60004210210

Amartya Mishra

COMPS – C31

ML Experiment 7

ML Experiment 7

60004210210
Amartya Mishra
COMPS - C31

Aim: To Implement Support Vector Machine

Theory
Support Vector machine (SVM) is used for linear/ non linear classification, regression & even outlier detection
It is supervised ML Algorithm.
The main objective of the SVM is to find the optimal hyperplane in a N-dimensional space that can separate the Data points in different classes in the feature space
The hyper plane tries that the margin between the closest points of different classes should be as maximum as possible

The Dimension of the hyper plane depends upon the number of features

Mathematical Intution

$$w^T x + b = 0$$

$$d_i = \frac{w^T x + b}{||w||}$$

$$\hat{y} = \begin{cases} 1 & : w^T x + b \geq 0 \\ 0 & : w^T x + b < 0 \end{cases}$$

Types of SVM

1) Linear
2) Non linear

Advantages

1) Effective in high dimensional cases
2) Its memory is efficient as it uses a subset of training points in the decision function called support vectors
3) Different kernal functions can be specified for the decision function & its possible to specify custom kernals.

Conclusion : Thus we implemented SVM.

Implementation:

```
from sklearn.datasets import load_breast_cancer
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import accuracy_score
data = load_breast_cancer()
X = data.data
y = data.target
X
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.3, random_state=42)
classifier = GaussianNB()
classifier.fit(X_train, y_train)
y_pred = classifier.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print(accuracy)
```

```
from sklearn.metrics import confusion_matrix
confusion_matrix(y_test, y_pred)
```