DSEM 6150 ASSIGNMENT 05

# SVM Classification Techniques

Presented by: Ruchi Kapadiwala

# Problem Statement

- Apply the Supervised Learning Model (SVM) on the loan dataset to find the hyperplane of the loan dataset
- Use different types of kernels in SVM to take the data and transform it into various forms
- Also showing the result on how different kernels produce different accuracies

# Overview

- Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.

- The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.

# Overview

- SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called support vectors, and hence algorithm is termed a Support Vector Machine.

# Technical Approach

01

**The dataset that you want to divide into two or more classes should be loaded.**

02

**Dividing the data**

Separate the data into a test set and a training set. The SVM model is trained using the training set, and its performance is assessed using the test set.

# Technical Approach

03

**Preprocess the data by scaling or standardizing the features, handling missing values, and removing outliers.**

04

## Pick a kernel

To convert the data into a higher-dimensional space, pick a kernel function. The kernel function can be linear or non-linear and can take the form of a polynomial, sigmoid, or radial basis function (RBF).

# Technical Approach

05

## Develop the model

Using the training set of data and the chosen kernel, train the SVM model. Finding the hyperplane that maximizes the gap between the two classes is the objective.

06

## Adjusting the hyperparameters

To enhance the performance of the SVM model, adjust its hyperparameters. The two most crucial hyperparameters are gamma, which governs the decision boundary's form, and C, which regulates the trade-off between maximizing the margin and minimizing the classification error.

# Technical Approach

07

## Review the model

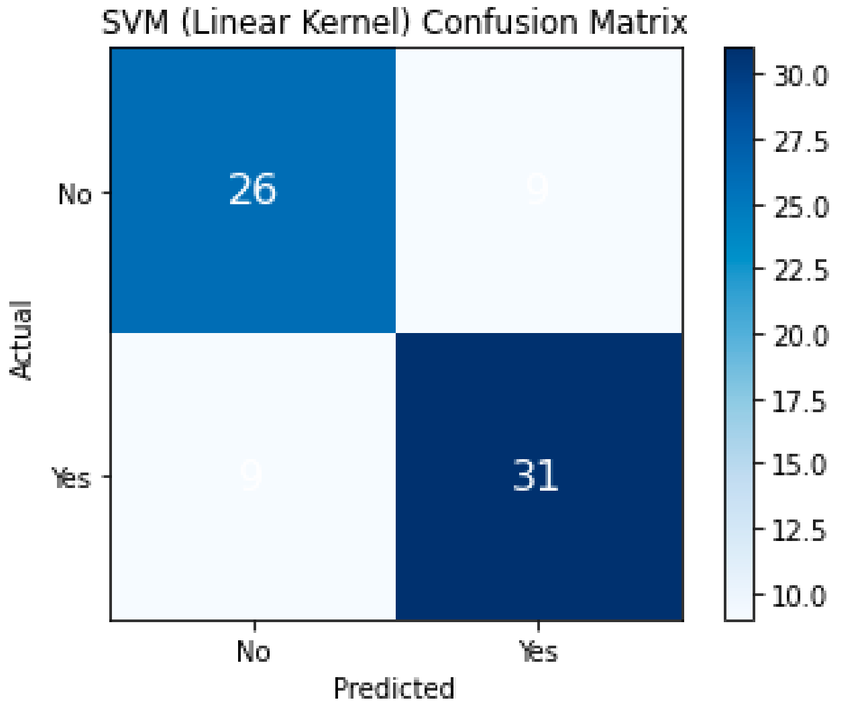
Utilize the test data to assess the SVM model's performance. Calculate metrics like F1 score, recall, accuracy, and precision.

