



FUNDAÇÃO EDSON QUEIROZ  
UNIVERSIDADE DE FORTALEZA  
Programa de Pós-Graduação em Informática Aplicada –  
PPGIA

## TRABALHO FINAL DA DISCIPLINA – MINERAÇÃO DE TEXTO E DA WEB – 2021.1

### **Objetivos:**

Desenvolver um classificador para a tarefa de detecção de inferência entre duas sentenças em linguagem natural. Dada duas sentenças em linguagem natural (Português), o classificador deve definir a relação de inferência entre ambas:

- Neutra
- Implicação (A  $\rightarrow$  B): posso inferir B a partir de A
- Contradição: A contradiz B, ou B contradiz A

Será fornecido um dataset com 9000 pares de sentenças rotuladas (com os resultados), para treinamento e validação.

As equipes serão avaliadas pelos entregáveis e pelos resultados a partir de um conjunto de teste (dados não vistos – Held-Out Data) com 1000 pares de sentenças.

### **Produtos Entregáveis:**

- Pré-processamento, *Steemer*, Lematização e POS Tagger, NER, etc (31/05/2021)
- Vetorização das Sentenças (*features*)
- Model Fitted e Resultados do processo de treinamento (28/06/2021)
- Execução com Held-Out-Data (Dados não vistos) (12/07/2021)
- Artigo resumo com 4 paginas (12/07/2021)

**Apresentação Oral e Entrega do artigo:** 05/07/2021 e 12/07/2021