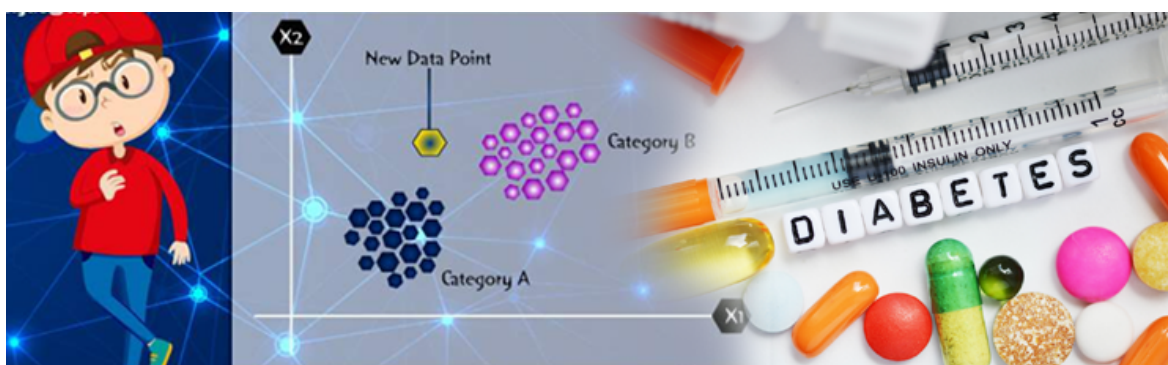


Exercício INDIVIDUAL AVALIATIVO

Envio: encaminhar arquivos via **SIGAA**

Assunto: EX09 - 1R / KNN / COLAB



QUESTÃO 01: Seja o cadastro de pacientes apresentados na tabela à seguir, faça:

- Construa** os seguintes classificadores capazes de distinguir entre pacientes potencialmente saudáveis de pacientes potencialmente doentes:
 - Classificador *1R* e;
 - Classificador *KNN*.
- Testar** cada classificador com **novos casos** a seguir e apresentar as classificações geradas:
 - (Luis, não, não, pequenas, sim);
 - (Laura, sim, sim, grandes, sim).

| Nome | Febre | Enjôo | Manchas | Dores | Diagnóstico |
|-------|-------|-------|----------|-------|-------------|
| João | sim | sim | pequenas | sim | doente |
| Pedro | não | não | grandes | não | saudável |
| Maria | sim | sim | pequenas | não | saudável |
| José | sim | não | grandes | sim | doente |
| Ana | sim | não | pequenas | sim | saudável |
| Leila | não | não | grandes | sim | doente |

QUESTÃO 02: Considerando o Classificador *KNN* (*K-Nearest Neighbor*) desenvolvido no COLAB¹ e disponível para download clicando AQUI, realize as alterações necessárias para a criação de um novo Classificador *KNN* para a doença de Diabetes.

Pontos importantes:

- Para o treinamento você deverá utilizar a base de dados disponível para *download* no *Kaggle*². Demais informações sobre este *dataset* clique AQUI;
- A validação deverá ser realizada a partir de uma base de dados de testes criada por você, crie quantos testes julgar necessário;
- As bases de dados (treinamento e testes) deverão ser carregadas direto do seu GitHub³.
- Por fim, como resolução deste exercício, encaminhe via SIGAA os seguintes arquivos:
 - código *Python Colab*;
 - *.csv* dos dados de treinamento;
 - *.csv* dos dados de testes.

¹O Colaboratory ou **Colab** é um produto do *Google Research*, área de pesquisas científicas do Google. O **Colab** permite que qualquer pessoa escreva e execute código *Python* arbitrário pelo navegador e é especialmente adequado para **Machine Learning**, **Análise de Dados** e educação. Mais tecnicamente, o **Colab** é um serviço de *notebooks* hospedados do *Jupyter* que não requer nenhuma configuração para usar e oferece acesso gratuito a recursos de computação como GPUs (<https://research.google.com/colaboratory/intl/pt-BR/faq.html>).

²**Kaggle**, uma subsidiária da *Google LLC*, é uma comunidade *online* de cientistas de dados e profissionais de aprendizado de máquina. O **Kaggle** permite que os usuários encontrem e publiquem conjuntos de dados, explorem e construam modelos em um ambiente de ciência de dados baseado na *web*, trabalhem com outros cientistas de dados e engenheiros de aprendizado de máquina e participem de competições para resolver desafios de ciência de dados (<https://en.wikipedia.org/wiki/Kaggle>).

³Existem outras formas, mas acredito que esta seja a mais simples.