08

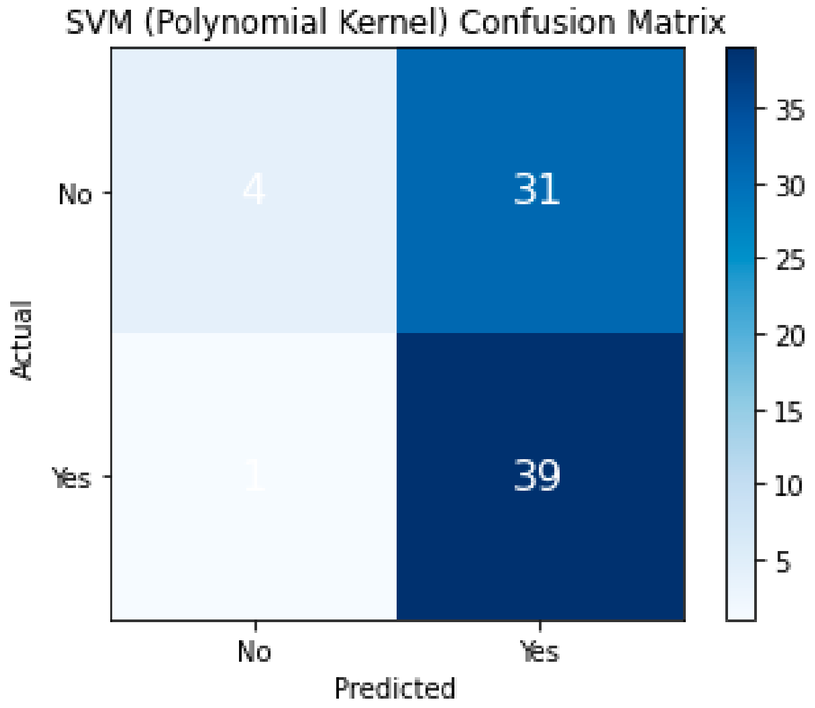
## Launch the model

nstall the SVM model in a setting where it can be used to categorize new data.

# Output

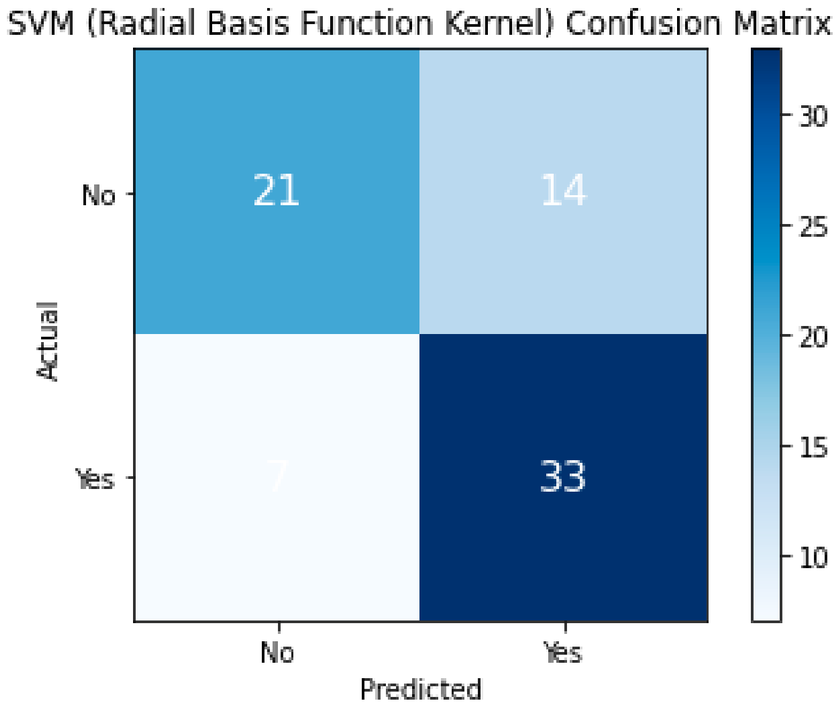



SVM (Linear Kernel) Accuracy: 0.76
Confusion Matrix:
[[26 9]
[ 9 31]]




SVM (Polynomial Kernel) Accuracy: 0.57
Confusion Matrix:
[[ 4 31]
[ 1 39]]

