

Date.....

Ensemble Learning

Q What is the meaning of Ensemble?

Ans Ensemble means \Rightarrow A group of anything.

In Machine Learning, Ensemble learning is defined as the collection of different Models and develop a new model through these Models.

Q What is the Core Idea of Ensemble Learning?

Ans ① Training \rightarrow For any ML Model...
② Prediction

② Prediction $M \Rightarrow$ ML Model

Way - First

Collection of

these

different
ML Models

KNN	SVM	LR	Log	DT
M_1	M_2	M_3	M_4	...	M_n

is termed as Ensemble learning...

"Note:- These all Models should be different"

Way - Second

In this

case, the
Model is fixed

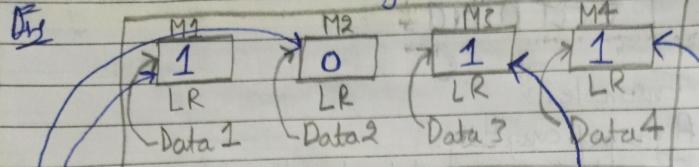
"Linear Regression"

But the data

'D' provided to all the Models are different...

LR	LR	LR	LR	LR
M_1	M_2	M_3	M_4	...	M_n
D_1	D_2	D_3	D_4	...	D_n

Q How the Ensemble Learning do Predictions?



Date...		
cgpa	ig	placement

Prediction { 8.1 (cgpa), 85 (ig) }

All the Models were send the Prediction:-

Prediction of M1 Model $\Rightarrow 1$

Prediction of M2 Model $\Rightarrow 0$

Prediction of M3 Model $\Rightarrow 1$

Prediction of M4 Model $\Rightarrow 1$

{ 1 \Rightarrow ? times }
0 \Rightarrow 1 times }

Majority Count...

So, as per the Ensemble learning, the Prediction will be "1".

Q What are the types of Ensemble learning?

Ans

Ensemble

Voting Ensemble

Bagging

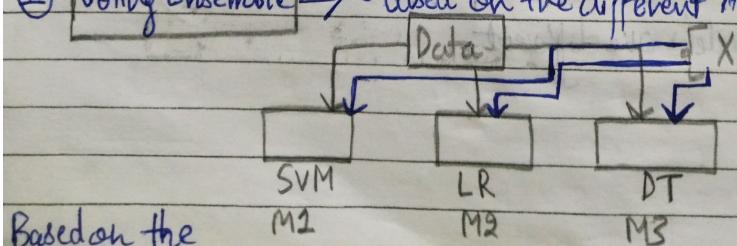
Boosting

Stacking

"Random Forest..."

- AdaBoosting
- Gradient Boosting
- Xg Boosting

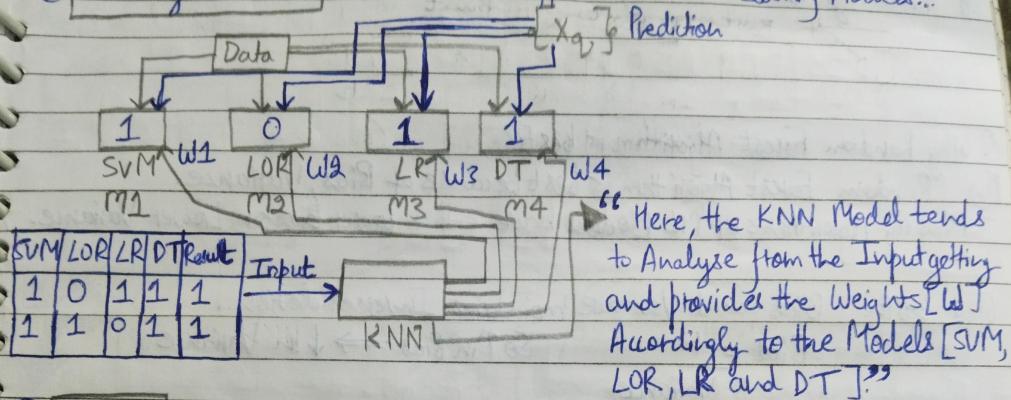
① Voting Ensemble \Rightarrow Based on the different Algorithms (Models).



Based on the Majority Count \Rightarrow Prediction will be done....

