

Scaling Your Machine Learning and Deep Learning Pipelines

In this course you'll learn how to take the machine learning pipelines developed on your desktop, train them on Anaconda Enterprise using big data sources, and deploy them to the cluster. You'll gain hands-on knowledge of how Anaconda Enterprise maintains reproducibility of the entire life-cycle of the model from development and training to production. You'll practice building machine learning pipelines using leading packages provided by Anaconda.

Morning Session

- Develop ML and DL models on the desktop
- Create Anaconda Project archives and upload to Anaconda Enterprise
- Run model training on high-performance computing hardware

Afternoon Session

- Building simple web apps and REST APIs for model prediction
- Choosing the right hardware resource profile to deploy the app
- Sharing and interacting with deployed apps

Preparation Requirements

Students will connect to an Anaconda Enterprise instance maintained by the Anaconda Training department.

This is an Intermediate to Advanced Data Science course and requires that students have taken the [Getting Started with Anaconda Enterprise](#) course. Further, students are expected to be proficient in the following topics:

- **Data ingest from databases, Spark/HIVE, and other sources**
 - [Introduction to Databases in Python](#)
 - [Importing Data in Python \(Part 1\)](#)
 - [Importing Data in Python \(Part 2\)](#)
 - [Introduction to PySpark](#)
- **Data wrangling with NumPy and Pandas**
 - [pandas Foundations](#)
 - [Manipulating DataFrames with pandas](#)
 - [Merging DataFrames with pandas](#)
 - Familiarity with [Dask](#) is recommended
- **Simple modeling tasks with Scikit-Learn**
 - [Supervised Learning with scikit-learn](#)
 - [Deep Learning in Python](#)

Relevant DataCamp courses:

- [Unsupervised Learning in Python](#)
- [Machine Learning with the Experts: School Budgets](#)
- [Extreme Gradient Boosting with XGBoost](#)