

# Home Assignment: Building a Real-Time Predictive Model Pipeline

As a Machine Learning Engineer, your ability to design and develop real-time predictive models is crucial. This home assignment will evaluate your technical skills in building such a model.

Your task is to build a real-time streaming predictive model that estimates the likelihood of a customer making a purchase on an e-commerce website. Consider a dataset which includes customer browsing behavior, such as the pages they visited and their past purchases. The dataset will be provided in real-time through a stream processing system such as Apache Kafka, RabbitMQ, or any other message broker.

## Model Building:

- Train and save a basic Machine Learning model that predicts the probability of a user making a purchase, based on real-time data (e.g., the last three pages visited) and offline features (e.g., mean of the last three purchases).
- You can generate the data synthetically.
- **Model's accuracy is not important for this assignment.**

## Real-time Prediction:

- Develop a microservice system that retrieves data from a message broker (e.g., Apache Kafka, RabbitMQ).
- Extract features from the real-time data.
- Extract features from an offline store (e.g., MySQL, BigQuery, etc.).
- Implement a data processing pipeline to clean and preprocess the data.
- Make a prediction and return the estimated probability.

## Deliverables:

- A detailed design of the system architecture.
- A Python program that implements the above specifications.
- A report that describes the technical design of your solution, including any design choices you made and any difficulties you faced. Explain what you would do differently if you had more time.

## Note:

- **You can mock-up any external dependency (e.g., Kafka, MySQL).**
- You can use any **Python** libraries or frameworks that you believe are appropriate, but you should explain your decisions in the report.

Good Luck !!!

## **Legal Disclaimer**

This Moon Active Machine Learning test includes without limitation any all attachments and/or linked assets and by extension includes any written or oral response or presentation provided by a candidate to such test ("Test"). The Test and all other materials, communications, information and documentations, in any form, disclosed by Moon Active Ltd. ("Company ") and/or anyone

on its behalf to you, or that is otherwise learned by you in the course of conducting the Test (including, without limitation, information regarding the Company, and/o any instructions provided by Company and/or other data relating to the Company's mobile games) ("Confidential Information") shall be deemed confidential and you shall not disclose or use it for any other purpose other than to conduct the Test, and shall take reasonable measures to prevent disclosure or use of such Confidential Information, and is intended exclusively for you. You hereby represent and warrant that you shall not use Test in any publicity or promotional or marketing publication or personal portfolio. Company is and shall at all times remain the sole and exclusive owner of the Test, and any part thereof (including, without limitation, any content, creatives and all intellectual property rights in the foregoing and in any enhancements, modifications, updates or derivatives thereof and related know-how and any and all logos, trade names, trademarks and service marks, whether or not registered). All rights which are not expressly granted herein are reserved by Company.