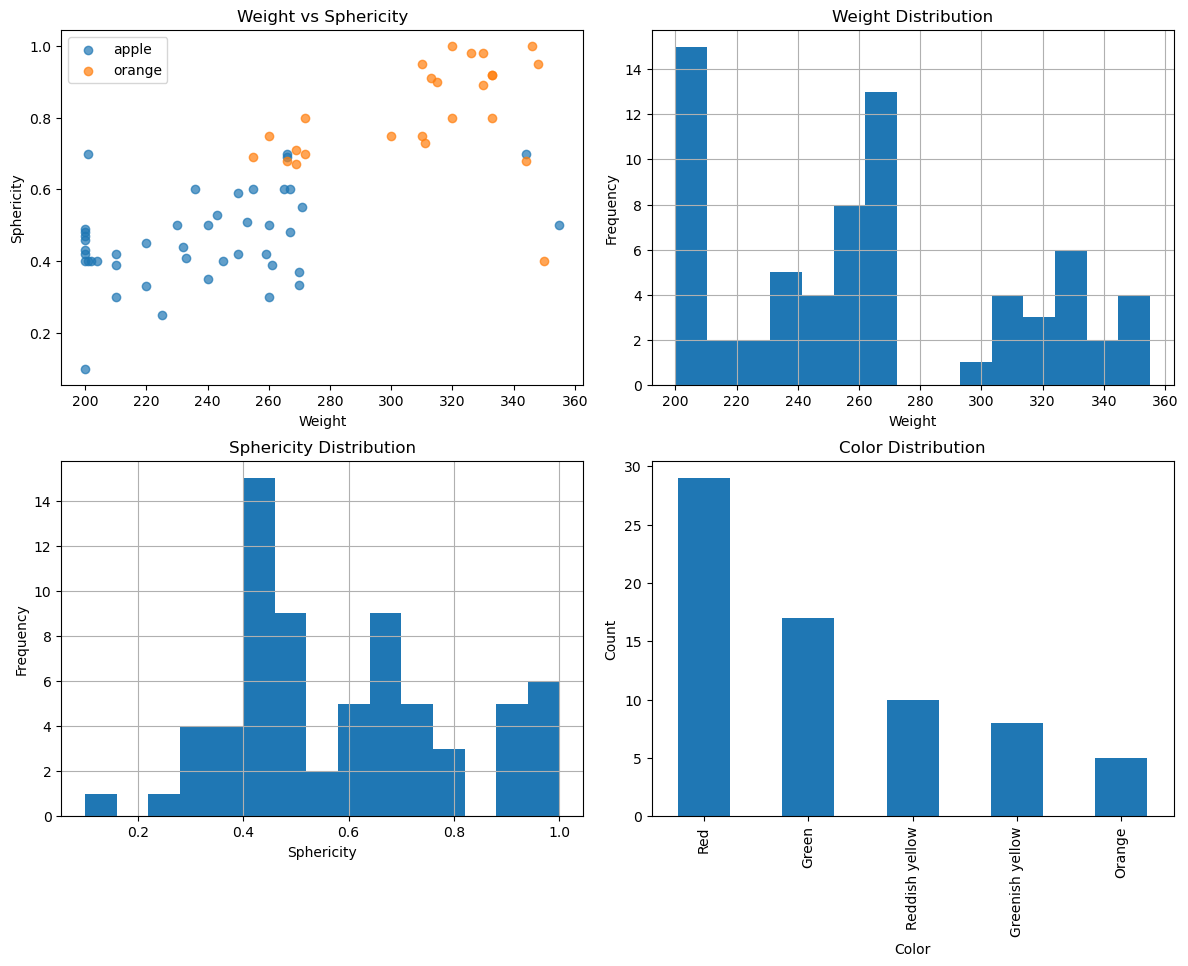
# ML Project Repoer

## 1. Fruit Dataset with model : Logistic Regression

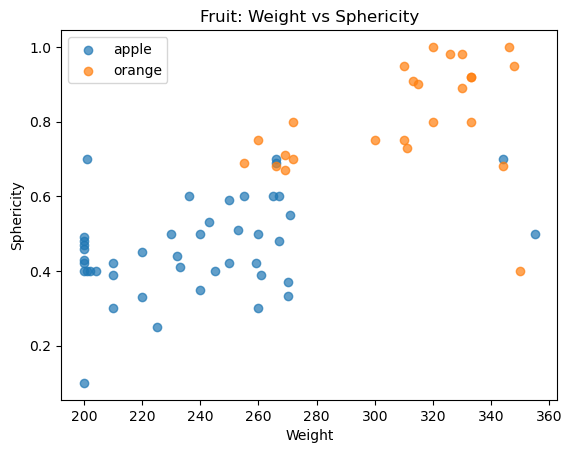
### Steps

* Loaded fruits\_weight\_sphercity.csv with mean imputation for Weight & Sphericity and one-hot encoding for Color
* Scaled numeric features (StandardScaler) and encoded categorical feature Color
* Visualized dataset with overview plots and Weight vs Sphericity scatter
* Trained Logistic Regression and evaluated performance
* Computed and plotted feature importances from model coefficients

