**Name: Brian Murgor**

**Student Number: 105055230**

**Project 2**

1. **Introduction**

This report shows a Performance analysis of regression and classification models that occur to predict house prices. The dataset used was from Kaggle and I chose to do the House Price prediction. The price prediction used Linear Regression and Decision Tree Regression models yet Logistic Regression and SVM methods served to classify houses as affordable or not

1. **Model Evaluation**

**Regression Models:**

* The model fits linear data with a Mean Squared Error of 47.28B and R² of 0.44 which explains 44% of variance.
* Decision Tree Regression: MSE = 81.89B, R² = 0.03 (High overfitting)

**Classification Models:**

* Logistic Regression: The Accuracy is 71.35%, Balanced precision and recall.
* SVM: The Accuracy is 71.13% and performs similarly to Logistic Regression.

1. **Results Interpretation**

* Linear Regression produces better reliability compared to Decision Tree which demonstrates excessive overfitting.
* Logistic Regression demonstrates minor superior performance over SVM although the mdels show comparable outcomes.
* The next phase includes modifications to features and adjustments of hyperparameter settings.

1. **Implications**

* The structured nature of house pricing data makes Linear Regression a suitable predictive model.
* The Decision Tree demands some changes to achieve better generalization abilities.
* Real estate pricing strategies benefit from adopting classification model systems.

1. **Reflection**

* **Challenges**: Overfitting in Decision Tree, feature selection.
* The combination of structured datasets functions well with linear models and classification models need additional adjustment before they reach optimal performance.

1. **Conclusion**

* Linear Regression provides the most accurate house price protocol yet Logistic Regression performs best for classifying homes. Additional optimization techniques have the power to boost final accuracy results.

1. **References**

* Python Libraries: Scikit-Learn, Pandas, Matplotlib
* Resources:Kaggle datasets