31/10 Ensemble learning

Date 20
M T W T F S S

Mid topics
- LR
- NN
- CNN
- Ensemble (partial)

- feedbacks, different opinions for how something is suitable for you
- we combine different opinions — cumulative
- e.g. sale a property — property value after 5 values

Take average of different opinions
This is Ensembling
- voting also
  - novice, professionals/experts → opinion inclined towards experts
  - this is weighted ensembling
    expert opinion is 3x that of novice

- A single complex model is sometimes not enough to recognize data
- Many simples models are trained to understand parts of data and results are combined. → better accuracy

→ computational power, resources required

- Ensemble is type of divide and conquer method
- Models can be same like Decision Tree
- Single dataset and different models ?
  LR DT

⇒ simple Ensemble Techniques: — different model on same data
1) Taking mode of the results:
   - voting scheme
   → maximum prediction
   - we keep it odd

2) Taking average of the results :
↳ average of predictions from all models → final prediction

For regression problem, this could be used

3) Taking weighted average of the results :
↳ models are assigned weights on basis of their importance
model weights are hyperparameters
↳ we ~~assigned~~ ourselves
Models are not equally powerful

⇒ Advanced Ensemble Techniques :
- Train same ~~data~~ model multiple times on different data sets
1) Bootstrapping aggregating / bagging :
Split in train / test
↓
random subsets of train data
- Apply same model on the data subsets
- Take mean

2) Random Forests
- train decision tree

3) Boosting :
subsets → apply weak learner → mistakes on data points
model now learns the mistakes ← re-train ← increase weights of such data points

Date 20
M T W T F S S

- Ensemble techniques reduce variance so no problem of overfitting
- Subsets of the samples } → subsets ways
- Subsets of the features }

→ Random Forest:
- it splits ~~feat~~ data in both ways
- it uses bagging on features
  e.g. 100 features subset → DT laga diya → ez pz

- trees will be different but depth should be same
- har ek DT ne different features se learn kiya hoga

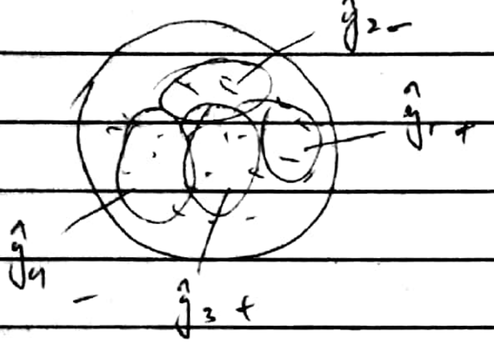
→ Boosting :
- focuses on classification problems
  XGBoost (Extreme Gradient Boost)

CamScanner

3/11 Ensemble learning

split a big problem into a set of simpler problems

→ Random Forest



we make odd no. of DTs

depth of all trees must be same

↳ so that outcome is at the same time

more depth → overfitting

less depth → underfitting

post-pruning , pre-pruning

→ Ada Boosting :

input → training example

classifier → function $f(x)$

~~wea~~

> 0 — it is threshold ↓ can be different in iterations

- initially equal weights — weights are probabilities
- iterations from 1 to T (no. of examples).
- Then we update the probabilities
- We see examples which have been misclassified
- We update threshold in each iterations
- Sum samples's probabilities which have been misclassified

E_t → error term ~~can be~~

α_t → can be a learning rate

* calculator

Stopping criteria → when no errors

Teacher's Signature ____________