PROJECT 5

DUE: 11:59PM, APRIL 25

1 Introduction

In this project, you will need to train a ML model to predict how likely an Indian student is admitted to a graduate school. You can use ML packages for this project. We are going to use the following dataset.

Graduate Admissions. This dataset was collected by Acharya et al. [1]. You can download it from https://www.kaggle.com/mohansacharya/graduate-admissions.

The dataset **Admission_Predict_Ver1.1.csv** contains 500 samples with several parameters which are considered important during the application for Masters Programs. The parameters included are: 1. GRE Scores (out of 340) 2. TOEFL Scores (out of 120) 3. University Rating (out of 5) 4. Statement of Purpose 5. Letter of Recommendation Strength (out of 5) 6. Undergraduate GPA (out of 10) 7. Research Experience (either 0 or 1) and The dependent variable - Chance of Admit (ranging from 0 to 1). So the goal here is to predict Chance of Admit using the left parameters.

For this project you need to implement **at least two** different ML algorithms to tackle the university admission prediction. Remember to compare their performance with clear documentation and provide the reasons for the prediction results..

Bonus Question. You will get extra **1 point** if you can build **at least three different** models and compare their performance using cross-validation.

2 Submission

You will need to submit one pdf file and one ipynb generated by the notebook. Fail to do so will make your final grade deducted. Make sure all codes are run before generating

the pdf file. In the report, you should specify your model details when necessary. Try to write your code clearly so that someone else reading the code can understand it without significant effort (i.e. structure it and put enough documentation). The final grade is based on the clarity of your report.

3 Collaboration

Note that this is an independent project, which means you are not allowed to make a group. However, discussion is allowed. If you have discussed with someone or got any help from others, you need to clearly specify their names in acknowledgement.

Principles of Machine Learning IUB CSCI B455 Spring 2021 References

[1] Aneeta S Antony Mohan S Acharya, Asfia Armaan. A comparison of regression models for prediction of graduate admissions. In *ICCIDS*, 2019.