Ciência da Computação

Instituto Federal do Ceará

Campus Maracanaú

Professor: Hericson Araújo

E-mail: hericson.araujo@ifce.edu.br

Disciplina: Reconhecimento de Padrões

3ª Lista de Exercícios - Regressão Logística

\*\*Instruções:\*\*

- Com exceção dos casos explicitamente indicados, os algoritmos e modelos devem ser implementados do início em qualquer linguagem de programação (Python, Swift, Octave...).

- Pacotes auxiliares (sklearn, matplotlib, etc) podem ser usados somente para facilitar a manipulação dos dados e criar gráficos.

- A entrega da solução pode ser feita via PDF ou Jupyter Notebook pelo Classroom. Os códigos também devem ser enviados.

Desenvolva um relatório incluindo a parte teórica, equações matemáticas, bem como imagens e metodologia utilizada de acordo com as questões abaixo.

---

\*\*Questão 1:\*\*

Considere o conjunto de dados disponível em breastcancer.csv, organizado em 31 colunas, sendo:

- As 30 primeiras colunas representam os atributos coletados de exames médicos.

- A última coluna representa a saída (classe positiva = 1; classe negativa = 0).

Detalhes dos dados:

Maiores informações sobre o dataset podem ser encontradas em:

https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load\_breast\_cancer.html#sklearn.datasets.load\_breast\_cancer

\*\*Subquestões:\*\*

(a) Implemente o algoritmo de Regressão Logística (treinado com GD ou SGD).

(b) Após a implementação do modelo, utilize o método holdout¹ e apresente:

- Valores de acurácia, revocação, precisão e F1-score.

- Matriz de confusão dos dados de teste.

(c) Considerando uma validação cruzada em 10 folds, avalie o modelo de classificação binária do item (a) e apresente:

- Valor médio e desvio padrão das métricas de acurácia, revocação, precisão e F1-score.

¹Divida o conjunto em 80% para treino e 20% para teste.

Fim do documento.