

# MODEL BUILDING

## TRAIN AND TEST THE MODEL USING CLASSIFICATION ALGORITHMS

Date	19 November 2022
Team ID	PNT2022TMID04221
Project Name	Project – Statistical Machine Learning Approaches to Liver Disease Prediction

There are several Machine learning algorithms to be used depending on the data you are going to process such as images, sound, text, and numerical values. The algorithms that you can choose according to the objective that you might have may be Classification algorithms are Regression algorithms.

You will need to train the datasets to run smoothly and see an incremental improvement in the prediction rate.

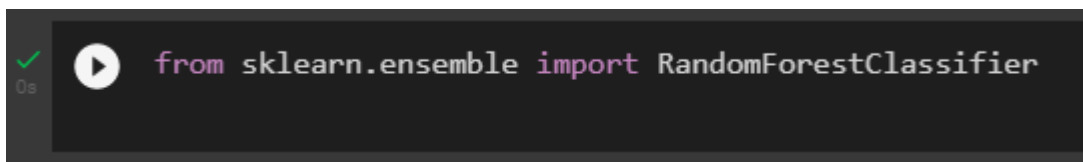
Now we apply classification algorithms on our dataset.

Random Forest Regression: Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes or mean/average prediction of the individual trees.

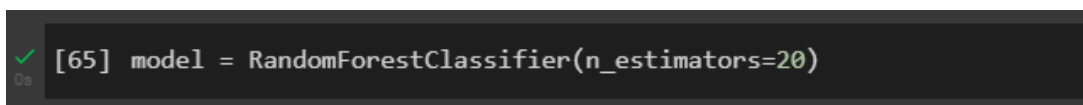
Build the model

We're going to use `x_train` and `y_train` obtained above in `train_test_split` section to train our regression model. We're using the `fit` method and passing the parameters as shown below.

1. Import the Classification algorithms

A screenshot of a Jupyter Notebook code cell. On the left, there is a green checkmark and a play button icon. The code text is `from sklearn.ensemble import RandomForestClassifier`.

2. Initialize the model

A screenshot of a Jupyter Notebook code cell. On the left, there is a green checkmark and a play button icon. The code text is `[65] model = RandomForestClassifier(n_estimators=20)`.

### 3. Training model with our data.

- Random Forest Model

