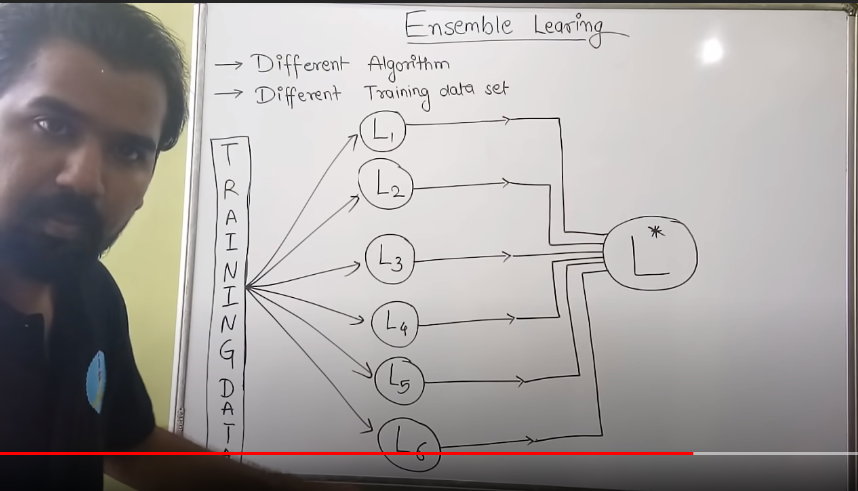
<https://www.youtube.com/watch?v=LNrBcDfUhq0>



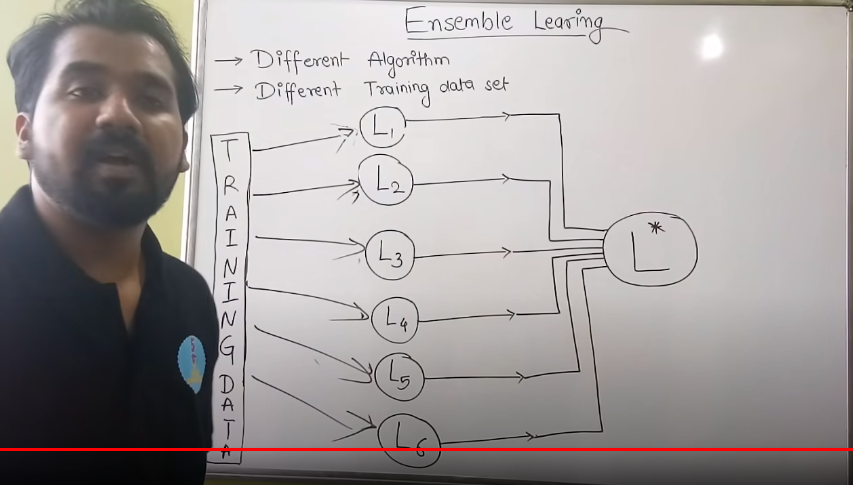
Two types of ensemble technique

* Heterogeneous – Here, **weak learners/models** are different i. e L1(Decision Tree), L2(SVM),L3(KNN)…so on
  + L1, L2, L3 …all takes complete/same data set to get trained itself (mentioned in snapshot above)
* Homogeneous - Here, weak learners/models are same i. e L1(KNN), L2(KNN),L3(KNN)…so on
  + L1, L2, L3 … all takes different set of random data. Let’s assume train data has 1000 records
    - L1 takes random 2/3 of 1000 records randomly (sampling with replacement), similarly L2,L3
    - L1, L2, L3 can have same records repeatedly I. e One record can come in all three L1, L2, L3 because we using technique called as **sampling with replacement**

**Final Classifier** – At the end we pass a testing record to test our complete ensemble method.

* Every model will give some prediction. Suppose L1 gives 1, L2 gives 1, L3 gives 0, L4 gives 1, and L5 gives 0. There is final classifier that aggregate prediction from all weak learners and take majority vote and give final prediction as 1.

**Boot Strapping** – The process of creating multiple weak learners/models having randomly selected data with sample with replacement technique and finally aggregating prediction from all models to get better result is called bootstrapping.



