Course Code: W7L156G

Course Title: Use IBM Watson APIs to Get Structured Data from Unstructured Text and Voice

Description:

In this course, the learner is guided through a realistic data science project. The project focuses on extracting business insights from large amounts of unstructured voice data, by using IBM Watson APIs. Starting from voice recorded customer reviews, the data are transcribed by using Watson Speech to Text and then are analyzed for emotion and keywords by using Watson Discovery. The structured results are stored in a tabular form and used to extract business insights in Watson Studio.

The course focuses on hands-on coding exercises that provide an opportunity to work with Watson Studio, Watson Discovery, and Watson Speech to Text. The exercises guide the learner on importing data and manipulating in a Watson Studio project, activate instances of installed services, and authenticate API connections by using Jupyter Notebooks. The exercises also introduce how to configure Watson Speech to Text and create custom language models, work with Watson Discovery, and introduce enrichments. Also, every exercise demonstrates suitable data manipulation that results in saved structured data.

The three hands-on exercises are included in this course, each focusing on a Jupyter Notebook and correspond to the three modules of the course. The necessary data, notebooks, and services are all available in the corresponding lab environment.

Objectives:

After completing this course, you should be able to:

- Utilize Watson Discovery and Watson Speech to Text API
- Outline the steps of creating custom model with Watson Speech to Text
- Dissect the process of enriching documents with Watson Discovery
- Design data schemas suitable for natural language processing
- Design a data science project based on product reviews

Prerequisites:

Before taking this course, you should have:

- Basic Linux knowledge
- Basic IBM Cloud Pak for data experience
- Basic knowledge of data manipulation
- Intermediate Python experience
- Working knowledge of Jupyter Notebook
- Basic knowledge of data science methods
- Basic knowledge of data visualization methods

Duration:

6.4 Hrs

Topics:

- Course Introduction
- Module 1. Speech to Text
- Exercise 1. Convert audio recordings to text - Minimize Word Error Rate
- Module 2. Text to Sentiment
- Exercise 2. Extract sentiment and keywords from text
- Module 3. Hidden Insights
- Exercise 3. Examine Discovery results to inform business decisions

Audience:

The course is primarily designed for data scientists, but this course would also be relevant to other enterprise roles that would like to utilize the powerful APIs of Watson Speech to Text and Watson Discovery to extract structured information from unstructured data.