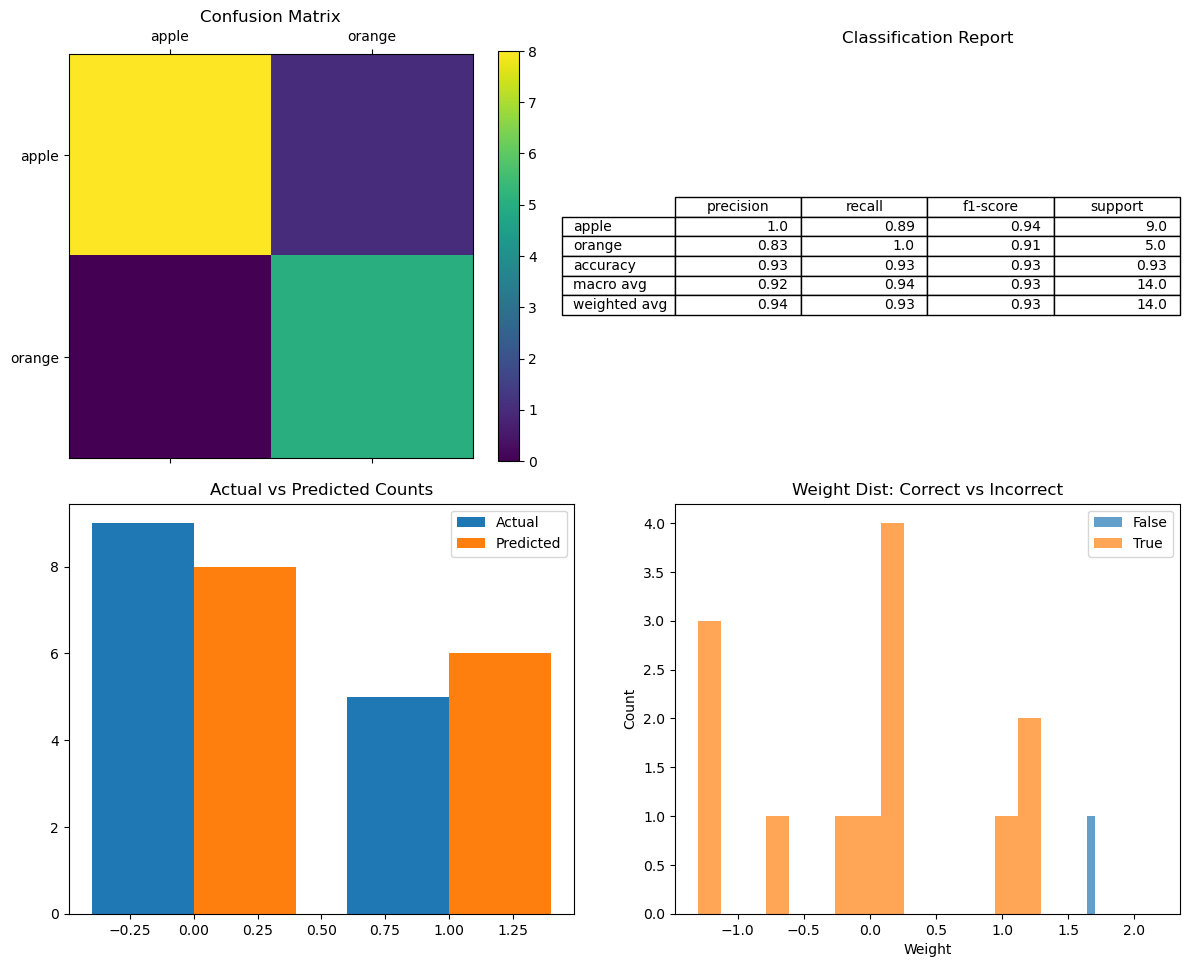
### Dataset Overview (2×2)



