

Software Analytics Canvas

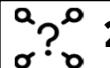
Project: Predição de Insuficiência Cardíaca



1. Question

What is it that we want to know about the software / processes / usage / organization / etc.?

- É possível prever se o paciente é propenso a sofrer uma parada cardiovascular com os dados disponíveis?
- Será que os níveis de colesterol apresentam uma forte correlação em relação a variável de insuficiência cardíaca?
- Será que pessoas que sofreram de insuficiência cardíaca apresentaram comportamentos semelhantes?
- Qual será a faixa etária das pessoas mais afetadas por insuficiência cardíaca?



2. Data Sources

Which data can possibly answer our question? What information do we need?

- Dados relacionados aos resultados dos exames dos pacientes, tais como: Idade, Nível de Colesterol, Dor no Peito, Pressão Arterial, entre outras variáveis
- Precisamos ter um dataset com dados consistentes para que o modelo treinado apresente bons resultados
- Precisamos da nossa variável-alvo para realizarmos inferências e treinarmos o modelo de Aprendizado de Máquina



3. Heuristics

Which assumptions do we want to make to simplify the answer to our question?

- Encontrar, por meio do estudo dos dados, características que possam resultar em casos de insuficiência cardíaca
- Procurar possíveis correlações nos dados em relação a nossa variável-alvo
- Documentar e realizar inferências estatísticas com base nos resultados encontrados



4. Validation

What results do we expect from our analysis, how are they reviewed and presented in an understandable way?

- É esperado uma base com dados consistentes para treinarmos o modelo da melhor forma possível e oferecermos previsões relevantes
- Os resultados serão analisados por meio métricas de classificação de modelos de Aprendizado de Máquina. Além disso, os resultados serão explicados de forma minuciosa na documentação do projeto, a fim de facilitar o entendimento



5. Implementation

How can we implement the analysis step by step and in a comprehensible way?

Para realizar o estudo, será aplicada a metodologia do CRISP-DM ((Cross Industry Standard Process for Data Mining). Essa metodologia é iterativa, então, apesar de termos as etapas definidas, existe a flexibilidade de retornarmos nas etapas de acordo com a necessidade do negócio e resultados.

O CRISP-DM pode ser dividido em algumas etapas::

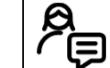
- Business understanding: Parte de entendimento do negócio e do problema a ser resolvido
- Data Understanding: Entendimento inicial dos dados para entendermos como os dados estão distribuídos e quais abordagens serão necessárias para desenvolver o projeto
- Data Preparation: Tratamento dos dados e criação de novas variáveis, a fim de potencializar o desempenho do modelo de acordo com as necessidades do negócio
- Modeling: Desenvolvimento de modelos de Aprendizado de Máquina
- Evaluation: Análise do desempenho dos modelos desenvolvidos. Nesta etapa, validamos se o modelo treinado é capaz de realizar previsões relevantes para solução do problema
- Deployment: Será feita a documentação dos resultados obtidos na pesquisa e as perguntas levantadas sendo respondidas baseadas nos estudos realizados



6. Results

What are the main insights from our analysis?

- Evidenciar que a Inteligência Artificial pode ser uma grande aliada na prevenção de casos de problemas cardiovasculares
- Entender quais variáveis foram mais relevantes para prever possíveis ataques de insuficiência cardíaca
- Entender se existe correlação entre as variáveis estudadas e nossa variável-alvo



7. Next Steps

What follow-up actions can we derive from the findings? Who or what do we need to address next?

- Como próximos passos, podemos solicitar novas bases de dados e entrar em contato com profissionais da área da saúde para ver como podemos melhorar mais o modelo preditivo e, como consequência, apoia-los na prevenção de possíveis casos de insuficiência cardíaca