Peyton Bailey **July 7th, 2025**

D602 – Deployment (Task 1)

A. Describe the objectives of an MLOps deployment architecture.

MLOps (Machine Learning Operations) deployment architecture aims to streamline and automate the lifecycle of machine learning models from development and training to deployment and maintenance (Salama et al., 2021). Kronkers is a regional supermarket chain that needs an MLOps deployment architecture to automate, standardize, and scale its machine learning workflows, which will enable faster, more reliable model development and deployment. This can be done by automating data ingestion, preprocessing, model training, validation, and deployment. The MLOps architecture strives to facilitate collaboration between teams and departments by standardizing tools and processes. Another goal of MLOps is continuously monitoring model performance by detecting drift and triggering alerts when necessary. When properly implemented, MLOps will achieve these goals for Kronkers and ensure security and compliance by protecting sensitive data and ultimately facilitating the deployment of models in any environment, including on-site or in the cloud. (Salama et al., 2021).

B. Identify constraints to implementing an MLOps solution.

Implementing MLOps within a business structure can face multiple challenges due to several constraints. For example, there can be a technical skill gap in which many key staff members lack expertise in MLOPs practices and tools. There could also be data management issues that have led to inconsistent data quality, resulting in difficulties integrating data from various sources. Furthermore, a lack of standardization could be an issue when no standard tools or processes exist for model development. A typical example could be when some data personnel utilized Python language while others used R. Finally, the business organization could have an overall resistance to change and be skeptical about implementing a new process.

C. Identify all functional and non-functional requirements for the MLOps solution.

Functional Requirements

Model Development and Experimentation – Support for building, testing, and validating ML models in multiple languages such as Python, R, and Julia (Curry, 2020).

Data Management – Ability to ingest, clean, and version datasets up to 500 MB in size with historical records tracked and used for training models (Sumo Logic, n.d).

Model Registry – Store and manage versions of models, code, and data by maintaining model documentation and reporting. (Salama et al., 2021).

Model Deployment and Serving – Enable deployment of models as APIs accessible by various departments.

Model Monitoring - Continuously monitor model performance, detect drift, and trigger alerts and/or retraining if necessary.

Model and Data Governance - Ensure compliance with internal and external policies and protect data insights as intellectual property and trade secrets (Salama et al., 2021).

Non-Functional Requirements

Scalability – System must handle increasing data volumes and user demand as the company expands (Smith, 2020).

Reliability- System should perform the correct function at the desired performance level even during hardware or software failure and human error (Smith, 2020).

Maintainability – The system should be easy to update, monitor, and troubleshoot.

Adaptability – The system should be capable of discovering aspects to improve its performance and allowing updates without interruption (Smith, 2020).

REFERENCES

Curry, D. (2020). *Forrester: 7 Key Requirements For Successful MLOps Deployment*. https://www.rtinsights.com/key-requirement-mlops/

Salama, K., Kazmierczak, J., & Schut, D. (2021). *Practitioners' guide to MLOps: A framework for continuous delivery and automation of machine learning*. https://services.google.com/fh/files/misc/practitioners_guide_to_mlops_whitepaper.pdf

Smith, J. A. (2020). *Designing machine learning systems. O'Reilly Media, Incorporated.* *https://ebookcentral.proquest.com/lib/westerngovernors-ebooks/reader.action?c=UERG&docID=6989361&ppg=51*

Sumo Logic. *(n.d.). What is MLOps?* *https://www.sumologic.com/glossary/mlops*