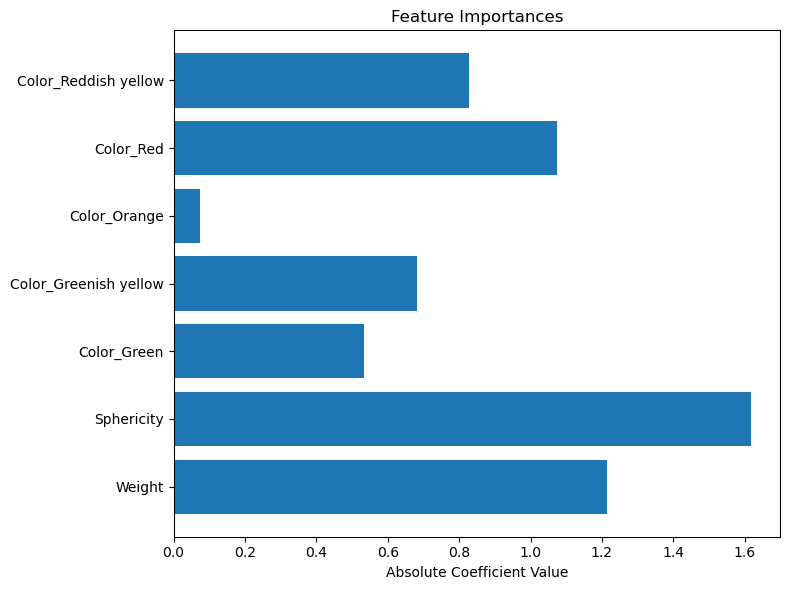
### Weight vs Sphericity



### Model Evaluation (4 subplots)



### Feature Importances

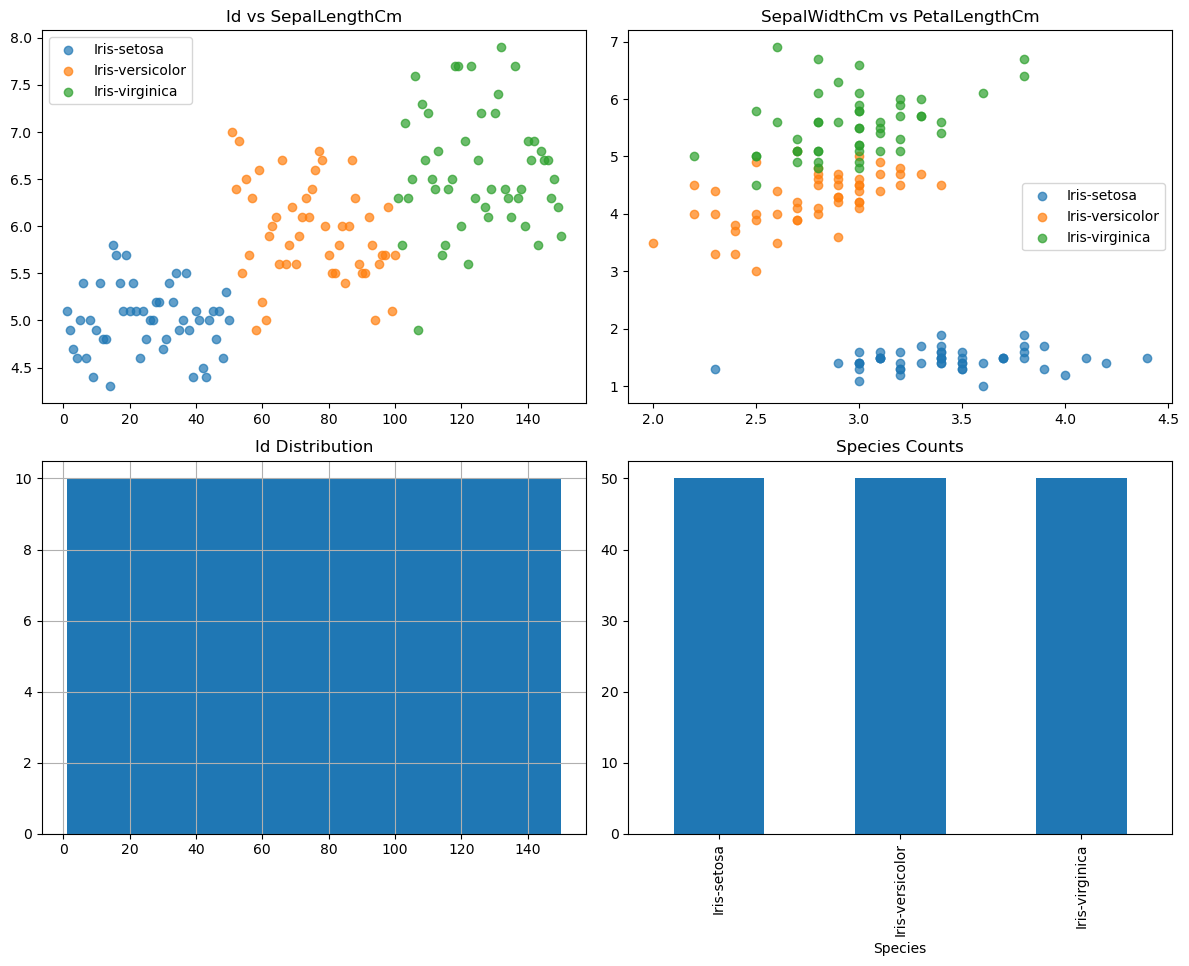


## 2. Iris Dataset with models : KNN, Decision Tree, Random Forest

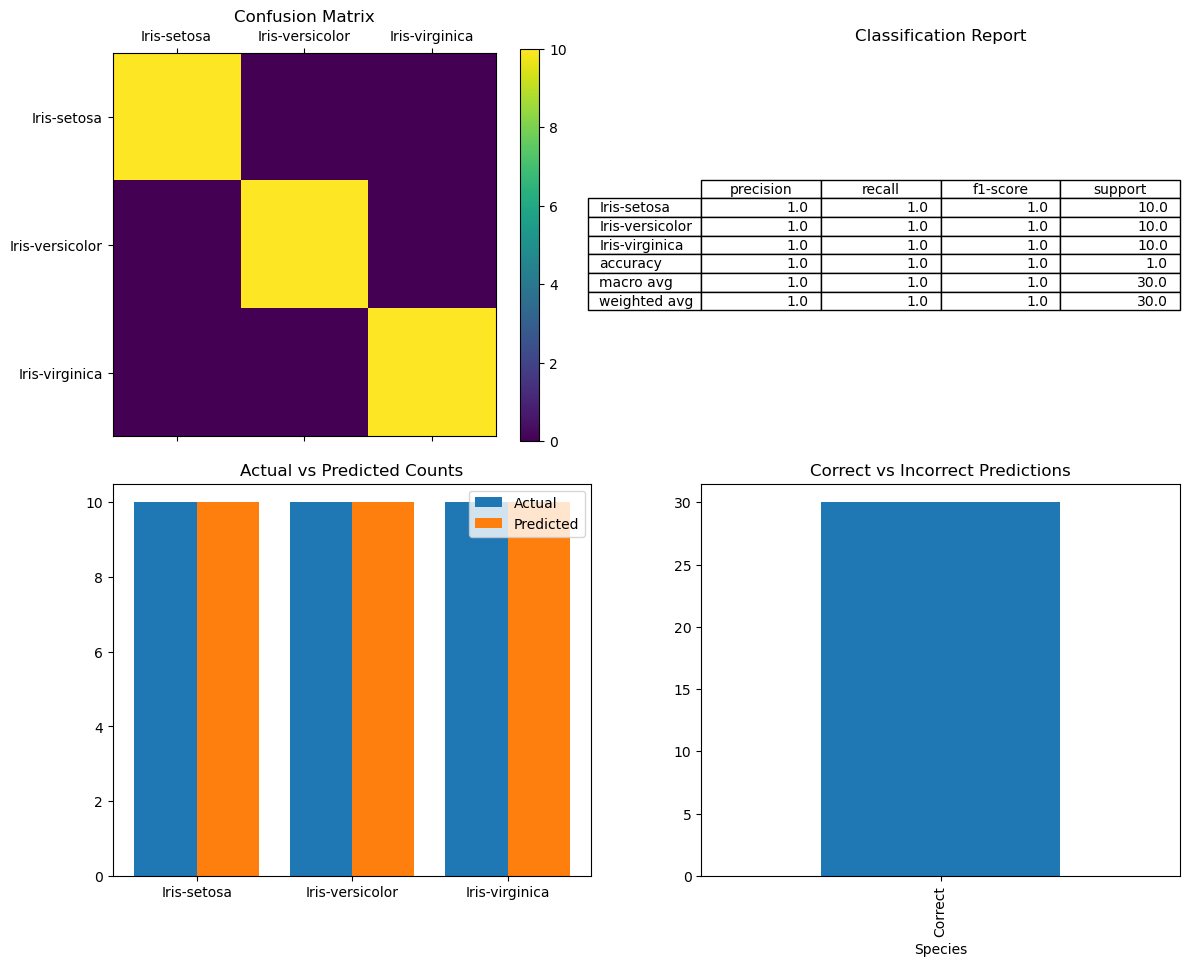
### Steps

* Loaded Iris.csv with mean imputation and scaling of numeric features
