

Trabalho final de Introdução Machine Learning

Alison - RM345377
Leonardo - RM345946

Modelo Supervisionado

Identificação de doenças afetadas por genes com um teste sanguíneo.

The logo for MIROCULUS, featuring the word "MIROCULUS" in a bold, white, sans-serif font. The letter "M" is stylized with a vertical line through its center. Below the main name, the tagline "SCIENCE SIMPLIFIED" is written in a smaller, white, sans-serif font. The entire logo is centered within a blue rectangular box.

MIROCULUS™
SCIENCE SIMPLIFIED

Solução do caso com modelo supervisionado

- Relação entre micro-RNA e genes.
- Banco de dados de artigos científicos para análise de textos
- Vetores de características
- Foram testados alguns outros modelos de SVM, Random Forest
- Regressão logística teve resultados mais rápidos e precisos



Modelo não supervisionado

Filtro de Spam

Marcar corretamente um e-mail recebido como spam ou não-spam.

Poderíamos utilizar técnicas supervisionadas para essa modelagem.



Resolução do Modelo não supervisionado

- Filtrar spam com K-means;
- Modelo avalia cabeçalho, remetente e conteúdo e agrupa os dados;
- A classificação dos grupos aumenta a precisão, podendo chegar a 97% de acerto.



Técnica que não precisa de machine learning

Restrição de conteúdos

- Uma empresa quer restringir o acesso a sites e determinados conteúdos;
- Seria possível fazer modelos de machine learning que aprendessem novos endereços, imagens e textos;
- No curto prazo uma lista com palavras-chave e endereços pré-determinados pode ser uma solução mais rápida, barata e ainda assim eficiente;



Apresentação final

Bike Sharing Demand Washington, D.C.

<https://www.kaggle.com/competitions/bike-sharing-demand/overview/description>

<https://www.kaggle.com/code/rajmehra03/bike-sharing-demand-rmsle-0-3194>

