenhas tenham qualidade de escrita, apresentando menos vícios de linguagem, gírias e acrônimos.

No contexto da análise de sentimento e mineração de opinião em revisões críticas de cinema, dois conjuntos de dados frequentemente utilizados são os de Pang e Lee (2002 e 2004). O *dataset* de Pang e Lee (2002) foi coletado do IMDb (Internet Movie Database), proprietário do RottenTomatoes.com e conta com 752 revisões negativas e 1301 revisões positivas. As classificações foram extraídas e convertidas em três categorias: positivas, negativas e neutras. O corpus está disponível em <http://www.cs.cornell.edu/people/pabo/-movie-review-data/>. O conjunto de dados de Pang e Lee (2004) consiste em 1000 resenhas positivas e 1000 negativas coletadas do site rottentomatoes.com. O conjunto de dados está disponível no mesmo site citado acima. Ambos são usualmente adotados como referências para comparação da eficácia dos métodos testados. Nota-se então a extensão do banco de dados criado exclusivamente para este trabalho.

Para a mineração de emoção em filmes poucos trabalhos foram encontrados, não tendo, portanto, uma base comparativa, apesar de que os dados utilizados para análise de sentimentos podem ser também empregados na mineração de emoção, desde que o contexto seja não supervisionado. A maioria dos trabalhos em computação afetiva coleta suas informações das mais variadas fontes e assuntos, principalmente daquelas onde a rotulação dos dados é facilitada ou pronta, como Facebook e Twitter (MOHHAMMAD,

2012; SARAVIA, ARGUETA e CHEN, 2015; DENG, WANG, LI e XU, 2015; CHUNG, HE, ZENG DAJUN e BENJAMIN, 2015; REN e KANG, 2013). O único trabalho da revisão bibliográfica de computação afetiva a utilizar revisões críticas de filmes é o de Subasic e Huettner (2001), que aplicam a mineração de emoção em conjuntos de revisões críticas para 5 sets de 15 filmes cada. O objetivo e os métodos dos autores, no entanto, são diferentes dos propostos aqui.

Para a computação afetiva foi feito um recorte dos dados coletados, de forma a utilizar somente as críticas dos consumidores. Conjunto de dados este, que foi processado pelo IBM Watson.

Das mais de 4 milhões de revisões críticas totais fez-se primeiramente um cruzamento entre o filme que recebeu as revisões no RottenTomatoes e a sua existência financeira no BoxOffice, uma vez que as estimativas não são independentes entre informações sobre a qualidade do filme (revisões críticas) e seus dados técnicos e financeiros. Além disso, foi necessário um corte no tempo que se dá a partir de 2004 pela disponibilidade de revisões críticas dos consumidores. O site RottenTomatoes, provedor das informações coletadas, só dispõe da participação da audiência após 2003, impossibilitando, portanto, a construção da variável proxy proposta para períodos anteriores a este. Este fator de datas também impediu que fossem feitas avaliações dos efeitos provocados pelo período após a propagação da Internet banda larga nos estados unidos, em 2004, visto que não se dispõe de informações anteriores para avaliação comparativa.

Adicionalmente, por um controle de efeitos, foi feito um outro corte nas revisões críticas dos consumidores de modo a limitar as avaliações. Foram consideradas apenas as revisões críticas com datas de publicação de até um ano após a data de lançamento do filme, uma vez que revisões críticas feitas muito tempo depois da saída do filme do cartaz dos cinemas não pode mais influenciar outros consumidores no mercado de cinema, objetivo proposto aqui, contudo o prazo estendido de até um ano é útil para o propósito de sinalização da qualidade do filme.

Desta forma, o conjunto de dados preparado para ser processado pela ferramenta de compreensão de linguagem natural do IBM Watson e, conseqüentemente, obter a análise afetiva, possui 777.951 registros, cada um deles contendo uma crítica a ser analisada. Os resultados das análises foram então agregados por filmes e compõem o banco de dados resultante, com aproximadamente 68 mil observações dispostas por finais

de semanas de 2004 a 2015, que está tabulado na tabela 1 no anexo B juntamente com as demais variáveis utilizadas nas estimativas.

4.5 Contexto: Economia

No contexto da literatura de revisão crítica em economia, muitos trabalhos abordam o estudo de filmes, conforme discutido no capítulo 3.

Eliashberg e Shugan (1997) construíram seu banco de dados a partir de um conjunto de vendas de bilheteria que inclui faturamento da bilheteria, distribuidores, gêneros, faturamento semanal e cinemas em apresentação semanal de Baseline Inc. e Entertainment Data Incorporated (EDI). Os dados se concentram entre os filmes lançados em 1991 e 1992 e apenas os que possuem dados completos das duas semanas pós lançamento. Os dados de revisões críticas foram obtidos da revista Variety. O *dataset* total conta com 2104 revisões críticas de críticos profissionais para 172 filmes.

Boatwright et al. (2007) analisam uma amostra de 200 filmes lançados entre dezembro de 1997 e março de 2001. Para cada filme foram coletados o faturamento semanal e dados cinematográficos da Variety *magazine*. Outras variáveis presentes no *dataset* são publicidade, estrelas, classificação MPAA, sequências, *movie appeal* e revisões críticas de críticos profissionais. Reinstein e Snyder (2005) focam nas revisões críticas de apenas dois críticos durante o final de semana de lançamento do filme. O conjunto final de dados possui 609 filmes com suas informações semanais coletadas do site BoxOffice e do site de Siskel e Ebert. Moon et al. (2010) utilizam um conjunto de dados semanais que abrange 246 filmes nas seis maiores categorias de gêneros. Os filmes possuem seus lançamentos entre maio de 2003 e outubro de 2005 e foram coletados do Yahoo Movies e do Web site RottenTomatoes.

Muitos outros bancos foram utilizados para diversas análises, como mostrado na revisão bibliográfica da literatura apresentado no capítulo 3. Kamakura et al (2006) utilizam 21.436 críticas para os 466 filmes de dezembro de 1997 a março de 2001 do cinema americano coletados da revista Variety. Moretti (2011) possui em sua coletânea informações de 4,992 filmes de 1982 a 2000 com 8 semanas de informação sobre a bilheteria americana coletados de ACNielsen-EDI. Liu (2006) usam dados do Yahoo! Movies para o cinema americano com uma amostra de 40 filmes e 12 mil resenhas de maio a setembro de 2002. Basuroy et al (2003) coletaram os dados da revista Variety para

200 filmes do cinema americano. Yuan-yuan et al (2008) utilizam os dados do Yahoo! Movies e do BoxOfficeMojo para os 200 melhores filmes do ano no cinema Americano.

Nota-se, portanto, que o *dataset* aqui construído segue a modelagem semanal, de revisões críticas e de variáveis descritivas como os trabalhos supracitados, no entanto, o volume de informação aqui apresentado é significativamente maior do que todos eles. Além disso, muitos autores citados controlam seus dados por tempo de exibição, cortando a visibilidade completa do filme em cartaz. Devido ao método de análise de sobrevivência proposto no capítulo 1, é necessário que os dados estejam completos, sem cortes em suas exibições, o que torna a amostra ainda mais rica em termos de qualidade.

4.6 Considerações Finais

Dispõe-se aqui então, a obtenção de variáveis proxies de avaliação de consumo por pares antes não disponível para análises econômicas, de mercado, e muitas outras, usando técnicas computacionais e a proposição de trabalho genuinamente interdisciplinar.

Este capítulo apresentou a criação do banco de dados inédito e muito abrangente sobre o mercado americano de cinemas disponível prontamente para vários estudos. Considerando os alcances das revisões bibliográficas realizadas tanto no contexto da computação afetiva quanto no contexto de influência de bilheteria de cinema, não se encontrou nenhum banco de dados tão abrangente e rico como o construído para este trabalho, justificando o ineditismo do mesmo e a necessidade de sua construção. A construção de um banco de dados novo se justifica ainda pela importância em se ter informações sejam atuais, de acordo com o período tecnológico que se deseja estimar.

5 Resultados

Neste capítulo é apresentado os resultados preliminares, observáveis a partir da computação afetiva, as estimativas não paramétricas e paramétricas realizadas, conforme proposto nos métodos do capítulo 1. Serão abordadas as decisões tomadas, na medida do possível, para cada resultado apresentado, bem como as discussões possíveis em torno deles.

5.1 Resultados Preliminares

A partir dos dados apresentados na seção 4.3 do capítulo 4, alguns tratamentos adicionais foram feitos para viabilizar as estimativas dos modelos paramétricos e não paramétricos propostos (Capítulo 1).

Destes dados, alguns filmes também foram excluídos por serem considerados *outliers* ou não estarem, de alguma forma, nos padrões das variáveis de interesse. Os gêneros “IMAX” e “documentary” foram retirados da nossa amostra por apresentarem naturezas diferentes comparado aos outros gêneros. Os filmes IMAX podem permanecer em cartaz por anos e de maneira intermitente. O filme Space Station 3-D (IMAX), por exemplo, foi lançado em abril de 2012 e já havia sido exibido, até a data da coleta, em 658 finais de semana. Da mesma forma, os filmes classificados como documentários também não entram na amostra por caracterizarem viés educacional e, portanto, se também se comportam de forma diferente dos demais. Os filmes que não estão rotulados quanto ao seu gênero também foram excluídos da amostra para que não influenciassem nos resultados estimados por gêneros. Para controlar os gêneros dos filmes, 82 no total, foram criadas as variáveis categorias: action, comedy, drama, romance & family, foreign, crime & war & western, horror, thriller, adventure & animation e others. Cada categoria criada é uma dummy que representa se o filme pertence ou não àquela categoria, sendo “others”, os filmes que não se enquadraram às demais categorias criadas. Contudo, um único filme pode pertencer a mais de uma categoria, o gênero “Romantic Comedy”, por exemplo é categorizado em romance & family e comedy (Anexo B).

Outro fator considerado foram as revisões críticas dos consumidores. O site RottenTomatoes, apesar de sua inauguração ser datada de 1998, a área de acesso aos usuários e, consequentemente, as postagens das críticas teve início apenas em 2004, ano

coincidente com a Web 2.0. As críticas profissionais que eram preenchidas no site com datas anteriores à 1998 são provenientes de boletins, revistas e televisão, porém, para as revisões críticas dos consumidores, há apenas informação a partir de 2004. Com isso, foi feito um corte no conjunto de dados formatado até então, resultando em 42.274 observações para 4.779 filmes. O conjunto de dados resultante da análise afetiva realizada nas revisões críticas dos consumidores está disposto no anexo C.

A partir deste ponto, algumas observações acerca da análise afetiva foram feitas. A figura 5 apresenta a mineração de emoção (médias de *joy*, *anger*, *fear*, *disgust*, *joy* e *sadness*) das resenhas críticas dos consumidores, através da presença em cada um dos 10 gêneros classificados:

Figura 5 - Média de emoções por gênero

Fonte: Thaís Luiza Donega e Souza, 2017

Na figura 5 (a) tem-se a distribuição média da emoção *fear* ao longo dos gêneros. Nota-se que *fear* é maior para os gêneros Horror e Thriller, respectivamente, em comparação aos demais gêneros e a média geral da presença da emoção nas críticas dos consumidores fica abaixo de 0.2. Em (b) tem-se a medida de *sadness* obtida na mineração de emoção. Tal emoção possui uma variação maior, de 0.0 a 0.9 e uma presença mais forte nas revisões críticas. A média entre os gêneros é de aproximadamente 0.42 e não é possível identificar com clareza qual gênero possui maior presença de *sadness*, apesar de “Drama” ter os maiores ranges da emoção. Para *disgust*, a figura 5 (c), destacam-se os gêneros de “Crime / War / Western” e “Horror”, comumente filmes mais sangrentos. A média geral da presença de *disgust* fica próximo de 0.1. Em (d), tem-se a média de *anger*. “Action”, “Horror” e “Thriller” possuem as maiores médias de *anger*, porém, as maiores variações estão em “Foreign” e “Romance / Family”. A presença geral da emoção também fica próximo de 0.1. Ainda na figura 5, o gráfico (e) mostra a média de *joy*, que tem uma presença geral maior depois de *sadness*, com aproximadamente 0.33. “Adventure / Animation” e “Romance / Family” possuem as maiores presenças da emoção, que tem seus máximos em “Adventure / Animation” e “Drama”. É importante notar que o gênero “Horror” possui a menor presença de *joy* entre todos os gêneros analisados.

A figura 6 apresenta a média da pontuação ou score dado pelos consumidores na resenha crítica, que varia de 1 a 5 estrelas. A média dos *ratings* fica em 3.5 e destacam-se positivamente os gêneros “Foreign”, “Drama” e “Crime / War/ Western”. Essa média de avaliação dos consumidores é semelhante à média dos críticos profissionais (Figura 8), indicando que o score dado por um consumidor tem uma característica maior de avaliação crítica técnica do filme, do que sua opinião, como expressado através da emoção minerada. Por exemplo, sendo *joy* uma emoção de características positivas, os gêneros destacados com a maior presença desta emoção (“Adventure / Animation” e “Romance / Family”) deveriam ser também os melhores avaliados pelos consumidores. Essa variação pode indicar que ao classificar um filme através dos *ratings* obrigatórios, o consumidor faz sua avaliação baseando-se nos seus conhecimentos técnicos e se expressando quanto à qualidade geral do filme – se o filme é bom ou não, enquanto que na sua avaliação escrita, o consumidor enfatiza sua experiência de assistir ao filme – se ele gostou do filme ou não.

Fonte: Thaís Luiza Donega e Souza, 2017

A figura 7 apresenta a análise de sentimentos realizada nas revisões críticas dos consumidores. Separado por positivo e negativo, as médias representam a polaridade encontrada também de acordo com os gêneros, variando de 0 a 1 para positivos e 0 a -1 para negativos. No geral, os gêneros mais bem avaliados quanto à polaridade são os mesmos indicados na mineração da emoção *joy*, “Adventure / Animation” e “Romance / Family”, mas com a inclusão de “Comedy”, neste caso. Os gêneros que tiveram a presença da polaridade negativa mais significativamente foram “Horror”, “Fantasy” e “Thriller”. Apesar das médias não diferirem muito entre os gêneros, as negativas ficam em 0.50 e as positivas em 0.60, o que mostra que é mais provável que se fale bem de um filme, do que se fale mal. Da mesma forma, a quantidade de notas máximas positivas (1) é maior do que a quantidade de notas máximas negativas (-1).

A figura 8 mostra a média dos críticos profissionais – Tops e Não Tops – distribuídas por gênero. Estas variáveis foram calculadas a partir das avaliações positivas sobre as avaliações positivas e negativas dos críticos profissionais, retiradas do site do RottenTomatoes (Capítulo 4). As melhores avaliações se concentram nos gêneros “Foreign”, “Crime / War / Western” e “Drama” que, como dito anteriormente, se assemelha aos scores dados pelos consumidores. A média geral da avaliação é de 0.57 para os críticos cadastrados como Tops e de 0.61 para os críticos regulares. Nota-se, pela

diferença entre as médias do gráfico que os críticos considerados Tops são mais exigentes em suas avaliações e têm uma média de avaliação menor que a dos críticos profissionais regulares.

Figura 7 - Média do sentimento por gênero

Fonte: Thaís Luiza Donega e Souza, 2017

Figura 8 - Média das Avaliações dos Críticos

Fonte: Thaís Luiza Donega e Souza, 2017

5.2 Estimativas

Tem-se como objetivo, avaliar o efeito das revisões críticas dos consumidores, no sentido do boca a boca, sobre o tempo de duração do filme em cartaz nos cinemas estadunidenses. Adicionalmente, verificar o impacto das críticas dos profissionais sobre a mesma medida. Desta forma, para atender os objetivos, as variáveis de interesse são as revisões críticas dos consumidores e as críticas profissionais, uma vez que elas medem o efeito do boca a boca e sinalizam a qualidade dos filmes. As revisões críticas dos consumidores são medidas conforme mostra a tabela 11.