* Visualized dataset with scatter and histogram plots for feature distributions and species counts
* Trained KNN, Decision Tree, and Random Forest classifiers
* Evaluated each model: confusion matrix, classification report, and count plots
* Computed feature importances: permutation for KNN and built-in for tree models

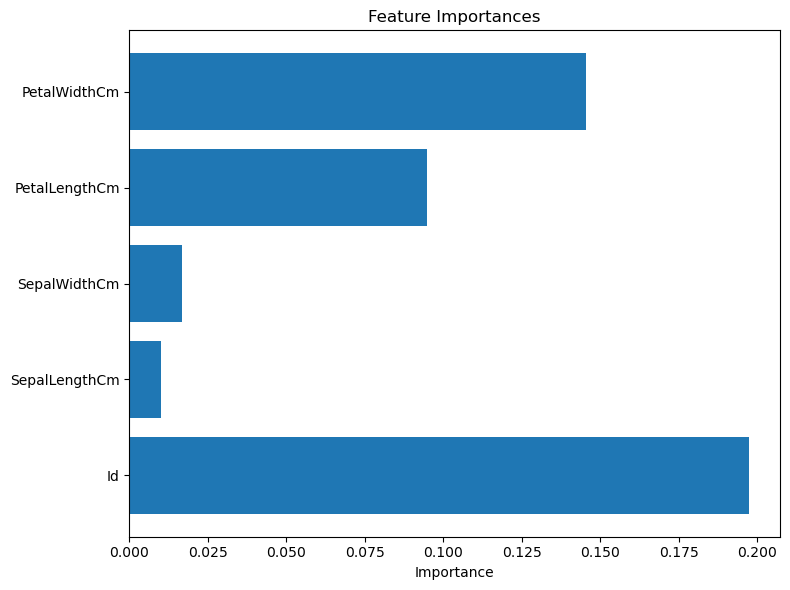
### Dataset Overview (2×2)



