**Advanced kNN Regression Report**

**Insights from Exploratory Data Analysis**

The exploratory data analysis of the Boston Housing dataset presented varied feature distributions with certain attributes like CRIM (crime rate), ZN (proportion of residential land zoned for lots over 25,000 sq.ft.) and B (proportion of black residents) being skewed while others such as RM (average number of rooms) are more normally distributed. The correlation matrix identified strong positive correlations between RAD (index of accessibility to radial highways) and TAX (full-value property-tax rate) at 0.91, and strong negative correlations between DIS (weighted distances to five Boston employment centres) and NOX (nitric oxides concentration) at -0.77. Standardization is successfully applied to the features centering them around zero mean and unit variance crucial for the subsequent modeling steps.

**Justification for Chosen k and Impact on Model Performance**

Hyperparameter tuning is conducted to identify the optimal k value for the kNN regressor. A range of k values (3, 5, 7, 10) are evaluated against a validation set. The k value of 3 yielded the lowest mean squared error (MSE) approximately 9.78 indicating it as the optimal choice for the number of neighbors balancing the bias-variance trade-off effectively. Larger k values resulted in increased MSE suggesting a loss in prediction accuracy thus affirming the selection of k = 3 for the model. The choice of k in the k-Nearest Neighbors algorithm significantly impacts model performance. A smaller k led to models that capture noise and overfit the training data whereas a larger k oversmooth and underfit failing to capture the complexities of the data.

**Evaluation of Model Performance Using MSE and R-squared Metrics**

The model performance is assessed using MSE which measures the average squared difference between the estimated values and the actual value. The performance of the final model with k=3 and distance weighting are evaluated using the MSE and R-squared metrics. An MSE of approximately 24.32 on the test set signifies the average squared prediction error and an R-squared of 0.69 indicates that the model explains 69% of the variance in the target variable which is a moderate level of predictive power.

**Discussion on the Limitations of kNN Regression and the Effect of Distance Weighting**

For the limitations of kNN regression, the algorithm suffers from the curse of dimensionality where performance degrades with an increase in feature space dimensionality. It is also sensitive to the scale of the data, making feature scaling crucial. The kNN model’s reliance on local information makes it susceptible to noise in the data. Distance weighting mitigates some of these issues by giving more weight to nearer neighbors potentially enhancing model performance by reducing the influence of outliers and noisy data points. However, the choice of the distance function and the way the distance is calculated significantly affect the model’s predictions and inappropriate weighting lead to biases towards some data regions.