

ADVANCED DATA SCIENCE & A.I. WITH PYTHON - MODULE 3

16-HOUR
CODING
CLASS

DATA ENGINEERING AND REGRESSION MODELING



Understand how to clean and model data for supervised **machine learning**. You will be able to use decision trees like Random Forest to build **classification models** and generalize new data and prevent overfitting in machine learning.

Course Fee: HKD12,800

Duration: please refer to the latest schedule on the website of RTP (Registered Public Courses)



Key Topic Highlights:

Data Cleaning and Supervised Learning

- Learn how to explore, clean and model data for supervised machine learning using preprocessing techniques. Implement linear and logistic regression, and **build classification models** with Random Forest to make predictions with Python machine learning library.

Unsupervised Learning

- Learn how to find associations, patterns, and relationships present in data using clustering techniques and dimensionality reduction techniques as well as matrix factorization to **break data set** in groups.

Generalization and Overfitting

- Determine whether a model is good or not by implementing both a training set and a test set to ensure its ability to adapt and **generalized to new, unseen data**. Learn how the reward-punishment mechanism works to penalized models to **prevent overfitting** using regularization and cross-validation methods to improve model accuracy.



Who is it for? Data Science & A.I. with Python Module 2 graduates who are interested to leverage Python further to derive value from data to research on various projects.

**66,000 NEW TECHNOLOGY AND DATA
SCIENCE JOBS WILL BE CREATED BY 2030.**

Source: Financial Times



Preface Coding



www.preface.ai/



+852 6634 0851

PREFACE
CODING



18/F, Tower 535, 535 Jaffe Road, Causeway Bay, Hong Kong