Hyperparameter tuning

HR ANALYTICS: PREDICTING EMPLOYEE CHURN IN PYTHON

Hrant Davtyan
Assistant Professor of Data Science
American University of Armenia

datacamp

GridSearch

Values for minimum samples in the leaf

Values for maximum depth	50	100	150	200	250	300	350	400	450
5	5, 50	5, 100	5, 150	5, 200	5, 250	5, 300	5, 350	5, 400	5, 450
6	6, 50	6, 100	6, 150	6, 200	6, 250	6, 300	6, 350	6, 400	6, 450
7	7, 50	7, 100	7, 150	7, 200	7, 250	7, 300	7, 350	7, 400	7, 450
8	8, 50	8, 100	8, 150	8, 200	8, 250	8, 300	8, 350	8, 400	8, 450
9	9, 50	9, 100	9, 150	9, 200	9, 250	9, 300	9, 350	9, 400	9, 450
10	10, 50	10, 100	10, 150	10, 200	10, 250	10, 300	10, 350	10, 400	10, 450
11	11, 50	11, 100	11, 150	11, 200	11, 250	11, 300	11, 350	11, 400	11, 450
12	12, 50	12, 100	12, 150	12, 200	12, 250	12, 300	12, 350	12, 400	12, 450
13	13, 50	13, 100	13, 150	13, 200	13, 250	13, 300	13, 350	13, 400	13, 450
14	14, 50	14, 100	14, 150	14, 200	14, 250	14, 300	14, 350	14, 400	14, 450
15	15, 50	15, 100	15, 150	15, 200	15, 250	15, 300	15, 350	15, 400	15, 450
16	16, 50	16, 100	16, 150	16, 200	16, 250	16, 300	16, 350	16, 400	16, 450
17	17, 50	17, 100	17, 150	17, 200	17, 250	17, 300	17, 350	17, 400	17, 450
18	18, 50	18, 100	18, 150	18, 200	18, 250	18, 300	18, 350	18, 400	18, 450
19	19, 50	19, 100	19, 150	19, 200	19, 250	19, 300	19, 350	19, 400	19, 450
20	20, 50	20, 100	20, 150	20, 200	20, 250	20, 300	20, 350	20, 400	20, 450

Cross-Validation

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Test	Train	Train	Train	Train
Train	**Test**	Train	Train	Train
Train	Train	**Test**	Train	Train
Train	Train	Train	**Test**	Train
Train	Train	Train	Train	**Test**

Let's practice!

HR ANALYTICS: PREDICTING EMPLOYEE CHURN IN PYTHON

Important features for predicting attrition

HR ANALYTICS: PREDICTING EMPLOYEE CHURN IN PYTHON

Hrant Davtyan
Assistant Professor of Data Science
American University of Armenia

datacamp

Feature Importances

- Importance is calculated as relative decrease in Gini due to the selected feature.
- Importances are scaled to sum up to 100%.
- Higher percentage, higher importance.

Let's practice!

HR ANALYTICS: PREDICTING EMPLOYEE CHURN IN PYTHON

Final thoughts

HR ANALYTICS: PREDICTING EMPLOYEE CHURN IN PYTHON

Hrant Davtyan
Assistant Professor of Data Science
American University of Armenia

Alternative methods

- Logistic Regression
- Tree based
  - Random Forest
  - Gradient Boosting
- Neural Networks

The End

HR ANALYTICS: PREDICTING EMPLOYEE CHURN IN PYTHON