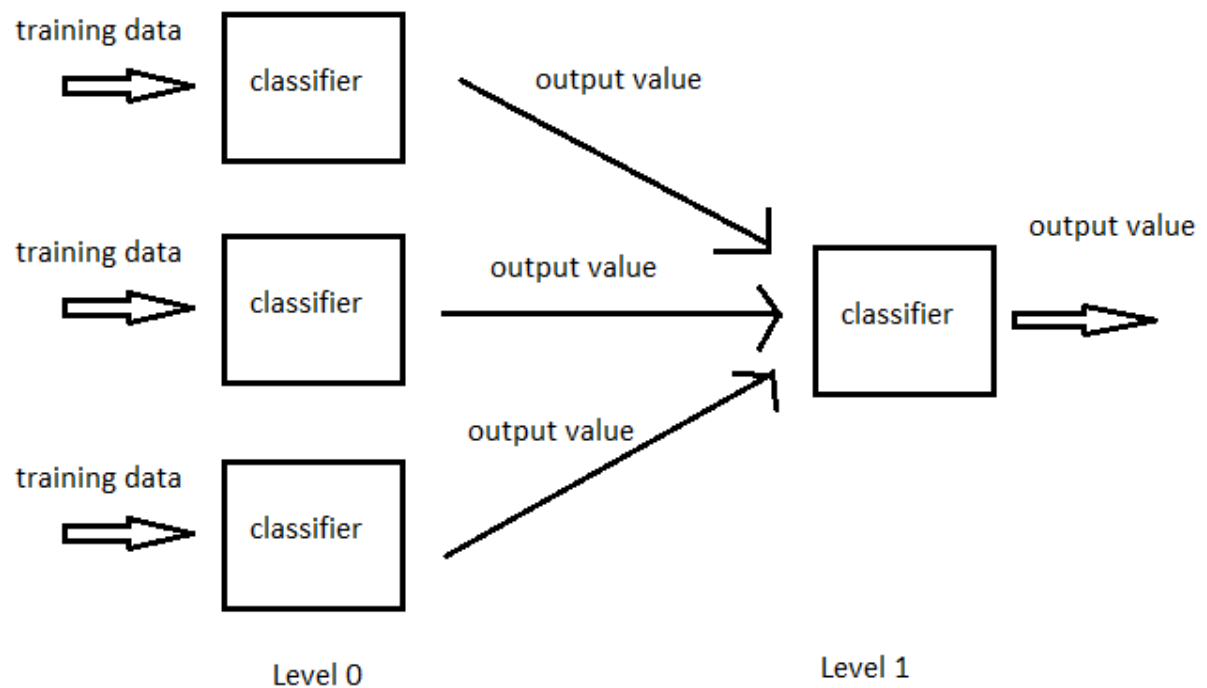


Concepts

Ensemble methods: *The main principle of ensemble is to group the weak learners output to make it stronger (by averaging results or choosing the majority). Generally accuracy of the model effected due to three particular reasons i.e., noise, variance and bias.*

- ❖ **Bootstrap sampling:** It is basically used for reducing the **variance** of the given model by splitting the given data(N) in to m baskets, where each basket size is equal to (N) , These baskets are filled with replacement techniques.
- ❖ **Bagging:** It trains the multiple models and aggregates the results. The training performs parallelly and **on all the parameters**
- ❖ **Boosting(Ada Boosting):** It is basically used for reducing the bias of the given model by working the failures of the previous model. Training performs sequentially.
- ❖ Random forest: It also works as a Bagging but the number of features for a modal training is the subset of the feature vectors. Bagging only equals using n_{tree} . Only Randomforests has n_{tree} and m_{try} . If $m_{try} = \text{all}$ then baggging == randomforests.
- ❖ Gradient boosting: Applied the gradient searching for optimizing the objective function.
- ❖ XGboost: Added the regularization to control the overfitting issue while applying the boosting technique.
- ❖ Stacking: It basically works just as a Bagging or Random forest or boosting, but at the end it uses another classifier/regressor to finalize the results.

Concept Diagram of Stacking



❖ Committee methods

Objective Function: In order to train any Machine Learning Algorithm need to define a objective function to understand how well your modal is performing on train and test datasets, the objective function is formulated based on Training loss and Regularization.

$$\text{obj}(\theta) = L(\theta) + \Omega(\theta)$$

Where L is the training Loss, squared error.

generally is considered as mean-

$$L(\theta) = \sum_i (y_i - \hat{y}_i)^2$$

Another commonly
function is logistic loss:

used loss

$$L(\theta) = \sum_i [y_i \ln(1 + e^{-y_i}) + (1 - y_i) \ln(1 + e^{y_i})]$$

Regularization is another factor, which used in objective function. It is used to reduce the complexity of model i.e., reduce the overfitting problem.