Tabela 11 - Variáveis de interesse consumidores

Variável	Obs	Média	SD	Mín	Máx
Media_audience	32.692	3,4191	0.5936	0.5	5
Geral_Audiencia (Bom/Ruim/Sem Avaliações)					
Audi_Bom	42.274	0,5245	0,4994	0	1
Audi_Ruim	42.274	0,3637	0,4811	0	1
Sentimento (Positivo/Negativo/Neutro)					
Senti_Bom	42.274	0,4938	0,4999	0	1
Senti_Ruim	42.274	0,1629	0,3693	0	1
Emoções (Anger / Joy / Fear / Sadness / Disgust)					
Media_anger	27.954	0.1205	0.0377	0.0025	0.5106
Media_joy	27.954	0.3271	0.0992	0.0019	0.9319
Media_fear	27.954	0.1348	0.0612	0.0027	0.7739
Media_sadness	27.954	0.4191	0.0702	0.0329	0.8833
Media_disgust	27.954	0.1102	0.0397	0.0044	0.5826

Fonte: Thaís Luiza Donega e Souza, 2017

A variável MEDIA_AUDIENCE é a média da pontuação – *rating score* – dada pelos consumidores no momento da revisão crítica. As medias AUDI_BOM e AUDI_RUIM são dummies construídas a partir da variável GERAL_AUDIENCIA, informação da classificação geral do filme atribuída pelos consumidores no site do RottenTomatoes e compreendem a classificação “Bom” e “Ruim”, respectivamente. Os filmes “Sem Avaliação” são a referência para tais dummies, visto que a ausência de um rótulo avaliativo faz com que o filme seja, automaticamente, não avaliado. Similarmente, as variáveis SENTI_BOM e SENTI_RUIM são dummies de positividade e negatividade obtidas pela análise de sentimento do IBM Watson, variável SENTIMENTO. As demais médias são medidas de intensidade das emoções anger, joy, fear, sadness e disgust,

obtidas pela mineração de emoção realizada nas revisões críticas dos consumidores, conforme discutido no capítulo 2. Ainda, conforme fundamentado no capítulo 2 e apontado por Hoelzl e Possas (2010), os indivíduos frequentemente consideram emoções durante a tomada de decisão de consumo, motivo este que inclui a mineração de emoção realizada neste trabalho como contribuição da computação aplicada.

Um consumidor que compartilha sua opinião na Internet pode ser selecionado de alguma maneira, provocando um viés de acordo com suas preferências de gênero do filme, suas características pessoais ou outros atributos. Contudo, utilizando a média de um grande volume destas revisões críticas, pode-se garantir uma boa medida de preferências espalhadas pelo efeito do boca a boca.

A sinalização de qualidade do filme dada pelos críticos profissionais são as variáveis descritas na tabela 12.

Tabela 12 - Variáveis de interesse críticos profissionais

Variável	Obs	Média	SD	Mín	Máx
Top_Media	32.497	0,5696	0,2723	0,025	0,98
NaoTop_Media	34.091	0,6038	0,2597	0,0114	0,99
Geral_Críticos (Muito Bom / Bom / Ruim / Sem Avaliações)					
Top_MuitoBom	42.274	0,3049	0,4604	0	1
Top_Bom	42.274	0,1511	0,3581	0	1
Top_Ruim	42.274	0,3657	0,4862	0	1

Fonte: Thaís Luiza Donega e Souza, 2017

Os críticos, no entanto, não atribuem notas como a audiência no site da fonte dos dados, RottenTomatoes. Cada revisão é rotulada entre “Bom” (*fresh*) e “Ruim” (*Rotten*) e, conforme explicado no capítulo 4, tais rótulos foram extraídos das revisões críticas dos críticos profissionais *tops* e dos críticos profissionais não *tops* na área dos *experts* no site. A média, TOP_MEDIA, consiste então na somatória das avaliações positivas – “Bom”, sobre o total de avaliações também tops – “Bom” e “Ruim”. Da mesma forma a NAOTOP_MEDIA foi construída a partir da avaliação positiva dos críticos não top. Para ambas as médias foram utilizadas as quatro variáveis seguintes: TOP_POSITIVO, TOP_NEGATIVO, NAOTOP_POSITIVO e NAOTOP_NEGATIVO, que são as somatórias das críticas positivas e negativas, tops e não tops, respectivamente (Anexo B).

A avaliação geral dos críticos, calculada pelo RottenTomatoes de acordo com as notas dadas em suas resenhas, compreende a variável GERAL_CRITICOS, que foi separada em três dummies: TOP_MUITOBOM, TOP_BOM, e TOP_RUIM. Os filmes

“Sem Avaliação” são a referência para tais dummies, uma vez que o filme que não consta um rótulo avaliativo é, automaticamente, não avaliado.

Para controlar as aberturas frias apresentada por Brown, Camerer e Lovallo (2012), quando produtores tornam seus filmes disponíveis aos críticos somente após a estreia em cartaz, separamos todas as variáveis de revisões críticas em cinco tempos: antes da estreia, durante a estreia, antes e durante a estreia, depois da estreia e o período completo. Os tempos e as nomenclaturas estão demonstrados na figura 9 e seguem o modelo de separação temporal de Reinstein e Snyder (2005).

Figura 9 - Distribuição dos tempos

Fonte: Thaís Luiza Donega e Souza, 2017

Desta forma, o tempo BEFORE compreende todas as revisões críticas realizadas antes da data de estreia do filme nos cinemas (RELEASE_DATE); o tempo DURING consiste no período entre a data de estreia, normalmente sextas ou quintas-feiras – considerando aberturas de final de semana prolongado, e se segue imediatamente até a próxima quinta-feira, respeitando a estreia do próximo filme; O período AFTER abrange todas as revisões críticas posteriores à primeira semana e tem sua última semana de exibição representada como tempo de falha, ou seja, quando o filme sai de cartaz. Adicionalmente, os tempos BD (BEFORE / DURING) e ALL foram criados sendo respectivamente a junção de BEFORE e DURING e o período todo compreendido da existência de revisões críticas para um filme. Desta forma podemos controlar os efeitos do momento em que as revisões foram realizadas sobre a sobrevivência, a exemplo de

Reinstein e Snyder (2005) e Eliashbeg e Shugan (1997), evitando assim problemas de endogeneidade.

Para atender as especificações do modelo proposto (Capítulo 1), utilizamos variáveis de controle que compõem a matriz X_i , descritas a seguir e disponibilizadas no anexo B.

MONTHS: Variável dummy do mês de lançamento do filme para controlar por gastos com marketing, uma vez que para filmes lançados nas férias de verão e no final do ano usualmente possuem gastos maiores com publicidades. Pretende controlar também por eventuais sazonalidades, quando há mais produções cinematográficas, como os períodos que antecedem as premiações (Reinstein e Snyder, 2005).

YEARS: Variável dummy do ano de lançamento do filme – de 2004 a 2015. Pretende controlar por renda doméstica.

LTHEATERS_FIRSTWEEK: Log do número de telas de cinema em que o filme é exibido no seu final de semana de estreia. A variável também é controle de publicidade e orçamento, uma vez que estas duas variáveis são altamente correlacionadas. O número de salas no momento da estreia revela a aposta do distribuidor quanto ao sucesso do filme, ao mesmo tempo que controla por filmes de baixa abertura, cuja variável comumente é menor que 50.

BIG_DISTRIBUTOR: Dummy para distribuidores que detêm grandes fatias da produção dos filmes da amostra, tais como Buena Vista, Fox, DreamWorks e MGM. Esta variável controla também por publicidade e orçamento, uma vez que grandes estúdios dispõem de um orçamento maior para suas produções.

STARS: Quantidade de atores e atrizes cujos nomes e recebimentos foram mencionados na lista da Forbes de atores mais bem pagos. Para cada ator listado, contou-se sua presença no filme em até dois anos anteriores ao ano da lista.

MPAA: Dummies de classificação indicativa do filme. As variáveis foram separadas entre até 16 anos, *MPAA_PG*, e acima de 16 anos, *MPAA_R*.

GENRES: Dummies para os gêneros de acordo com o RottenTomatoes.

Algumas variáveis, apesar de disponíveis (Anexo B), não foram utilizadas. Os tickets da semana, por exemplo, não foram utilizados por serem fonte endógena ao modelo. Quer-se que X_i expliquem as variáveis de interesse Y (sobrevivência). A presença do volume de tickets de cada semana pode causar viés, uma vez que a diminuição do volume ao longo do tempo indica que o filme “está perdendo a força” ao mesmo tempo que o crescimento das semanas indica uma diminuição de tickets vendidos.

O Volume de revisões críticas também não foi utilizado devido às limitações enfrentadas no momento da extração de dados do Rotten Tomatoes. A coleta das revisões críticas dos consumidores se limitou a 50 páginas (de 1 a 51) da área de revisões críticas publicadas pelos usuários cadastrados como audiência, sendo 20 postagens por página. Desta forma, filmes de grande audiência como Avatar, apesar de constar em tal área o total de 5016 páginas, o alcance da coleta foi de apenas 50, limitando o número máximo dos textos dos consumidores em 1000 revisões independentemente das suas características de sucesso ou fracasso. Note que esta limitação de volume não afeta o aspecto qualitativo extraído das revisões dos consumidores em forma de análise afetiva. Ainda, há uma literatura significativa que aborda esta questão sobre o boca a boca (KIM et al, 2013; liu, 2006; DUAN et al, 2008) e não compreende o foco qualitativo deste trabalho.

Ainda, foram estimados modelos considerando os orçamentos (BUDGET). Nota-se que ele é estatisticamente significativo, ou seja, quanto maior o budget, maior sua influência sobre o tempo de duração do filme em cartaz. Isso está associado às grandes produções, publicidade, estúdios/distribuidores, atores famosos, etc. Desta forma, as estimativas com budget eliminam a significância do número de cinemas do final de semana de lançamento do filme, ao passo que esta variável é estatisticamente significativa na estimativa sem o budget. Como as observações de budget não estão completas, o número de cinemas no final de semana de estreia do filme passa a ser proxy de budget porque expressam as mesmas informações, como publicidade e custo de produção.

A expectativa de vida média da amostra final é de 8,92 semanas. O maior tempo de sobrevivência é de 62. Todos os 4479 filmes falham, não existindo, portanto, censuras.

As correlações entre as variáveis contínuas consideradas estão relatadas na tabela 13.

Tabela 13 - Matriz de correlação entre as variáveis contínuas

Variável	1	2	3	4	5	6	7	8
1 TOP_ MEDIA_BD	1							
2 MEDIA_ AUDIENCE	0,610	1						
3 MEDIA_ ANGER	-0,194	-0,317	1					
4 MEDIA_ JOY	0,182	0,434	-0,531	1				
5 MEDIA_ FEAR	0,079	-0,054	0,125	-0,449	1			
6 MEDIA_ SADNESS	-0,174	-0,297	0,241	-0,634	-0,103	1		
7 MEDIA_ DISGUST	0,010	-0,143	0,293	-0,424	0,225	0,196	1	
8 STARS	-0,046	-0,064	0,034	-0,001	-0,059	0,011	-0,104	1
9 LTHEATERS_ FIRSTWEEK	-0,307	-0,185	-0,010	0,095	-0,087	-0,110	-0,297	0,268

Fonte: Thaís Luiza Donega e Souza, 2017

5.2.1 Resultados não paramétricos

Como resultados não paramétricos, foram estimados modelos de duração Kaplan-Meier e a Taxa de Risco. A figura 10 estima a taxa de risco de um filme sair de cartaz e mostra que ela não é constante conforme o tempo, evoluindo após o seu lançamento e sugerindo que funções mais flexíveis devem ser exploradas. A estimativa também não paramétrica Kaplan-Meier, figura 11, mostra os filmes que sobrevivem de acordo com a primeira, segunda e assim por diante semana de exibição. Note que com 6 semanas em cartaz, 50% dos filmes já saíram das salas de semana.

Figura 10 - Intervalo de confiança para o risco dos filmes saírem das telas de cinema

Fonte: Thaís Luiza Donega e Souza, 2017

Figura 11 - Sobrevivência dos filmes norte-americanos nas salas de cinema

Fonte: Thaís Luiza Donega e Souza, 2017

A figura 12 apresenta uma estimativa de sobrevivência quanto ao sentimento dos consumidores, obtido pela análise de sentimentos. É possível observar que os filmes avaliados positivamente possuem maior expectativa de vida, quando comparado aos sentimentos negativos ou neutros. A figura sugere também que o sentimento negativo tem maior taxa de duração que os filmes classificados como neutros até a décima semana, quando estes dois praticamente se igualam.

No que se refere às revisões dos críticos profissionais, a avaliação “Muito Bom” aumenta consideravelmente o tempo de duração do filme em relação às demais avaliações. A avaliações “Bom” e “Ruim” possuem um tempo de sobrevivência semelhante. Pode-se notar que a avaliação negativa tem mais efeito benéfico sobre o tempo de duração do filme do que a ausência de uma avaliação, sugerindo que mesmo que as avaliações não sejam positivas, a existência de uma já é melhor, em termos de sobrevivência, do que a não avaliação em si. Esta análise está especificada na figura 13 abaixo.

Figura 12 - Sobrevivência conforme a prevalência de sentimentos

Fonte: Thaís Luiza Donega e Souza, 2017

Figura 13 - Sobrevivência conforme as avaliações das críticas de profissionais

Fonte: Thaís Luiza Donega e Souza, 2017

Finalmente, a figura 14 retrata a estimativa de sobrevivência por emoção predominante:

Figura 14 - Estimativa de sobrevivência por emoção predominante

Fonte: Thaís Luiza Donega e Souza, 2017

Para esta última figura apresentada (14) utilizou-se da informação de emoção predominante, ou seja, a emoção com maior média para o filme foi escolhida como sendo a principal. Apesar de se poder notar que as emoções *sadness* e *joy* aparentam ter uma taxa de sobrevivência maior, a avaliação fica incompleta ou incoerente devido as emoções *anger* e *disgust*. Isso ocorre porque ambas emoções são minoritárias nesta classificação de dominância, correspondendo a 0,15 e 0,09% da amostra. Além do mais, nenhuma emoção existe sozinha (conforme descrito no capítulo 4 e resultados no anexo C) e, portanto, justifica-se o uso das pontuações médias.

5.2.2 Resultados Paramétricos

O modelo especificado no capítulo 1, $\ln S_i = \alpha + \beta C_i + \delta M_i + \theta X_i + e_i$, foi estimado em duas versões. No primeiro modelo (i), C_i contém a variável média das notas dos críticos top TOP_MEDIA (Tabela 12) referente ao período BD, que compreende o tempo anterior e durante a semana de estreia do filme; M_i contém as médias das notas dos consumidores, e as médias de cada um dos cinco sentimentos minerados em todos os períodos: MEDIA_AUDIENCE, MEDIA_ANGER, MEDIA_JOY, MEDIA_FEAR, MEDIA_SADNESS e MEDIA_DIGUST (Tabela 11). No segundo modelo (ii), C_i contém 3 *dummies* de avaliação de críticos top, muito bom, bom e ruim: TOP_MUITOBOM, TOP_BOM, TOP_RUIM (Tabela 12), e “sem avaliações” como referência; M_i contém 2 *dummies* derivadas da análise de sentimento dos consumidores, sendo bom e ruim e não avaliado a referência: AUDI_BOM e AUDI_RUIM (Tabela 11).

Para as duas versões, X_i contém a) *dummies* de mês (MONTH); b) *dummies* de ano (YEAR); c) *dummies* de gênero (GENRES); d) logaritmo do número de salas de cinema na estreia (LTHEATER_FIRSTWEEK); e) *dummies* de atores famosos (STARS); f) *dummy* de grande distribuidor (BIG_DISTRIBUTOR); g) *dummy* de classificação indicativa (MPAA).

A versão dos modelos para diferentes distribuições, estão expostos nas tabelas a seguir. A tabela 14 apresenta os resultados para modelo (i), enquanto a tabela 15 mostra os resultados do modelo (ii). As estimativas de modelos de duração foram realizadas em versões aceleradas (AFT), de modo que a parametrização deve ser lida da seguinte maneira: nas versões AFT modela-se o tempo para ocorrer a falha, assim coeficientes positivos são bons indicativos de que o filme deverá demorar para sair de cartaz.

