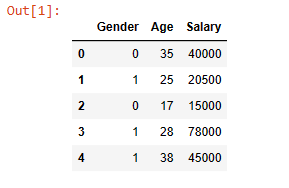
**Scaling Techniques:-**

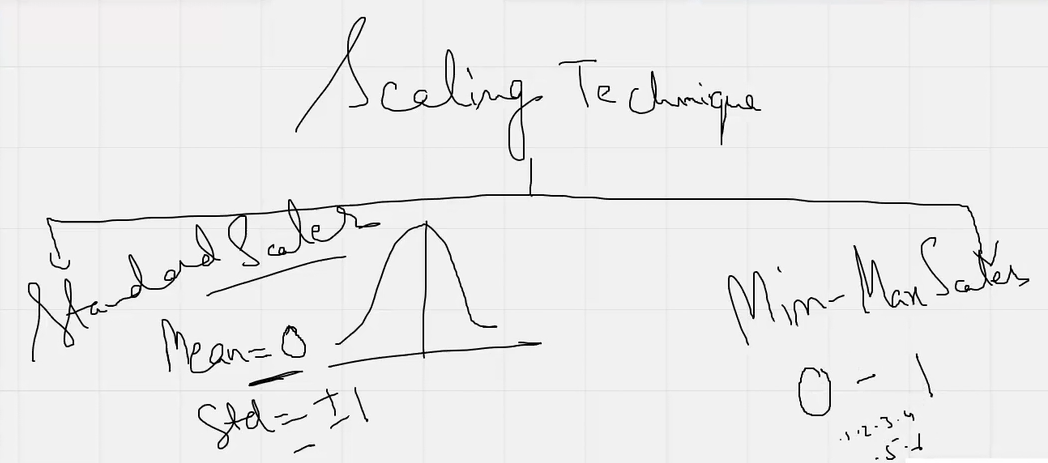
Scaling scales your dataset and transform your dataset with small values that model can learn easyly.

See the table below:



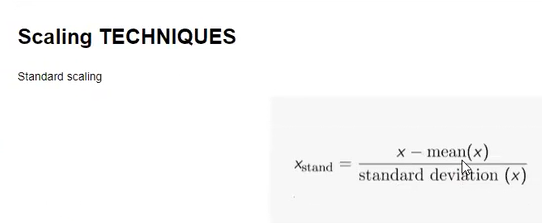
Suppose I have a dataset like this.So in this dataset you can see Gender column has only binary values,Age column has only two digit values and Salary column has upto 5 digit values.You can pass this dataset to machine learning there is no problem but you could increase your accuracy score if you apply scalling techniques to your dataset.

There are two types of scaling techniques 1> StandardScaler, 2>MinMaxScaler.



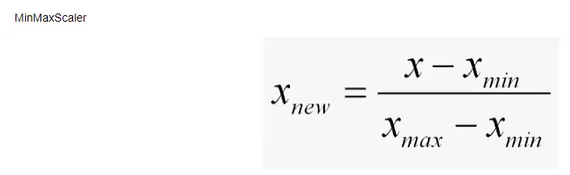
**1> StandardScaler:-**

When you apply Standard scaler to your dataset it organise your dataset like its mean value should 0 and std should +-1.That means it converts your dataset into normal distribution.



**2>MinMaxScaler:-**

When you apply MinMaxScaler to your dataset it sets your dataset column values to 0 to 1 where 1 is maximum and 0 is minimum. So if you apply MinMaxScaler to your dataset so every columns data of your dataset is converted to 0 to 1 values like 0.2,0.256,0.84522,0.99.



**Ensemble Technique:-**Ensemble technique is used for larger dataset like 1000 rows and 300,400 columns.So if I give this dataset to machine then machine will never understand this and cant give you the proper prediction.So ensemble technique divide the whole dataset into some chunks and those chunks are provided differed different models ,these models are also called weak learners and after that we calculate the mean accuracy from all the models.

There are two ensemble technique is present :- 1> Bagging and 2> Boosting.

**1> Bagging:-** In Bagging technique the larger dataset first first split into subsets and those subsets are assign to different different models,these models are basically DecisionTrees and after that it calculates the average accuracy or mean accuracy from all DecisionTrees.This is why bagging is also known as parallal.

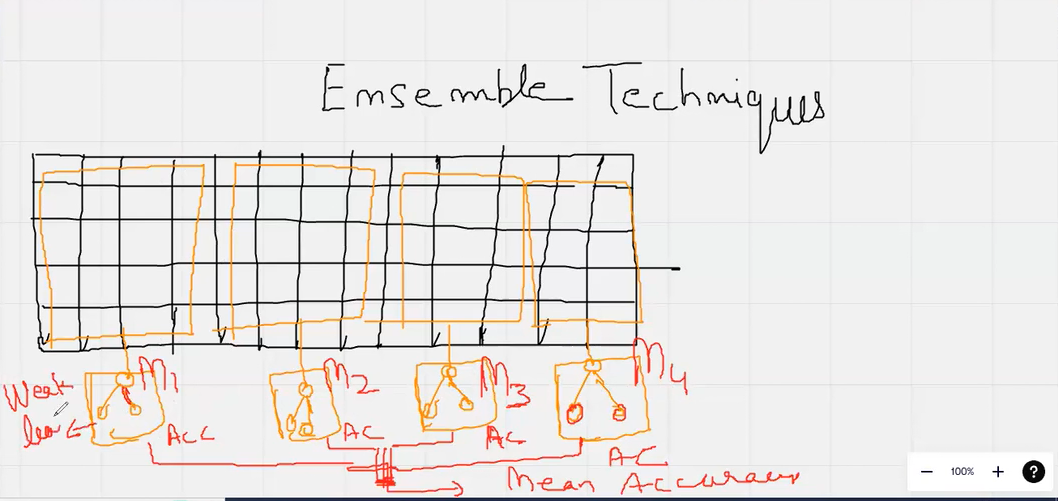
Bagging has two methods :- 1> RandomForest and 2> ExtraTrees.

