# Assignment 4-Classifier

# Group 5

# Submitted by Maisha Labiba-501191497

# Mehjabin Rahman -501200106

# Problem 1) Function to implement an AND function using a single layer neural network.

Text

Description automatically generated

Output:

A picture containing shape

Description automatically generated

Function to implement an OR function using a single layer neural network

Text

Description automatically generated

Text

Description automatically generated

Output:

Text

Description automatically generated

Text

Description automatically generated

Text

Description automatically generated

Problem 2)

# Use Neural network for implantation of AND and OR functions For SVM, a margin is used to indicate the distance between the two-margin hyperplane .

It is used to indicate the distance from the hyperplane (solid line) and the observations/closest points from either class (which are support vectors). A larger and more elastic margin produces a more stable result, by adding bias for less variance error while training, reducing the risk of overfitting as well.

SVM finds a linear function in where x in the input vector, w is the vector representing the weight and b is the bias. Positive or a negative class would represent input vector xi as follows:

The equation for H+ is

The equation of H- is

Chart, line chart

Description automatically generated2) The meaning of margin in SVM and the equations of the two margin hyperplanes H+ and H- :

The two positive points and one negative point from the graph represents support vectors, with the hyperplane margin passing through the two positive points, in parallel with the hyperplane margin passing through the negative point. The ideal decision boundary would be a line which passes halfway between the two margins, which is illustrated by the red line passing through the and margins below. The equation for the decision boundary is -x + 2 = 0, as the x intercept is 2 and the line is linearly going down along the x-axis.

# 3) kernel function :

The kernel function is represented by the equation where *x* and *y* represent the input vectors and represents the function used to transform the data from the input space to the feature space. The kernel function *K* is derived through the dot product of the two input vectors which have been transformed. The transformed vectors and their dot product are used to derive the kernel function associated with two input vectors.

# 3. Compare Neural Network and SVM in Classification of heart disease data set in R or Python language:

To answer this question, we have used Python language. First, we compared different kernels for SVM and then we compared Stochastic Gradient Descent and Adam for ANN. Finally, we compared ANN and SVM.

Table

Description automatically generated with medium confidence

Table

Description automatically generated

Table

Description automatically generated

In above codes, we can see the accuracy on the test set. As we can see, the linear  kernels perform similarly well, with an accuracy of 0.804. However, the Gaussian and sigmoid kernels perform significantly worse, with an accuracy of only 0.68 and 0.49 respectively.

Text

Description automatically generated with medium confidence

Graphical user interface, text, application

Description automatically generated

On the heart disease dataset, we have applied Stochastic Gradient Descent and Adam Gradient Descent to build the multi-layer Neural Network and note down the model accuracy for each. We can see from the above evaluation that the Adam model performs better than the SGD with an accuracy of 0.8.

As we can see, the SVM linear kernels perform similarly as ANN with Adam optimizer, with an accuracy of 0.8 .

# 4. Evaluation of the performance of a system using natural language processing and comparing two classifiers:

To answer question 4, we have used the provided dataset of musical review. Dataset looks as following,

Text

Description automatically generated

Table

Description automatically generated

After loading the dataset, we have applied different natural language processing techniques to preprocess the data. We have converted the text into lower case, removed punctuations, numbers and stop words. We have also performed tokenization and stemming. After all these the dataset appears as following,

Graphical user interface, text, application, email

Description automatically generated

Then we performed lemmatization.

Graphical user interface, text, application, email

Description automatically generated

Then we converted the tokens into string format to train the model.

Graphical user interface, text, email

Description automatically generated

We have performed Decision Tree and Naive Bayes classification on the preprocessed dataset.

Graphical user interface, text, application, email

Description automatically generated

A screenshot of a computer

Description automatically generated with low confidence

Text

Description automatically generated

Comparison of the evaluation metrics for the four models:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| Decision Tree | 0.67 | 0.67 | 0.67 | 0.67 |
| Naive Bayes | 0.81 | 0.81 | 0.81 | 0.81 |

We can see that Naive Bayes has the highest accuracy, precision, recall, and F1-score among the two models, followed by Decision Tree having the lowest evaluation metrics.