← Go Back to Model Tuning

Course Content

Course Outline

Overview

Predicted outcomes of a model are associated with business costs, and to reduce the cost, it is necessary to maximize a model's performance. In order to maximize a model's performance, we need to first get a clear picture of how a given model is performing, and using cross-validation helps us better use our data to get that picture. Once this clarity is obtained, we can use one of the most common methods to maximize a model's performance, called hyperparameter tuning, to build a good model.

When dealing with any classification problem, there will be situations where one has to deal with datasets having an unequal distribution of the target class, and the value of finding the minority class is much higher than that of finding the majority. Such datasets are called imbalanced datasets and one can use specialized techniques to deal with the imbalance and build a robust predictive model.

We will be covering the concepts of cross-validation, imbalance handling, and hyperparameter tuning in this course. We will also cover how to automate standard workflows in a Machine Learning process using a Python pipeline, which will enable robust implementation of the processes and make work more reproducible.

Course Objectives

The objective of the course is to understand how to standardize the model-building process using pipelines and optimize the business costs associated with model predictions by performing the following:

- Feature engineering
- Handling imbalanced data
- Regularization to reduce overfitting

- Model hyperparameter tuning

Topics Covered

Week	Module	Name of the topic
1	Feature Engineering and Cross-Validation	• Feature Engineering • Cross-validation • Oversampling and Undersampling • Regularization
2	ML Pipeline and Hyperparameter Tuning	• Machine Learning Pipeline • Model Tuning and Performance • Hyperparameter Tuning • Grid Search • Random Search

Learning Material

Week	Module	No. of Videos	Total Duration	No. of Graded Quizzes	No. of Test Your Understanding Quizzes	No. of Practice Assignments
1	Feature Engineering and Cross-Validation	8	~2 hours	1	8	1
2	ML Pipeline and Hyperparameter Tuning	8	~1.5 hours	1	8	1

Project

A graded project is to be submitted by the learners at the end of the course.

The project is about “ReneWind”, a company working on improving the machinery/processes involved in the production of wind energy using machine learning by analyzing sensor data from wind turbine generators. The objective is to build various classification models, tune

them, and find the best one that will help identify failures so that the generators could be repaired before failing/breaking to reduce the overall maintenance cost.

Power Ahead!

Next >