Tabela 14 – Estimativa paramétrica modelo (i)

Variáveis	AFT Weib. Coeff	AFT Exp. Coeff	AFT Lognor. Coeff	AFT Loglog. Coeff	Gamma Coeff
Média binária de críticas positivas dos críticos Top antes e durante o final de semana de estreia	0.570*** (0.070)	0.685*** (0.077)	0.820*** (0.090)	0.735*** (0.086)	0.505*** (0.071)
Nota média consumidores total (score)	0.161*** (0.039)	0.155*** (0.044)	0.138*** (0.051)	0.153*** (0.050)	0.160*** (0.036)
Média anger total nos textos de consumidores	0.886 (0.560)	0.802 (0.632)	0.686 (0.694)	0.530 (0.727)	0.910* (0.517)
Média joy total nos textos de consumidores	0.465 (0.428)	0.655 (0.453)	0.859* (0.482)	0.628 (0.471)	0.337 (0.411)
Média fear total nos textos de consumidores	0.799* (0.447)	0.889* (0.458)	0.962* (0.501)	0.826* (0.469)	0.747* (0.444)
Média sadness total nos textos de consumidores	0.102 (0.435)	0.377 (0.467)	0.614 (0.513)	0.449 (0.499)	-0.083 (0.417)
Média disgust total nos textos de consumidores	0.038 (0.524)	0.199 (0.557)	0.376 (0.564)	0.151 (0.615)	-0.067 (0.497)
Stars	0.045 (0.029)	0.070** (0.031)	0.099*** (0.037)	0.083** (0.034)	0.030 (0.028)
Grandes Distribuidores	0.103*** (0.026)	0.130*** (0.029)	0.181*** (0.035)	0.145*** (0.033)	0.093*** (0.025)
Log dos Número de cinemas no final de semana de estreia	0.013** (0.006)	0.030*** (0.006)	0.054*** (0.007)	0.039*** (0.007)	0.005 (0.006)
Gênero Ação	-0.102*** (0.039)	-0.111** (0.044)	-0.126** (0.053)	-0.107** (0.049)	-0.098*** (0.037)
Gênero Comedy	-0.011 (0.032)	-0.003 (0.035)	0.001 (0.041)	0.009 (0.038)	-0.017 (0.030)
Gênero Drama	-0.069** (0.033)	-0.058 (0.037)	-0.034 (0.044)	-0.046 (0.040)	-0.071** (0.031)
Gênero Romance / Family	0.008 (0.039)	0.020 (0.047)	0.030 (0.058)	0.044 (0.055)	-0.005 (0.036)
Gênero Foreign	0.148** (0.059)	0.198*** (0.067)	0.257*** (0.081)	0.268*** (0.077)	0.117** (0.056)
Gênero Crime / War / Western	-0.043 (0.047)	-0.020 (0.056)	-0.009 (0.070)	0.010 (0.064)	-0.064 (0.043)
Gênero Horror	-0.239*** (0.077)	-0.222*** (0.082)	-0.205** (0.096)	-0.190** (0.086)	-0.253*** (0.074)
Gênero Adventure / Animation	0.124*** (0.043)	0.112** (0.048)	0.074 (0.060)	0.113** (0.051)	0.123*** (0.041)
Gênero Thriller	-0.091** (0.044)	-0.086* (0.050)	-0.087 (0.061)	-0.081 (0.056)	-0.096** (0.041)
Classificação MPAA PG	0.421*** (0.079)	0.515*** (0.092)	0.606*** (0.104)	0.689*** (0.130)	0.376*** (0.075)
Classificação MPAA R	0.256*** (0.076)	0.343*** (0.089)	0.427*** (0.100)	0.509*** (0.127)	0.214*** (0.071)
Constante	0.753* (0.457)	0.152 (0.480)	-0.467 (0.489)	-0.232 (0.506)	1.125** (0.458)
Observações	17,201	17,201	17,201	17,201	17,201

Fonte: Thaís Luiza Donega e Souza, 2017

Tabela 15 – Estimativa paramétrica modelo (ii)

Variáveis	AFT Weib. Coeff	AFT Exp. Coeff	AFT Lognor. Coeff	AFT Loglog. Coeff	Gamma Coeff
Críticas Profissionais “Muito Bom”	0.742*** (0.039)	0.804*** (0.037)	0.923*** (0.038)	0.923*** (0.042)	0.888*** (0.040)
Críticas Profissionais “Bom”	0.294*** (0.040)	0.309*** (0.039)	0.310*** (0.039)	0.334*** (0.041)	0.315*** (0.039)
Críticas Profissionais “Ruim”	0.030 (0.041)	0.026 (0.037)	0.037 (0.033)	0.049 (0.035)	0.032 (0.034)
Análise de Sentimento “Positivo”	-0.044 (0.029)	-0.014 (0.027)	0.029 (0.027)	0.034 (0.028)	0.018 (0.027)
Análise de Sentimento “Negativo”	-0.132*** (0.037)	-0.118*** (0.034)	-0.100*** (0.033)	-0.097*** (0.034)	-0.105*** (0.032)
Stars	0.070*** (0.026)	0.084*** (0.027)	0.122*** (0.031)	0.104*** (0.031)	0.108*** (0.030)
Grandes Distribuidores	0.232*** (0.022)	0.267*** (0.022)	0.336*** (0.024)	0.320*** (0.025)	0.314*** (0.024)
Log dos Número de cinemas no final de semana de estreia	0.042*** (0.005)	0.058*** (0.005)	0.094*** (0.005)	0.094*** (0.005)	0.082*** (0.006)
Gênero Ação	-0.056 (0.034)	-0.061* (0.034)	-0.065* (0.035)	-0.065* (0.035)	-0.066* (0.034)
Gênero Comedy	0.024 (0.042)	0.015 (0.036)	0.010 (0.031)	0.012 (0.031)	0.011 (0.031)
Gênero Drama	0.062 (0.044)	0.057 (0.039)	0.066** (0.033)	0.056* (0.034)	0.061* (0.034)
Gênero Romance / Family	-0.013 (0.038)	-0.010 (0.038)	0.005 (0.041)	0.016 (0.041)	-0.001 (0.040)
Gênero Foreign	0.089** (0.044)	0.115*** (0.043)	0.168*** (0.042)	0.206*** (0.043)	0.153*** (0.042)
Gênero Crime / War / Western	0.025 (0.050)	0.041 (0.051)	0.061 (0.054)	0.050 (0.056)	0.056 (0.053)
Gênero Horror	-0.207*** (0.043)	-0.200*** (0.043)	-0.186*** (0.045)	-0.166*** (0.047)	-0.190*** (0.044)
Gênero Adventure / Animation	0.173*** (0.048)	0.147*** (0.045)	0.087* (0.046)	0.097** (0.045)	0.108** (0.045)
Gênero Thriller	-0.027 (0.042)	-0.027 (0.040)	-0.024 (0.040)	-0.023 (0.042)	-0.024 (0.040)
Classificação MPAA PG	0.660*** (0.047)	0.682*** (0.043)	0.636*** (0.040)	0.715*** (0.044)	0.666*** (0.042)
Classificação MPAA R	0.498*** (0.047)	0.517*** (0.042)	0.469*** (0.037)	0.545*** (0.041)	0.499*** (0.039)
Constante	1.100*** (0.070)	0.876*** (0.069)	0.424*** (0.069)	0.391*** (0.072)	0.566*** (0.083)
Observações	42,273	42,273	42,273	42,273	42,273

Fonte: Thaís Luiza Donega e Souza, 2017

Estas duas versões também foram estimadas, no modelo de Weibull, para cada um dos gêneros separadamente. Os resultados estão apresentados nas tabelas 16 e 17, para o modelo (i) e (ii), respectivamente.

Tabela 16 - Estimativa modelo paramétrico (i) por gênero

Variáveis	Weibull Coeff Action	Weibull Coeff Comedy	Weibull Coeff Drama	Weibull Coeff Romance Family	Weibull Coeff Foreign	Weibull Coeff Crime War Western	Weibull Coeff Horror	Weibull Coeff Adventure Animation	Weibull Coeff Thriller
Média binária de críticas positivas dos críticos Top antes e durante o final de semana de estreia	0.205 (0.252)	0.547*** (0.119)	0.810*** (0.127)	0.590*** (0.210)	0.871** (0.419)	1.014*** (0.313)	0.321 (0.353)	-0.009 (0.239)	0.621*** (0.234)
Nota média consumidores total (score)	0.253* (0.144)	0.252*** (0.077)	0.208*** (0.065)	0.234* (0.137)	0.251 (0.190)	0.132 (0.257)	-0.096 (0.180)	0.284*** (0.075)	0.174 (0.128)
Média anger total nos textos de consumidores	0.578 (1.678)	1.049 (1.104)	2.517** (1.097)	5.061*** (1.935)	-2.216** (1.048)	1.766 (3.645)	16.317*** (3.655)	-2.653 (3.072)	-0.398 (1.363)
Média joy total nos textos de consumidores	0.904 (1.019)	0.149 (1.004)	1.942** (0.947)	2.527 (1.852)	0.991 (1.197)	1.712 (2.560)	10.159*** (1.946)	-1.528 (2.151)	-1.313 (1.644)
Média fear total nos textos de consumidores	-0.267 (1.296)	-0.080 (1.030)	2.708*** (0.939)	1.863 (1.647)	0.044 (0.836)	2.243 (2.902)	9.672*** (1.795)	-1.392 (2.278)	-2.458* (1.435)
Média sadness total nos textos de consumidores	0.798 (1.538)	-0.864 (0.994)	1.341 (1.007)	1.913 (1.650)	1.507 (1.222)	2.541 (3.114)	3.942* (2.140)	-2.260 (1.897)	-1.647 (1.658)
Média disgust total nos textos de consumidores	-1.755 (2.396)	0.503 (1.141)	0.471 (1.169)	4.420** (2.178)	-0.470 (1.792)	-1.392 (1.430)	6.997*** (1.592)	2.096 (2.839)	0.934 (1.560)
Constante	0.662 (1.280)	1.011 (1.078)	-1.172 (0.969)	-4.081** (1.844)	-0.272 (1.186)	-1.186 (2.844)	-8.637*** (1.931)	3.016 (2.320)	1.422 (1.459)
Observações	1,742	5,013	5,222	1,505	1,490	906	656	2,127	1,575

Fonte: Thaís Luiza Donega e Souza, 2017

Tabela 17 - Estimativa modelo paramétrico (ii) por gênero

Variáveis	Weibull Coeff Action	Weibull Coeff Comedy	Weibull Coeff Drama	Weibull Coeff Romance Family	Weibull Coeff Foreign	Weibull Coeff Crime War Western	Weibull Coeff Horror	Weibull Coeff Adventure Animation	Weibull Coeff Thriller
Críticas Profissionais “Muito Bom”	0.556*** (0.099)	0.748*** (0.082)	0.583*** (0.073)	0.744*** (0.137)	1.032*** (0.077)	0.606*** (0.134)	0.760*** (0.144)	0.502*** (0.107)	0.663*** (0.114)
Críticas Profissionais “Bom”	0.274** (0.112)	0.329*** (0.086)	0.119 (0.077)	0.371** (0.157)	0.343*** (0.072)	0.214 (0.157)	0.278** (0.126)	0.172* (0.105)	0.425*** (0.116)
Críticas Profissionais “Ruim”	0.122 (0.081)	0.005 (0.080)	-0.180** (0.076)	0.059 (0.096)	0.057 (0.082)	-0.082 (0.140)	0.102 (0.095)	0.143 (0.097)	0.071 (0.097)
Análise de Sentimento “Positivo”	-0.015 (0.077)	-0.046 (0.049)	0.020 (0.050)	0.019 (0.090)	0.041 (0.060)	0.003 (0.107)	-0.180* (0.100)	-0.139* (0.080)	-0.028 (0.083)
Análise de Sentimento “Negativo”	-0.089 (0.079)	-0.191*** (0.060)	-0.212*** (0.066)	-0.156 (0.105)	-0.072 (0.086)	-0.146 (0.139)	-0.153* (0.085)	-0.104 (0.117)	-0.189** (0.088)
Constante	1.342*** (0.159)	1.174*** (0.131)	1.093*** (0.125)	0.691*** (0.190)	1.332*** (0.155)	1.309*** (0.234)	0.664*** (0.170)	1.175*** (0.195)	0.869*** (0.179)
Observações	4,021	11,124	12,027	3,748	6,199	1,849	2,371	4,829	3,632

Fonte: Thaís Luiza Donega e Souza, 2017

As taxas de risco de cada um dos modelos estão no anexo E, tabelas 23, 24, 25 e 26.

A partir do modelo (i) pode-se inferir que tanto a média da aprovação dos críticos tops (TOP_MEDIA) quanto a média da pontuação estabelecida pelas notas revisões críticas dos consumidores (MEDIA_AUDIENCE) afetam positivamente o tempo de duração dos filmes no mercado de cinema americano. Ambas variáveis são estatisticamente significativas, porém o coeficiente da média dos críticos profissionais é, em geral, quase três vezes maior que o coeficiente da média da audiência, sugerindo o tamanho do impacto de especialistas. Para a análise afetiva usou-se as médias das emoções, que obtiveram resultados marginalmente significantes estatisticamente para *anger*, *joy* e *fear*.

No segundo modelo (ii) pôde-se visualizar um resultado mais específico de acordo com a polarização das revisões críticas. Assim como no modelo anterior, os coeficientes da revisão crítica dos críticos profissionais são estatisticamente significativos e afetam positivamente o tempo de vida dos filmes. No entanto, com a separação entre “Muito Bom”, “Bom” e “Ruim” e tendo como referência os filmes “Sem Avaliação”, observa-se que o tamanho do coeficiente e, portanto, do seu efeito, está relacionado ao nível de polaridade da avaliação. Assim, a avaliação “Muito Bom” tem coeficiente maior que “Bom”, enquanto que avaliações “Ruins” não exercem influência. Para as revisões críticas dos consumidores foi utilizado, no modelo (ii), a análise de sentimento rotulada enquanto a sua polaridade positiva / negativa. Os resultados sugerem que enquanto o sentimento positivo dos consumidores não influencia no tempo de sobrevivência, o sentimento negativo afeta negativamente, com coeficientes estatisticamente significativos em todas as especificações.

Para ambos os modelos as demais variáveis obtiveram resultados semelhantes. STARS, BIG_DISTRIBUTOR, LTHEATERS_FIRSTWEEK e MPAA mostraram exercer influência positiva para o tempo de duração dos filmes em cartaz. Para entender melhor estes resultados, foram feitas estimativas adicionais para cada um dos gêneros, como mostram os modelos (i) e (ii) por gênero dispostos nas tabelas 16 e 17.

O resultado por gênero do modelo (i) apresenta que as revisões críticas dos profissionais tops não são significativas para ação, horror e aventura e animação, enquanto permanecem altamente significativos para os demais gêneros testados. Por outro lado, as notas da audiência importam para os gêneros ação, romance / família, comédia, drama e aventura / animação, sendo estes três últimos com maiores coeficientes. Para as emoções, observa-se que todas as emoções afetam positivamente em algum

gênero específico. O gênero drama, por exemplo, sofre influências de *anger*, *joy* e *fear*, que exercem impacto positivo.

Para os gêneros do modelo (ii) as avaliações “Muito Bom” dos críticos profissionais são altamente significativas e positivas para todos os gêneros, similar ao resultado da tabela 14. Ainda, as críticas “Ruim” dos especialistas continuam não importando, exceto para drama, onde exercem influência negativa com nível de significância em 5%. Do mesmo modelo, os sentimentos negativos impactam o tempo de sobrevivência negativamente para os gêneros comédia e drama e também para horror e thriller, mas com coeficientes menores nestes dois últimos.

Assim, observa-se que os resultados dos modelos sugerem que as críticas boas de profissionais tops afetam positivamente o tempo de duração e que quanto melhor a avaliação, maior o coeficiente e, consequentemente, maior o impacto. As notas avaliativas dos consumidores também afetam positivamente, porém possuem coeficientes menores que a dos especialistas. No modelo (ii) também é notável que a polarização boa/ruim das variáveis construídas gera um resultado mais significativo que as médias usadas no modelo (i). Para as estimativas por gênero, o gênero drama se mostrou mais sensível às avaliações, tanto de críticos profissionais quanto de consumidores nos dois modelos. O gênero terror mostrou-se dependente das críticas dos consumidores.

6 Conclusão

Um filme é um exemplo excelente de bens diferenciados que competem uns com os outros em um período curto de tempo. Os atributos que levam um filme ser mais bem-sucedido que outros não é de fácil mensuração e definir como medir exatamente sucesso não é tão trivial quanto parece. Focar no tempo de vida fornece uma alternativa à métrica tradicional de faturamento por considerar a interação entre o fornecimento e a demanda em uma configuração dinâmica onde a expectativa de vida do filme é a agregação das decisões de maximização dos lucros dos exibidores em resposta às demandas semanais reveladas.

Dado que é um bem de experiência, a decisão de ver um filme recai em outros fatos que não a qualidade vivenciada por aquele filme em questão. Estes "outros fatores" podem ser, inclusive, formas de sinais de qualidade do filme, como o boca a boca, considerado um dos sinais de qualidade mais importante, a medida que se espalha mais rapidamente conforme a base de consumidores aumenta.

A teoria da informação defende que consumidores buscarão por informações sobre o produto de forma a diminuir o risco associado à sua compra e consumo (Stigler, 1961). O papel desta informação é especialmente crítico para bens de experiência, uma vez que sua qualidade é desconhecida antes do seu consumo (Eliashberg & Shugan, 1997).

Dispõe-se aqui então, a obtenção de variáveis proxies de avaliação de consumo por pares antes não disponível para análises econômicas, de mercado e muitas outras, usando técnicas computacionais e a proposição de trabalho genuinamente interdisciplinar.

O objetivo deste trabalho, simplistamente, é investigar os efeitos do boca a boca das revisões críticas de consumidores controladas pelo efeito das críticas profissionais sobre o tempo de sobrevivência dos filmes em cartaz nos cinemas americanos. Foram coletados dados sobre filmes do BoxOfficeMojo e as críticas do RottenTomatoes, ambos publicamente disponíveis. Conduzimos então a computação afetiva, realizando a análise de sentimentos e a mineração de emoção com o IBM Watson nas postagens dos consumidores. Com o banco de dados montado, foram realizadas estimativas sobre dois modelos paramétricos de análise de sobrevivência para identificar os efeitos de boca a boca e demais influenciadores.

