Earthquake prediction model using python

MODEL BUILDING

- The yield of an ML algorithm is a 'model'.
- To begin with, the target variable and feature variable are comprehended and fetched.
- Second, the data-set is partitioned into training and testing data-set and third, the regressor/classifier model is constructed and fitted to training data-set.
- In python, scikit-learn is a simple, basic, efficient open source library that executes a range of machine learning algorithms featuring various classification, regression and clustering algorithms using a unified interface.
- Step by step building is as follows:
- Building A Random Forest Regression Model :
- Random forests are an ensemble learning method that can be fabricated for both regression as well as classification chore.
- It takes on the task of constructing multiple of decision trees during training and outputs the class that is mean prediction (regression) of each individual tree or the mode of the classes (classification).
- This huge number of trees represents a forest. Decision trees are rule based models; on a given training data-set with targets and features, the decision tree algorithm will come up with rules to carry out classification

DATA PRE-PROCESSING

- Data Pre-processing is a technique that converts given and regression.
- Features will be nodes and their presence and absence will represent likeliness.
- This helps in constructing a path of rules to work with.
- The root and splitting node is based on information gain or gini index.
- In Random Forest, the root and splitting nodes are calculated in a random manner
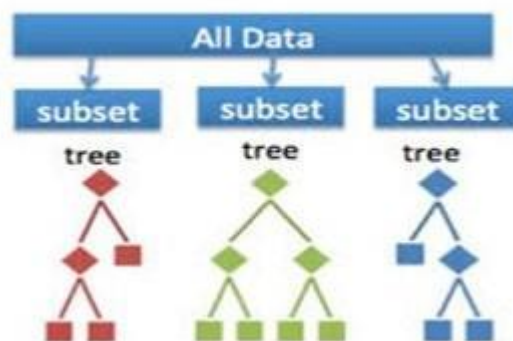

Fig. Random Forest

Therefore random forest is a model comprising of various streets with the capability of making decision based on rule and the procedure of choosing root nodes and parent node sis random.

Building A Support Vector Machine Regression Model:

- SVM segregates different data classes using a decision line named hyperplane.
- When predicting a numerical value, SVR attempts to find a function f(x) in the form of decision boundary at a certain deviation from €, which is a threshold value for all prediction to be within, from obtained targets value Yi, the original hyperplane, such that data points are within the boundary line.
- This decision boundary is the Margin of tolerance - a boundary that allows errors under given range.
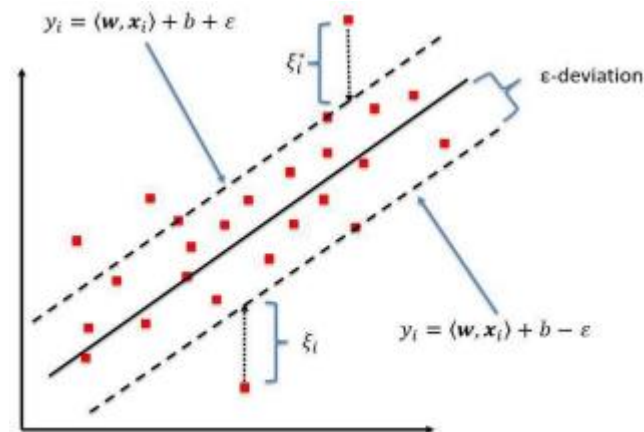

$$y_i = \langle w, x_i \rangle + b + \varepsilon$$

$$y_i = \langle w, x_i \rangle + b - \varepsilon$$

Fig. 2. Support Vector Regressor

Building A Stacking Regress or Model :

Stacking regression is an ensemble learning method. Several regression models collaborate, as a result, meta-regress or is build & itself finds its best fit by making use of output of
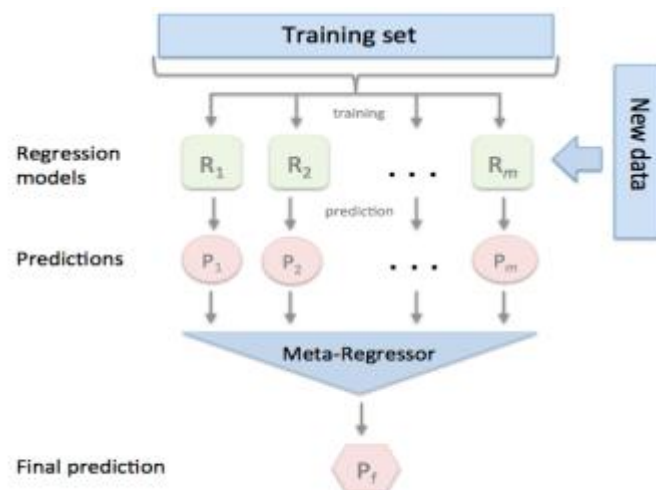

Fig. 3. Stacking