Home

Answers

Explore

Live Events ^

ALL EVENTS

ARCHITECTURAL KATAS

AI & ML

DATA SCI & ENG

PROGRAMMING

INFRA & OPS

SOFTWARE ARCH

Interactive

Certifications

< VIEW ALL EVENTS

((•)) MACHINE LEARNING

Learn XGBoost for Machine Learning

What you'll learn   Is this course for you?   Schedule

XGBoost is a powerful machine learning algorithm that is designed for speed and performance and rates along-side neural networks and deep learning models. Hence, it is being considered a new one-stop machine learning algorithm. Large-scale cloud providers such as AWS support this algorithm and heavily use it.

This training is important as it teaches XGBoost from the ground up. The course begins with the basic intuition behind how XGBoost works and it gradually moves to intermediate and advanced levels. You will learn how to install the XGBoost package in Python and use it for classification, regression and time-series forecasting. The training also involves training XGBoost models locally as well as on the AWS cloud.

Familiarity with XGBoost and how to optimally use it provides a big plus in any data scientist's skill set. This course is rich with ideas and hands-on exercises that involve solving real-world machine learning problems using XGBoost.

What you'll learn and how you can apply it

- Develop an understanding of the power of the XGBoost algorithm

- Become a competent user of this powerful machine learning algorithm

- Learn how to solve real-world classification, regression, and time-series forecasting problems using this algorithm

- Learn how to run XGBoost hyperparameter optimization to select the best hyperparameter combination for this algorithm

- Learn how extract and interpret feature importance using XGBoost

- Learn how to use XGBoost for feature selection

- Learn how to do all of the above locally and on the AWS cloud

This course is for you because...

- You have machine learning experience and would like to be familiar with a powerful algorithm

- You are keen to extend your skill-set and learn how to use XGBoost locally and on AWS

- You would like to learn how to perform hyperparameter tuning for XGBoost and other algorithms locally and on AWS

- You would like to learn a single algorithm that can help you solve real-world classification, regression and time-series problems

- You would like to increase your employability by gaining these skills

- You would like to become a competent machine learning specialist

- You would like to work as a well-versed data scientist

Prerequisites

- Familiarity with Python. Students should be relatively comfortable with Python coding practices (i.e. intermediate Python level).

- Basic knowledge of classical machine learning algorithms such as decision trees, random-forest, and artificial neural networks

- Basic knowledge of AWS

Course Set-up

- Any operating system is fine

- Speedy internet connection

- Python 3.6 or above (Anaconda distribution is highly recommended) Intermediate level Python knowledge is required (e.g. knowledge of how to install and use packages, dealing with objects and so on)

- Installation of boto3 Python package

- An account on AWS

June 3, 2021
6:00 p.m. - 10:00 p.m. GMT+2

You're signed up!

Add to Calendar

Cancel registration

YOUR INSTRUCTOR

Noureddin Sadawi

Dr. Noureddin Sadawi is a consultant in machine/deep learning and data science. He has several years' experience in various areas involving data manipulation and analysis. He received his PhD from the University of...

Read more

Recommended Preparation

- Video: Introduction to Python. By: Arianne Dee.
  https://learning.oreilly.com/videos/introduction-to-python/9780135707333/

- Book: Learning Amazon Web Services (AWS): A Hands-On Guide to the Fundamentals of AWS Cloud. By Mark Wilkins. https://www.oreilly.com/library/view/learning-amazon-web/9780135301104/

Recommended Follow-up

- Video: Essential Machine Learning and AI with Python. By: Noah Gift.
  https://learning.oreilly.com/videos/essential-machine-learning/9780135261118

- Video: AWS Certified Machine Learning Specialty. By: Noah Gift.
  https://learning.oreilly.com/videos/aws-certified-machine/9780135556597

Schedule

The timeframes are only estimates and may vary according to how the class is progressing.

Part 1: Introduction to XGBoost and its intuition (60 minutes)

- Overview of decision trees, how they work and why they are useful

- Overview of boosting and gradient boosting: intuition behind them

- Boosted trees and XGBoost: how XGBoost uses decision trees in model building

- Intuition of XGBoost and what gives it its power

- XGBoost vs RandomForest: master the difference between them

- XGBoost installation in Python

- Characteristics of XGBoost: what makes it special

Q&A (10 minutes)

Break (10 minutes)

Part 2: XGBoost for Classification, Regression and Time-series Forecasting (60 minutes)

- Binary and multi-class classification with XGBoost: learn how to make accurate classifications by solving a real-world problem

- Regression and time-series forecasting with XGBoost: solve real-world problems and become comfortable with predicting real-value outcomes (even if data is ordered by time)

- XGBoost feature importance: learn how to identify the most important features

- Feature selection with XGBoost: select the most informative features to optimize XGBoost

- Model parameters vs hyperparameters: learn the difference between them

- Explanation of XGBoost hyperparameters: learn why they mean so you can use them correctly

- Hyperparameters tuning for XGBoost: how to select the best combination of hyperparameters

Q&A (10 minutes)

Break (10 minutes)

Part 3: XGBoost on AWS's SageMaker (70 minutes)

- Overview of AWS's SageMaker and its implementation of XGBoost

- Overview of the boto3 SDK and SageMaker SDK Python packages

- Training XGBoost (for classification, regression or time-series analysis) on AWS's SageMaker

- Deploying a model endpoint and invoking it using Python (model hosting and invocation)

- Running a hyperparameter tuning job on AWS for XGBoost

- Why XGBoost can be better than deep learning (if time permits)

Q&A (10 minutes)