Os resultados obtidos mostram que as críticas dos profissionais são mais importantes sobre o tempo de exibição dos filmes no cinema americano. Os coeficientes ligados a essas críticas são positivos e possuem um valor três vezes maior, em geral, em

comparação às notas dos consumidores. Observa-se também que a polaridade da crítica do crítico profissional não importa, desde que exista alguma crítica o impacto é sempre positivo e o tamanho do coeficiente é proporcional à intensidade da avaliação. Este resultado vai de encontro com Shrum (1991, p. 368), que diz que “Mesmo uma avaliação medíocre ou negativa é melhor do que avaliação nenhuma (...)”.

Por sua vez, a média da avaliação dos consumidores (score) demonstrou-se também de grande valia para aumentar a expectativa de vida dos filmes em cartaz. Essa medida, da média da avaliação dos consumidores, é mais significativa do que as medidas de afetividade (sentimento e emoção) também dos consumidores. Este resultado pode caracterizar em uma dupla atuação do consumidor, quando os scores diferem da intenção da polaridade do texto da sua revisão.

Quanto aos resultados envolvendo as análises afetivas nota-se que o sentimento negativo dos consumidores, obtido pela análise de sentimento, afeta o tempo de sobrevivência negativamente, enquanto o sentimento positivo não exerce nenhuma influência. As emoções estudadas, *anger*, *disgust*, *fear*, *joy* e *sadness*, afetam positivamente o tempo de sobrevivência em algum nível, principalmente considerando os resultados por gênero, onde se destacam as emoções *anger*, *joy* e *fear* e os gêneros drama e horror.

Nossos resultados também concordam com os apresentados por Walls (1998) e De Vany e Walls (1996) em que a função de risco é uma função crescente no tempo e que está fortemente associada ao faturamento inicial da bilheteria, que no nosso caso, se reflete no número de cinemas do final de semana de estreia.

Similarmente aos resultados de Rui et al (2013), trabalho que utiliza análise de sentimento em tweets (capítulo 3) para identificar efeitos nas vendas da bilheteria americana, também encontramos que a valência do sentimento dos consumidores afeta o modelo estimado, com sentimentos negativos impactando negativamente sobre o tempo de sobrevivência. Por outro lado, discordamos quanto aos sentimentos positivos, que em nossos modelos não foi observada a mesma influência relatada no trabalho dos autores. Isso pode ocorrer devido a própria definição do modelo, que diverge entre metodologias, mas também pela fonte de dados, sendo o Twitter uma fonte passiva de boca a boca – com maior poder de conscientização, enquanto o RottenTomatoes, por ser uma comunidade apenas para filmes, pode possuir no boca a boca um efeito maior de persuasão.

Legoux et al (2016) apontam que somente as críticas excelentes dos críticos profissionais afetam positivamente o tempo de duração, quando comparado aos demais rótulos avaliativos (*poor*, *fair* e *good*), que não exercem nenhuma influência. Os resultados apresentados aqui sugerem que os críticos profissionais (tops) possuem sim grande influência em críticas “excelentes”, porém, críticas “boas” também possuem alta significância estatística. De qualquer forma, ambos resultados estão em acordo, uma vez que foi identificada a importância do crítico para o tempo de sobrevivência do filme em cartaz no mercado americano de cinema. Os autores controlaram também os possíveis efeitos de endogeneidade e o fazem, em seus modelos, com a média das críticas dos consumidores. Os resultados de Legoux et al (2016) para esta variável é insignificante. Neste ponto, discordamos dos autores, uma vez que tanto as médias de score quanto as variáveis provenientes da análise afetivas, afetam sim a sobrevivência de alguma forma. Essa divergência talvez seja devido ao mercado alvo: Legoux et al (2016) usam dados do mercado cinematográfico de Quebec, enquanto utilizamos dados estadunidenses de cinema.

Quanto a Reinstein e Snyder (2005), trabalho que embasou boa parte dos nossos modelos, é possível concordar que os críticos profissionais tops exercem sim papel de *influenciadores* além do papel de *preditores* (ELIASHBERG e SHUGAN, 1997). Porém, nossos modelos mostram que o efeito dos críticos profissionais é bem maior do que os apresentados por Reinstein e Snyder (2005), mesmo controlando pelo tempo das críticas profissionais e pelas críticas dos consumidores.

Assim, objetivando propiciar uma medida do efeito boca a boca sobre o consumo de um determinado bem, este trabalho estuda bens de informação, sejam filmes sejam as revisões críticas de todos os tipos, que associados à sua extrema disseminação na Internet pode ser visualizado como um estudo de sistemas de informação para consumo de mercados específicos de bens de informação, mas que pode ainda ser generalizado para qualquer bem de consumo que seja de experiência. Os resultados levantados podem ajudar a determinar como as revisões críticas de especialistas exercem influência na sobrevivência de outros bens de experiência e suas respectivas cadeias de distribuição.

6.1 Principais contribuições

Este trabalho enriqueceu as pesquisas anteriores de algumas formas, apresentaremos algumas delas a seguir.

A construção de um banco de dados inédito e muito abrangente sobre o mercado americano de cinemas disponível prontamente para vários estudos. Considerando os alcances das revisões bibliográficas realizadas tanto no contexto da computação afetiva quanto no contexto de economia e revisão crítica, não se encontrou nenhum banco de dados tão abrangente e rico como o construído para este trabalho, além deste ser, para o momento, o mais atual possível, o que corresponde com as mudanças tecnológicas.

Do ponto de vista de Reinstein e Snyder (2005), o controle de viés de variável não observável através do proxy de boca a boca consta como um problema resolvido para a literatura. Adicionalmente, do ponto de vista do McKenzie (2009) e de Eliashberg e Shugan (1997) os tempos das revisões críticas dos profissionais são controlados tal qual Reinstein e Snyder (2005) e o tempo das revisões críticas dos consumidores são dispostos de acordo com Legoux et al (2016), separando assim possíveis vieses entre os efeitos de influência e predição.

Ainda, e talvez o principal ineditismo deste trabalho, a abordagem diferenciada sobre a sinalização de qualidade do filme, utilizando computação afetiva, que compõe uma análise mais profunda e abrangente através das medidas de emoção e sentimento, além das tradicionais notas objetivas dos consumidores.

6.2 Trabalhos futuros

Dada as características do banco de dados construído e dos “intrigantes quebra-cabeças que ainda existem no mercado cinematográfico” (ELIASHBERG et al, 2006, p. 638) muitos problemas podem ser explorados. Algumas sugestões são apresentadas a seguir.

Agregar dados de consumidores no período anterior à 2003 utilizando banco de dados já prontos de outros trabalhos ou outras fontes de informação que incluam publicações que, a priori, não eram digitais. Desta forma, pode-se observar o efeito da Internet, mais precisamente da banda-larga e, posteriormente da Web 2.0 no comportamento do boca a boca sobre as medidas de sucesso do filme.

Análise afetiva das revisões críticas dos críticos profissionais. Uma vez que foi identificada sua extrema importância sobre a expectativa de vida do filme, identificar se as notas objetivas são mais eficazes do que a análise subjetiva textual proferida por eles.

Utilizar de outros recursos de compreensão de linguagem natural do IBM Watson e sua capacidade cognitiva, como extração de características e *targets*, podendo assim identificar sobre o que ou quem se fala a respeito quando se resenha sobre um filme e testar os efeitos do boca a boca considerando a teoria de *superstars* (ADLER, 2006) – quando os indivíduos gostam mais de artista ou de produtos culturais a medida que sabem mais sobre eles e o comportamento de manada – quando indivíduos, sob incerteza, imitam os comportamentos dos demais agentes supostamente mais bem informados.

Outros contextos computacionais também podem ser utilizados, como ferramentas de mineração para identificar padrões de sucesso e/ou fracasso, explorando a quantidade de informações sobre os filmes, ou identificar efeitos tecnológicos, explorando a extensão temporal dos dados.

Referências

ABBASI, A; FRANCE, S; ZHANG, Z; CHEN, H. **Selecting attributes for sentiment classification using feature relation networks**. IEEE Transactions on Knowledge and Data Engineering, 23(3), 447-462. 2011.

ADLER, M.; **Stardom and talent**. Handbook of the Economics of Art and Culture, 1, 895-906. 2006.

AGARWAL, S., & SUREKA, A. **But I did not Mean It!—Intent Classification of Racist Posts on Tumblr**. In Intelligence and Security Informatics Conference (EISIC), 2016 European (pp. 124-127). IEEE. 2016.

AL-HAJJAR, D.; & SYED, A. Z.; **Applying Sentiment and Emotion Analysis on Brand Tweets for Digital Marketing**. In Applied Electrical Engineering and Computing Technologies (AEECT), 2015 IEEE Jordan Conference on (pp. 1-6). IEEE. 2015.

ALLISON, P. D.; **Event history analysis: regression for longitudinal event data**. California: Sage, 87p. 1984.

ALM, E. C. O.; **Affect in text and speech**. ProQuest, 2008.

ALM, E. C. O.; ROTH, D.; SPROAT, R.; **Emotions from text: Machine learning for text-based emotion prediction**. Vancouver, British Columbia, Canada. In Proceedings of the Joint Conference on Human Language Technology/Empirical Methods in Natural Language Processing (HLT/EMNLP-2005), Vancouver, Canada, pp. 579–586. 2005.

Bagozzi, R. P., Gopinath, M., & Nyer, P. U.; The role of emotions in marketing. Journal of the academy of marketing science, 27(2), 184-206. 1999.

BAI, X.; **Predicting consumer sentiments from online text**. Decision. Support System 50-4, 732–742p. 2011.

BALAHUR, A., MIHALCEA, R., & MONTOYO, A.; **Computational approaches to subjectivity and sentiment analysis: Present and envisaged methods and application**. Published by Elsevier Ltd. Computer Speech and Language. 2014.

BALAHUR, A; JESUS, M; HERMIDA, A; MONTOYO, A.; **Detecting implicit expressions of emotion in text: a comparative analysis**. Decision Support Systems, 53(4), 742-753. 2013.

BALAZS, Jorge A.; VELÁSQUEZ, Juan D. **Opinion mining and information fusion: A survey**. Information Fusion, v. 27, p. 95-110, 2016.

BASUROY, S.; CHTTERJEE, S.; RAVID, S. A., 2003. **How Critical Are Critical Reviews? The Box Office Effects of Films Critics, Star Power, and Budgets.** Journal of Marketing, Vol. 67, pp. 103-117. 2003.

BATOOL, R., KHATTAK, A. M., MAQBOOL, J., & LEE, S.; **Precise tweet classification and sentiment analysis.** In Computer and Information Science (ICIS), 2013 IEEE/ACIS 12th International Conference on (pp. 461-466). IEEE. 2013.

Bellegarda, J. R.; **A Data-Driven Affective Analysis Framework Toward Naturally Expressive Speech Synthesis.** IEEE Transactions on Audio, Speech, and Language Processing, 19(5), 1113-1122. 2011.

BOATWRIGHT, P.; BASUROY, S.; KAMAKURA, W.; **Reviewing the reviewers: The impact of individual film critics on box office performance.** Quantitative Marketing and Economics. pp. 401-425. DOI 10.1007/s11. 2007.

BOIY, E.; MOENS, M.-F.; **A machine learning approach to sentiment analysis in multilingual Web texts.** Information retrieval 12, no. 5 (2009): 526-558. 2009.

BRADLEY, M. M., & LANG, P. J.; **Affective norms for English words (ANEW): Instruction manual and affective ratings** (pp. 1-45). Technical report C-1, the center for research in psychophysiology, University of Florida. 1999.

CALVO R. A.; D'MELLO S.; **Affect detection: An interdisciplinary review of models, methods, and their applications.** IEEE Transactions on affective computing 1.1: 18-37. 2010.

CALVO, R. A., & MAC KIM, S.; **"Emotions in text: dimensional and categorical models."** Computational Intelligence 29.3 (2013): 527-543. 2013.

CALVO, R. A.; & D'MELLO, S.; **Affect detection: An interdisciplinary review of models, methods, and their applications.** IEEE Transactions on affective computing 1.1 (2010): 18-37. 2010.

CAMBRIA, E., LIVINGSTONE, A., & HUSSAIN, A.; **The hourglass of emotions.** In Cognitive behavioural systems (pp. 144-157). Springer Berlin Heidelberg. 2012.

CAMBRIA, E.; SCHULLER, B.; XIA, Y.; HAVASI, C.; **New Avenues in Opinion Mining and Sentiment Analysis.** IEEE Intell. Syst., vol. 28, no. 2, pp. 15–21, Mar 2013. 2013.

CAMPBELL, M.; HOANE, A. J.; & HSU, F. H.; **Deep blue.** Artificial intelligence. 134(1-2), 57-83. 2002.

CAMPOS, E.; MATSUBARA, E.; **An Experimental Evaluation of Sentiment Analysis on Financial News Using Prior Polarity Words**. In: Ibero-American Conference on Artificial Intelligence. Springer International Publishing, 2014. p. 218-228. 2014.

CATTELL, R.; **Sentiment or attitude? The core of a terminology problem in personality research**. Journal of Personality 9, no. 1 (1940): 6-17. 2006.

CHAFFAR, S.; & INKPEN, D.; **Using a Heterogeneous Dataset for Emotion Analysis in Text**. In Canadian Conference on Artificial Intelligence (pp. 62-67). Springer Berlin Heidelberg. 2011.

CHEN H.; ZIMBRA D. **AI and Opinion Mining**. IEEE Intell. Syst., vol. 25, no. 3, pp. 74–76, May 2010.

CHEN, Y.; XIE, J.; **Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix**. Management Science, 54. 477–91p. 2008.

CHINTAGUNTA, P. K., GOPINATH, S.; VENKATARAMAN, S.; **The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets**. Marketing Science, 29(5), 944-957. 2010.

CHISHOLM, D. C., & NORMAN, G.; **When to exit a product: Evidence from the US motion-picture exhibition market**. The American economic review, 96(2), 57-61. 2006.

CHOMSKY, N.; **A review of BF Skinner's Verbal Behavior**. Language, 35(1), 26-58. 1959.

CHOMSKY, N.; **Three models for the description of language**. IRE Transactions on information theory, 2(3), 113-124. 1956.

CHUNG, W.; HE, S.; ZENG, D. D.; & BENJAMIN, V.; **Emotion Extraction and Entrainment in Social Media: The Case of US Immigration and Border Security**. In Intelligence and Security Informatics (ISI), 2015 IEEE International Conference on (pp. 55-60). IEEE. 2015.

CIULLO, F., ZUCCO, C., CALABRESE, B., AGAPITO, G., GUZZI, P. H., & CANNATARO, M.; **Computational challenges for sentiment analysis in life sciences**. In High Performance Computing & Simulation (HPCS), 2016 International Conference on (pp. 419-426). IEEE. 2016.

COVINGTON, M.; **Natural Language Processing for Prolog Programmers**. Englewood Cliffs (NJ). Prentice-Hall. 1994.

COVINGTON, M.; NUTE, D.; VELLINO, A.; **Prolog Programming in Depth**. USA: Artificial Intelligence Programs The University of Georgia, 1995. Prentice-Hall 2^a Ed.. 1997.

DANG, Y; ZHANG, Y; CHEN, H. **A lexicon-enhanced method for sentiment classification: an experiment on online product reviews**. IEEE Intell. Syst, 25. 46–53p. 2010.

DARWIN, C; **Expression of the Emotions in Man and Animals**. Oxford University Press, Inc., 2002.

DASH, A. K.; ROUT, J. K.; JENA, S. K.. **Harnessing Twitter for Automatic Sentiment Identification Using Machine Learning Techniques**. In: Proceedings of 3rd International Conference on Advanced Computing, Networking and Informatics. Springer India, 2016. p. 507-514. 2016.

DE VANY, A. S.; WALLS, W. D.; **The market for motion pictures: rank, revenue, and survival**. Economic Inquiry, 35(4), 783-797. 1997.

DE VANY, A., & WALLS, W. D.; **Uncertainty in the movie industry: Does star power reduce the terror of the box office?**. Journal of cultural economics, 23(4), 285-318. 1999.

DELLAROCAS, C.; NARAYAN, R., and others. **A statistical measure of a population's propensity to engage in post-purchase online word-of-mouth**. Statistic Measure Science. 21, 2, 277–285. 2006.

DENG, S.; WANG, D.; LI, X.; & XU, G.; **Exploring user emotion in microblogs for music recommendation**. Expert Systems with Applications, 42(23), 9284-9293. 2015.

DEUCHERT, E.; ADJAMAH, K.; PAULY, F; **For Oscar glory or Oscar money?**. Journal of Cultural Economics, 29(3), 159-176. 2005.

DOBRESCU, L; LUCA, M.; MOTTA, A. **What makes a critic tick? Connected authors and the determinants of book reviews**. Working Paper 12-080, August 6, 2013.

DODDS, P. S., & DANFORTH, C. M.; **Measuring the happiness of large-scale written expression: songs, blogs, and presidents**. Journal of Happiness Studies 11 (2010) 441–456. 2010.

DUAN, W.; GU, B.; WHINSTON, A. B; (A) **Do online reviews matter? — An empirical investigation of panel data**. Decision Support Systems, 45 (1007-1016). 2008.

DUAN, W.; GU, B.; WHINSTON, A. B.; (B) **The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry.** Journal of retailing, 84(2), 233-242. 2008.

