Assignment 1: MLOps fundamentals - Understanding Principles and Team Roles

Satvir Kaur Mehra

Machine Learning Diploma Program

Assignment 1

CMPT 2500

Mario Soriano Morales

February 07, 2025

Part 1: Understanding MLOps Principles

MLOps (Machine Learning Operations) integrates software engineering practices to improve the development, deployment, and maintenance of machine learning models. It emphasizes automation, reliability, and collaboration to ensure efficient, scalable ML systems. Key principles include continuous integration, model monitoring, retraining, and team collaboration.

1. Continuous Integration and Delivery in ML

Continuous integration and delivery (CI/CD) in MLOps ensure that models are updated seamlessly with minimal disruptions. This involves automating the ML pipeline, including data processing, model training, validation, and deployment. By implementing version control for data, code, and models, organizations can efficiently manage model updates and prevent inconsistencies. Automated testing and validation help maintain performance standards before deployment, reducing errors and enhancing reliability.

2. Model Monitoring and Retraining

Model monitoring is essential for detecting performance degradation due to data drift or changes in real-world conditions. Continuous tracking mechanisms assess model accuracy, input data quality, and operational efficiency. When significant deviations are detected, retraining is triggered to update the model with new data, ensuring its relevance and effectiveness. Automating this process helps maintain long-term model performance and minimizes risks associated with outdated predictions.

3. Collaboration Between Teams

MLOps promotes collaboration among data scientists, ML engineers, software developers, and IT operations teams. Establishing standardized workflows, shared repositories, and automated pipelines helps streamline communication and coordination. By fostering a structured approach to model development and deployment, MLOps enables teams to work efficiently, ensuring that ML models are well-maintained, scalable, and aligned with business objectives.

Part 2: Comparison Between MLOps and DevOps

	ML Ops	**DevOps**
Tools Used	TensorFlow, MLflow, Kubeflow, Airflow, TFX	Jenkins, Docker, Kubernetes, Ansible, Terraform
Handling of Data & Models	Manages large datasets, version control for data and models, continuous model training & validation	Focuses on code versioning and application dependencies, but does not require continuous data updates
Deployment Workflows	Involves model training, retraining, validation, and continuous monitoring before deployment	Focuses on CI/CD pipelines to build, test, and deploy applications with minimal human intervention

Challenges in Production	Model drift, data bias, reproducibility issues, need for continuous retraining, explainability concerns	Infrastructure scaling, software bugs, security vulnerabilities, and CI/CD failures

Part 3: Analyzing Team Roles in MLOps

In an MLOps workflow, Data Scientists and Machine Learning (ML) Engineers have different **responsibilities** and they have complementary jobs to make sure ML models work well and stay updated. Data Scientists mainly work with data and cleaning it, choosing useful features, and building models. They focus on analyzing data, picking the right machine learning techniques, and training models to make accurate predictions. They use programming languages like Python and R, along with tools such as TensorFlow and Scikit-learn.

The **tools** and technical skills used by Data Scientists and ML Engineers differ based on their roles. They ensure models can handle large-scale operations, integrate into existing systems, and run efficiently. Their job includes deploying models, improving performance, and setting up monitoring systems. They work with cloud computing, automation tools like Docker and Kubernetes, and CI/CD pipelines for continuous updates.

Good **collaboration** between these roles is essential for successfully managing ML projects. By combining their skills, they ensure models stay accurate, efficient, and useful, leading to better AI-powered solutions. Both roles must work together to solve issues like data drift, where new data differs from the training data, making models less accurate. Data Scientists

identify these issues, while ML Engineers set up automatic updates and monitoring systems to keep models working correctly.

Part 4: Explore MLOps Tools – MLflow

MLflow is an open-source MLOps tool that streamlines the management, tracking, and deployment of machine learning models throughout their lifecycle.

- **Experiment Tracking:** MLflow enables the tracking of experiments by recording parameters, metrics, and artifacts throughout the ML pipeline. This feature helps to manage and compare multiple models, ensuring efficient experimentation and reproducibility, which is a core principle of MLOps.
- **Model Versioning:** MLflow supports versioning of machine learning models, which ensures that data scientists and engineers can track different model iterations, improve their performance, and safely deploy models without confusion.
- **Pipeline Management:** MLflow integrates with various tools to streamline the management of machine learning pipelines, allowing for automation, consistent workflows, and monitoring. This feature supports the MLOps principle of efficient, automated model deployment and management.
- **Model Deployment and Monitoring:** MLflow facilitates easy deployment of models into production environments and allows continuous monitoring of model performance. This supports MLOps principles by ensuring that models in production are well-maintained, and their performance is tracked over time.

References

DataCamp**.** (2025, January 21). *MLOps vs DevOps: Differences, overlaps, and use cases*. DataCamp. https://www.datacamp.com/blog/mlops-vs-devops

Eken, B., Pallewatta, S., Tran, N. K., Tosun, A., & Babar, M. A. (2024). *A multivocal review of MLOps practices, challenges, and open issues. arXiv preprint arXiv:2406.09737*. https://arxiv.org/abs/2406.09737

Kreuzberger, D., Kühl, N., & Hirschl, S. (2023). Machine learning operations (MLOps): Overview, definition, and architecture. IEEE Access, 11, 35275–35291. https://ieeexplore.ieee.org/abstract/document/10081336

Spadari, V., Cerasuolo, F., Bovenzi, G., & Pescapè, A. (2023). An MLOps framework for explainable network intrusion detection with MLflow. *Proceedings of the 2023 IEEE/ACM 1st International Conference on MLOps (ICMLoPs)*, 1-9. https://www.researchgate.net/profile/Francesco-Cerasuolo-2/publication/380759913_An_MLOps_Framework_for_Explainable_Network_Intrusion_Detection_with_MLflow/links/664da6de479366623a05ce09/An-MLOps-Framework-for-Explainable-Network-Intrusion-Detection-with-MLflow.pdf

Warnett, S. J., & Zdun, U. (2024). On the understandability of MLOps system architectures. *IEEE Transactions on Software Engineering, 50*(5), 1015–1031. https://ieeexplore.ieee.org/abstract/document/10440483