

UNIVERSIDADE DE SÃO PAULO  
HOSPITAL DE REABILITAÇÃO DE ANOMALIAS CRANIOFACIAS

PATRICK PEDREIRA SILVA

**Application of data mining to support knowledge discovery  
in the context of the Randomized Clinical Trial - Florida  
Project**

**Aplicação da mineração de dados para apoio à descoberta  
de conhecimento no contexto do Estudo Clínico  
Randomizado - Projeto Florida**

BAURU

2021

## RESUMO

**Objetivo:** O objetivo geral deste trabalho foi ampliar as parcerias entre as áreas de computação e ciências da saúde, sobretudo no escopo das fissuras labiopalatinas, buscando otimizar processos relacionados às práticas dessas ciências, por meio da aplicação de conceitos relacionados à inteligência artificial, ao aprendizado de máquina e à mineração de dados no contexto do HRAC (Hospital de Reabilitação de Anomalias Craniofaciais). Mais especificamente, objetivou-se identificar correlações e padrões nos dados que permitam fornecer insights ao tratamento das fissuras labiopalatinas. **Método:** Foram propostas duas tarefas de pesquisa que são: a comparação de algoritmos de aprendizado de máquina para a predição de ocorrência de fístulas após a palatoplastia primária em pacientes com fissura transforame unilateral (FTU) e a utilização de técnicas de mineração de dados para a descoberta de conhecimento sobre fatores associados à ocorrência de fístulas após a palatoplastia primária. **Resultados:** A análise dos dados revelou que a ausência de alguns sintomas (febre, tosse, infecção) bem como características associadas à cirurgia em si (cirurgião, técnica, retalho de vômer) e ao paciente (hipernasalidade e sinais sugestivos de disfunção velofaríngea), podem ajudar a predizer o sucesso ou insucesso da palatoplastia. Também revelou que fatores associados às complicações pós-operatórias (infecção, vômito, tosse e febre) bem como características associadas à cirurgia em si (duração, técnicas cirúrgicas, retalho de Vômer, incisão relaxante) e ao paciente (idade na época da palatoplastia), podem ajudar a predizer o sucesso ou insucesso da palatoplastia com relação à ocorrência de fístulas consideradas complicações. Quando se considera a capacidade de predição (correta e do maior número de casos) o algoritmo de melhor desempenho foi o de Máquina de Vetores-Suporte (SVM), cuja métrica de f-measure dentro da classe de insucessos foi a mais alta. **Conclusão:** Além das contribuições pontuais desta pesquisa salientadas nos artigos que compõem esta tese, podem ser evidenciadas outras contribuições potenciais deste trabalho, principalmente, a ampliação das parcerias entre as áreas de ciência da computação e ciências da saúde, sobretudo no escopo das fissuras labiopalatinas, o que permite otimizar processos relacionados às práticas dessas ciências, por meio das tecnologias da informação e comunicação, fomentando projetos interdisciplinares.

Palavras-chave: Fissuras. Fístulas. Inteligência Artificial. Mineração de Dados.

## ABSTRACT

**Objective:** the general objective of this work was to expand the interface between the areas of computer science and health sciences, especially in the scope of cleft lip and palate, seeking to optimize processes related to the practices of these sciences, through the application of concepts related to artificial intelligence, to machine learning and data mining in the context of the HRAC (Hospital for the Rehabilitation of Craniofacial Anomalies). More specifically, the objective was to identify correlations and patterns in the data that provide insights into the treatment of cleft lip and palate.

**Method:** two research tasks were proposed in this project, which are: a comparison of machine learning algorithms for predicting the occurrence of fistulas after primary palatoplasty in patients with unilateral transforame cleft (UTC) and the use of data mining techniques for discovery of knowledge about factors associated with the occurrence of fistulas after primary palatoplasty.

**Results:** Data analysis revealed that the absence of some symptoms (fever, cough, infection) as well as characteristics associated with the surgery itself (surgeon, technique, vomer flap) and the patient (hypernasality and signs suggestive of velopharyngeal dysfunction), can help predict the success or failure of palatoplasty. It also revealed factors associated with postoperative complications (infection, vomiting, cough and fever) as well as characteristics associated with the surgery itself (duration, surgical techniques, vomer flap, relaxing incision) and with the patient (age at the time of palatoplasty), can help predict the success or failure of palatoplasty in relation to the occurrence of fistulas considered complications. When considering the prediction capacity (correct and the largest number of cases), the best performing algorithm was the Support Vector Machine (SVM), whose f-measure metric within the failure class was the highest.

**Conclusion:** In addition to the specific contributions of this research highlighted in the articles that make up this thesis, other potential contributions of this work can be evidenced, especially, by the expansion of partnerships between the areas of computer science and health sciences, particularly in the scope of cleft lip and palate, which allows optimizing processes related to the practices of these sciences, contributing with information and communication technologies, that foster future multidisciplinary projects.

Keywords: Cleft. Fistula. Artificial Intelligence. Data Mining.

