About this Course

Welcome to the **IBM AI Enterprise Workflow Certification course!**

*The six courses in this Coursera specialization will immerse you in the experience of working as a data scientist for an [over the top online media company](#) called **AAVAIL***.*

Quick background on AAVAIL

AAVAIL provides a streaming service similar to Netflix, Amazon Prime, and Disney+. Their value-add is a unique technology that allows them to offer local, national and international news to its subscribers in 12 languages. Subscribers can point the service at any video news feed. The news feed is piped through the service and both the speaker's voice and the movement of their lips are modified to match the language the subscriber has selected. There are separate deep-learning models for the image and audio portions of the service, but the experience is seamless from the standpoint of the user. They have been operating in stealth mode in the United States, with some initial success in urban markets.

Your role with AAVAIL

You've just been hired as part of a larger team to help AAVAIL deal with several business challenges. At the top of the list, AAVAIL's eccentric billionaire CEO, Josefina Echeverri, has made it clear that she wants to leverage her company's unique technology to expand aggressively into new markets worldwide.

Echeverri's team believes that AAVAIL's streaming service would be popular with expats wanting to watch the local news, so they have been experimenting with this concept in Singapore. The Singapore market demonstrated a strong initial adoption, with increasing revenue and traction from subscribers in the first 4 months. However, subscriptions and engagement have stagnated. The data show that after about 4 months there is a major decrease in watch-time and subscribers are quickly dropping off the platform. To deal with these challenges, Echeverri's management team discussed launching various marketing campaigns, modifying the pricing model, refining the product and more, with the goal of driving the product's growth in the new markets. Ultimately they wish to improve engagement and sales. Your first job is to help solve this problem and second, to drive a new phase of growth for AAVAIL.

Note: AAVAIL company and business idea were created for this course and any similarity in name or otherwise to an existing entity was unintended and purely coincidental.

Specialization Overview

This course is the first of six courses that make up the entire specialization for the IBM AI Enterprise Workflow Certification. These courses will guide you through the entire AI enterprise workflow, applying the workflow methods to help solve some of AAVAIL's business challenges:

- **Course 1 -- Business Priorities and Data Ingestion**: Covers the major starting points for a successful data scientist, scientific thinking, understanding the data, and identifying business opportunities.

- **Course 2 -- Data Analysis and Hypothesis Testing**: Covers exploratory data analysis and data visualization, hypothesis testing and how to handle missing data and outliers.

- **Course 3 -- Feature Engineering and Bias Detection**: Covers the next stage of the workflow, including dealing with class imbalances, dimension reduction, outlier detection, and using unsupervised learning. It also covers use of the IBM AI Fairness Toolkit.

- **Course 4 -- Machine Learning, Visual Recognition, and NLP**: Covers the evaluation of pipelines and different machine learning models, as well as the IBM Watson Visual Recognition service and IBM Watson Natural Language Understanding service.

- **Course 5 -- Enterprise Model Deployment**: Covers the deployment of machine learning models in a large enterprise, using Apache Spark and Docker containers. IBM Watson Machine Learning and IBM Watson Studio are also covered in this course.

- **Course 6 -- AI in Production**: Covers performance monitoring and unit testing, including the use of IBM Watson Openscale. It also introduces the capstone project for this specialization which involves the end-to-end execution of the AI enterprise workflow to deploy a model for a specific business challenge.

What makes this course different from other data science courses?

You will be focused on implementing AI solutions grounded in [design thinking](#). Design thinking is a process framework for managing the building of business solutions. It is IBM's preferred approach to managing complex AI deployments, and you will be introduced to design thinking at a high level throughout this course.

Many examples and case studies with a wide variety of different data sets and business problems are threaded throughout. You will build recommender and predictions systems, but

more importantly you will learn to solve problems in a systematic and comparative way. Data cleaning procedures, data transformations, and models will soon become an exchangeable set of building blocks for myriad possible solutions.

The narrative will also include mentorship and guidance. Sometimes a data scientist will start the process of solving a problem before taking the time to carefully define and investigate the subtleties of that problem. Experienced data scientists that have a process in place have learned the benefits of understanding the business problem before diving into the solution. That is only a small piece in the overall workflow and it is one most data scientists are already familiar with. This course will take you much deeper and present a process that will help guide you from business opportunity to production.

Most data science courses are focused on the details of the methods themselves. Obviously, it is important for you to thoroughly understand those methods. However this course is meant for practitioners who wish to go beyond the basics. We assume that you are a data science professional, and are looking to sharpen your skills with regard to real-world business challenges, to understand the *when, where and why* we choose to use the tools of data science.

Chief among the tools of data science is the [Python programming language](#). Python has become the *de facto* standard for data science in the industry, and these courses will make heavy use of the Python development environment and tooling.

[IBM Watson Studio](#) on the [IBM Cloud](#) will be introduced in these courses as a platform for collaborative work. Watson Studio is designed from the ground up to use Python-based tools, so you will not be required to learn a new framework or toolset in order to take advantage of the power and features of IBM Watson Studio.

Note: This course is specifically aligned with the contents of the IBM X professional certification provided by Pearson VUE.