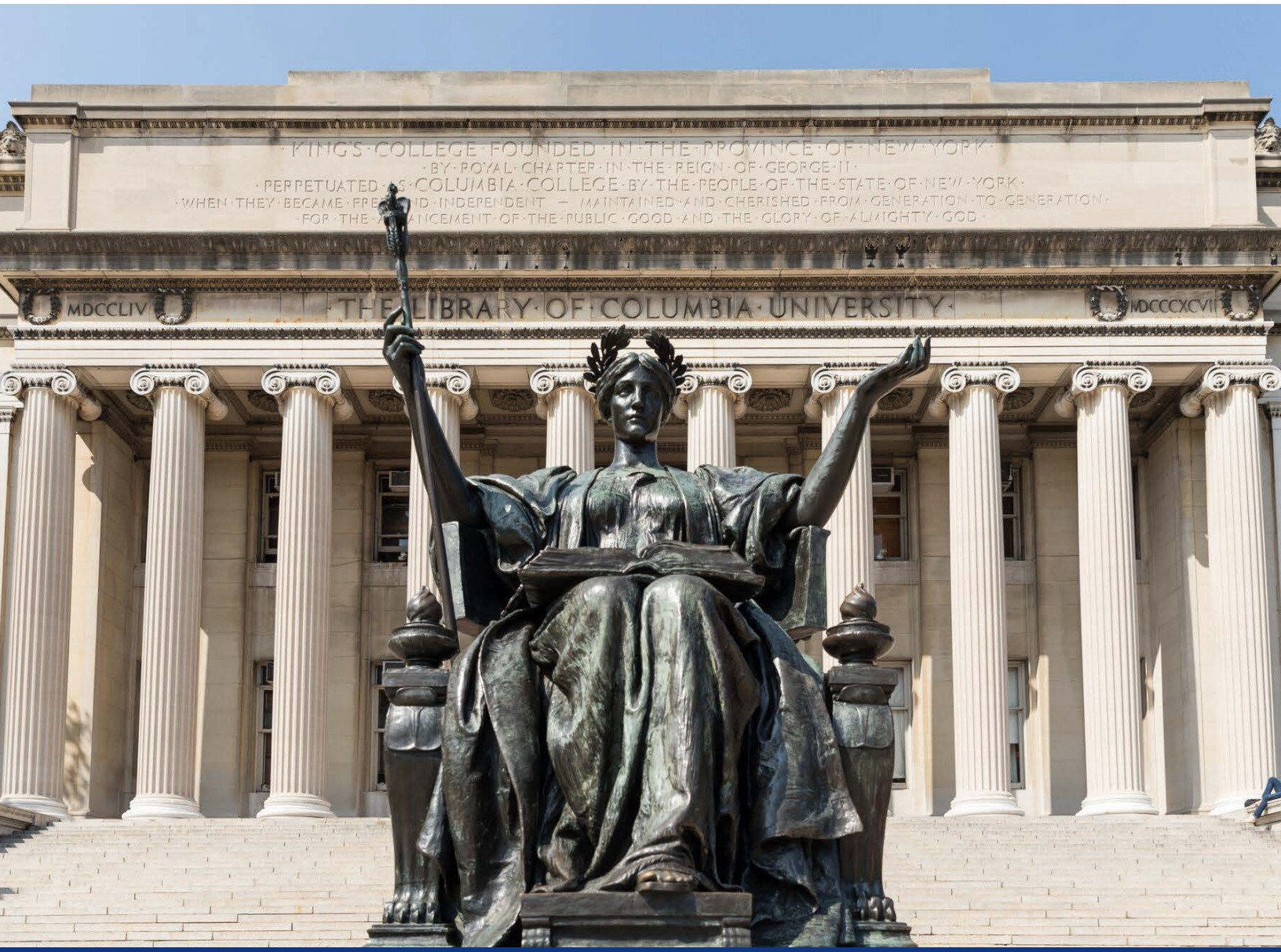
COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science

APPLIED MACHINE LEARNING COURSE



OVERVIEW

Machine Learning has become an entrenched part of everyday life. The books we buy, the movies we watch, the sports we follow, the driving directions we get are driven by Machine Learning algorithms. It is one of the most exciting fields of computing today. And Machine Learning practitioners are in high demand, with a shortfall of 250,000 data scientists forecast.

At Columbia Engineering, we are fascinated by the possibilities of Machine Learning. We have created the Applied Machine Learning course, in partnership with EMERITUS, to help students across the world apply Machine Learning to improve every aspect of human life.

