

SESSION PLAN

Session Name

Ensembling and Random Forest

Learning Outcomes

- Understand the intuition behind ensemble methods
- Know about bias-variance trade-off
- Know how different types of ensemble methods work

Prerequisites for the Student

- Ensembling and Random Forest - Go through the concept and solve the tasks and assessments.

Student Activities

- Discuss with the Mentor what you have learned.
- Overview of Ensembling and Random Forest
 - Aggregation
 - Hyper-parameter tuning
- Intuition on Random Forest:- https://www.youtube.com/watch?v=D_2LkhMJcfY
- Medium blog on Ensemble methods.
<https://medium.com/diogo-menezes-borges/ensemble-learning-when-everybody-takes-a-guess-i-guess-ec35f6cb4600>
- Medium blog on Random Forest.
<https://medium.com/@williamkoehrsen/random-forest-simple-explanation-377895a60d2d>
- Consider 5 different models who try to predict the certain values.
Model1 = 30, model2 = 35, model3 = 37, model4 = 37, model5 = 38
(Actual model predicted value is 36 but we see none of the models predicted that)
 - Combine all the 5 models by calculating the average?
 - Use the voting method to find the value that has been predicted by the majority of the model.
 - Based on the model's accuracy, we have now given weights to each model. Let's now calculate the Weighted average: $(2 * \text{model1} + 2 * \text{model2} + 2 * \text{model3} + 2 * \text{model4} + 5 * \text{model5}) / 13$
- What's the Bias and variance tradeoff in different ensembling methods?
- Practice problem on Ensembling and Random Forest
 - Refer the GitHub repo for problems
- Quiz on Ensembling and Random Forest.
- Code Along.
- Questions and Discussion on doubts - AMA

Next Session

- Concept - Gradient Boosting Machines
- Key topics to be highlighted - highlight where they would need to spend more time and importance w.r.t Data Science.
 - Intuition of Boosting
 - Adaboost
 - Gradient Boosting
 - XGBoost