EIRINAKI, M.; PISAL, S; SINGH, J. **Feature-based opinion mining and ranking.** Journal of Computer and System Sciences 78, no. 4, pp. 1175–1184. 2012.

EKMAN, P.; & FRIESEN, W. V.; **Constants across cultures in the face and emotion.** Journal of personality and social psychology, 17(2), 124. 1971.

EKMAN, P.; **An argument for basic emotions.** Journal Cognition & emotion, 6(3-4), 169-200. 1992.

ELBERSE, Anita. **The power of stars: Do star actors drive the success of movies?** Journal of Marketing 71, 4, 102–120. 2007.

ELIASHBERG, J.; ELBERSE, A.; LEENDERS, M. A. A. M.; **The motion picture industry: Critical issues in practice, current research, and new research directions.** Marketing Science, 25, pp. 638–661. 2006.

ELIASHBERG, J.; HUI, Sam K.; ZHANG, Zhongwei J. **Assessing box office performance using movie scripts: A kernel-based approach.** Knowledge and Data Engineering, IEEE Transactions on 26, pp. 2639–2648. 2014.

ELIASHBERG, J.; SHUGAN, S. M.; **Film Critics: Influencers or Predictors?** Journal of Marketing, Vol. 61, 2 (Abril, 1997), pp. 68-78. 1997.

ESULI, A., & SEBASTIANI, F.; **SENTIWORDNET: A high-coverage lexical resource for opinion mining.** Evaluation, 1-26. 2007.

FEHR, B.; RUSSELL, J. A.; **Concept of emotion viewed from a prototype perspective.** Journal of experimental psychology. General, 113(3), 464.1984.

FELDMAN, R.; **Techniques and applications for sentiment analysis.** Communications of the ACM. ACM 56 (4) pp. 82–89. 2013.

FERRUCCI, D., BROWN, E., CHU-CARROLL, J., FAN, J., GONDEK, D., KALYANPUR, A. A., & SCHLAERFER, N.; **The AI behind Watson - the technical article.** The AI Magazine. 2010a. Acessado em: Fev/2017. Disponível em: <http://www.aaai.org/Magazine/Watson/watson.php>

FERRUCCI, D., BROWN, E., CHU-CARROLL, J., FAN, J., GONDEK, D., KALYANPUR, A. A., & SCHLAERFER, N.; **Building Watson: An overview of the**

DeepQA project. AI magazine, 31(3), 59-79. 2010b. Acessado em: Fev/2017. Disponível em:

FERRUCCI, D., LEVAS, A., BAGCHI, S., GONDEK, D., & MUELLER, E. T.; **Watson: beyond jeopardy!.** Artificial Intelligence, 199, 93-105. doi:10.1016/j.artint.2012.06.009. 2013. Acessado em: Fev/2017. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0004370212000872>.

FERRUCCI, D.; **An Introduction to “this is watson”.** IBM Journal of Research and Development 56.3.4 (2012): 1-1. 2012.

FERRUCCI, D.; **How It All Began and What's Next - Computer History Museum.** IBM Research. December 1, 2011. Acessado em: Fev/2017. Disponível em: <https://www.ibm.com/blogs/research/2011/12/dave-ferrucci-at-computer-history-museum-how-it-all-began-and-whats-next>.

FIGUEIREDO F.; ROCHA L.; COUTO T.; SALLES T.; GONÇALVES M. A.; MEIRA W. Jr.; **Word co-occurrence features for text classification.** Inform. Syst., 36 pp. 843–858. 2011.

FU, W. W.; **Screen survival of movies at competitive theaters: Vertical and horizontal integration in a spatially differentiated market.** Journal of Media Economics, 22(2), 59-80. 2009.

GARCÍA-CRESPO, Á., COLOMO-PALACIOS, R., GÓMEZ-BERBÍS, J. M., & RUIZ-MEZCUA, B.; **SEMO: a framework for customer social networks analysis based on semantics.** Journal of Information Technology, 25(2), 178-188. 2010.

GODES, D.; MAYZLIN, D.; CHEN, Y.; DAS, S.; DELLAROCAS, C.; PFEIFFER, B.; LIBAI, B.; SEN, S.; SHI, M.; VERLEGH, P.; **The Firm’s Management of Social Interactions.** Marketing Letters, 16 (3–4), 415–28. 2005.

GODES, D.; MAYZLIN, D.; **Using Online Conversations to Study Word-of-Mouth Communication.** Marketing Science, 23 (4), 545–60. 2004.

GUNDECHA, P. **"IBM Watson just got more accurate at detecting emotions".** October, 2016. IBM Bluemix Blog. Acessado em: Fev/2017. Disponível em: <https://www.ibm.com/blogs/bluemix/2016/10/watson-has-more-accurate-emotion-detection/>

HAO, Y. Y., LI, Y. J., YE, Q., & ZOU, P.; **Dynamic impacts of online reviews and other information sources on sales in panel data environment: Evidence from movie industry.** In Management Science and Engineering, 2008. ICMSE 2008. 15th Annual Conference Proceedings, International Conference on (pp. 493-500). IEEE. 2008.

HIGH, R. **Interview at The Montgomery Summit. Rob High, IBM Fellow and Vice President and Chief Technology Officer at IBM Watson.** Montgomery Summit Conference. 2016. Acessado em: Fev/2017. Disponível em: https://www.youtube.com/watch?v=g_tn2vtn87I.

HOESMER, D.; LEMESHOW, S.; **Applied Survival Analysis.** New York: Wiley, pp. 58-65, 90. 1999.

HOFMANN, S., BEVERUNGEN, D., RÄCKERS, M., & BECKER, J.; **What makes local governments' online communications successful? Insights from a multi-method analysis of Facebook.** Government Information Quarterly, 30(4), 387-396. 2013.

HORRIGAN, John A. **Online shopping.** Pew Internet & American Life Project. Report, 2008. Disponível em: http://www.pewinternet.org/files/old-media/Files/Reports/2008/PIP_Online%20Shopping.pdf.pdf. Acessado em 17/06/2016.

HORRIGAN, John A. **The Internet and Consumer Choice: Online Americans use different search and purchase strategies for different goods.** Pew Internet & American Life Project. Report, 2008. Disponível em: http://www.pewinternet.org/files/old-media/Files/Reports/2008/PIP_Consumer.Decisions.pdf.pdf. Acessado em 17/06/2016.

https://www.ibm.com/support/knowledgecenter/en/SSWTQQ_1.2.0/com.ibm.swg.ba.cognos.sifs_solution_guide.1.2.0.doc/c_trd_emotion_detection_library.html

HU, N., LIU, L., & ZHANG, J. J.; **Do online reviews affect product sales? The role of reviewer characteristics and temporal effects.** Information Technology and Management, 9(3), 201-214. 2008.

HU, N.; LIU, L.; ZHANG, J.J. **Do online reviews affect product sales? The role of reviewer characteristics and temporal effects.** Information Technology and Management 9, 3, 201–214. 2008.

IBM Research. **"How Watson 'sees', 'hears', and 'speaks' to play Jeopardy!'"**. David Gondek. Janeiro de 2011. 2011b. Acessado em Fev/2017. Disponível em: <https://www.ibm.com/blogs/research/2011/01/how-watson-sees-hears-and-speaks-to-play-jeopardy/>

IBM Research. **"The DeepQA Project"**. 2011a. Acessado em: Jan/2017. Disponível em: <https://www.pdf-archive.com/2015/06/13/ibm-journal-of-r-d-this-is-ibm-watson/ibm-journal-of-r-d-this-is-ibm-watson.pdf>

IBM Research. **"This is Watson"**. IBM Journal of Research and Development, Issue 3.4 - Maio/Junho, 2012. Acessado em: 02/02/2017. Disponível em:

http://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=6177717&cm_mc_uid=78650730606414880590983&cm_mc_sid_50200000=1490238292

IBM Watson. **"AlchemyLanguage Overview"**. Junho de 2016. 2016a. Acessado em: Dez/2016. Disponível em: <https://www.youtube.com/watch?v=pQZObFvou40>

IBM Watson. **"Cognitive APIs with Watson: Sentiment Analysis"**. Junho de 2016. 2016b. Acessado em: Dez/2016. Disponível em: <https://www.youtube.com/watch?v=xU6nknznr34>

IBM Watson. **"How it Works"**. Outubro de 2014. Acessado em: Dez/2016. Disponível em: https://www.youtube.com/watch?v=_Xcmh1LQB9I

IBM Watson. **"Introducing Our Cognitive Insights Chalk Talk Series"**. Mackenzie Sedlak. Julho de 2015. Acessado em: Dez/2016. Disponível em: <https://developer.ibm.com/watson/blog/2015/07/23/introducing-our-cognitive-insights-chalk-talk-series/>

IBM Watson. **"Watson Natural Language Understanding Service Overview"**. Fevereiro de 2017. 2017b. Acessado em: Dez/2016. Disponível em: <https://www.youtube.com/watch?v=J5IIKj7H8T8>

IBM Watson. **"Why you should use Watson Natural Language Understanding"**. Fevereiro de 2017. 2017a. Acessado em: Dez/2016. Disponível em: <https://www.youtube.com/watch?v=MB4ynlq52C4>

IBM. **"DeepQA Project: FAQ"**. February, 2011. 2011b. Acessado em: Fev/2017. Disponível em: http://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=6177717&cm_mc_uid=78650730606414880590983&cm_mc_sid_50200000=1490238292

IBM. **"IBM Acquires AlchemyAPI, Enhancing Watson's Deep Learning Capabilities"**. Jay Stankiewicz. Março de 2015. New York, NY. 2015. Acessado em: Jan/2017. Disponível em: <https://www-03.ibm.com/press/us/en/pressrelease/46205.wss>

IBM. **"IBM Watson: How it Works"**. Outubro de 2014. 2014. Acessado em 27/01/2017. Disponível em: https://www.youtube.com/watch?v=_Xcmh1LQB9I

IBM. **"IBM Watson: The Face of Watson"**. Fevereiro de 2011. 2011a. Acessado em 02/02/2017. Disponível em: <https://www.youtube.com/watch?v=WIKM732oEek&t=15s>

IBM. **"Sentiment, emotion, attitude, and personality, via Natural Language Processing"**. Julho de 2016. 2016. Acessado em 17/01/2017. Disponível em:

<https://www.ibm.com/blogs/watson/2016/07/sentiment-emotion-attitude-personality-via-natural-language-processing/>

IBM. **"Today with Watson"**. Março de 2016. 2016. Acessado em 16/03/2017. Disponível em: <https://www.youtube.com/watch?v=JtgCcvzQstk>

IBM. **"Watson – A System Designed for Answers: The future of workload optimized systems design. IBM Systems and Technology"**. February, 2011. 2011c. An IBM White Paper. Acessado em: Fev/2017. Disponível em: http://www-03.ibm.com/innovation/us/engines/assets/9442_Watson_A_System_White_Paper_POW03061-USEN-00_Final_Feb10_11.pdf

IZARD, C. E.; **The face of emotion**. Appleton-CenturyCrofts. 1971.

JAMES, W.; **What is an emotion?**. Mind. 9, 188-205. 1884.

JENNINGS, K.; **"Ken Jennings: Watson, Jeopardy e eu, o sabe-tudo obsoleto"**. Fevereiro de 2013. TEDx Seattle. Acessado em: Jan/2017. Disponível em: https://www.ted.com/talks/ken_jennings_watson_jeopardy_and_me_the_obsolete_know_it_all?language=pt-br#t-1382

JOSHI M.; PENSTEIN-ROSÉ C. **Generalizing dependency features for opinion mining**, in: Proceedings of the 4th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2009), Association for Computational Linguistics, Suntec, Singapore, 2009, pp. 313–316.

JURAFSKY, D.; JAMES, H.; **Speech and language processing an introduction to natural language processing, computational linguistics, and speech**. Pearson Education, Inc. Tahan Terbit, 2009. 2000.

KAMAKURA, W.A.; BASUROY, S.; BOATWRIGHT, P. **Is silence golden? An inquiry into the meaning of silence in professional product evaluations**. Quantitative Marketing and Economics 4, 2 (2006), 119–141. 2006.

KANG D.; PARK, Y.; **Review-based measurement of customer satisfaction in mobile service: sentiment analysis and VIKOR approach**, Expert Systems with Application. 2013.

KELLY III, J., & HAMM, S.; **Smart Machines: IBM's Watson and the Era of Cognitive Computing**. Columbia University Press. 2013.

KIM, S. H., PARK, N., & PARK, S. H.; **Exploring the effects of online word of mouth and expert reviews on theatrical movies' box office success**. Journal of Media Economics, 26(2), 98-114. 2013.

KIM, Y.; KWON, Do Y.; JEONG, S. R. **Comparing machine learning classifiers for movie WOM opinion mining**. KSII Transactions on Internet and Information Systems, v. 9, n. 8, p. 3178-3190. 2015.

KING, T; **Does film criticism affect box office earnings? Evidence from movies released in the U.S. in 2003**. Springer Science + Business Media B.V. Journal of Cultural Economics N° 31, 171–186. 2007.

KLEINGINNA, P. R.; KLEINGINNA, A.; **A categorized list of emotion definitions, with suggestions for a consensual definition**. Motivation Emotion, vol. 5, no. 4, pp. 345–379, 1981.

KULSHRESTHA, J., ZAFAR, M. B., NOBOA, L. E., GUMMADI, K. P., & GHOSH, S.; **Characterizing Information Diets of Social Media Users**. In ICWSM (pp. 218-227). 2015.

Legoux, R.; Larocque, D.; Laporte, S.; Belmati, S.; Boquet, T.; **The effect of critical reviews on exhibitors' decisions: Do reviews affect the survival of a movie on screen?**. International Journal of Research in Marketing, 33(2), 357-374. 2016.

LI, W; XU, H. **Text-based emotion classification using emotion cause extraction**, Expert Systems with Applications. 2013.

LI, Y.M.; LI, T.-Y. **Deriving market intelligence from microblogs**, Decis. Support Syst. 55, pp. 206–217. 2013.

LIM, J.; KIM, J. **An Empirical Comparison of Machine Learning Models for Classifying Emotions in Korean Twitter**. J. Korea Multimed. Soc, vol. 17, no. 2, pp. 232–239. 2014.

LIU B.; **Sentiment analysis and opinion mining**, Synth. Lect. Hum. Lang. Technol. 5 (1) 1–167. 2012.

LIU, B.; **Sentiment analysis and subjectivity**, in: R. Dale, H. Moisl, H. Somers (Eds.), Handbook of Natural Language Processing, second ed., Machine Learning and Pattern Recognition, CRC Press, New York, NY, USA, pp. 627–666. 2010.

LIU, Yong. **Word of mouth for movies: Its dynamics and impact on box office revenue**. Journal of marketing 70, 3, 74–89. 2006.

LOIA, V., & SENATORE, S.; **A fuzzy-oriented sentic analysis to capture the human emotion in Web-based content**. Knowledge-based systems, 58, 75-85. 2014.

LU, Y.; KONG, X.; QUAN, X.; LIU, W.; XU, Y. **Exploring the sentiment strength of user reviews**, In International Conference on Web-Age Information Management. Springer Berlin Heidelberg. WAIM 2010, LNCS 6184, pp. 471–482. 2010.

LU, Yafeng; WANG, Feng; MACIEJEWSKI, Ross. **Business intelligence from social media: A study from the vast box office challenge**. Computer Graphics and Applications, IEEE 34, 5 (2014), 58–69. 2014.

LUCA, M. **Reviews, Reputation, and Revenue: The Case of Yelp.com**. Harvard Business School, Working Paper. 2011.

MA, Y., ZENG, Y., REN, X., & ZHONG, N.; **User interests modeling based on multi-source personal information fusion and semantic reasoning**. In International Conference on Active Media Technology (pp. 195-205). Springer Berlin Heidelberg. 2011.

MANNING, C. D.; RAGHAVAN P.; SCHÜTZE, H.; **Introduction to information retrieval**. Vol. 1, No. 1. Cambridge: Cambridge university press. 2008.

MANNING, C.D.; RAGHAVAN, P.; SCHUTZE, H. **Introduction to Information Retrieval**, Cambridge University Press, New York, USA, 2008.

MAYZLIN, D.; **Promotional Chat on the Internet**, Marketing Science, 25 (2), 155–63. 2006.

McKENZIE, J. **The Economics Of Movies: A Literature Survey**. Journal of Economic Surveys Vol. 26, N°1, 42–70. 2012.

McKENZIE, J.; **Revealed word-of-mouth demand and adaptive supply: survival of motion pictures at the Australian box office**. Journal of Cultural Economics, 33(4), 279-299. 2009.

MEDHAT, W. HASSAN A.; KORASHY, H.; **Sentiment analysis algorithms and applications: a survey**, Ain Shams Engineering Journal, 5(4), 1093-1113. 2014.

MEEHAN, K., LUNNEY, T., CURRAN, K., & MCCAUGHEY, A.; **Context-aware intelligent recommendation system for tourism**. In Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on (pp. 328-331). IEEE. 2013.

