

CENTRO UNIVERSITÁRIO DO PARÁ - CESUPA
ESCOLA DE NEGÓCIOS, TECNOLOGIA E INOVAÇÃO - ARGO
CURSO DE ENGENHARIA DA COMPUTAÇÃO

CAUÊ FARIAS MARTINS
GABRIEL GALEGO MAIA

DATASET DE SOBREVIVÊNCIA DE HABERMAN

BELÉM
ANO 2022

1. INTRODUÇÃO

A introdução tem o propósito de conduzir o leitor de um tema geral para um campo particular de pesquisa. Ela estabelece o contexto da pesquisa que está sendo conduzida, resumindo o entendimento atual e informações básicas sobre o tópico, declarando o propósito do artigo na forma de hipótese, questão ou problema de pesquisa. Ela explica brevemente sua justificativa e abordagem metodológica, destacando os potenciais resultados que o estudo pode revelar, e pode descrever a estrutura restante do artigo. A introdução deve seguir um roteiro mental que responde “o que se estava estudando?”, “por que foi importante investigar o tópico?”, “o que se sabia sobre este tópico antes de desenvolver o estudo?”, e “como o estudo aumentará o conhecimento sobre o tópico?”. Serve para apresentar a importância, novidade, contexto e objetivos do trabalho. Uma introdução bem escrita é importante porque, simplesmente, não existe uma segunda chance de causar uma boa primeira impressão. A Introdução não serve para falar sobre tudo o que se sabe ou se quer dizer a respeito tema do artigo, não é uma revisão da literatura, e não detalha métodos, a não ser aqueles necessários à expressão do ineditismo do artigo. Recomenda-se o uso da forma clássica de introdução de 4 parágrafos, apresentando “O porquê o tema geral do trabalho e sua importância” (primeiro parágrafo); o contexto específico do trabalho (segundo parágrafo); principais lacunas de conhecimento (terceiro parágrafo); e novidades do projeto (que atendem as lacunas de conhecimento), seus objetivos e hipóteses (quarto parágrafo).

2. METODOLOGIA

Utilizando-se o site UCI Machine Learning Repository, é possível se deparar com datasets disponibilizados para uso público, variando em diversos temas nos quais se encontram datasets. Ao se deparar com o dataset "Haberman's Survivability Data Set" foi possível instalar por meio do baixa o seu arquivo em formato “.DATA”, uma vez que isto foi feito, nos deparamos com uma banco de dados de 4 colunas. A primeira coluna se refere a idade dos pacientes avaliados, variando desde por volta dos 30 anos até além dos 60 anos, a segunda coluna apresenta o ano no qual se foi feito a análise de sobrevivência dos pacientes, válido notar que dentro do banco ele se apresenta como números de 2 casas decimais tomando como referência o século 20 (ex: 64 = 1964; 99 = 1999; 00 = 1900).

As seguintes colunas se referem ao número de nós axilares que os pacientes possuíam, uma vez que isto é uma informação crucial uma vez que o estudo se baseia no estudo do câncer nas áreas mamárias, e a uma numeração de forma a qualificar o óbito dos pacientes em “1” ou “2”, onde 1 representa pacientes que sobreviveram mais de 5 anos após a cirurgia de câncer de mama, e 2 para pacientes que vieram ao óbito anteriormente a isso.

Após essa avaliação inicial dos valores apresentados, utilizou-se da linguagem Python e, o IDE Google Collab para um dos membros e Visual Code para o outro, para poder gerar tabelas, gráficos e cálculo de valores para poder compreender os resultados como um todo, executado por meio de livrarias digitais pertencentes ao Python como Pandas, Numpy e Matplotlib.

3. RESULTADOS

É possível concluir que a efetividade das cirurgias avaliadas estudadas por Haberman são de um calibre muito positivo e satisfatório, mesmo com o número grande de pacientes sendo avaliados, a grande maioria destes conseguiram viver por mais de 5 anos após tal cirurgia, comportamento este que foi avaliado pelo banco de dados utilizado, em o relatório aponta para uma enorme quantidade de pacientes avaliados com o status de sobrevivência “1” (Sobrevivência de mais de 5 após a cirurgia). Também é notável como a idade dos pacientes não estabeleceu uma relação direta com o índice de sucesso da cirurgia, tanto os mais jovens avaliados quanto os mais velhos apresentando sobrevivência de mais de 5 anos após a cirurgia.

No entanto, nos casos avaliados em que houvesse a infortuna sobrevivência menor a 5 anos, podemos observar que, com consistência, todos estes possuíam um número de nós axilares extremamente elevados, similares a extremos fora da curva em comparação a outros pacientes, podendo apontar para este alto valor como um fator que aumenta o risco de sobrevivência ao processo. Ainda assim, foi possível observar casos relativamente curiosos, em que pacientes jovens e com pouco número de nós ou sequer algum, ainda tiveram uma sobrevivência baixa inferior aos 5 anos.

Em conclusão, esta operação demonstra resultados bem positivos para a área do combate ao câncer de mama, ainda que o alto número de nós axilares reduzisse a eficácia do processo e também de casos falhos em que jovens pacientes e que a saúde axilar em dia ainda sucumbissem precocemente, é de suma importância lembrar que as operações avaliadas dentro do banco datam, em média, a cirurgias efetuadas nos anos 1960, ou seja, uma medicina

ainda em desenvolvimento e de ainda eficácia precária. Por fim, os resultados deste estudo provam desde de tal época um futuro promissor para cirurgias de câncer de mama e um índice de sobrevivência que somente tende a crescer com o avanço da ciência.

3 REFERÊNCIAS BIBLIOGRÁFICAS

Papéis utilizados pelo estudo de Haberman:

Haberman, S. J. (1976). Generalized Residuals for Log-Linear Models, Proceedings of the 9th International Biometrics Conference, Boston, pp. 104-122.

Landwehr, J. M., Pregibon, D., and Shoemaker, A. C. (1984), Graphical Models for Assessing Logistic Regression Models (with discussion), Journal of the American Statistical Association 79: 61-83.

[\[Web Link\]](#)

Lo, W.-D. (1993). Logistic Regression Trees, PhD thesis, Department of Statistics, University of Wisconsin, Madison, WI.

[\[Web Link\]](#)

Conteúdo utilizado pela equipe:

AVALIAÇÃO e remoção dos linfonodos axilares no tratamento do câncer de mama.
G1-Mato Grosso, [S. l.], p. 00-01, 14 maio 2021