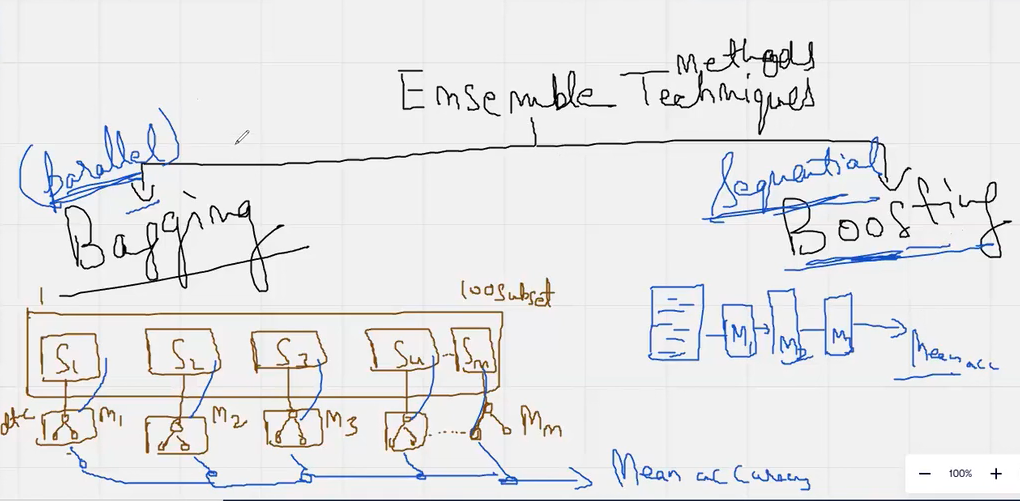
These two methods have their own Regressor and Classifiers like RandomForestRegressor , RandomForestClassifier and ExtraTreesRegressor , ExtraTreesClassifier.These are present in sklearn.

**2> Boosting:-** In Boosting technique the whole dataset passed through a model and then its output again passed through the another model and its goes on.Default number for this 50 but you can change that.After all this passing we get mean accuracy.So this is why Boosting is also called Sequential technique.

Boosting has three method:- 1> Adaboost , 2> GradientBoosting and 3> Voting classifier

These two methods have their own Regressor and Classifier.Those are like AdaboostRegressor , AdaboostClassifier and GradientBoostingRegressor, GradientBoostingClassifier.These all are present in sklearn.



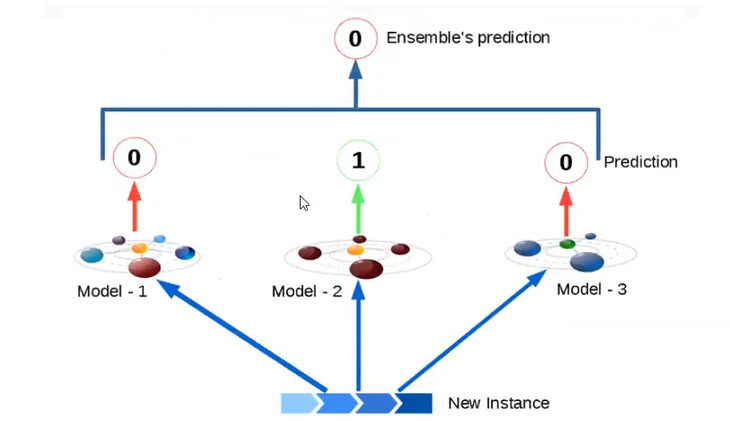


# Voting classifier:-

A Voting Classifier is a machine learning model that trains on an ensemble of numerous models and predicts an output (class) based on their highest probability of chosen class as the output. It simply aggregates the findings of each classifier passed into Voting Classifier and predicts the output class based on the highest majority of voting. The idea is instead of creating separate dedicated models and finding the accuracy for each them, we create a single model which trains by these models and predicts output based on their combined majority of voting for each output class.

Voting Classifier supports two types of votings.

Hard Voting: In hard voting, the predicted output class is a class with the highest majority of votes i.e the class which had the highest probability of being predicted by each of the classifiers. Suppose three classifiers predicted the output class (A, A, B), so here the majority predicted A as output. Hence A will be the final prediction.



Soft Voting: In soft voting, the output class is the prediction based on the average of probability given to that class. Suppose given some input to three models, the prediction probability for class A = (0.30, 0.47, 0.53) and B = (0.20, 0.32, 0.40). So the average for class A is 0.4333 and B is 0.3067, the winner is clearly class A because it had the highest probability averaged by each classifier.

