Ensemble Learning

***Wisdom of Crowd***

The wisdom of crowds and collective decision-making are important tools for integrating information between individuals, which can exceed the capacity of individual judgments

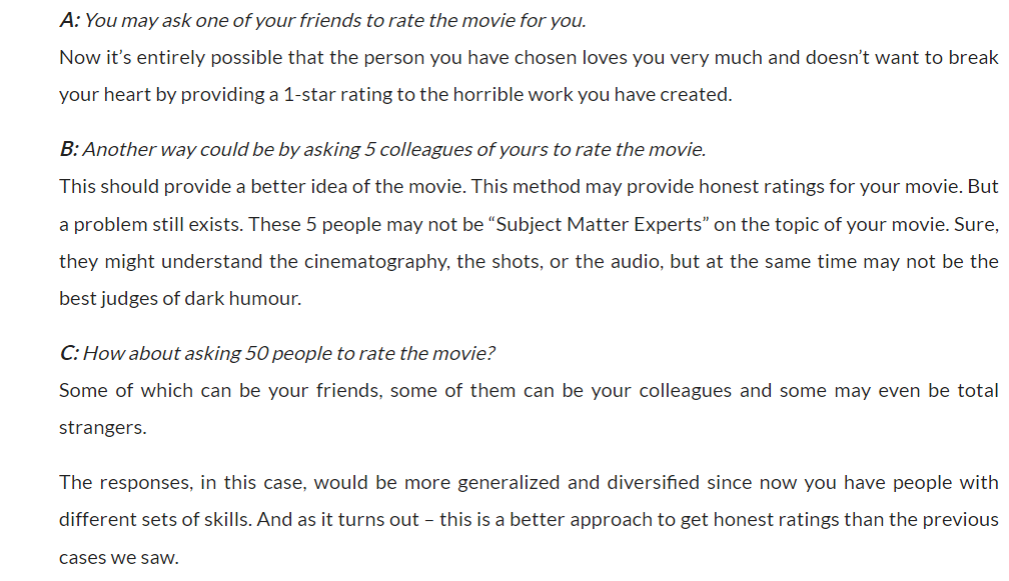
Basically Collective Decision jo rehta vahi,

E.g: Deomocracy mai PM Election etc etc

Ish Basis pe Ensemble Learning aaya hai

They combine the decisions from multiple models to improve the overall performance.

The underlying concept behind ensemble learning is to combine the outputs of diverse models to create a more precise prediction. By considering multiple perspectives and utilizing the strengths of different models, ensemble learning improves the overall performance of the learning system. This approach not only enhances accuracy but also provides resilience against uncertainties in the data



[Ensemble Learning | Ensemble Techniques (analyticsvidhya.com)](https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/)

<https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/>

Ensemble Learning

alag Algos use krna , Data alag hai, 3rd Hybrid Approach hai

1. Different Algos hoo alag alag
2. Algo 1 hie hoo but training saari model ki alag alag dataset pe huyi hai
3. Hybrid way 🡪 upar ka dono ka combination

*Classification Way*

1 new query point aaya

e.g iq & cgpa aaya placement Predict krna hai uska

1st Model : yes

2nd Model : no

3rd Model :yes

4th Model : yes

5th Model : no

Fir yaha Majority Count lagega

3 model says Yes

2 model says no

So majority count ke hisabse Placement yes hoga

**Regression Way**

1. Different Model

Or else

1. Same model trained on different dataset

Sbh Models ko bhejdenge Query Point

Jo Prediction aayega

Unsbhka Mean nikallega & vahi humara answer hoga

E.g:

Model 1 = 9

Model 2 = 10

Model 3 = 10

Model 4 = 9

Mean of all is like 9.5. So the Answer is 9.5

2 Ways hai jaise

1. Alag Models use kro
2. Alag Data alag Data use kro ek hie model mai

*Types of Ensemble Learning*

1. Voting Ensemble (Easy)
2. Bagging (usefull ) 🡪 Random Forest
3. Boosting (Sbhse zyaada usefull) 🡪

Ada Boost, Gradient Boosting, XgBoost

1. Stacking
2. **Voting Ensemble**

Models with Different Algorithms

Ispe base hai yeah

Ese hum Variety Create krrhe hai

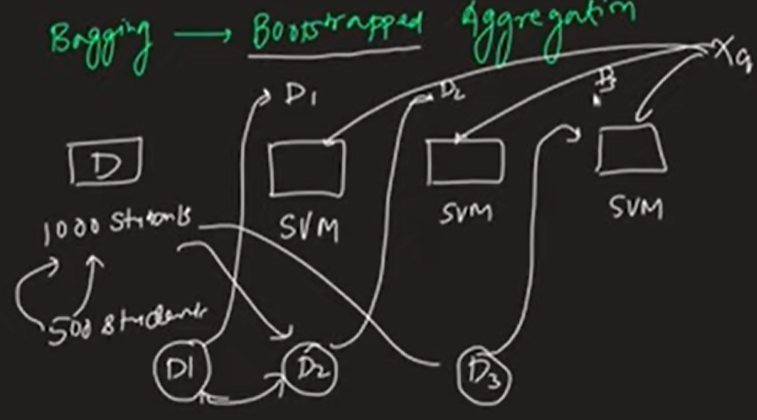
1. **Stacking**

Yeah bhi Voting ke bhaut Close hai

1. **Bagging**

Bootstrap Aggregation

Har Model ko alag Data dete hai



Yaha Hum kya krte hai ki

Har Model ko alag Data dete hai

We Give random different Data to every model

**So in Above Image:**

There is a Dataset of thousand rows

& there are 3 Models

Then what we do We Takeout 500 Random Data from Main Dataset

& give it to Model 1

Then we do same things with all the remaining 2 Models to

We just take out Random Data & give them both

So because of this there is a Huge Variation of Data for the Model

& then We give them a Query point or unseen Data to Predict or Classify about it

Then we get x result From all the 3 models & now what we do is ,

If it is Classification then we apply Majority Count & jiska Majority is higher lyk whether it is yes or no if there is 2 yes & 1 No then we will take Result as 2

If it is Regression then we take out Mean as a result.