> “Over the past decades computers have broadly automated tasks that programmers could describe with clear rules and algorithms. Modern machine learning techniques now allow us to do the same for tasks where describing the precise rules is much harder.”
>
> ***- Jeff Bezos, Founder, Amazon***

ABOUT COLUMBIA ENGINEERING

Transcending Disciplines, Educating Leaders, Transforming Lives

Columbia Engineering (Columbia University Fu Foundation School of Engineering and Applied Science), is committed to pushing the frontiers of knowledge and shaping discoveries to meet the needs of society. These aspirations have been fundamental since its early origins in 1864 as a school devoted to metallurgy and mining. Over the years, Columbia's faculty and students have made remarkable contributions that have spurred technological and social progress. Today, Columbia carries the tradition of innovation as engineering transforms nearly every aspect of life.

Faculty at Columbia Engineering have won 10 Nobel Prizes in physics, chemistry, medicine, and economics.

COURSE FACULTY

John W. Paisley

Columbia University
Associate Professor, Electrical Engineering
Affiliated Member, Data Sciences Institute

John has a PhD from Duke and has been a postdoctoral researcher in the Computer Science departments at Princeton University and UC Berkeley.

John Paisley's research focuses on developing models for large-scale text and image processing applications. He is particularly interested in Bayesian models and posterior inference techniques that address the big data problem.

COURSE LEADERS

Tom Dougherty

Course Leader, EMERITUS

Tom has spent his career in increasingly senior roles developing and managing advanced analytics teams across a range of industries. Most recently Tom spent 12 years as Head of Advanced Analytics for the Institutional Division of Fidelity Investments. He holds both bachelor's and master's degrees in mathematical sciences from Binghamton University.

Carleton Smith

Course Leader, EMERITUS

Carleton Smith is a Data Science educator and practitioner in Chicago, IL. Carleton has taught several data science courses at General Assembly and has served as the Lead Data Science Instructor for a number of corporate training engagements. Before becoming an Instructor, Carleton worked on data projects in the finance, health care, and medical research domains. Carleton has a MS in Predictive Analytics from DePaul University and an BS in Entrepreneurship from Indiana University's Kelley School of Business.

COURSE HIGHLIGHTS

Our approach to this course is to teach the underlying concepts and math of Machine Learning. Going beyond the theory, our approach invites participants into a conversation, where learning is facilitated by live subject matter experts and enriched by practitioners in the field of machine learning. We expect learners would be required to put in 6-8 hours per week.

Faculty Video Lectures

Quizzes / Assignments

Q&A Sessions with Course Leaders

Moderated Discussion Boards

Application Projects

Live Online Teaching

SYLLABUS

Supervised Learning

Module 1

Regression

Maximum Likelihood, Least Squares, Regularization

Module 2

Bayesian Methods

Bayes Rule, MAP Inference, Active Learning

Module 3

Foundational Classification Algorithms

Nearest Neighbors, Perceptron, Logistic Regression

Module 4

Refinements to Classification

Kernel Methods, Gaussian Process

Module 5

Intermediate Classification Algorithms

SVM, Trees, Forests and Boosting

Unsupervised Learning

Module 6

Clustering Methods

K-Means Clustering, E-M, Gaussian Mixtures

Module 7

Recommendation Systems

Collaborative Filtering, Topic Modeling, PCA

Module 8

Sequential Data Models

Markov and Hidden Markov Models, Kalman Filters

Module 9

Association Analysis

Module 10

Clustering Methods

Model Comparisons, Analysis Considerations

All assignments and application projects will be done using the Python programming language.

APPLICATION PROJECTS

The course requires learners to work on application projects. These projects require learners to apply the Machine Learning concepts they have learned to datasets and derive inferences. These application projects are intentionally made to be challenging. We expect learners to spend substantial time and effort solving the application projects. At the end of the course, we expect learners to be able to apply Machine Learning to solve many of the business problems they face at their workplace.

Movie Recommendation Engine

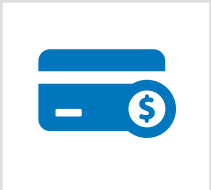
You will build a movie recommendation engine by applying collaborative filtering and topic modelling techniques. You use a dataset which contains 20 million viewer ratings of 27,000 movies.

House Price Prediction

You will write code to predict house prices based on several parameters available in the Ames City dataset compiled by Dean De Cock using least squares linear regression and Bayesian linear regression.

Human Activity Recognition

You will predict the human activity (walking, sitting, standing) that corresponds to the accelerometer and gyroscope measurements by applying the nearest neighbours technique.