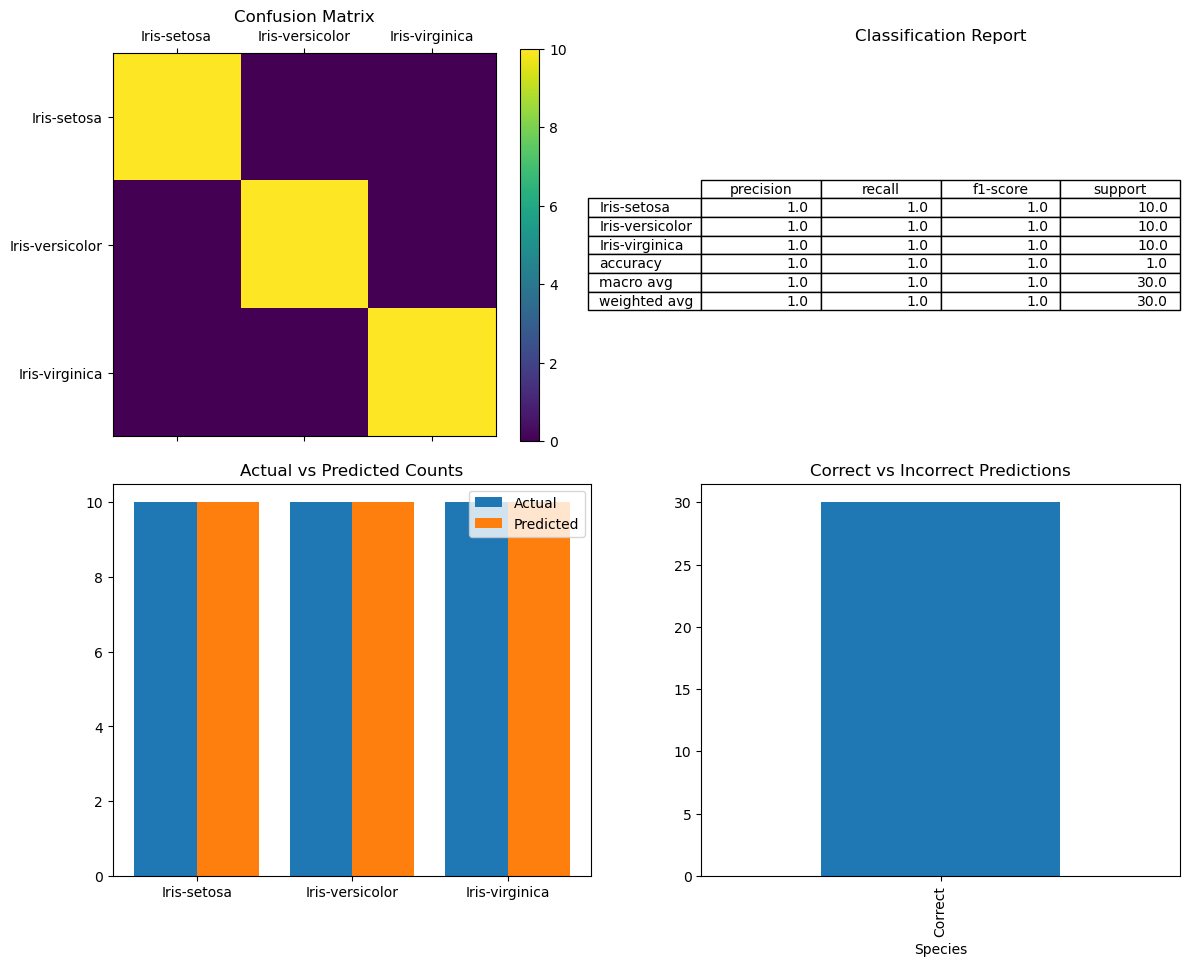
### KNN Model Evaluation



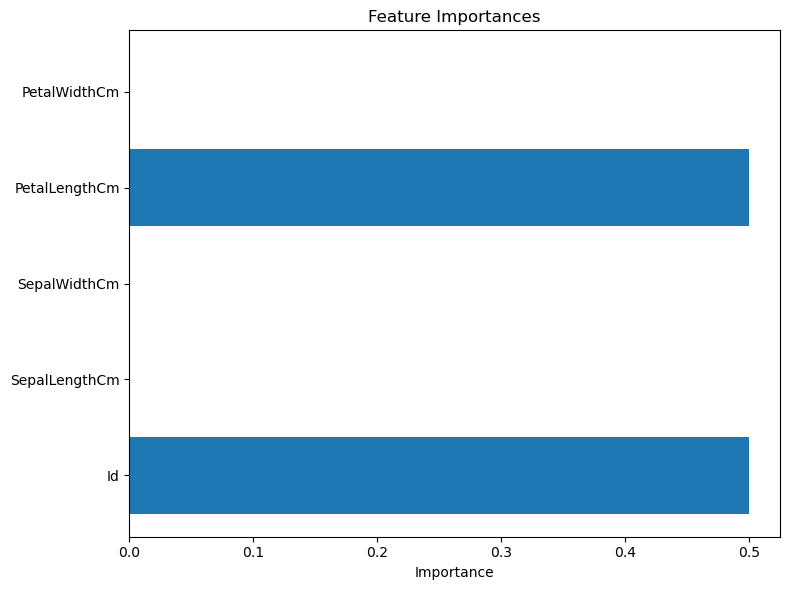
#### KNN Feature Importance



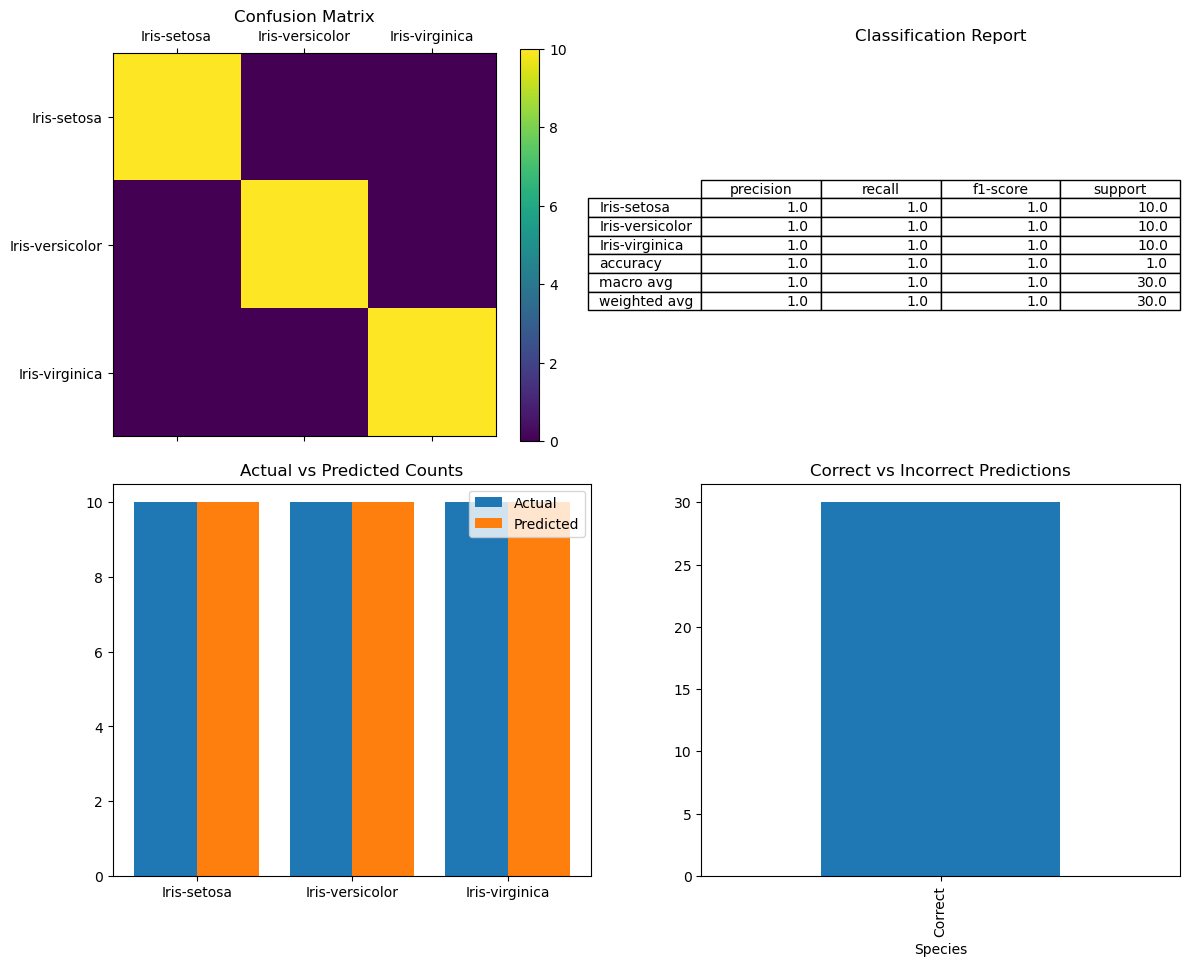
### Decision Tree Model Evaluation



