**Iris Flower Classification Using Support Vector Machine**

**Abstract:** This project aims to develop a classification model for predicting the species of Iris flowers using the Support Vector Machines (SVM) algorithm. The report describes the problem, dataset, classification algorithm, justification for using SVM, and visualization of the results.

**Section 1: Problem Overview**

The Iris flower dataset is well-known in the fields of machine learning and pattern recognition. The primary goal of this project is to develop a classification model that can accurately predict the species of an Iris flower based on its features. By accomplishing this, the model can assist botanists, researchers, and enthusiasts in identifying Iris flowers without prior knowledge of their species.

**Section 2: Description of the Dataset**

The Iris flower dataset was used in this project, and it consists of 150 samples of Iris flowers, each with four attributes: sepal length, sepal width, petal length, and petal width (in centimetres). The dataset includes three Iris flower species: Iris setosa, Iris versicolor, and Iris virginica. The dataset is balanced because each species has 50 samples. The dataset is available as a built-in dataset in R, called "iris."

**Section 3: Classification Algorithm Specifications**

Support Vector Machines (SVM) are the classification algorithms used in this project. SVM is a supervised learning algorithm for classification and regression. It seeks the best hyperplane that separates the classes by the greatest margin. SVM transforms the input space into a higher-dimensional space using kernel functions such as linear, polynomial, and radial basis function (RBF), allowing it to find a suitable hyperplane for non-linearly separable data. The RBF kernel with a cost parameter of 10 is used in this project.

**Section 4: Classification Algorithm Justification**

SVM was chosen for this project because of its ability to handle both linear and non-linear classification problems. The algorithm is robust and accurate, especially when combined with the appropriate kernel functions. Furthermore, when compared to other classification algorithms such as decision trees and k-nearest neighbours, SVM is less prone to overfitting. SVM can adapt well to different data distributions and complexities due to the flexibility of choosing kernel functions.

**Section 5: Visualization of Results**

A confusion matrix and accuracy score can be used to visualise the SVM model's performance. The confusion matrix displays the number of correct and incorrect predictions for each class, whereas the accuracy score represents the proportion of correct predictions to total predictions. Scatter plots can also be used to visualise the relationships between Iris flower features and species. By plotting the SVM model's decision boundaries on these scatter plots, we can see how well the model separates the different classes.

In conclusion, this project demonstrates the effectiveness of using the Support Vector Machines algorithm for classifying Iris flower species. The model achieves high accuracy and provides a reliable method for predicting the species of Iris flowers based on their features.