



Universidade Federal do Ceará
Centro de Ciências
Departamento de Computação
Bacharelado em Engenharia de Computação

CKP9011 – Introdução à Ciência de Dados

CK0268 – Introdução à Ciência de Dados

2022.1

Trabalho 3: Algoritmos de Machine Learning

1. Descrição:

Este trabalho consiste na utilização de algoritmos de machine learning com a finalidade de criar modelos (preditivos, descritivos ou híbridos) que possibilitem extrair padrões ou conhecimento do dataset utilizado nos trabalhos anteriores.

2. Escolha uma das categorias de tarefas/problemas a seguir:

- a. Regressão
- b. Classificação
- c. Agrupamento (Clustering)
- d. Regras de Associação
- e. Detecção de Outlier
- f. Redução de Dimensionalidade e Seleção de Features

3. Escolha uma métrica para avaliação desempenho. Justifique a escolha.

4. Avaliação

Espera-se com a realização deste trabalho que cada estudante elabore e entregue (de forma digital) os seguintes documentos:

- Arquivo CSV contendo os dados do Dataset;
- Jupyter Notebook contendo as etapas de treinamento e teste do modelo, bem como os resultados obtidos em termos de desempenho do modelo.
- Vídeo (disponibilizado no Youtube) apresentando e descrevendo as atividades desenvolvidas.

A avaliação deste trabalho se dará em duas etapas:

1ª. Vídeo de Apresentação do Dataset: Cada estudante irá disponibilizar um vídeo (no Youtube) apresentando o dataset gerado e o processo utilizado para gerar e avaliar o modelo preditivo/descritivo. O estudante pode utilizar slides e notebooks.

2ª. Avaliação do Notebook: O professor da disciplina irá avaliar a qualidade do notebook gerado pelo estudante, bem como dos códigos implementados e análises realizadas.

A avaliação do trabalho irá envolver os seguintes quesitos:

- Abrangência e Organização do Notebook
- Qualidade dos Códigos Utilizados
- Clareza do Texto Utilizado para Descrever as Atividades Realizadas e os Resultados Obtidos
- Domínio do Tema

3. Data da Entrega: 15/07/2022

- PS. Não serão aceitos trabalhos que não forem apresentados (por meio de vídeo disponibilizado no Youtube).
- PS. Cada estudante será responsável pela disponibilização do ambiente (software e hardware) necessário para a gravação da apresentação do seu trabalho.
- Os Notebooks e Arquivos CSV deverão ser disponibilizados, em formato .ZIP, no SIGAA ou em um repositório público (GitHub ou GitLab).
- Utilize o formulário a seguir para indicar as URLs do vídeo e do arquivo ZIP contendo os códigos e os dados utilizados:

<https://forms.gle/dRUSF1n9TwctxSBs7>

“Se não posso estimular sonhos
impossíveis, não devo negar o direito
de sonhar com quem sonha.”.

Paulo Freire