Credit Card Fraud Detection

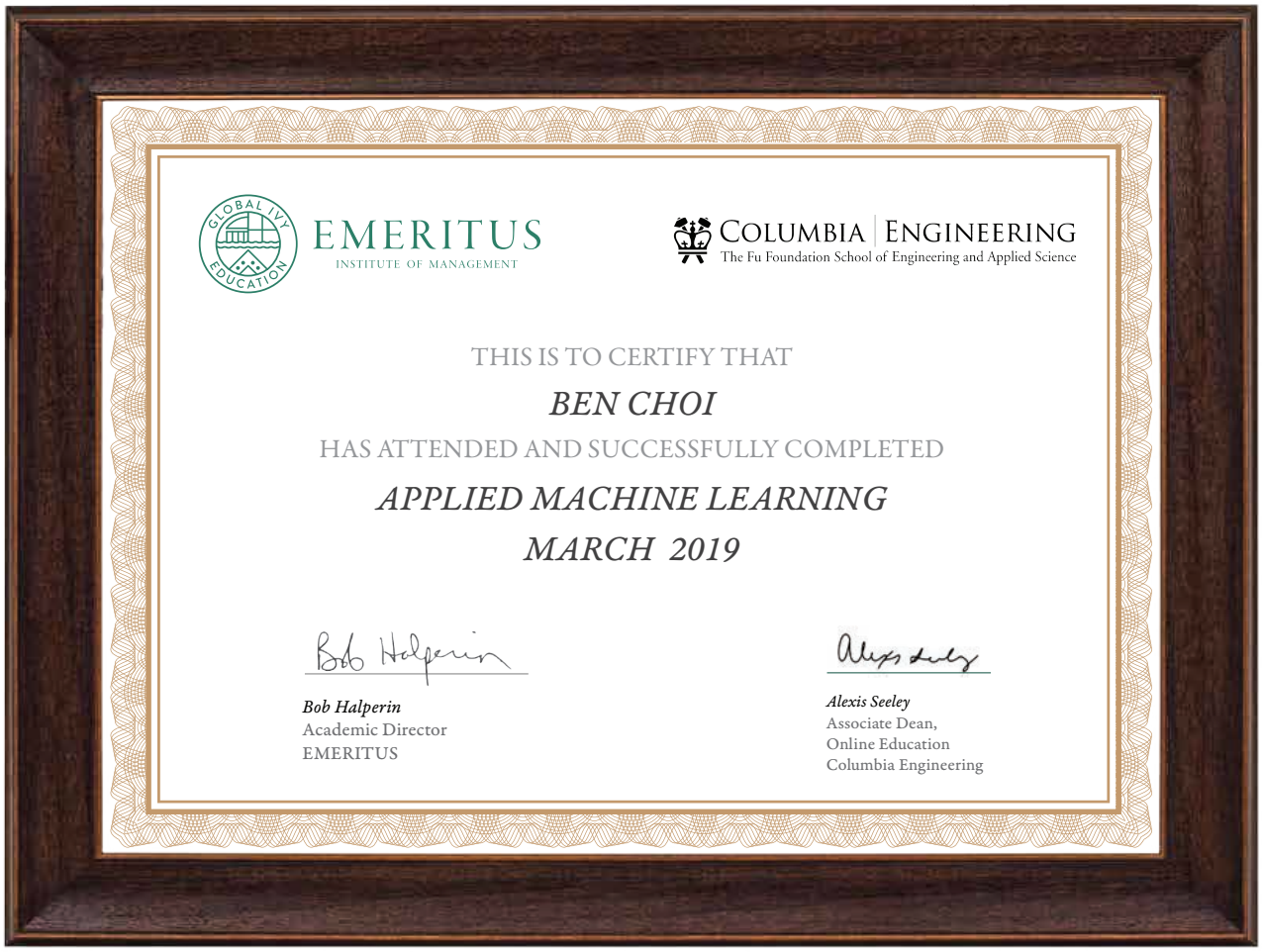
You will detect potential frauds using credit card transaction data. You will apply the random forest method to identify fraudulent transactions.

Market Segmentation

You will create market segments using the US Census dataset and by applying the k-means clustering method.

CERTIFICATE

Upon successful completion of the course, participants will receive a verified digital certificate from EMERITUS in collaboration with Columbia Engineering Executive Education.



COURSE FEE AND DURATION

$1,200 USD

Singapore residents who wish to enroll for this course will be charged GST.

3 Months, Online

PRE-REQUISITES

The course requires prior exposure to calculus, linear algebra and probability. Familiarity with a programming language is expected.

ABOUT EMERITUS

Columbia Engineering Executive Education is collaborating with online education provider EMERITUS Institute of Management to offer a portfolio of high-impact online courses. These courses leverage Columbia's thought leadership in management practice developed over years of research, teaching, and practice. By collaborating with EMERITUS, we are able to broaden access beyond our on-campus offerings in a collaborative and engaging format that stays true to the quality of Columbia. EMERITUS' approach to learning is formulated on a cohort-based design to maximize peer-to-peer sharing and includes live teaching with world-class faculty and hands-on project-based learning. In the last year, more than 10,000 students from over 120 countries have benefitted professionally from EMERITUS'

CONTACT EMERITUS

E-mail: columbiauniversity@emeritus.org
Phone: +1 (315) 888-4310