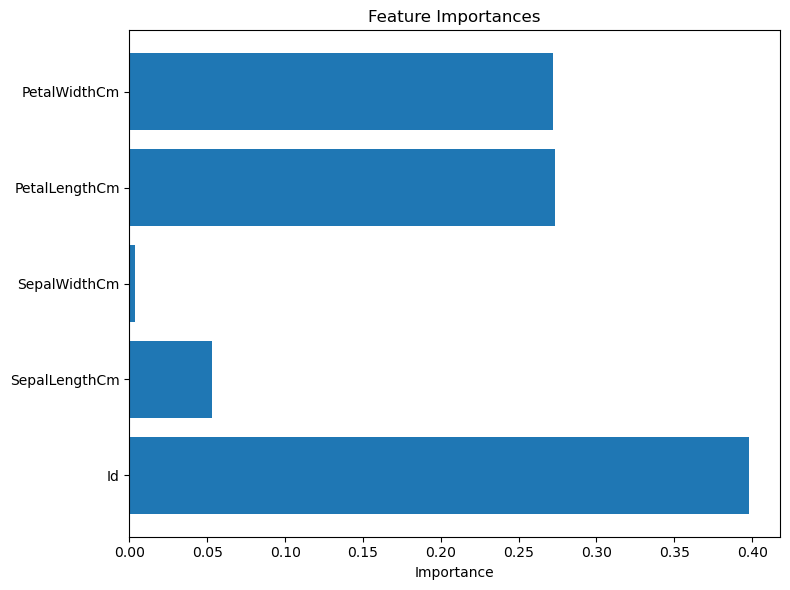
#### Decision Tree Feature Importance



### Random Forest Model Evaluation



#### Random Forest Feature Importance

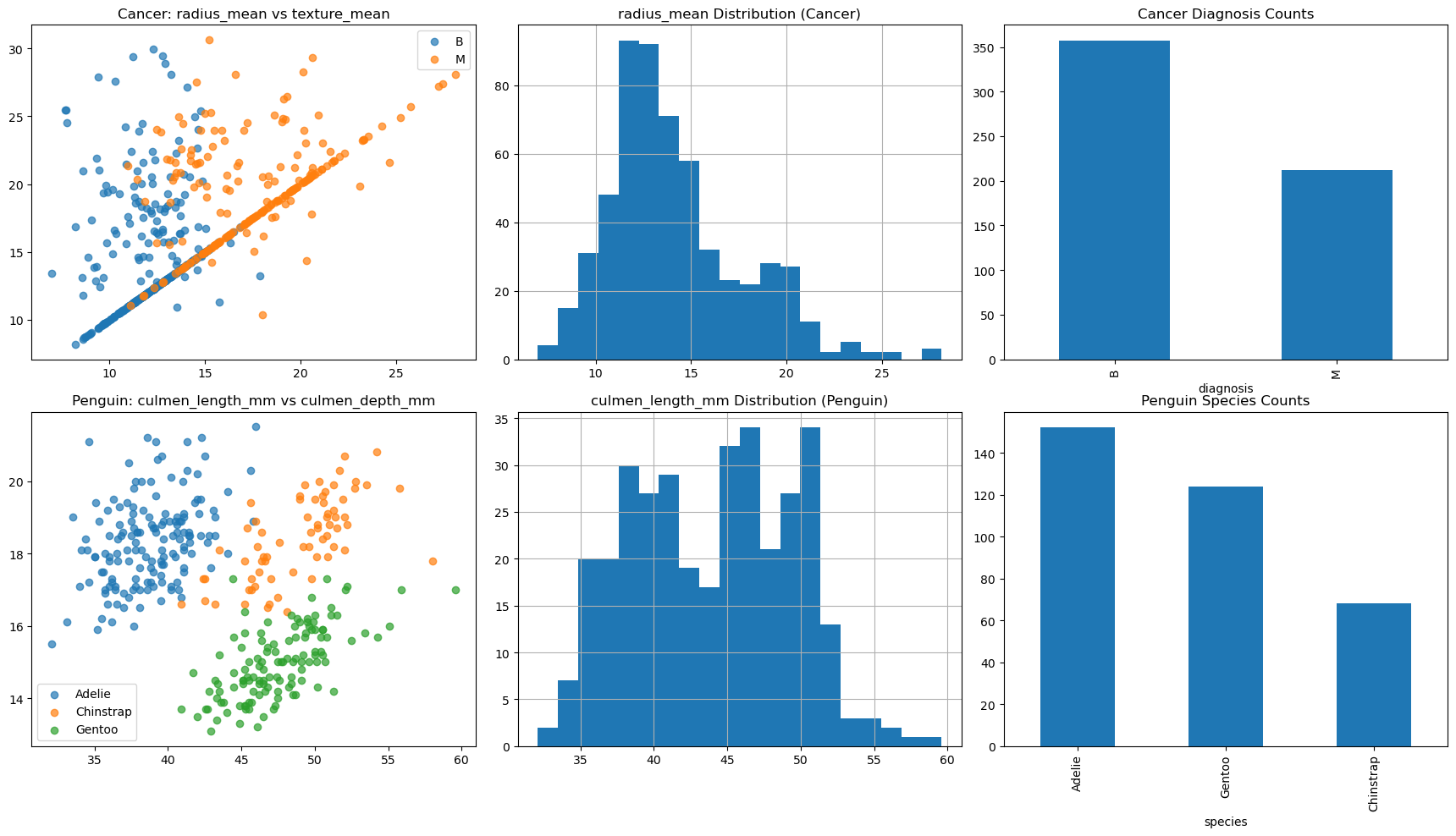