MITCHELL, T. **Machine Learning**, McGraw-Hill, 1997

MOHAMED, M.; **More than words: social networks' text mining for consumer brand sentiments**, Expert Syst. Appl. 40, 4241–4251. 2013.

MOHAMMAD, 2012: From once upon a time to happily ever after: Tracking emotions in mail and books Mohammad, Saif M. Crown Copyright (C) 2012 Published by Elsevier B.V. All rights reserved. 2012 DECISION SUPPORT SYSTEMS

MOHAMMAD, S. M.; & TURNEY, P. D.; **Crowdsourcing a word–emotion association lexicon**. Computational Intelligence, 29(3), 436-465. 2013.

MOHAMMAD, S. M.; **Sentiment analysis: Detecting valence, emotions, and other affectual states from text**. Emotion Measurement: 201-238. 2015.

MOHAMMAD, S.; **From once upon a time to happily ever after: Tracking emotions in novels and fairy tales**. In Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (pp. 105-114). Association for Computational Linguistics. 2011.

MOON, S.; BERGEY, P. K.; IANCOBUCCI D.; **Dynamic Effects Among Movie Ratings, Movie Revenues, and Viewer Satisfaction**. Journal of Marketing Vol. 74, 108–121, American Marketing Association. 2010.

MORAES, R; VALIATI, J.F; GAVIAO, W.P. **Document-level sentiment classification: an empirical comparison between SVM and ANN**, Expert Systems with Applications. 40, 621–633. 2013.

MORETTI, E. **Social learning and peer effects in consumption: Evidence from movie sales**. The Review of Economic Studies 78, 1, 356–393. 2011.

MPAA Motion Picture Association of America. **The Economic Contribution of the Motion Picture and Television Industry: to the United States**. Technical Report, 2013.

MUNEZERO, M.; MONTERO, C.; SUTINEN, E.; PAJUNEN, J.; **Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text**. IEEE Transactions on Affective Computing, 5(2), 101–111. 2014.

NIE, P., ZHAO, X., YU, L., WANG, C., & ZHANG, Y.; **Social Emotion Analysis System for Online News**. In Web Information System and Application Conference (WISA), 2015 12th (pp. 43-48). IEEE. 2015.

ORTONY, A.; & TURNER, T. J.;. **What's basic about basic emotions?**. Psychological review 97.3 (1990): 315. 1990.

PANG, B; LEE, L. **A sentiment education: sentiment analysis using subjectivity summarization based on minimum cuts**, in: Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics, p. 271. 2004.

PANG, B; LEE, L. **Opinion mining and sentiment analysis. Foundations and trends in information retrieval**, v. 2, n. 1-2, p. 1-135, 2008.

PANG, B; LEE, L.; VAITHYANATHAN, S. **Thumbs up?: sentiment classification using machine learning techniques**, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2002), vol. 10, Association for Computational Linguistics, Philadelphia, PA, USA, pp. 79–86. 2002.

PANG, L., ZHU, S., & NGO, C. W.; **Deep multimodal learning for affective analysis and retrieval**. IEEE Transactions on Multimedia, 17(11), 2008-2020. 2015.

PENNEBAKER, J. W., FRANCIS, M. E., & BOOTH, R. J.; **Linguistic inquiry and word count: LIWC 2001**. Mahway: Lawrence Erlbaum Associates, 71(2001). 2001.

PICARD R. W.; **"Affective Computing"**. MIT Technical Report #321. 1995.

PICARD R. W.; **Affective Computing**. Cambridge, MA: MIT Press, Vol. 252. 1997.

PLUTCHIK, R.; **A general psycho evolutionary theory of emotion. Emotion**. Theory, Research, and Experience. American Press, New York. (3):3–33. 1980.

PLUTCHIK, R.; **The Emotions**. Random House: New York. 1962.

PLUTCHIK, R.; **The Nature of Emotions Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice**. American Scientist, 89(4), 344-350. 2001.

PORIA, S., CAMBRIA, E., HUSSAIN, A., & HUANG, G. B.; **Towards an intelligent framework for multimodal affective data analysis**. Elsevier 2014. Neural Networks, 63, 104-116. 2015.

PRABOWO, R.; THELWALL, M.; **Sentiment analysis: a combined approach**, J. Informet. 3, 143–157. 2009.

RAVI, K.; RAVI, V.; **A survey on opinion mining and sentiment analysis: Tasks, approaches and applications**. Knowledge-Based Systems, v. 89, p. 14-46, 2015.

REBOLLEDO, V.L.; L'HUILLIER, G.; VELÁSQUEZ, J.D.; **Web pattern extraction and storage**, in: Advanced Techniques in Web Intelligence-I, Springer, Berlin, Heidelberg, pp. 49–77. 2010.

REINSTEIN, D. A.; SNYDER, C. M. **The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics**. The Journal of Industrial Economics, 53, 1. 2005.

REN, F., & KANG, X.; **Employing hierarchical Bayesian networks in simple and complex emotion topic analysis**. Elsevier Ltd. Computer Speech & Language, 27(4), 943-968. 2013.

REN, F.; & QUAN, C.; **Linguistic-based emotion analysis and recognition for measuring consumer satisfaction: an application of affective computing**. Information Technology and Management 13.4 (2012): 321-332. 2012.

REN, F.; QUAN, C.; **Linguistic-based emotion analysis and recognition for measuring consumer satisfaction: an application of affective computing**. Information Technology and Management, 13(4), 321-332. 2012.

RODRIGUES, R. G., DAS DORES, R. M., CAMILO-JUNIOR, C. G., & ROSA, T. C.; **SentiHealth-Cancer: A sentiment analysis tool to help detecting mood of patients in online social networks**. International journal of medical informatics, 85(1), 80-95. 2016.

RUI, H.; LIU, Y.; WHINSTON, A. **Whose and what chatter matters? The effect of tweets on movie sales**, Decision Support Systems. 55, 863–870. 2013.

RUSSELL, S.; NORVIG, P.; **Artificial Intelligence - a modern approach**. Prentice-Hall, Englewood Cliffs, 25, 27. 1995.

SAIF, H., HE, Y., & ALANI, H.; **Alleviating data sparsity for twitter sentiment analysis**. CEUR Workshop Proceedings (CEUR-WS. org). 2012.

SAIF, H., HE, Y., & ALANI, H.; **Semantic sentiment analysis of twitter**. The Semantic Web–ISWC 2012, 508-524. 2012.

SAINI, A., BHATIA, N., SURI, B., & JAIN, S.; **EmoXtract: Domain Independent Emotion Mining Model for Unstructured Data**. In Contemporary Computing (IC3), 2014 Seventh International Conference on (pp. 94-98). IEEE. 2014.

SALEH, M. R., MARTÍN-VALDIVIA, M. T., MONTEJO-RÁEZ, A., & UREÑA-LÓPEZ, L. A.; **Experiments with SVM to classify opinions in different domains**. Expert Systems with Applications, v. 38, n. 12, p. 14799-14804, 2011.

SARAVIA, E., ARGUETA, C., & CHEN, Y. S.; **Emoviz: Mining the world's interest through emotion analysis**. In Advances in Social Networks Analysis and Mining (ASONAM), 2015 IEEE/ACM International Conference on (pp. 753-756). IEEE. 2015.

SERRANO-GUERRERO, J., OLIVAS, J. A., ROMERO, F. P., & HERRERA-VIEDMA, E.; **Sentiment analysis: a review and comparative analysis of web services**. Information Sciences, 311, 18-38. 2015.

SHAPIRO, C.; VARIAN, H. R.; **Information rules: a strategic guide to the network economy**. Boston: Harvard Business School Press. 1999.

SHRUM, W.; **Critics and publics: Cultural mediation in highbrow and popular performing arts**. American Journal of Sociology, 97(2), 347-375. 1991.

SILVER, J. D. **Hollywood's dominance of the movie industry: How did it arise and how has it been maintained**. 612 p. Tese (Doutor em Filosofia) – Queensland University of Technology – Department of Advertising, Marketing and Public Relations, Australia. 2007.

SIMONTON, D. K. **Cinematic Success Criteria and Their Predictors: The Art and Business of the Film Industry**. Psychology & Marketing. v. 26, 5, p. 400-420. 2009.

SINGH, V. K., PIRYANI, R., UDDIN, A., & WAILA, P.; **Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification**. In Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on (pp. 712-717). IEEE. 2013.

SINGH, V. K., PIRYANI, R., UDDIN, A., & WAILA, P.; **Sentiment analysis of Movie reviews and Blog posts**. In Advance Computing Conference (IACC), 2013 IEEE 3rd International (pp. 893-898). IEEE. 2013.

STIGLER, G. J. **The economics of information**. The Journal of Political Economy, 213–225. 1961.

SUBASIC, P.; & HUETTNER, A.; **Affect analysis of text using fuzzy semantic typing**. IEEE Transactions on Fuzzy Systems, 9(4):483–496. 2001.

TAO, J.; TAN, T.; **Affective Computing: A Review**. In International Conference on Affective Computing and Intelligent Interaction. LNCS 3784. Springer. pp. 981–995. doi:10.1007/11573548. 2005.

TOMKINS, S. S.; **Affect theory**. In K. R. Scherer & P. Ekman (Eds.). Approaches to emotion, 163, 195. 1984.

TRUSOV, M.; BUCKLIN, R. E.; PAUWELS, K. **Effects of Word-of-Mouth Versus Traditional Marketing: Findings from an Internet Social Networking Site**, Journal of Marketing, 73 (September), 90–101. 2009.

Turkle, S; **The Second Self: Computers and the Human Spirit**. Simon & Schuster, Australasian Journal of Philosophy, 63, 520. 1984.

TURNEY, P. **Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews**, in: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics ACL'02, Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 417–424. 2002.

VENTURA, D.; LINS, M.A.D.T. **Educação superior e complexidade: integração entre disciplinas no campo das relações internacionais**. (Fundação Carlos Chagas. Impresso), 2014. Cadernos de Pesquisa (Fundação Carlos Chagas. Impresso), v. V.151, p. 104-131. 2014.

VILARES, D; ALONSO, M. A; GOMÉS-RODRÍGUES, C. **A syntactic approach for opinion mining on Spanish reviews**, Nat. Lang. Eng. 21 (01) 139–163. 2015.

WALLS, W. D.; **Product survival at the cinema: Evidence from Hong Kong**. Applied Economics Letters, 5(4), 215-219. 1998.

WEICHSELBRAUN, A.; GINDL, S.; SCHARL, A. **Extracting and grounding contextualized sentiment lexicons**, IEEE Intelligent Systems 28, no. 2 (2013): 39-46. 2013.

WOOLDRIDGE, J. M.; **Econometric Analysis of Cross Section and Panel Data**. MIT Press, Cambridge MA. 2002.

WU, X., BARTRAM, L., & SHAW, C.; **Plexus: An Interactive Visualization Tool for Analyzing Public Emotions from Twitter Data**. arXiv preprint arXiv:1701.06270. 2017.

WU, Y.; KITA, K.; MATSUMOTO, K.; **Three predictions are better than one: Sentence multi-emotion analysis from different perspectives**. IEEJ Transactions on Electrical and Electronic Engineering, 9(6), 642-649. 2014.

XIA, R.; ZONG, C; LI, S. **Ensemble of feature sets and classification algorithms for sentiment classification**, Inform. Sci. 181, 1138–1152. 2011.

YU, Y.; DUAN, W.; CAO, Q. **The impact of social and conventional media on firm equity value: a sentiment analysis approach**, Decis. Support Syst. 55, 919–926. 2013.

YUAN-YUAN, H.; YI-JUN, L.; QIANG, Y.; PENG, Z. **Dynamic impacts of online reviews and other information sources on sales in panel data environment: Evidence from movie industry**. In Management Science and Engineering. ICMSE

ZHANG, X.; ZHOU, Y.; **Holistic approaches to identifying the sentiment of blogs using opinion words**. In International Conference on Web Information Systems Engineering (pp. 15-28). Springer Berlin Heidelberg. 2011.

ZHANG, Zhu. Weighing stars: Aggregating online product reviews for intelligent e-commerce applications. *Intelligent Systems, IEEE* 23, 5, 42–49. 2008.

ZHU, F.; ZHANG, X. **Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics.** *Journal of marketing* 74, 2, 133–148. 2010.

Anexo A – Script de análise afetiva com API Watson

Quadro 3 - Algoritmo para análise afetiva com API Watson

```
# Thaís Luiza Donega e Souza; 11/2016  
# IBM Watson AlchemyLanguage API  
# Análise Afetiva usando método combinado AlchemyLanguage  
  
import json  
import pandas as pd  
import csv  
from watson_developer_cloud import AlchemyLanguageV1  
from watson_developer_cloud import WatsonException  
from requests.exceptions import ConnectionError  
import requests  
  
# Registra API Key  
APIKEY = ['8ebec74c78cbb4c2bfc8bca6f7e6599880a48961']  
alchemy_language = AlchemyLanguageV1(api_key=APIKEY[id_key])  
  
# Definição das variáveis  
colunas = ['ID', 'TITULO', 'ROTTEN_ID', 'BOXOFFICE_ID', 'ROTULO',  
'DATA', 'CRITICA', 'DIA', 'MES', 'ANO', 'ANO_ROTTEN', 'REL_DATE',  
'ANO_BOXOFFICE', 'MES_BOXOFFICE', 'DIA_BOXOFFICE']  
linha = []  
erros = []  
  
# Leitura e Escrita dos arquivos de Entrada e Saída  
nome_file = 'CriticasAudiencia' + '.csv'  
df = pd.read_csv(nome_file, names = colunas, delimiter = ',',  
encoding='latin-1')  
nomearq = 'CriticasAudienciaMinerado' + '.csv'  
nomearqerro = 'CriticasAudiencia_UnsupportedLanguage' + '.csv'  
resultFile = open(nomearquivo, 'w')  
errorfile = open(nomearquivoerro, 'w')  
wr = csv.writer(resultFile, dialect='excel')  
wr2 = csv.writer(errorfile, dialect='excel')  
  
# Para cada linha leia a crítica, passe para a API, escreva o  
# retorno da análise afetiva.  
for index, row in df.iterrows():  
    critica = row['CRITICA']  
    print(row['ID'])  
  
    try:  
        response = alchemy_language.combined (text=critica,  
        extract='doc-emotion, doc-sentiment', sentiment=1,  
        max_items=1)
```

```

except WatsonException as err:
    erros.append([row['ID'], critica])
    wr2.writerows(erros)
    errorfile.flush()
    continue

language = response['language']
sentscore = response['docSentiment']['type']
senttype = response['docSentiment']['type']
anger = response['docEmotions']['anger']
joy = response['docEmotions']['joy']
fear = response['docEmotions']['fear']
sadness = response['docEmotions']['sadness']
disgust = response['docEmotions']['disgust']
status = response['status']

if language == 'english':
    linha.append ([row['ID'], row['TITULO'],
    row['ROTTEN_ID'], row['BOXOFFICE_ID'],
    row['ROTULO'], row ['DATA'], row['DIA'], row['MES'],
    row['ANO'], row['ANO_ROTTEN'], row['REL_DATE'],
    row['ANO_BOXOFFICE'], row['MES_BOXOFFICE'],
    row['DIA_BOXOFFICE'], sentscore, senttype, anger,
    joy, fear, sadness, disgust, status])
else:
    linha.append ([row['ID'], row['TITULO'],
    row['ROTTEN_ID'], row['BOXOFFICE_ID'],
    row['ROTULO'], row ['DATA'], row['DIA'], row['MES'],
    row['ANO'], row['ANO_ROTTEN'], row['REL_DATE'],
    row['ANO_BOXOFFICE'], row['MES_BOXOFFICE'],
    row['DIA_BOXOFFICE'], sentscore, senttype, NULL,
    NULL, NULL, NULL, NULL, NULL])

# Salva arquivo
wr.writerows(linha)
resultFile.flush()
linha = []

# Fim da análise, feche os arquivos
resultFile.close()
errorfile.close()

