Analysis

Impact of Train-Test Split
- Generally, a larger training set leads to better model performance because the model has more data to learn from.
- However, with very small training sets (e.g., 95:5 split), the model may not generalize well to new data.

Impact of K Value
- Lower K values can lead to overfitting, where the model is too closely fitted to the training data.
- Higher K values can lead to underfitting, where the model is too simple to capture the underlying patterns in the data.
- The optimal K value often lies between these extremes, balancing bias and variance.

Conclusion
- The performance of the KNN model depends on both the train-test split and the value of K.
- A balanced approach is necessary to ensure that the model generalizes well to new data.