ADMN5016 Assignment: Proof of Concept for Machine Learning Application

Introduction

In this assignment, you will prepare a proposal to create a product or service or improve a company's process using a machine learning application. The project will have two parts, the proposal i.e showing the value of your machine learning application; this includes a proposal and analysis of your results. The second part is the implementation of the application, i.e how well your application works on data. Your final product must be a notebook on GitHub with each part of the project, the report must be a GitHub Page .

Given the timeline of the project it's suggested you work backwards i.e find a dataset that works, and reverse engineer a product or service, some dataset sources include.

kaggle

https://www.kaggle.com/

https://archive.ics.uci.edu/ml/datasets.php

Your application can be similar to existing companies, but you must modify their business plan; for example, Open Door is a company that uses A. I to predict the price of the house and buy it from you. If your A.I the application predicts the price of the home, the application must be different.

Show the value of the Machine learning algorithm 40 points

1. What problem your application solves? If it's a new product, what is the size of the market it can address? If it's I solution how much money can it save and what are the risks.
2. Explain your results: what was the performance of your method using metrics in class? Compare the results with other models example, Linear Regression vs Ridge Regression.
3. What is the monetary value and Risks of your application after its performance? How much money can you save? For example, if you build an application to determine if someone would default on their loan, how much money would you save if your application prevented 50 people from getting a loan who defaulted? How much money did you lose from rejecting people who would have paid back your loan? Would you save money as you need fewer employees? You can estimate values if you can't find the data.
4. Other risks and benefits?

AI Applications 35 points

In the Jupyter lab you must write your A. I pipeline, including data analysis (10), feature engineering (5), preprocessing (5), training at least two models (5), hyperparameter selection (5) and test results (5). If your dataset is challenging, you can get 15 bonus points, but the max you can get in the assignment is 100.

Presentation 25 points

Present results in class using your Jupyter notebook, don't forget to practice.

Note

You will lose marks if you do not answer questions, like what hyperparameters do, why you used a specific model, and what a line of code does.

if you do not have you notebook on Github you will get zero