```

Fonte: Thaís Luiza Donega e Souza, 2016

Anexo B – Descrição das Variáveis

Tabela 18 - Descrição das Variáveis

Origem	Descrição	Obs	Média	Std. Dev.	Min	Max
Coleta	Filmes	4,779				
Coleta	Total Observações	42,274				
Construção	Dummy para Atores mais bem pagos dos últimos 2 anos	42,274	.1226759	.3280687	0	1
Coleta	Orçamento Deflacionado	19,85	61.98647	59.94059	.01	344.51
Coleta	Bilheteria Nacional Total Deflacionada	42,274	4.90e+07	8.41e+07	35.44	8.83e+08
Coleta	Ano de Lançamento	42,274	2009.449	3.423224	2004	2015
Coleta	Mês de Lançamento	42,274	6.779794	3.359014	1	12
Coleta	Número da Semana	42,274	7.666414	5.922543	1	62
Construção	Nota média consumidores total (score)	32,692	3.419089	.5935512	.5	5
Construção	Nota média consumidores antes do final de semana de estreia	24,150	3.465681	.6561232	.5	5
Construção	Nota média consumidores durante do final de semana de estreia	22,851	3.420872	.6958814	0	5
Construção	Nota média consumidores após o final de semana de estreia	31,679	3.409802	.5886869	.5	5
Construção	Nota média consumidores antes e durante o primeiro final de semana	21,555	3.450435	.5794844	.375	5
Construção	Quantidade Total de críticas que foram mineradas emocionalmente	27,954	261.6504	256.6031	0	1830
Construção	Quantidade Total de críticas Negativas - Consumidores	27,954	92.26587	106.0496	0	1010
Construção	Quantidade Total de críticas Positivas - Consumidores	27,954	150.2802	158.1903	0	730
Construção	Quantidade Total de críticas Neutras - Consumidores	27,954	28.16795	28.77868	0	178
Construção	Pontuação média das críticas Negativas - Consumidores	27,954	-.5061494	.1505475	-.972257	0
Construção	Pontuação média das críticas Positivas - Consumidores	27,954	.6104342	.1389889	0	.980984

Construção	Média anger total nos textos de consumidores	27,954	.1205309	.0376861	.002496	.510592
Construção	Média joy total nos textos de consumidores	27,954	.327059	.0992197	.001923	.931992
Construção	Média fear total nos textos de consumidores	27,954	.1348077	.061186	.002744	.7739
Construção	Média sadness total nos textos de consumidores	27,954	.4190776	.070212	.032999	.883265
Construção	Média disgust total nos textos de consumidores	27,954	.1102156	.0396856	.004396	.582632
Construção	Número Total de Semanas	42,274	14.30787	7.350252	1	62
Construção	Log do número de telas no final de semana de estreia	42,273	4.715597	3.167536	0	8.404696
Construção	Log do Orçamento deflacionado	19,854	3.556511	1.293029	-6.71366	5.844968
Construção	Dummy de falha para a última semana de exibição	42,274	.1130482	.3166555	0	1
Construção	Gênero Action	42,274	.0951176	.2933807	0	1
Construção	Gênero Comedy	42,274	.2631641	.4403559	0	1
Construção	Gênero Drama	42,274	.2845011	.4511818	0	1
Construção	Gênero Romance / Family	42,274	.0886597	.2842553	0	1
Construção	Gênero Foreign	42,274	.1466386	.3537495	0	1
Construção	Gênero Crime / War / Western	42,274	.0437385	.204515	0	1
Construção	Gênero Horror	42,274	.0560865	.2300914	0	1
Construção	Gênero Fantasy	42,274	.114231	.3180953	0	1
Construção	Gênero Adventure / Animation	42,274	.0263282	.1601114	0	1
Construção	Gênero Thriller	42,274	.0859157	.2802428	0	1
Construção	Classificação MPAA PG - Até 16 anos	42,274	.4892605	.4998906	0	1
Construção	Classificação MPAA R - Acima de 16 anos	42,274	.3940247	.488646	0	1
Construção	Classificação MPAA R - Não classificado	42,274	.1167148	.3210839	0	1
Construção	Grandes Distribuidores	42,274	.5803331	.4935102	0	1

Construção	Média binária (1=bom; 0=ruim) de críticas positivas dos críticos Top antes do final se semana de estreia	26,745	.5660072	.2499436	.0357143	.972973
Construção	Média binária (1=bom; 0=ruim) de críticas positivas dos críticos Top durante o final se semana de estreia	26,454	.5272645	.2545385	.0344828	.9642857
Construção	Média binária (1=bom; 0=ruim) de críticas positivas dos críticos Top após o final se semana de estreia	15,696	.64388	.2217716	.027027	.9615384
Construção	Média binária (1=bom; 0=ruim) de críticas positivas dos críticos Top antes e durante o final se semana de estreia	22,498	.5456813	.2305966	.0512821	.9574468
Construção	Média binária (1=bom; 0=ruim) de críticas positivas dos críticos Top	32,497	.569559	.2723002	.025	.98
Construção	Dummy para Sentimento (Negativo=-1; Neutro=0; Positivo=1)	27,954	.5003935	.8619209	-1	1
Construção	Sentimento Consumidores de acordo com a análise de sentimentos (Missing / Positivo / Negativo / Neutro)	27,954	4			
Construção	Críticas Profissionais (Muito Bom=1; Bom=2; Ruim=3; Ref: "Sem Avaliação)	34,738	1926104	.9003892	1	3
Construção	Críticas Consumidores (Bom=1; Ruim=0; Ref: "Sem avaliação")	37,548	.5905241	.4917437	0	1
Coleta	Rótulo dos Críticos (Missing / Bom / Muito Bom / Ruim / Sem Avaliações)	37,173				
Coleta	Rótulo da Audiência (Missing / Bom / Ruim / Sem Avaliações)	38,471				
Construção	Classificação por predominância de Anger (1 a 5)	27,954	3.789011	.8038736	1	5
Construção	Classificação por predominância de Joy (1 a 5)	27,954	1.891107	.6863639	1	5
Construção	Classificação por predominância de Fear (1 a 5)	27,954	3.660442	.8836535	1	5
Construção	Classificação por predominância de Sadness (1 a 5)	27,954	1.262109	.4648611	1	5
Construção	Classificação por predominância de Disgust (1 a 5)	27,954	4.397331	.7903706	1	5

Fonte: Thaís Luiza Donega e Souza, 2017

Anexo C – Script de coleta do banco de dados

Os próximos quadros apresentam alguns dos scripts desenvolvidos para a coleta de dados no Box Office Mojo e no Rotten Tomatoes. O quadro 4 apresenta o script de indexação do Crawler.

Quadro 4 - Script de indexação dos filmes

```
# Thaís Luiza Donega e Souza; 10/2015
# Indexação dos filmes de acordo com seu ano de lançamento
# Year's Release BoxOffice

from scrapy.selector import Selector
from scrapy.spiders import CrawlSpider, Rule
from scrapy.linkextractors import LinkExtractor
from LinksPorAno.items import LinksPorAnoItems
import re

class RecursiveSpider(CrawlSpider):
    name = "Links"
    allowed_domains = ["boxofficemojo.com"]
    start_urls = [
        "http://www.boxofficemojo.com/yearly/chart/?page=1&view=releasedate&view2=domestic&yr=1980&p=.htm",
        "http://www.boxofficemojo.com/yearly/chart/?page=2&view=releasedate&view2=domestic&yr=1980&p=.htm"]

    rules = (
        Rule(LinkExtractor(allow=('boxofficemojo\.com/\movies/\?id=*\.'),
            deny=('boxofficemojo\.com/\oscar/\movies/\*\.'),
            callback='parse_item', follow=True), )

    def parse_item(self, response) :
        sel = Selector (response)
        items = []

        item = LinksPorAnoItems()

        #TITLE = ''.join(sel.xpath('//td[2]/b/font/a/text()').extract())
        #LINKWEEK = ''.join(sel.xpath('//td[2]/b/font/a/text()').extract())
        #LINKFILME = ''.join(sel.xpath('//td[2]/b/font/a/@href').extract())
```

```

        item['LINKFILME'] = response.request.url

        items.append(item)
    return items

## Execução em "scapy spider links -o BoxOfficeMoviesList.csv"

```

Fonte: Thaís Luiza Donega e Souza, 2017

O quadro 5 mostra o script de coleta da página de informação semanal do BoxOffice, não há função spider, apenas scrapper. O quadro 6 apresenta o spider que mapeia e visita as páginas de revisões críticas da audiência no Rotten Tomatoes.

Quadro 5 - Script de Coleta Week BoxOffice

```

# Thaís Luiza Donega e Souza; 10/2015
# Coleta do BoxOfficeMojo Utilizando Spider Scrapy

from scrapy.selector import Selector
from scrapy.spiders import Spider
from Week.items import WeekSpiderItems
#from scrapy import optional_features

class WeekendSpider(Spider):
    # Define o Spider (Links já indexados)
    name = "week"
    allowed_domains = ["boxofficemojo.com"]
    moviesfile = open("/ubuntu/Week/spiders/file_week.csv")
    moviesfileread = moviesfile.read()
    movieslist = moviesfileread.split("\r\n")
    print movieslist
    start_urls = movieslist
    # Inicia o parse para Scrap
    def parse(self, response):
        sel = Selector(response)
        items = []
        # Coleta Items por filme!
        for tr in sel.xpath('//table[@class="chart-wide"]//tr'):
            item = WeekSpiderItems()
            MOVIEID = response.request.url
            WEEK
            =
''.join(tr.xpath('./td[1]/nobr/font/a/b//text()').extract())

```

```

        GROSSWEEK
''.join(tr.xpath('./td[3]/font//text()').extract())
        WEEKNUM
''.join(tr.xpath('./td[9]/font//text()').extract())
        RANK
''.join(tr.xpath('./td[2]/font//text()').extract())
        THEATERS
''.join(tr.xpath('./td[5]/font//text()').extract())
        MovieIni = MOVIEID.find("id=") + 3
        MovieFim = MOVIEID.find(".htm")
        MOVIEID = MOVIEID[MovieIni:MovieFim]
# Armazena no Dicionário
        item['WEEK'] = WEEK
        item['GROSSWEEK'] = GROSSWEEK
        item['WEEKNUM'] = WEEKNUM
        item['RANK'] = RANK
        item['THEATERS'] = THEATERS
        item['MOVIEID'] = MOVIEID

        items.append(item)

    return items
## Execução em "scrapy spider week -o BoxOfficeWeekMovies.csv"

```

Fonte: Thaís Luiza Donega e Souza

Quadro 6 - Script de Coleta Revisões Críticas Consumidores RottenTomatoes

```

# Thaís Luiza Donega e Souza; 10/2015
# Coleta das críticas dos consumidores do RottenTomatoes
# Utilizando Spider Scrapy e Crawler

from scrapy.selector import Selector
from scrapy.spiders import Spider
from scrapy.contrib.spiders import CrawlSpider, Rule
from scrapy.contrib.linkextractors.sgml import SgmlLinkExtractor
from scrapy.selector import HtmlXPathSelector
from RottenAudience.items import RottenAudienceItems

# Inicia Crawler e Spider
class RottenAudienceSpider(CrawlSpider):
    movies = []
    Ano = [from 1980 to 2015]
    name = "RottenAudience"+str(Ano)
    allowed_domains = ["rottentomatoes.com"]
    moviesfile = open("/ubuntu/RottenConsumers"+Ano+".csv")
    moviesfileread = moviesfile.read()
    movieslist = moviesfileread.split("\n")

# Para cada filme na lista de filmes do BoxOffice
for filme in movieslist:
    movies.append(filme+"/reviews/?type=user")

```

```

start_urls = movies
rules = (Rule (SgmlLinkExtractor (allow=(), restrict_xpaths =
('//*[ @id="reviews"]/div[2]/a[2]')), callback="parse_items", follow=
True), )
def parse_items(self, response):
sel = Selector(response)
items = []
divs = sel.xpath('//*[ @id="reviews"]/div[3]/div')

# Inicia Scrapper
for div in divs:
    item = RottenAudienceItems()
    NOMEFILME =
sel.xpath('//*[ @id="main_container"]/div[1]/div[1]/div/div[1]/h2/a/t
ext()').extract()
    LINKFILME = response.request.url
    NOMEUSUARIO =
div.xpath("./div[1]/div[3]/a/text()").extract()
    ROTULO = ''.join(div.xpath("./div[2]/span[1]").extract())
    ROTULO_MEIO =
''.join(div.xpath("./div[2]/span[1]/text()").extract())
    CRITICA =
''.join(div.xpath('//*[ @class="user_review"]/text()').extract())
    DATA =
''.join(div.xpath("./div[2]/span[2]/text()").extract())
    ESTRELAS = ROTULO.count('star')
    if ROTULO_MEIO != " ":
        ESTRELAS = int(ESTRELAS) + 0.5

# Salva itens no dicionário
    item['LINKFILME'] = LINKFILME
    item['NOMEUSUARIO'] = NOMEUSUARIO
    item['ROTULO'] = ESTRELAS
    item['CRITICA'] = CRITICA
    item['DATA'] = DATA
    item['NOMEFILME'] = NOMEFILME

    items.append(item)

return items

## Execução em "scapy spider RottenAudience -o RottenAudience.csv"

```

Fonte: Thaís Luiza Donega e Souza

Anexo D – Resultados da Computação Afetiva

De forma prática, o funcionamento do *script* e, conseqüentemente, a chamada à API serão mostradas a seguir. O quadro 7 apresenta o método de chamada à função de mineração de emoção e de análise de sentimento através da combinação, ou seja, para se obter as 2 análises simultaneamente:

Quadro 7 - Chamada Combinada da API

```
alchemy_language.combined (text=critica, extract='doc-emotion,  
doc-sentiment', max_items=1)
```

Em que, *response* é a variável que armazena o retorno da chamada combinada (*combined*) da API *alchemy_language*; *text* é a variável que recebe o texto da crítica a ser analisado, *extract* recebe a informação de extração desejada, neste caso, emoção (*doc-emotion*) e sentimento (*doc-sentiment*) e *max_items* é a quantidade máxima de respostas que esta chamada deve retornar. A ausência de métodos de extração de entidades, palavras-chave ou conceitos implica, necessariamente, na análise no âmbito do documento. A resposta da API, já tratada pelo *script*, é dada pelas tuplas:

Quadro 8 - Leitura do retorno da chamada da API

```
['docSentiment'] ['score'], ['docSentiment']  
['type'], ['docEmotions'] ['anger'], ['docEmotions']  
['joy'], ['docEmotions'] ['fear'], ['docEmotions']  
['sadness']), ['docEmotions'] ['disgust'],  
['language'], ['status']
```

Em que, *score* e *type* de *doc-sentiment* são informações da pontuação e da polaridade do sentimento extraído do texto; *anger*, *joy*, *fear*, *sadness* e *disgust* de *doc-emotion* são as pontuações de cada uma das emoções; *language* retorna a informação da detecção automática do idioma¹² do texto e; *status* a informação de sucesso (OK) ou erro (ERRO) na chamada.

¹² A lista de idiomas suportados é extensa e está disponível em:
<https://www.ibm.com/watson/developercloud/alchemy-language/api/v1/?python#language-support>

Por exemplo, a crítica retirada do filme Capitão Fantástico¹³ (*Captain Fantastic*); crítica avaliada pelo usuário como positiva (5 estrelas).

Figura 15 - Exemplo de crítica retirada do Rotten Tomatoes

lives up to its name in every way. I have never seen anything like it.
The best of movies, one that makes you question what you believe
long after it's done

Fonte: https://www.rottentomatoes.com/m/captain_fantastic/reviews/?type=user

Analizando o texto da revisão crítica pelo AlchemyLanguage Demo¹⁴, figura 16, e chamada da API para a mesma revisão crítica via código, quadro 9:

Figura 16 - Requisição de análise do Documento

Analyze Text

Try the sample content, or paste your own into the text box or URL field.

Body of Text URL

lives up to its name in every way. I have never seen anything like it. The best of movies, one that makes you question what you believe long after it's done

Reset

Analyze

Fonte: <https://alchemy-language-demo.mybluemix.net>

Quadro 9 - Chamada da API AlchemyLanguage para emoção e sentimento

```
alchemy_language.combined (text="lives up to its name in every  
way. I have never seen anything like it. The best of movies, one  
that makes you question what you believe long after it's done",  
extract='doc-emotion, doc-sentiment', max_items=1)
```

As respostas da análise de sentimento via AlchemyLanguage Demo e via API, estão apresentadas na figura 17 e quadro 10, respectivamente. A figura 18 mostra o retorno do

¹³ https://www.rottentomatoes.com/m/captain_fantastic

¹⁴ <https://alchemy-language-demo.mybluemix.net/>

processamento do idioma, bem como seu retorno via *script* da API, no quadro 11. Na figura 19 tem-se a resposta da mineração de emoção, apresentada via código no quadro 12.