# Output



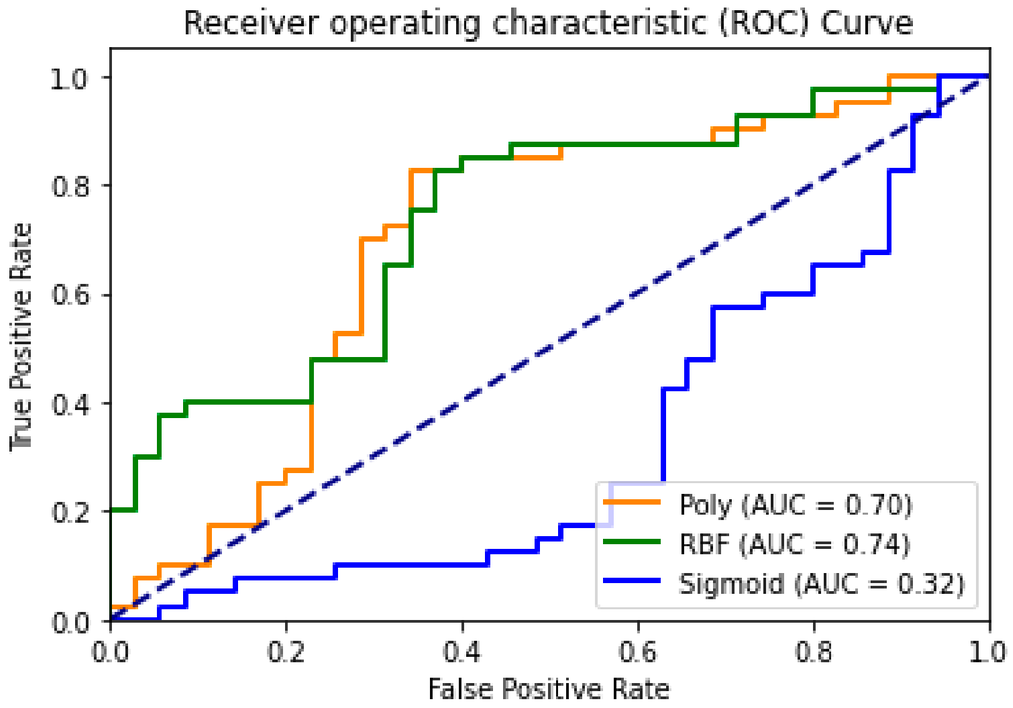

SVM (Radial Basis Function Kernel) Accuracy: 0.72

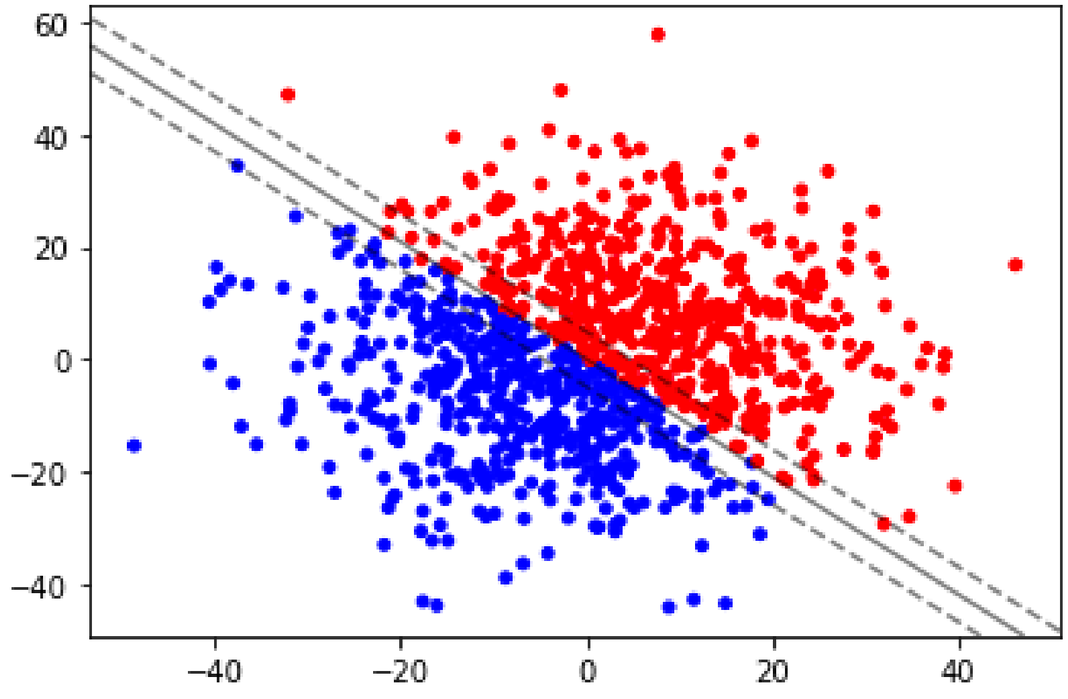
Confusion Matrix:

[[21 14]

[ 7 33]]

# Output


Receiver operating characteristic (ROC) Curve
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Poly (AUC = 0.70)
RBF (AUC = 0.74)
Sigmoid (AUC = 0.32)


60
40
20
0
−20
−40
−40
−20
0
20
40