Machine Learning Extended Project

Machine Learning Extended Project		
Criteria	**Ratings**	**Pts**
This criterion is linked to a Learning Outcome1.1) Read the dataset. Describe the data briefly. Interpret the inferences for each. Initial steps like head() .info(), Data Types, etc . Null value check, Summary stats, Skewness must be discussed.	*This area will be used by the assessor to leave comments related to this criterion.*	6.0 pts
This criterion is linked to a Learning Outcome1.2) Perform EDA (Check the null values, Data types, shape, Univariate, bivariate analysis). Also check for outliers. Interpret the inferences for each Distribution plots(histogram) or similar plots for the continuous columns. Box plots, Correlation plots. Appropriate plots for categorical variables. Inferences on each plot. Outliers proportion should be discussed, and inferences from above used plots should be there. There is no restriction on how the learner wishes to implement this but the code should be able to represent the correct output and inferences should be logical and correct.	*This area will be used by the assessor to leave comments related to this criterion.*	8.0 pts
This criterion is linked to a Learning Outcome1.3) Encode the data (having string values) for Modelling. Is Scaling necessary here or not?, Data Split: Split the data into train and test (70:30). The learner is expected to check and comment about the difference in scale of different features on the bases of appropriate measure for example std dev, variance, etc. Should justify whether there is a necessity for scaling. Object data should be converted into categorical/numerical data to fit in the models. (pd.categorical().codes(), pd.get_dummies(drop_first=True)) Data split, ratio defined for the split, train-test split should be discussed.	*This area will be used by the assessor to leave comments related to this criterion.*	6.0 pts
This criterion is linked to a Learning Outcome1.4) Apply Logistic Regression . Interpret the inferences of the model. Successful implementation of each model. Logical reason behind the selection of different values for the parameters involved in each model. Calculate Train and Test Accuracies for each model. Comment on the validness of models (over fitting or under fitting)	*This area will be used by the assessor to leave comments related to this criterion.*	6.0 pts
This criterion is linked to a Learning Outcome1.5) Apply KNN Model. Interpret the inferences of the model. Successful implementation of each model. Logical reason behind the selection of different values for the parameters involved in each model. Calculate Train and Test Accuracies for each model. Comment on the validness of models (over fitting or under fitting)	*This area will be used by the assessor to leave comments related to this criterion.*	6.0 pts
This criterion is linked to a Learning Outcome1.6) Model Tuning , Bagging and Boosting . Apply grid search on each model (include all models) and make models on best_params. Define a logic behind choosing particular values for different hyper-parameters for grid search. Compare and comment on performances of all. Comment on feature importance if applicable. Successful implementation of both algorithms along with inferences and comments on the model performances.	*This area will be used by the assessor to leave comments related to this criterion.*	12.0 pts
This criterion is linked to a Learning Outcome1.7 Performance Metrics: Check the performance of Predictions on Train and Test sets using Accuracy, Confusion Matrix, Plot ROC curve and get ROC_AUC score for each model, classification report Final Model - Compare and comment on all models on the basis of the performance metrics in a structured tabular manner. Describe on which model is best/optimized, After comparison which model suits the best for the problem in hand on the basis of different measures. Comment on the final model.	*This area will be used by the assessor to leave comments related to this criterion.*	8.0 pts
This criterion is linked to a Learning Outcome1.8) Based on your analysis and working on the business problem, detail out appropriate insights and recommendations to help the management solve the business objective. There should be at least 3-4 Recommendations and insights in total. Recommendations should be easily understandable and business specific, students should not give any technical suggestions. Full marks should only be allotted if the recommendations are correct and business specific.	*This area will be used by the assessor to leave comments related to this criterion.*	8.0 pts
Total Points: 60.0		