## 3. Breast Cancer & Penguin Datasets with model : Linear SVM

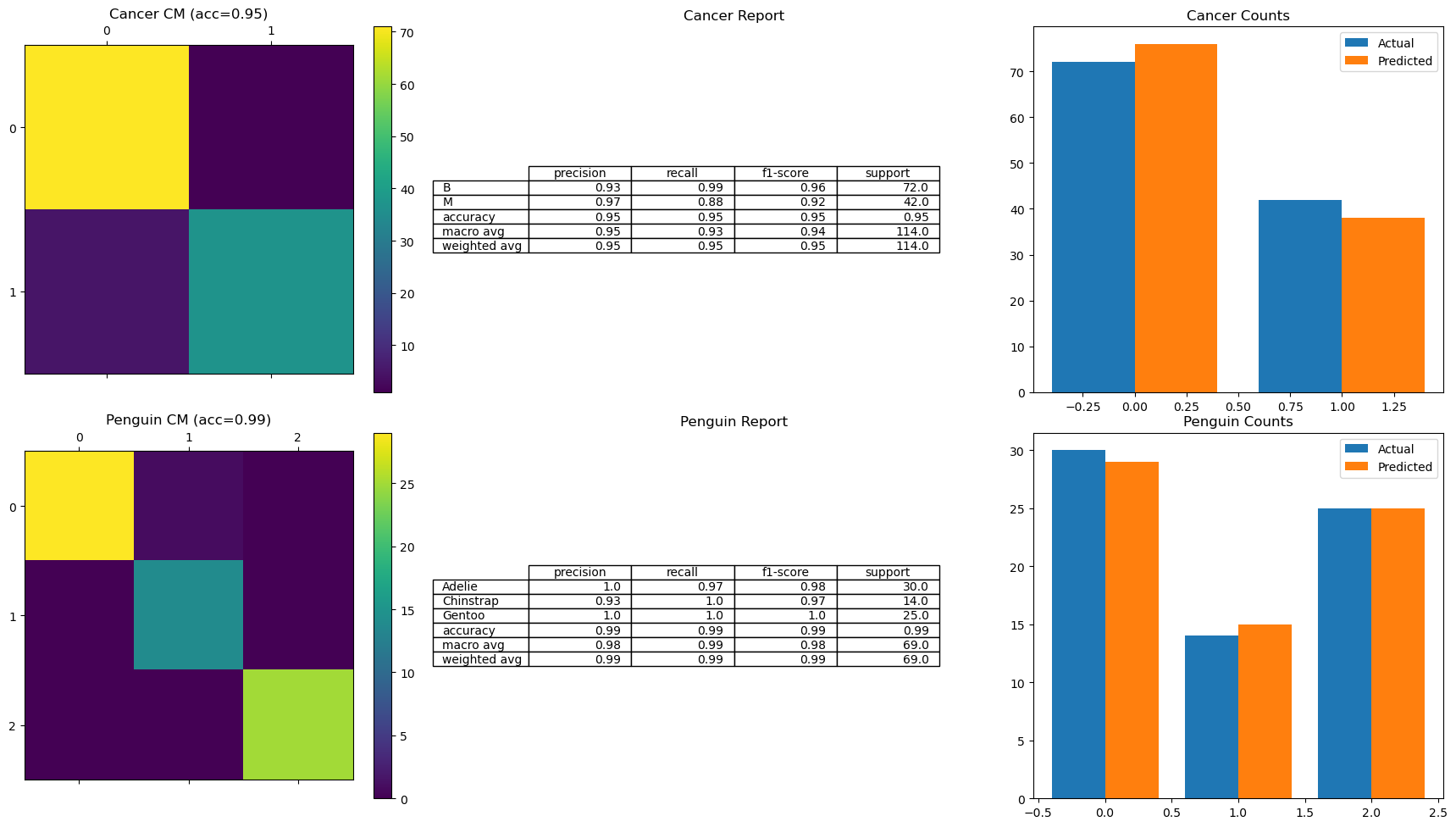
### Steps

* Loaded and preprocessed Breast Cancer and Penguin datasets with imputation and encoding
* Created a 2×3 overview comparison for both datasets
* Trained Linear SVM models and evaluated performance
* Extracted and plotted feature importances from SVM coefficients

### Overview (2×3)



### Model Evaluations (2×3)



### Feature Importances (2×)

