



Universidade Federal do Ceará  
Centro de Ciências  
Departamento de Computação

## **CKP9011 e CK0268 – Introdução à Ciência de Dados 2024.1**

### **Lista 6**

Exercício: Regressão

Objetivos: Exercitar os conceitos referente à regressão.

Data da Entrega: 31/08/2024

OBS 1: Exercício Individual.

OBS 2: A entrega da lista deverá ser executada utilizando-se o SIGAA.

### **Questão 1**

Utilizando os dados referente a postagens no WhatsApp, crie um modelo preditivo (regressor) para, dado os dados de uma postagem, prever a quantidade de compartilhamentos dessa mensagem, o que é denominado potencial de “viralização”.

A avaliação experimental deverá considerar:

- a) O modelo de regressão: regressão linear multivariada, regressão polinomial multivariada;
- b) Explorar funções: exponencial, seno, cosseno;
- c) Regularização: L2 (Ridge), Lasso e ElasticNet;
- d) Normalização dos dados: sem normalização, Z-Score, Min-Max;
- e) Pré-processamento de dados: sem pré-processamento, com pré-processamento;
- f) Embedding: BOW, TF-IDF, Word2Vec;
- g) N-Gramas: unigramas, bigramas, trigramas;
- h) Treinamento, Validação e Teste: Outer K-Fold Cross-Validation;

“A Educação, qualquer que seja ela, é sempre uma teoria  
do conhecimento posta em prática”.

**Paulo Freire**