Experiment - 6

Project Report.

AIM : Implementation of K-Nearest Neighbors Algorithm from scratch.

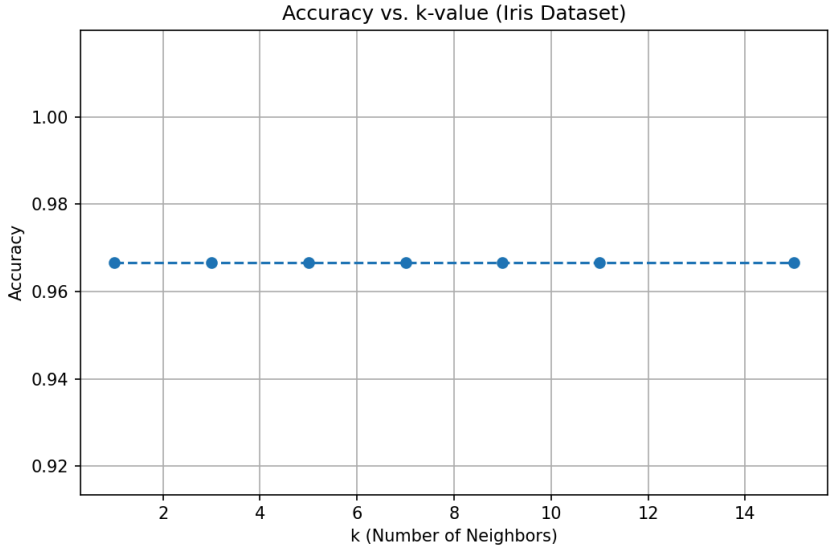
1. Exploratory Data Analysis (EDA) :

Setosa is clearly and linearly separable from the other two species across all feature pairs. **Versicolor** and **Virginica**, however, exhibit overlap—particularly in sepal-based features and slightly in the petal dimensions—leading to mild ambiguity. The **petal length × petal width** combination provides the **best class separation**, showing distinct clusters for all three classes.

2. Accuracy in Iris Dataset (k=1) :

97 %.

3. Accuracy vs k-value Plot.



4. Analysis of K.

The accuracy is the same even in increasing the value of K. In k =1 , the accuracy is 97%.

5. Accuracy on Wine Dataset :

72.22 %.

6. Conclusion

The KNN algorithm was implemented from scratch using NumPy. EDA revealed distinct clusters in the iris dataset, especially for setosa. Hyperparameter tuning confirmed that the accuracy is coming the same for every k , hence k=1 is write choice. The model generalizes well to the WIne dataset , demonstrating string adaptability. This experiment provided hands - on understanding of instance - based learning and distant metrics.