Amrita School of Computing

Dept of CSE

19CSE304 Foundations of Data science

Assignment4

* Marks: Submission shall contribute 05marks towards the Evaluation Lab3, which is allotted 15 marks.
* Individual work: Due date for submission: 01Dec23

------------------------------------------------------------------------------------------------------------------------------

**A. Classifier**

*Datasets: ckd.csv and breast-cancer.csv - Ref Text Book Adhikari.*

Q1. Consider the Chronic Kidney Disease (CKD) dataset ckd.csv to develop a KNN classifier model. *Given that a new patient Alice has Haemoglobin* ***0*** *and Glucose levels* ***1.1*** *standard deviations above average, find if Alice is likely to have CKD.*

Q2. Develop three different classifier models for the breast-cancer.csv dataset. Analyse your results with the standard performance measures, and compare them employing RoC curves.

**B. Regression**

Q3. Predict Home prices employing multiple regression. Draw a residual plot for these predictions and offer your comments.

*Dataset: house.csv -Ref Text Book Adhikari.*

==================================================================================