**This is How Bagging Works**

Bagging has 1 Special Case Called as Random Forest

Jo Random Forest hota hai

**That is Special Case of Random Forest Jaha Model Hamesa Decision Trees hote hai**

**That’s why Hum ishe Random Forest bolte hai**

**Suppose Hum Kisi aur Algorithm ka agar Yeah kaam krrhe hai jaise above Example mai SVM tha so ushe hum Normal Bagging kahenge**

**But if We are using Collection of Decision Tree then we call it as Random Forest.**

***Boosting***

In Boosting also we use Same Model

At time of Training Model note down ki what mistake it is doing

& Jo vo galti krta hai

Vo Aageh Vale Model ko Batadeta hai ki yaha yaha Galti huyi hai

So vo thoda Improve krta hai

**E.g:**

Model 1 Traing hote time Note Down krta hai ki where it is making Mistakes

Then vo jo Mistakes hoti hai he share that mistake with

Model 2 & then Model 2 thoda improved result dene lagta hai

& then Model 2 bhi note down krta hai ki where he is making mistake then it transfers that mistake to Model 3 & Model 3 make sure krta hai

Ki Model 2 vali galti na hoo & model 3 aur Improved Result dene lagta hai

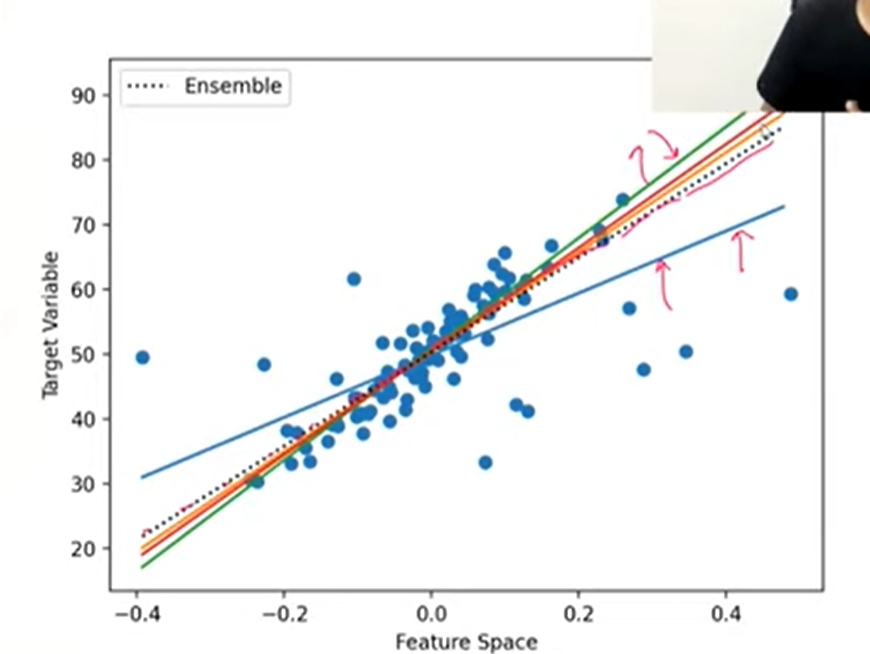
Then ese Loop mai chalte rehta hai

It is One of the Most Powerfull technique of ML

**Like this in Chain it is how It all Works**

Why Ensemble Learning Works

***In Case of Regression***



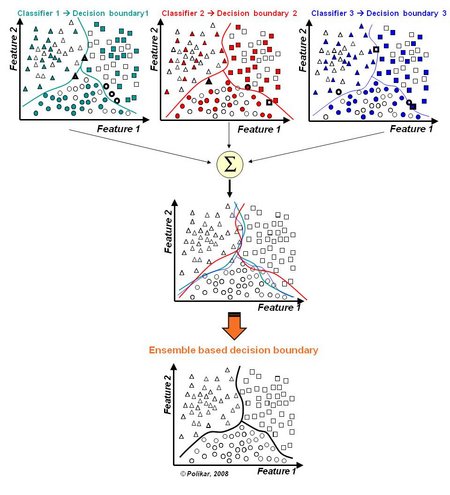
As we can see ki itne saare lines hai

But Ensemble Kya krta hai Humme Mean Dedeta hai Insbh Line ka

So Best Result kindoff Milljata hai

*Case of Classification*

3 Models hai & if you observe in the Image Teeno ka Decision Boundary alag hai & At the End Majority Count ke through Hum Decision Boundary Separate krderahe hai .



*Benefit of Ensemble Learning*

**First let us know 1 Disadvantage of it**

If you are using Ensemble Learning then you have to use & run Multiple Models which means Lot of Computation.

**If using Ensemble Learning has this disadvantage then why we use it ??**

**Simple Because of its Game Changing Advantages**

1. **Improvement in Performance**
2. **Bias & variance Reducer**

**(End Goal -> Low Bias + Low Variance)**

**Becoz of Ensemble we are able to achieve this**

1. **Robustness**

**When to Use?**

Always use it.

There is no Reason to not use it

1. **Industry**
2. **Kaggle Compedition Xd**

**So yeah use krnekebaad Even Deep Learning ka use bhi nahi krna padta if Dataset Chota hai tou**

**Data set chota hai tou Ensemble Learning Beats Deep Learning**