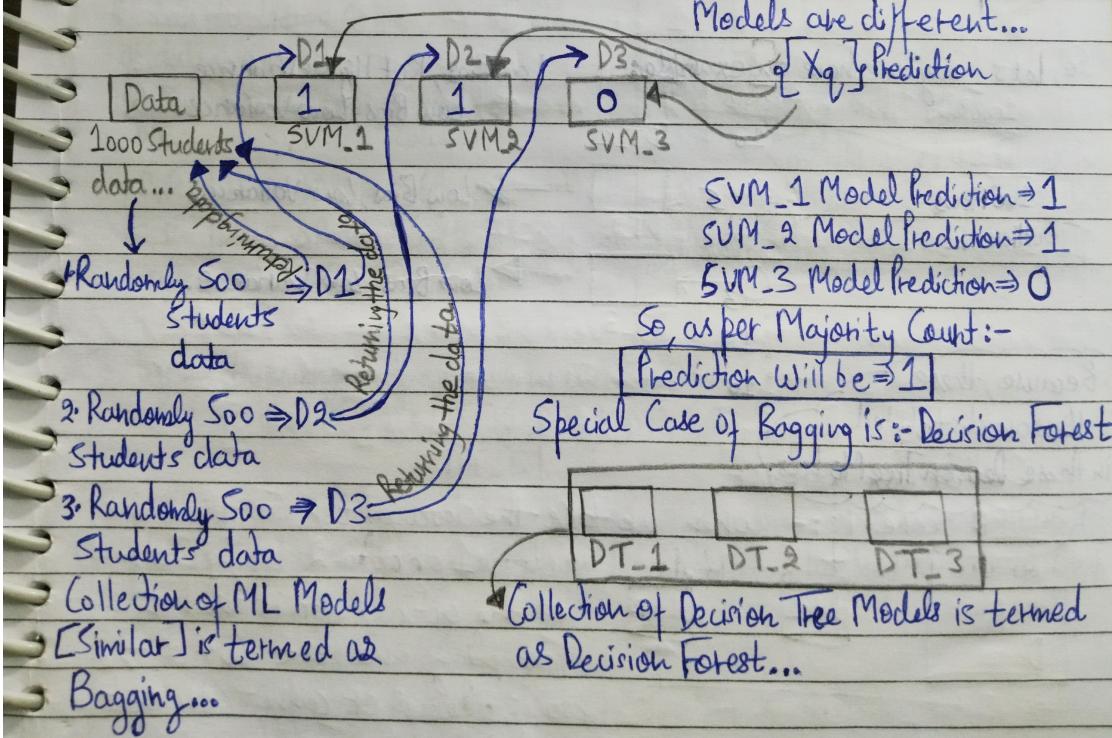
Date.....

② Stacking Ensemble \Rightarrow Based on the different Machine Learning Models..



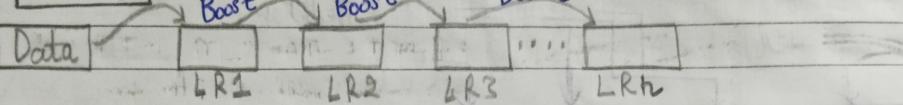
③ Bagging \Rightarrow Bagging is also defined as Bootstrapped Aggregation.

In Bagging, the Criteria is \Rightarrow The Base Models will be same but the data provided to those Models are different...



Date.....

④ Boosting \Rightarrow In Boosting Criteria \Rightarrow The Base Model is Same...



Q Why Random Forest Algorithm is best?

A₃ Random Forest Algorithm is best because of Bias, Variance.

Mostly Algorithms of ML Models have either lower Bias or Lower Variance.

But, Both Bias and Variance maintains Inverse Relation..

$$\text{Bias} = \frac{1}{\text{Variance}}$$

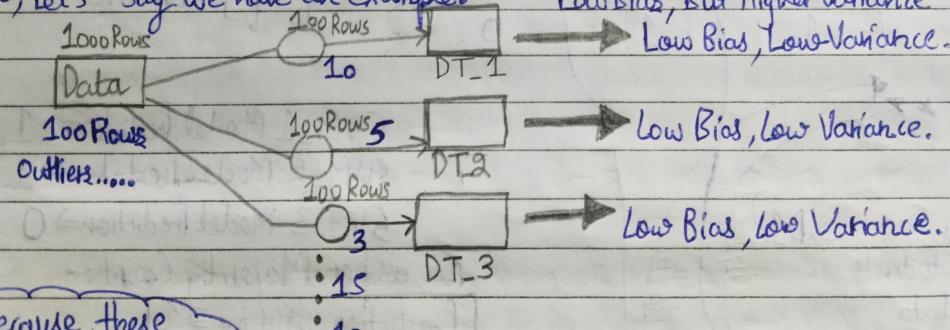
So ↑ in Bias \Rightarrow ↓ in Variance.

① Higher Bias is not good for the Model.

LB LV

② Higher Variance is also not good for the Model,
as it provides fluctuation in the Prediction.

So, Let's say we have an example:-



Because, these
Outliers are distributed
in these Decision Tree Models.