NAME- RISHI KUMAR
CWID- 20015656
HOME WORK #5

Part 1: Prepare a 1-page slide/poster to review/summarize the concept of Support Vector Machine(SVM)

Support Vector Machine

What is Support Vector Machine(SVM)???

A support vector machine (SVM) is a supervised machine learning model that uses classification algorithms for two-group classification problems. After giving an SVM model set of labeled training data for each category, they're able to categorize new text.

How does it work??

SVM works by mapping data to a high-dimensional feature space so that data points can be categorized, even when the data are not otherwise linearly separable. A separator between the categories is found, then the data are transformed in such a way that the separator could be drawn as a hyperplane. Following this, characteristics of new data can be used to predict the group to which a new record should belong.
For example, consider the following figure, in which the data points fall into two different categories.

Figure 1. Original dataset

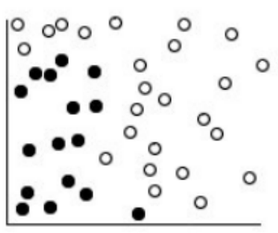


Figure 2. Data with separator added

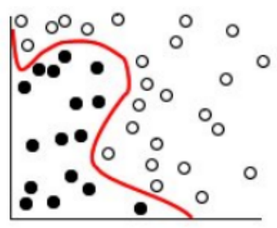
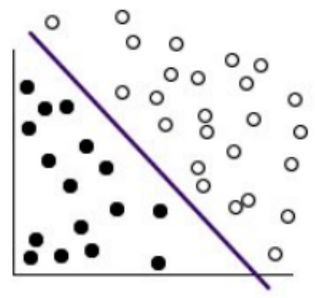


Figure 3. Transformed data



- The two categories can be separated with a curve, as shown in the following figure. 2
- After the transformation, the boundary between the two categories can be defined by a hyperplane, as shown in the following figure.3
- The mathematical function used for the transformation is known as the kernel function

Advantages:

- SVM works relatively well when there is a clear margin of separation between classes.
- SVM is more effective in high-dimensional spaces.
- SVM is effective in cases where the number of dimensions is greater than the number of samples.
- SVM is relatively memory efficient

Disadvantages:

- The SVM algorithm is not suitable for large data sets.
- SVM does not perform very well when the data set has more noise i.e. target classes are overlapping.
- In cases where the number of features for each data point exceeds the number of training data samples, the SVM will underperform.
- As the support vector classifier works by putting data points, above and below the classifying hyperplane there is no probabilistic explanation for the classification.

references:

1. *https://www.ibm.com/docs/en/spss-modeler/saas?topic=models-how-svm-works*
2. *https://dhirajkumarblog.medium.com/top-4-advantages-and-disadvantages-of-support-vector-machine-or-svm-a3c06a2b107*