Mandar Angchekar
386916341

HW06
Introduction to Machine Learning

In the comparative study of machine learning methodologies for citation data prediction and classification, we examined different computational strategies across three homework assignments. In HW4, we implemented a neural regression model that resulted in a Mean Absolute Error (MAE) of 68.7 for forecasting 2022 citation counts. Contrastingly, HW6 employed a linear regression algorithm, which surprisingly outperformed the neural network with a lower MAE of 34.12, suggesting the effectiveness of simpler models for this type of regression task.

For classification purposes, HW5's deployment of a neural network classifier showed superior accuracy over HW6's logistic regression model. The neural network model approach was more effective in distinguishing between different citation categories.

The insights derived from this exercise clearly underscore the importance of model selection in machine learning tasks. For quantitative predictions, such as citation counts, linear regression may offer a more efficient and accurate solution. However, for qualitative tasks that involve categorizing data into discrete classes, the complexity of neural networks may provide a more nuanced understanding, as evidenced by the performance in HW5.

In summary, the decision on which machine learning model to apply is contingent upon the specific nature of the predictive or classification challenge at hand. Linear regression can be more suitable for straightforward predictive analyses, while neural networks may be better equipped to handle complex classification problems, underscoring the necessity of aligning the machine learning approach with the dataset characteristics and the intended analytical objectives.