Figura 17 - Análise de Sentimento do Documento

Entities	Document Sentiment Identifies the overall positive or negative sentiment within any document or webpage.	
Keywords	View JSON	
Concepts	Sentiment	Score
Taxonomy	positive	0.663396
Document Emotion		

Quadro 10 - Resposta via API da Análise de Sentimento

```
{
  "status": "OK",
  "usage": "By accessing AlchemyAPI or using information
generated by AlchemyAPI, you are agreeing to be bound by
the AlchemyAPI Terms of Use:
http://www.alchemyapi.com/company/terms.html",
  "totalTransactions": "1",
  "language": "english",
  "docSentiment": {
    "score": "0.663396",
    "type": "positive"
  }
}
```

Figura 18 - Resposta da detecção de Idioma

Entities	Language Identifies the language of the document or webpage.			
Keywords	View JSON			
Concepts	Language	Native Speakers	Ethnologue	Wikipedia
Language	english	309-400 million	Ethnologue	Wikipedia

Quadro 11 - Resposta via API da detecção automática do Idioma

```
{
  "status": "OK",
  "usage": "By accessing AlchemyAPI or using information
generated by AlchemyAPI, you are agreeing to be bound by
the AlchemyAPI Terms of Use:
http://www.alchemyapi.com/company/terms.html",
  "url": "",
  "language": "english",
  "iso-639-1": "en",
  "iso-639-2": "eng",
  "iso-639-3": "eng",
  "ethnologue":
"http://www.ethnologue.com/show_language.asp?code=eng",
  "native-speakers": "309-400 million",
  "wikipedia":
"http://en.wikipedia.org/wiki/English_language"
}
```

Figura 19 - Resposta da Mineração de Emoção

Entities	Document Emotion Analyzes the emotions in the entire document or webpage.	
Keywords	View JSON	
Concepts	Emotion	Score
Taxonomy	Anger	0.060962
Document Emotion	Disgust	0.039237
Targeted Emotion	Fear	0.036663
Document Sentiment	Joy	0.444491
Targeted Sentiment	Sadness	0.251976
Typed Relations		

Quadro 12 - Resposta via API da Mineração de Emoção

```
{
  "status": "OK",
  "usage": "By accessing AlchemyAPI or using information
generated by AlchemyAPI, you are agreeing to be bound by the
AlchemyAPI Terms of Use:
http://www.alchemyapi.com/company/terms.html",
  "totalTransactions": "1",
  "language": "english",
}
```

```

"docEmotions": {
  "anger": "0.060962",
  "disgust": "0.039237",
  "fear": "0.036663",
  "joy": "0.444491",
  "sadness": "0.251976"
}
}

```

A partir do exemplo dado, pode-se montar a representação a ser lida pelo *script* através dos retornos das análises solicitadas pela API. O retorno da chamada tratado para entrada de dados está demonstrado no quadro 13, que por sua vez, é a leitura do quadro 7.

Quadro 13 - Retorno completo da chamada combinada da API

```

[ 'SentScore' ][0.663396], [ 'SentType' ][ 'positive' ],
  [ 'Anger' ][0.060962], [ 'Disgust' ][0.444491],
    [ 'Fear' ][0.036663], [ 'Joy' ][0.251976],
      [ 'Sadness' ][0.039237], [ 'english' ], [ 'OK' ]

```

No exemplo apresentado, o Watson AlchemyLanguage identificou corretamente o domínio do texto apresentado. Conforme mostra a figura 20, a taxonomia extraída de maior pontuação foi arte e entretenimento / filmes e TV / filmes.

Figura 20 - Taxonomia do Documento

Entities	Taxonomy Classifies your content into a hierarchy of categories that's five levels deep. View JSON		
Keywords			
Concepts			
Taxonomy			
Document Emotion			
Targeted Emotion			
Document Sentiment			

Label	Score	Confident?
/art and entertainment/movies and tv/movies	0.64871	
/technology and computing/hardware/computer components/memory	0.304537	no
/society/sex/pornography	0.29655	no

Essa identificação é fundamental para que ele possa reconhecer as palavras e frases com conotações afetivas e calcular o sentimento e a emoção do documento completo baseando-se nos treinamentos de dados prévios da mesma área da taxonomia apontada.

A análise obtida condiz, em termos da polaridade, com a nota dada pelo usuário em sua revisão crítica. A tabela 19 traz outros exemplos das análises afetivas aplicadas contendo as críticas e seus respectivos elementos de sentimento e emoção.

O script criado tem chamada para todas as cerca de 800 mil revisões críticas e gera um novo documento com os respectivos resultados. Apesar das revisões críticas serem majoritariamente em inglês, cerca de 2% do total destas revisões foram descartados durante a análise afetiva por estarem em outras línguas não suportadas pela mineração de emoção da API AlchemyLanguage ou estarem em caracteres ilegíveis (tabela 20), resultando em um total de 776.328 revisões críticas analisadas das 777.951 revisões críticas totais.

Tabela 19 - Exemplos da Análise Afetiva das Revisões Críticas

★	SentScore	SentType	Anger	Joy	Fear	Sadness	Disgust
4	0	Neutral	0.077808	0.636968	0.094051	0.078952	0.204069
		Crítica	luvd it esp madhuri & her dancing in the title track!!				
5	0.483694	Positive	0.330451	0.501859	0.050082	0.123864	0.031214
		Crítica	This astonishing, strange documentary is one of the most disturbing films I've ever seen and might be the best film of the year. There's an ongoing discussion in the internet film-o-sphere about whether it should even have been made. I get it but, when the result is art of this calibre, the question answer is obvious. Amazing.				
4,5	0.813781	Positive	0.15196	0.007685	0.283828	0.289597	0.409938
		Crítica	OK, so this is by far the strangest,, weirdest and the most revolting documentary I've ever seen. Seeing this monsters laughing and joking about killing thousands of people is uncomfortable and gut wrenching. I'm strongly debating about the rating on this one it's powerful movie making and a subject matter that everyone should know about but at the same time is so hard to watch.				
2	-0.438432	Negative	0.099627	0.189654	0.190192	0.541317	0.069805
		Crítica	This movie starts well and makes u anticipate something is going to happen... Alas nothing happens and after 20 minutes u wished you have left the theater. U get the point early on which is repeated over and over and over again... OVERRATED				
5	-0.016239	Negative	0.127005	0.157853	0.245586	0.168669	0.379204
		Crítica	Jesus... the most horrifying and greatest documentary I've ever seen. Words cannot express this film, only your eyes and thoughts can.				

Fonte: Thaís Luiza Donega e Souza, 2017

Anexo E – Taxas de Risco Modelos (i) e (ii)

Tabela 23 - Taxas de risco modelo (i)

Variáveis	AFT Weib.	AFT Exp.	AFT Lognor.	AFT Loglog.	Gamma
Média binária de críticas positivas dos críticos Top antes e durante a estreia	1.768*** (0.124)	1.984*** (0.154)	2.270*** (0.205)	2.086*** (0.179)	1.658*** (0.117)
Nota média consumidores total (score)	1.174*** (0.046)	1.168*** (0.052)	1.148*** (0.059)	1.166*** (0.058)	1.174*** (0.042)
Média anger total nos textos de consumidores	2.426 (1.358)	2.230 (1.410)	1.987 (1.378)	1.698 (1.234)	2.486* (1.286)
Média joy total nos textos de consumidores	1.591 (0.680)	1.926 (0.873)	2.361* (1.138)	1.873 (0.883)	1.400 (0.576)
Média fear total nos textos de consumidores	2.224* (0.995)	2.434* (1.115)	2.616* (1.312)	2.284* (1.070)	2.110* (0.936)
Média sadness total nos textos de consumidores	1.107 (0.481)	1.458 (0.681)	1.848 (0.949)	1.568 (0.782)	0.920 (0.384)
Média disgust total nos textos de consumidores	1.039 (0.545)	1.220 (0.679)	1.456 (0.821)	1.163 (0.715)	0.935 (0.465)
Stars	1.046 (0.030)	1.072** (0.034)	1.104*** (0.040)	1.086** (0.037)	1.030 (0.029)
Grandes Distribuidores	1.108*** (0.028)	1.139*** (0.033)	1.199*** (0.042)	1.156*** (0.038)	1.098*** (0.027)
Log dos Número de cinemas na estreia	1.013** (0.006)	1.030*** (0.006)	1.055*** (0.008)	1.040*** (0.007)	1.005 (0.006)
Gênero Ação	0.903*** (0.036)	0.895** (0.040)	0.882** (0.047)	0.898** (0.044)	0.906*** (0.033)
Gênero Comedy	0.989 (0.031)	0.997 (0.035)	1.001 (0.041)	1.009 (0.038)	0.983 (0.030)
Gênero Drama	0.933** (0.031)	0.943 (0.035)	0.967 (0.043)	0.955 (0.039)	0.931** (0.029)
Gênero Romance / Family	1.008 (0.040)	1.020 (0.048)	1.030 (0.060)	1.045 (0.058)	0.995 (0.036)
Gênero Foreign	1.160** (0.068)	1.219*** (0.082)	1.293*** (0.104)	1.307*** (0.100)	1.125** (0.063)
Gênero Crime / War / Western	0.958 (0.045)	0.980 (0.055)	0.991 (0.069)	1.010 (0.065)	0.938 (0.041)
Gênero Horror	0.787*** (0.060)	0.801*** (0.066)	0.815** (0.078)	0.827** (0.071)	0.777*** (0.057)
Gênero Adventure / Animation	1.132*** (0.049)	1.119** (0.054)	1.077 (0.065)	1.119** (0.057)	1.131*** (0.047)
Gênero Thriller	0.913** (0.040)	0.917* (0.046)	0.916 (0.056)	0.923 (0.052)	0.909** (0.037)
Classificação MPAA PG	1.524*** (0.121)	1.673*** (0.154)	1.833*** (0.190)	1.992*** (0.259)	1.456*** (0.109)
Classificação MPAA R	1.292*** (0.098)	1.410*** (0.125)	1.533*** (0.153)	1.664*** (0.212)	1.239*** (0.088)
Constante	2.122* (0.970)	1.164 (0.559)	0.627 (0.307)	0.793 (0.401)	3.081** (1.410)
Observações	17,201	17,201	17,201	17,201	17,201

Fonte: Thaís Luiza Donega e Souza, 2017

Tabela 24 - Taxas de risco modelo (ii)

Variáveis	AFT Weib.	AFT Exp.	AFT Lognor.	AFT Loglog.	Gamma
Críticas Profissionais “Muito Bom”	2.100*** (0.081)	2.235*** (0.084)	2.516*** (0.096)	2.516*** (0.105)	2.430*** (0.098)
Críticas Profissionais “Bom”	1.342*** (0.054)	1.362*** (0.053)	1.363*** (0.053)	1.397*** (0.058)	1.370*** (0.053)
Críticas Profissionais “Ruim”	1.031 (0.042)	1.027 (0.038)	1.038 (0.034)	1.051 (0.037)	1.032 (0.035)
Análise de Sentimento “Positivo”	0.957 (0.027)	0.986 (0.027)	1.029 (0.028)	1.035 (0.029)	1.018 (0.027)
Análise de Sentimento “Negativo”	0.876*** (0.032)	0.889*** (0.030)	0.904*** (0.030)	0.908*** (0.031)	0.901*** (0.029)
Stars	1.072*** (0.027)	1.087*** (0.029)	1.129*** (0.035)	1.109*** (0.034)	1.114*** (0.033)
Grandes Distribuidores	1.261*** (0.027)	1.306*** (0.029)	1.400*** (0.034)	1.378*** (0.035)	1.369*** (0.033)
Log dos Número de cinemas no final na estreia	1.043*** (0.005)	1.060*** (0.005)	1.099*** (0.005)	1.098*** (0.006)	1.086*** (0.007)
Gênero Ação	0.945 (0.032)	0.941* (0.032)	0.937* (0.033)	0.937* (0.033)	0.937* (0.032)
Gênero Comedy	1.025 (0.043)	1.015 (0.037)	1.010 (0.031)	1.012 (0.031)	1.011 (0.031)
Gênero Drama	1.063 (0.047)	1.059 (0.041)	1.068** (0.035)	1.057* (0.036)	1.063* (0.036)
Gênero Romance / Family	0.987 (0.037)	0.990 (0.038)	1.005 (0.041)	1.017 (0.042)	0.999 (0.040)
Gênero Foreign	1.093** (0.049)	1.122*** (0.048)	1.183*** (0.049)	1.228*** (0.053)	1.166*** (0.049)
Gênero Crime / War / Western	1.026 (0.051)	1.041 (0.053)	1.063 (0.057)	1.051 (0.059)	1.058 (0.056)
Gênero Horror	0.813*** (0.035)	0.819*** (0.035)	0.830*** (0.037)	0.847*** (0.040)	0.827*** (0.036)
Gênero Adventure / Animation	1.188*** (0.057)	1.159*** (0.053)	1.091* (0.050)	1.102** (0.050)	1.114** (0.050)
Gênero Thriller	0.974 (0.041)	0.973 (0.039)	0.977 (0.039)	0.977 (0.041)	0.976 (0.039)
Classificação MPAA PG	1.936*** (0.092)	1.979*** (0.086)	1.889*** (0.076)	2.044*** (0.090)	1.946*** (0.081)
Classificação MPAA R	1.646*** (0.077)	1.678*** (0.070)	1.599*** (0.058)	1.725*** (0.071)	1.647*** (0.064)
Constante	3.004*** (0.212)	2.400*** (0.167)	1.528*** (0.105)	1.479*** (0.107)	1.760*** (0.146)
Observações	42,273	42,273	42,273	42,273	42,273

Fonte: Thaís Luiza Donega e Souza, 2017

Tabela 25 - Taxas de risco modelo (i) por gênero – estimação Weibull

Variáveis	Action	Comedy	Drama	Romance Family	Foreign	Crime War Western	Horror	Adventure Animation	Thriller
Média binária de críticas positivas dos críticos Top antes e durante a estreia	1.227 (0.310)	1.728*** (0.206)	2.248*** (0.285)	1.804*** (0.379)	2.389** (1.000)	2.755*** (0.861)	1.378 (0.486)	0.991 (0.237)	1.861*** (0.436)
Nota média consumidores total (score)	1.287* (0.185)	1.287*** (0.099)	1.231*** (0.080)	1.264* (0.173)	1.285 (0.244)	1.142 (0.293)	0.909 (0.163)	1.328*** (0.099)	1.189 (0.153)
Média anger total nos textos de consumidores	1.782 (2.991)	2.856 (3.152)	12.392** (13.591)	157.763*** (305.324)	0.109** (0.114)	5.849 (21.320)	12199482.222*** (44589889.582)	0.070 (0.216)	0.672 (0.915)
Média joy total nos textos de consumidores	2.469 (2.515)	1.160 (1.165)	6.973** (6.603)	12.517 (23.177)	2.694 (3.224)	5.538 (14.176)	25,816.882*** (50,244.429)	0.217 (0.467)	0.269 (0.442)
Média fear total nos textos de consumidores	0.766 (0.993)	0.924 (0.951)	15.004*** (14.092)	6.441 (10.608)	1.045 (0.874)	9.417 (27.329)	15,866.369*** (28,485.913)	0.249 (0.566)	0.086* (0.123)
Média sadness total nos textos de consumidores	2.221 (3.416)	0.421 (0.419)	3.825 (3.852)	6.774 (11.177)	4.514 (5.515)	12.689 (39.515)	51.513* (110.256)	0.104 (0.198)	0.193 (0.320)
Média disgust total nos textos de consumidores	0.173 (0.414)	1.653 (1.887)	1.602 (1.872)	83.061** (180.900)	0.625 (1.120)	0.248 (0.355)	1,093.772*** (1,741.663)	8.135 (23.099)	2.544 (3.970)
Constante	1.938 (2.480)	2.748 (2.961)	0.310 (0.300)	0.017** (0.031)	0.762 (0.903)	0.306 (0.869)	0.000*** (0.000)	20.416 (47.363)	4.146 (6.051)
Observações	1,742	5,013	5,222	1,505	1,490	906	656	2,127	1,575

Fonte: Thaís Luiza Donega e Souza, 2017

Tabela 26 - Taxas de risco modelo (ii) por gênero – estimação Weibull

Variáveis	Action	Comedy	Drama	Romance Family	Foreign	Crime War Western	Horror	Adventure Animation	Thriller
Críticas Profissionais “Muito Bom”	1.743*** (0.172)	2.113*** (0.172)	1.791*** (0.130)	2.104*** (0.287)	2.807*** (0.216)	1.833*** (0.246)	2.139*** (0.309)	1.653*** (0.177)	1.940*** (0.221)
Críticas Profissionais “Bom”	1.315** (0.147)	1.389*** (0.119)	1.126 (0.087)	1.449** (0.228)	1.410*** (0.101)	1.239 (0.195)	1.320** (0.166)	1.188* (0.124)	1.530*** (0.178)
Críticas Profissionais “Ruim”	1.130 (0.091)	1.005 (0.081)	0.835** (0.063)	1.061 (0.101)	1.059 (0.087)	0.921 (0.129)	1.107 (0.105)	1.154 (0.112)	1.073 (0.104)
Análise de Sentimento “Positivo”	0.985 (0.076)	0.955 (0.047)	1.020 (0.051)	1.019 (0.091)	1.042 (0.063)	1.003 (0.107)	0.835* (0.084)	0.870* (0.070)	0.972 (0.081)
Análise de Sentimento “Negativo”	0.915 (0.072)	0.826*** (0.050)	0.809*** (0.053)	0.855 (0.090)	0.931 (0.080)	0.864 (0.120)	0.858* (0.073)	0.901 (0.106)	0.827** (0.073)
Constante	3.826*** (0.606)	3.236*** (0.423)	2.982*** (0.373)	1.996*** (0.378)	3.788*** (0.586)	3.701*** (0.868)	1.943*** (0.330)	3.239*** (0.631)	2.385*** (0.427)
Observações	4,021	11,124	12,027	3,748	6,199	1,849	2,371	4,829	3,632

Fonte: Thaís Luiza Donega e Souza, 2017