MULTIPLE DISEASE PREDICTION SYSTEM

1. **Language Used:**
2. Python
3. **Libraries Used:**
4. **Support Vector Machine (SVM) for Diabetes Prediction**

* Supervised learning model based
* We feed glucose level, BMI, and insulin level data to the model and train them based on it for what is diabetic and what is not.
* And then the model placed the data in either positive hyperplane or negative hyperplane where positive means one is diabetic and the negative means one is not.

A diagram of a positive hyperplane

AI-generated content may be incorrect.

Figure : SVM (Support vector Machine) Working Mechanism

1. **Logistic Regression (Heart Disease Prediction)**

* Using it because as it’s a binary classification case
* And Logistic regression is best for binary classification datas
* Due to its ability to predict probabilities instead of just class labels and its relatively simple implementation and interpretation.
* And it uses the sigmoid function to map predictions to probabilities between 0 and 1, allowing for nuanced understanding of the uncertainty associated with classifications.

1. **Support Vector Machine (SVM ) for Parkinson detection**

* Supervised learning model base
* We feed the data with symptoms .
* And then the model placed the data in either positive hyperplane or negative hyperplane where positive means one has parkinson and the negative means one does not.

#How SVM works

* X axis mentions feature 1 , and y axixs represents feature 2
* Then the data is plotted, and the model tries to find a separating line between those points called hyperplane
* The point near to the hyperplane is called support vector
* If the location of those points changes the hyperplane changes as well
* Then tyeei anusar the new inputted data is analyzed and placed in one of the sides of hyperplane stating either one is infected, or one is not infected.

\*If badi features haru xa vane tesari naai 3d ma multiple features haru separate hune gari hyperplane khojera banauna parxa\*

A diagram of a plane and a support vector

AI-generated content may be incorrect.

Figure : How SVM Works

1. **Workflow of the Project:**

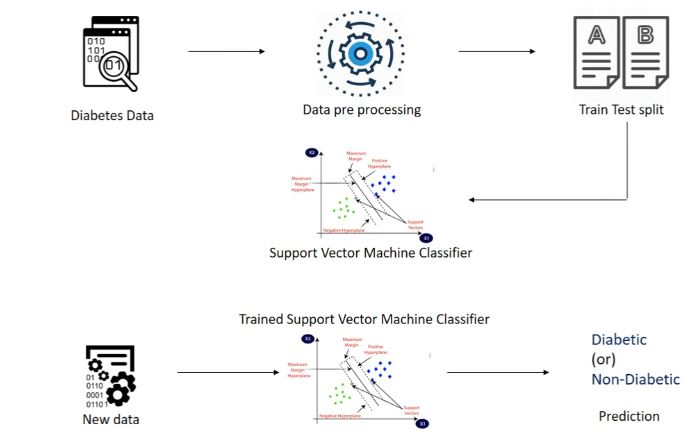


Figure : Project Working Model

* Data search and collection are done
* The data is processed i.e. cleaned, and the null values are replaced, and the only required data are kept
* The data is then split into testing and training sets where the training set is used to train the model to make it work as per need and the testing set is used to test if the model is trained as per needed or not.
* Then the trained model is fed up with the actual data which would predict if one was suffering from the disease or not.
* **Code explained:**
* Train\_test\_split will give 4 outputs which will be stored as xtrain , x test , y train , y test
* Test size = 0.2 represents 20% is test data and 80 percent will be left as training data
* Stratify=y means we are stratifying based on the values in Y as it consists of 0 and 1 data and we want our data set to be split based on it . if we don’t do so there are chances of all the diabetic data to go to one variable and non-diabetic to go to another variable.