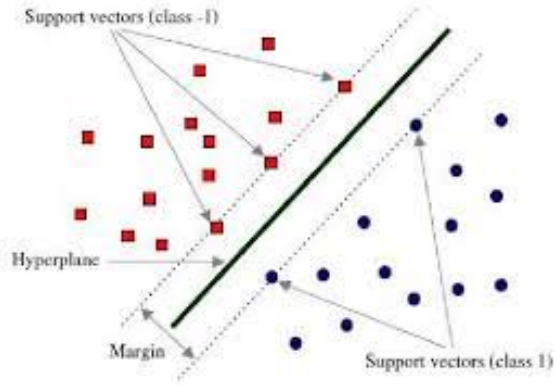
SUPPORT VECTOR MACHINE

Support Vector Machine Theory

Support Vector Machine is a supervised learning classifier that segregates the data by using a separative hyperplane such that the distance between the hyperplane and support vectors is maximum. It can be used for both classification and regression.



SVM can be used to classify non-linear data by using the kernel function, that transforms the data into another dimension so that a hyperplane can be easily drawn between classes of the data. It is effective in high dimensional spaces.

Support Vector Machine with Python

The Heart Disease dataset has been taken from Kaggle. This database contains 76 attributes, but all published experiments refer to using a subset of 14 of them. It has a total number of 303 rows and 14 columns among which 165 have a heart disease.

age: age in years

sex: (1 = male; 0 = female)

cp: chest pain type

trestbps: resting blood pressure (in mm Hg on admission to the hospital)

chol: serum cholestoral in mg/dl

fbs: (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)

restecg: resting electrocardiographic results

thalach: maximum heart rate achieved

exang: exercise induced angina (1 = yes; 0 = no)

oldpeak: ST depression induced by exercise relative to rest

slope: the slope of the peak exercise ST segment

ca: number of major vessels (0-3) colored by flourosopy

thal: thalassemia (1 = normal; 2 = fixed defect; 3 = reversable defect)

target: (1= heart disease or 0= no heart disease)