# Assignment02.

# IRIS DATASET

**Conclusion**

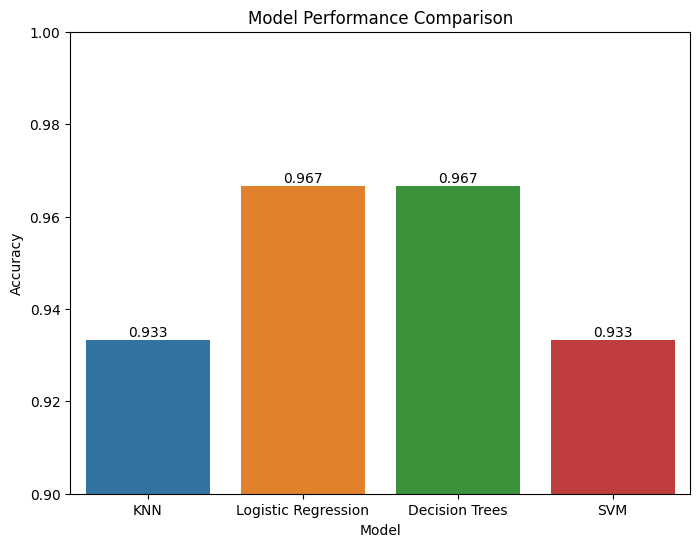
**Model Selection:**

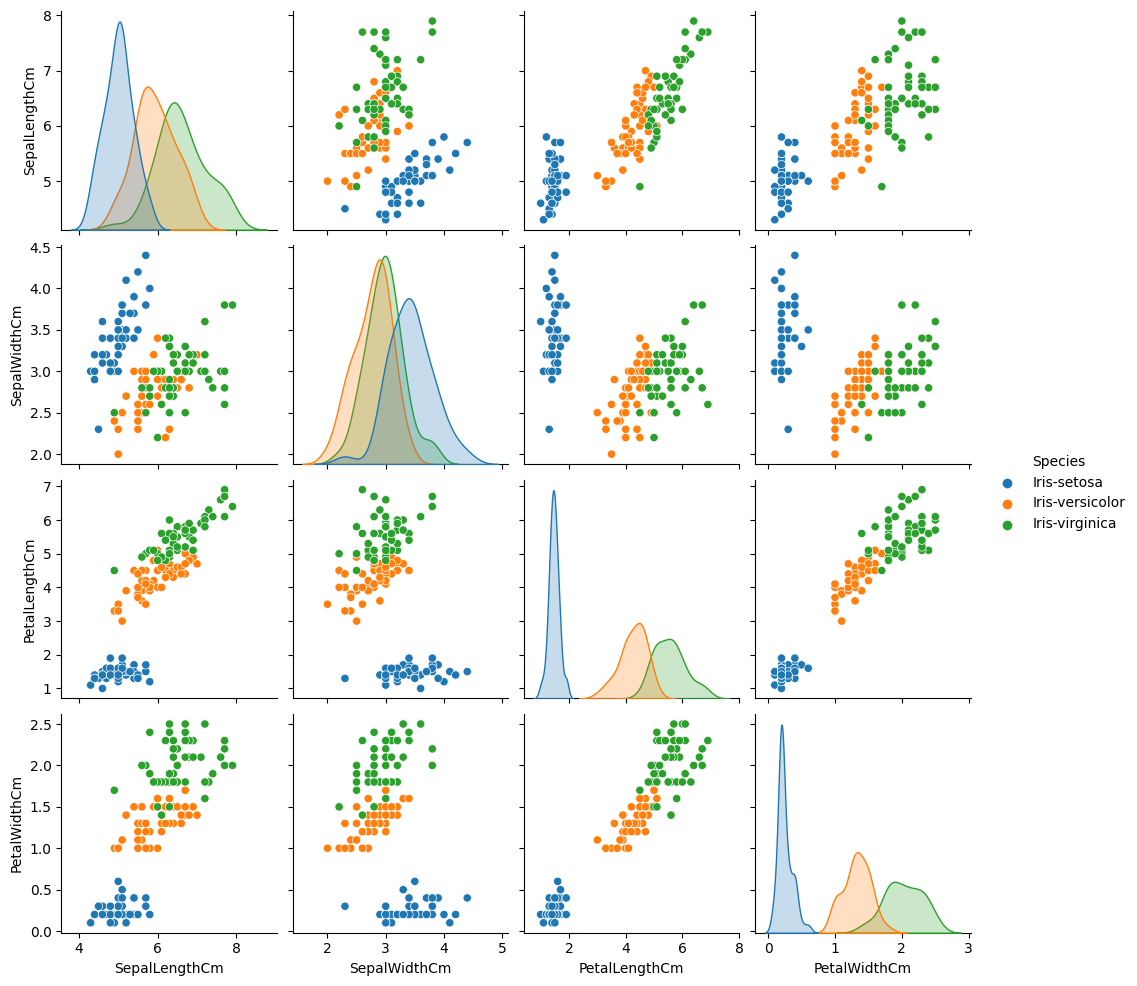
All models demonstrate high accuracy and good performance on the Iris dataset.

Consider other factors for model selection, such as interpretability and computational complexity.

**The insights gained from the analysis.**

* **For the Iris dataset, Logistic Regression & Decision Tree with a linear kernel generally performs well and achieves high accuracy. Below image see Model Performance Comparison.**

****

* **The pair plot provides insights into the relationships between features and how they separate different species.** ****

**Recommendations:**

* For a simpler model and interpretability, use Logistic Regression or Decision Trees.
* For more complex datasets, consider SVM or KNN.
* Ensemble methods like Random Forest may further enhance performance.

**Further Improvements:**

* **Further improvements can be made by exploring other classification algorithms, feature engineering, or hyperparameter tuning.**
* **For more complex datasets, advanced algorithms like Random Forest, Gradient Boosting, or Neural Networks could be considered.**
* **Feature selection techniques or dimensionality reduction methods like PCA might be helpful for high-dimensional data.**
* **Cross-validation: Perform cross-validation for more robust evaluation.**
* **Ensembles: Explore combining multiple models using ensemble methods.**