<https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205>

Two type of ensembles methods -

* Bagging
* Boosting

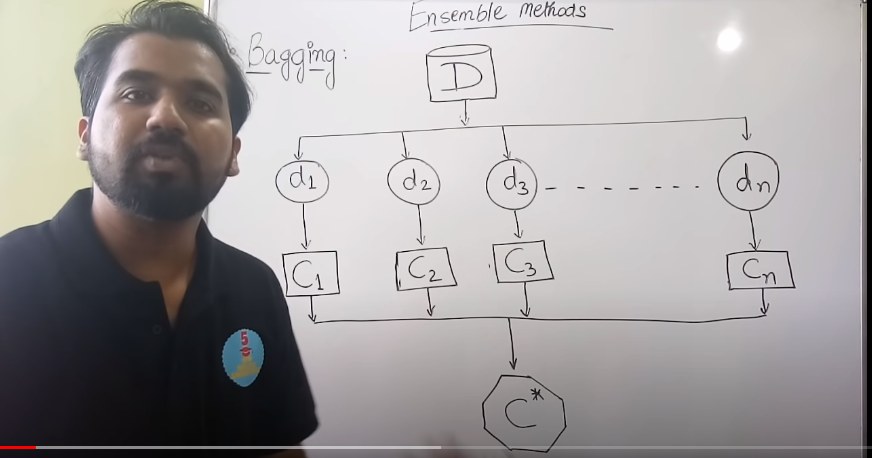
**Bagging** - It often considers homogeneous weak learners, learns them independently from each other in parallel and combines them following some kind of deterministic averaging process

**Boosting** – It often considers homogeneous weak learners, learns them sequentially in a very adaptive way (a base model depends on the previous ones) and combines them following a deterministic strategy

**Bagging vs Boosting** - Bagging will mainly focus at getting an ensemble model with less variance than its components whereas boosting will mainly try to produce strong models less biased than their components (even if variance can also be reduced).

**Bagging** - <https://www.youtube.com/watch?v=kW4XSvuBtDw>

* Create boot strap samples D1, D2, D3…from main data D.
  + Records, Observations in bootstrap samples D1, D2, D3…etc. will always be less than D. On an average it will be 2/3 of D.
  + Observation might get repeated as is based on sampling with replacement. (One record can come in all three
* C1, C2, C3… are the homogeneous classifiers/models/weak learners (Decision Trees) in case we are using random forest.
* All models will get trained independently from each other in parallel
* **Final Classifier** – Once training gets completed, we pass a testing record to test our complete ensemble method.
  + Every model will give some prediction. Suppose L1 gives 1, L2 gives 1, L3 gives 0, L4 gives 1, and L5 gives 0. There is final classifier that aggregate prediction from all weak learners and take majority vote and give final prediction as 1.



**Boosting -** <https://www.youtube.com/watch?v=CV9PE3iTjPI>

* Suppose main data set ‘D’ some records and we need to create boot strap samples D1, D2, D3…samples with replacement
* Initially, we initialise same weight to all records/observation in main data set ‘D’
  + In D1, we took some observations randomly 1,3,6,7 and train the model.
    - We pass entire data set ‘D’ as testing data set and make prediction. Suppose it misclassify the 4 & 6 observation
    - Now weight of misclassified observation (4&6) will get updated (assign higher weights)
  + In D2, we will take 4 & 6 observations with updated weights along with some other observation (7,3) and train the model
    - We pass entire data set ‘D’ as testing data set and make prediction. Suppose this time it misclassify the 2 & 3 observation
    - Now weight of misclassified observation (2&3) will get updated (assign higher weights)
  + In D3 – Same process repeats as D2
  + This process will keep on repeating till all classifier get absorbed or there will be no more misclassification
* **Final Classifier** – Once training gets completed of all classifiers, we pass a testing record to test our complete ensemble method.
  + Every model will give some prediction. Suppose M1 gives 1, M2 gives 1, M3 gives 0, There is final classifier that aggregate prediction from all weak learners and take majority vote and give final prediction as 1.

