**IBM AI Engineer/Data Scientist Assessment**

Exercise: Healthcare Data Analysis with Watsonx LLM Models

**Introduction to the Use Case**

In this exercise, you will work on a healthcare data analysis project using Watsonx LLM models. The goal is to translate natural language questions/requests into JSON format requests that can be processed by a healthcare statistics API. These requests will be used to retrieve healthcare-related metrics such as patient satisfaction scores, daily medication administration rates, and treatment success rates.

Your task involves two main parts:

1. **Prompt Engineering**: Designing prompts to translate user questions into valid API requests.
2. **Evaluation Pipeline**: Creating a pipeline to evaluate the accuracy of the generated JSON requests.

**1. Prompt Engineering**

Your first task is to create a prompt that will effectively translate natural language questions into the correct JSON format required by the healthcare statistics API.

**Resources Provided:**

* **Examples File** (healthcare\_examples.csv): This file contains examples of natural language questions and their corresponding JSON requests. You can use these examples as references for constructing your prompt.
* **API Function Definition File** (Function\_file.json): This file contains the definition of the healthcare statistics API, including the available functions and parameters. You will need to ensure that your prompt generates JSON requests that conform to this functions.
* **Metrics Encoding File** (Metrics\_encoding.txt): This file lists the healthcare metrics that can be requested via the API, along with their corresponding identifiers.
* **Context File** (context.txt): This file provides context about the healthcare domain, including common information that may be relevant to generating accurate JSON requests.

**Task Description:**

1. **Understand the Inputs**: Review the provided examples, function definition, and metrics encoding. Understand how the natural language questions are translated into JSON requests.
2. **Create the Prompt**: Using the Watsonx LLM models provided in the notebook assessment connection (Assessment connection.ipynb), design a prompt that effectively translates user questions into the required JSON format. Your prompt should take into account the structure of the function and the metrics available.

**Deliverable:**

* **Prompt**: A well-crafted prompt that can be used with the Watsonx LLM models to generate JSON requests from natural language questions.
* **Healthcare\_examples**.csv with new generated answer column.

**2. Evaluation Pipeline**

Once you have produced the JSON requests using your prompt, your next task is to evaluate their accuracy by comparing them to the expected outputs provided in the examples file.

**Task Description:**

1. **Build the Evaluation Pipeline**: Create a pipeline that compares the JSON requests generated by your prompt to the expected outputs in the examples file (healthcare\_examples.csv).

You are free to define and use metrics that you believe are relevant for this comparison. These metrics could include standard metrics or custom metrics you defined.:

1. **Implement the Pipeline**: Implement your evaluation pipeline in a notebook. The notebook should systematically evaluate each example in the dataset and produce a CSV file with the evaluation scores for every metric.

**Deliverables:**

* **Notebook**: A Jupyter notebook containing the evaluation pipeline. This notebook should clearly document your process, including any assumptions or decisions you made when defining the evaluation metrics.
* **CSV File**: A CSV file that includes the evaluation scores for each example in the dataset (healthcare\_examples.csv).
* **Summary slide. (ppt)**

*In this step, The focus is on building a robust evaluation process, not necessarily achieving high scores.*

Good luck.

**Access to watsonx.ai for prompt lab if needed.**

IBM Cloud Login

<https://cloud.ibm.com/authorize/itzwatsonx9>

Username

user\_f0qxf

Password

4510eh40pdc1smg