

## Relatório 2 - Projeto.

### Grupo 2: Bárbara Melo, Juan Gomes, Leonardo Fila e Rebecca Campos

Primeiro notebook - Hta\_predict.ipynb:

O notebook realiza uma análise de risco de ataque cardíaco usando um sistema de lógica fuzzy. O processo se inicia com a importação de bibliotecas essenciais para a manipulação de dados, como pandas e numpy, além da biblioteca skfuzzy para a implementação da lógica fuzzy. Em seguida, ele carrega ou simula um conjunto de dados sobre diabetes, o que garante que o código possa ser executado e testado mesmo se os dados reais não estiverem disponíveis.

O sistema de lógica fuzzy é então construído. Para isso, são definidas as variáveis de entrada, que são a pressão arterial, a idade e a glicose. A variável de saída é o risco de ataque cardíaco. O próximo passo é criar as funções de pertinência, que são representadas por conjuntos fuzzy. Essas funções classificam os valores de entrada em categorias como "baixa", "normal" ou "alta" para as variáveis de entrada, e "baixo", "médio" ou "alto" para a variável de saída.

A lógica central do sistema é definida por quatro regras fuzzy, que estabelecem a relação entre as variáveis de entrada e a saída. Por exemplo, uma das regras determina que se a pressão arterial for alta ou a glicose for alta, o risco de ataque cardíaco é alto. Finalmente, o sistema de controle fuzzy é construído e simulado com um ponto de dados de exemplo do conjunto de dados de diabetes, e o resultado do cálculo do risco de ataque cardíaco é exibido.

Com base nos dados de entrada de pressão arterial de 92.00 mmHg, idade de 30.00 anos e glicose de 110.00 mg/dL, o sistema de lógica fuzzy calculou um risco de ataque cardíaco de 78.31%.

Para a análise, a pressão arterial de 92 mmHg é considerada um valor alto, enquadrando-se no Estágio 2 de Hipertensão, o que pode ser um fator de risco. A idade, 30 anos, é classificada como jovem no sistema fuzzy, enquanto o valor da glicose, 110 mg/dL, é considerado normal.

Apesar da idade e da glicose não serem fatores de alto risco neste exemplo, a pressão arterial elevada é um fator de peso no sistema de lógica fuzzy. No código, uma das regras determina que "se a pressão arterial for alta ou a glicose for alta, o risco de ataque cardíaco é alto." Como a pressão arterial de 92 mmHg é categorizada como "high" pelo sistema, isso aciona a regra, resultando em um alto risco.

A combinação desses fatores leva a um resultado final de 78.31% de risco de ataque cardíaco, demonstrando como o sistema de lógica fuzzy pode pesar a importância de cada variável para chegar a uma conclusão final.

Segundo notebook - Health\_Monitoring.ipynb:

O notebook realiza uma análise de desempenho de modelos de aprendizado de máquina para detecção de diabetes e estresse, com foco na otimização de hiperparâmetros usando um algoritmo genético.

Inicialmente, o modelo de diabetes, um classificador Random Forest, já demonstrava um desempenho robusto com 99% de acurácia, apresentando alta precisão e recall para as duas classes. O modelo de estresse, uma máquina de vetores de suporte, teve uma acurácia de 90%, mas mostrou uma fraqueza no recall da classe "estresse", que foi de apenas 0.62, apesar de uma alta precisão de 0.99 para a mesma classe.

Para aprimorar esses resultados, o notebook utilizou um algoritmo genético para otimizar os hiperparâmetros dos modelos. A otimização do modelo de diabetes resultou nos melhores hiperparâmetros como  $n\_estimators=161$  e  $max\_depth=5$ , elevando a precisão de validação cruzada para 0.9896. O desempenho final do modelo otimizado de diabetes foi perfeito, atingindo 100% de acurácia, precisão, recall e f1-score.

Para o modelo de estresse, os melhores hiperparâmetros encontrados foram  $C=2.0031$  e  $gamma=0.0151$ , com uma precisão de validação cruzada de 0.9860. Após a otimização, o modelo de estresse alcançou 99% de acurácia, e o recall para a classe "estresse" melhorou significativamente para 0.97, resolvendo o principal problema do modelo inicial.