Project Overview

**Instructions**

Now that you are equipped with the knowledge and skills to work with Spark Structured Streaming and Spark ML you have the opportunity in the final project to practice and apply your skills to create your own Apache Spark application as an end-to-end use case from data acquisition, transformation, and model training to deployment.

**Scenario**

In this scenario, you will import data from an external dataset and create a DataFrame. You'll save the data to a Parquet file and follow the steps to train the module using with Apache Spark. You'll complete the project by deploying the model to the IBM Watson Machine Learning (WML) Service. IBM WML is a scalable, scale-to-zero cloud service that supports training and serving of machine learning and deep learning models, providing a HTTP(S) endpoint for seamless consumption from third-party applications. Detailed instructions provide guidance throughout the project.

**Grading Criteria**

**There are a total of 20 points possible for this final project.**

Your final assignment will be graded by your peers who are also completing this assignment within the same session. Your grade will be based on the following tasks:

* Task 1: Import the component library. (1 pt)
* Task 2: Explore component library transformations. (1 pt)
* Task 3: Convert .csv to Parquet. (1 pt)
* Task 4: Perform model training. (1 pt)
* Task 5: Complete the model training. (1 pt)
* Task 6: Deploy the model to Watson Machine Learning. (1 pt)
* Task 7: Perform model inference. (2 pts)
* Task 8: HyperParameter Tuning. (3 pts)
* Task 9: Resample data splits. (3 pts)
* Task 10: RandomForest Classification. (6 pts)

**How to submit**

A screenshot in JPEG or PNG format is required to be submitted for most tasks. The screenshots will be uploaded in the submission step of the final project. You will be prompted to save screenshots, reports and files throughout the lab and these will be the files you submit during the Project Submission and Peer Review section of this course.