

## Atividade

### 1. Predição de doença cardíaca coronária

Crie um notebook no Google Colab para realizar o estudo em torno do dataset do estudo de doenças do coração realizado na população de Framingham, Massachusetts.

Um estudo completo de análise e estudo dos dados deve ser realizado, incluindo seleção de atributos, limpeza da base e visualização dos dados.

Toda a decisão tomada acerca de escolhas realizadas durante o código deve ser detalhadamente explicada e explicitada.

Utilize uma rede neural e mais dois algoritmos de Machine Learning para realizar o estudo e comparar os resultados utilizando métricas como: Acurácia, precisão, Recall e ROC. Deverá ser justificado a escolha do modelo de Machine Learning escolhido pelo grupo como modelo final.

Trabalhos com copiados do Kaggle ou plagio entre os grupos receberão zero, sem direito a trabalho substitutivo.

#### Infos do dataset:

Demographic: • Sex: male or female(Nominal) • Age: Age of the patient;(Continuous - Although the recorded ages have been truncated to whole numbers, the concept of age is continuous) Behavioral • Current Smoker: whether or not the patient is a current smoker (Nominal) • Cigs Per Day: the number of cigarettes that the person smoked on average in one day.(can be considered continuous as one can have any number of cigarettes, even half a cigarette.) Medical( history) • BP Meds: whether or not the patient was on blood pressure medication (Nominal) • Prevalent Stroke: whether or not the patient had previously had a stroke (Nominal) • Prevalent Hyp: whether or not the patient was hypertensive (Nominal) • Diabetes: whether or not the patient had diabetes (Nominal) Medical(current) • Tot Chol: total cholesterol level (Continuous) • Sys BP: systolic blood pressure (Continuous) • Dia BP: diastolic blood pressure (Continuous) • BMI: Body Mass Index (Continuous) • Heart Rate: heart rate (Continuous - In medical research, variables such as heart rate

though in fact discrete, yet are considered continuous because of large number of possible values.) • Glucose: glucose level (Continuous) Predict variable (desired target) • 10 year risk of coronary heart disease CHD (binary: "1", means "Yes", "0" means "No")

Outras referências: [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))