CMPS 460 – Project Grading Rubrics

Requirement	Poor (0-50%)	Fair (51-65%)	Good (66-85%)	Excellent (86-100%)
Exploratory Data Analysis (EDA) (**10%**)	- Little to no utilization of appropriate summary statistics or visualizations. - Limited understanding of the dataset's structure, primary distributions, and relationships. - Minimal or absent interpretation of findings; lacks depth and thoroughness.	- Conducts a basic exploration of the dataset with some summary statistics and visualizations. - Demonstrates a partial understanding of the dataset's structure, main distributions, and relationships. - Offers limited or basic interpretations of findings.	- Conducts a comprehensive exploration of the dataset using summary statistics and visualizations. - Shows a good understanding of the dataset's structure, main distributions, and relationships. - Provides reasonable interpretations of findings and identifies some potential patterns or anomalies based on the analysis.	- Conducts a thorough exploration of the dataset employing appropriate summary statistics alongside clear and concise visualizations. - Demonstrates a deep understanding of the dataset's structure, distributions, and relationships. - Offers insightful interpretations of findings and identifies potential patterns or anomalies based on the analysis.
Data Cleaning and Preprocessing (**10%**)	- Little to no addressing of missing values, outliers, and duplicates in the dataset. - Limited investigation of correlations between attributes, resulting in simplistic feature engineering decisions.	- Addresses most missing values, outliers, and duplicates in the dataset using basic techniques. - Conducts some investigation of the correlations between attributes to guide feature engineering decisions, but may lack depth or thoroughness.	- Effectively addresses missing values, outliers, and duplicates in the dataset using mostly appropriate techniques. - Conducts comprehensive investigation of correlations between attributes, providing valuable insights for feature engineering decisions.	- Demonstrates mastery in addressing missing values, outliers, and duplicates in the dataset by employing appropriate techniques. - Conducts extensive and insightful analysis of correlations between attributes, leading to highly informed feature engineering decisions.
Model Selection and Training (**45%**)	- Poorly design and train two or three ML models. - Limited or no justification provided for each model choice that mostly ignores the dataset characteristics, and the desired performance metrics.	- Successfully design and train two or three basic ML models but lacking either an ensemble or deep learning model. - Present basic justifications for each model choice with limited consideration of the dataset characteristics, and the desired performance metrics.	- Thoroughly design and train ML models: one traditional, one ensemble, and one deep learning. - Present a comprehensive and well-reasoned justification for each model mostly based on the dataset characteristics and the desired performance metrics.	- Expertly design and train three ML models: one traditional, one ensemble, and one deep learning. - Provide insightful justifications for each model based on the dataset characteristics and the desired performance metrics.

Model Evaluation and Comparison (**15%**)	- Poor model evaluation with limited or no comparison of the performance of the models. - Limited or no analysis of strengths, weaknesses for each model.	- Performs basic model evaluation by comparing the performance of the models using some evaluation metrics. - Provides basic analysis of strengths, weaknesses for each model.	- Performs complete model evaluation by comparing the performance of the models using appropriate evaluation metrics. - Provides detailed analysis of strengths, weaknesses, and some areas for improvement for each model.	- Performs comprehensive model evaluation by comparing the performance of the models using various evaluation metrics. - Provides in-depth analysis of strengths, weaknesses, and areas for improvement for each model.
Model Optimization (**15%**)	- Limited or no attempt to improve the model performance, with no observed enhancements.	- Shows moderate improvement in model performance by applying basic techniques like feature ranking or regularization but lacks thorough hyperparameter tuning.	- Achieves significant improvement in model performance by applying appropriate techniques such as feature ranking/selection, regularization, or hyperparameter tuning.	- Achieves superior model performance by applying multiple techniques such as feature ranking/selection, regularization, or hyperparameter tuning.
Jupyter notebook writing quality (**5%**)	- Poorly structured with no clear sections or headings, making the notebook difficult to navigate. - Untidy and confusing Code and Markdown cells hinder understanding. - Variable and function names are poorly chosen or inconsistent.	- Some structure and basic headings exist, but further organization is needed. - Code and Markdown cells are moderately commented, making them somewhat readable and tidy. - Naming conventions show some consistency but could be improved for clarity.	- Well-organized with clear sections and headings, facilitating easy navigation. - All Code and Markdown cells are properly commented, readable, and tidy. - Naming conventions are consistently followed.	- Exceptionally well-structured notebook with intuitive flow. - Engaging, readable, and impeccably tidy throughout, elevating the overall presentation. - Meticulous adherence to naming conventions further enhances clarity and professionalism.