MLOps Implementation in Azure

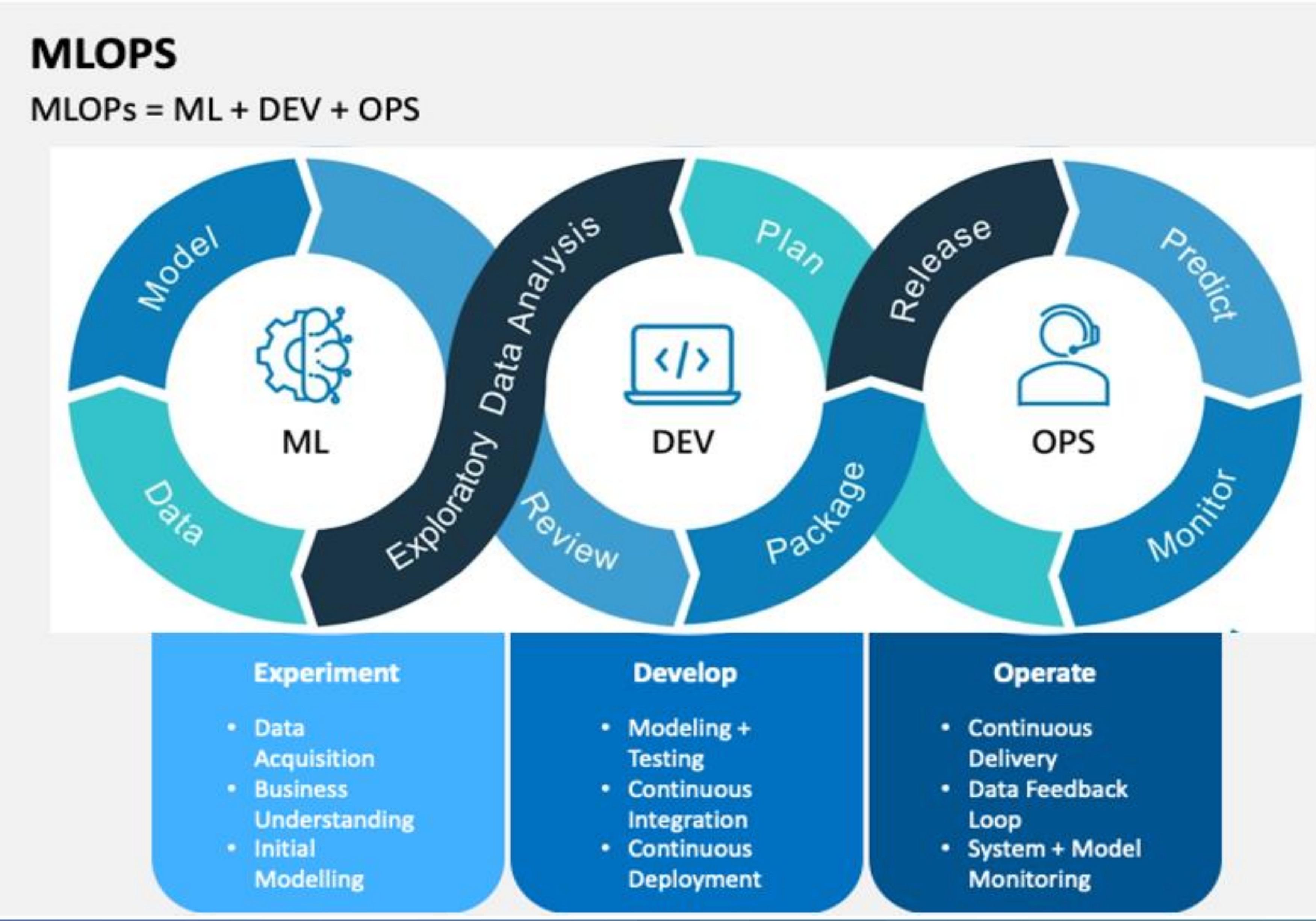
Insurance cross-selling use case

Patricia Inyang, MSc Data Analytics

Aston University

Introduction

An estimated 85% of artificial intelligence (AI) and machine learning (ML) projects fail to produce return for businesses. This is due to the intricate processes, infrastructures and cross-functional discipline necessary to deploy and maintain optimal ML models in real-world production environments. This project addresses this challenge by implementing Machine Learning Operations (MLOps) which applies DevOps practices to automate the end-to-end machine learning cycle within Azure Machine Learning (AML). The use case involves the development of a cross-selling Insurance model to predict existing health insurance policyholders responses ('Interested' or 'Not interested') to newly offered vehicle insurance policy.



Project Objectives

To successfully Implement MLOps in Azure, advancing through the MLOps maturity levels:

- **Close the business value gap in ML models**

Transit from development to deployment of the cross-selling ML model

- **Surmount MLOps Implementation Challenges**

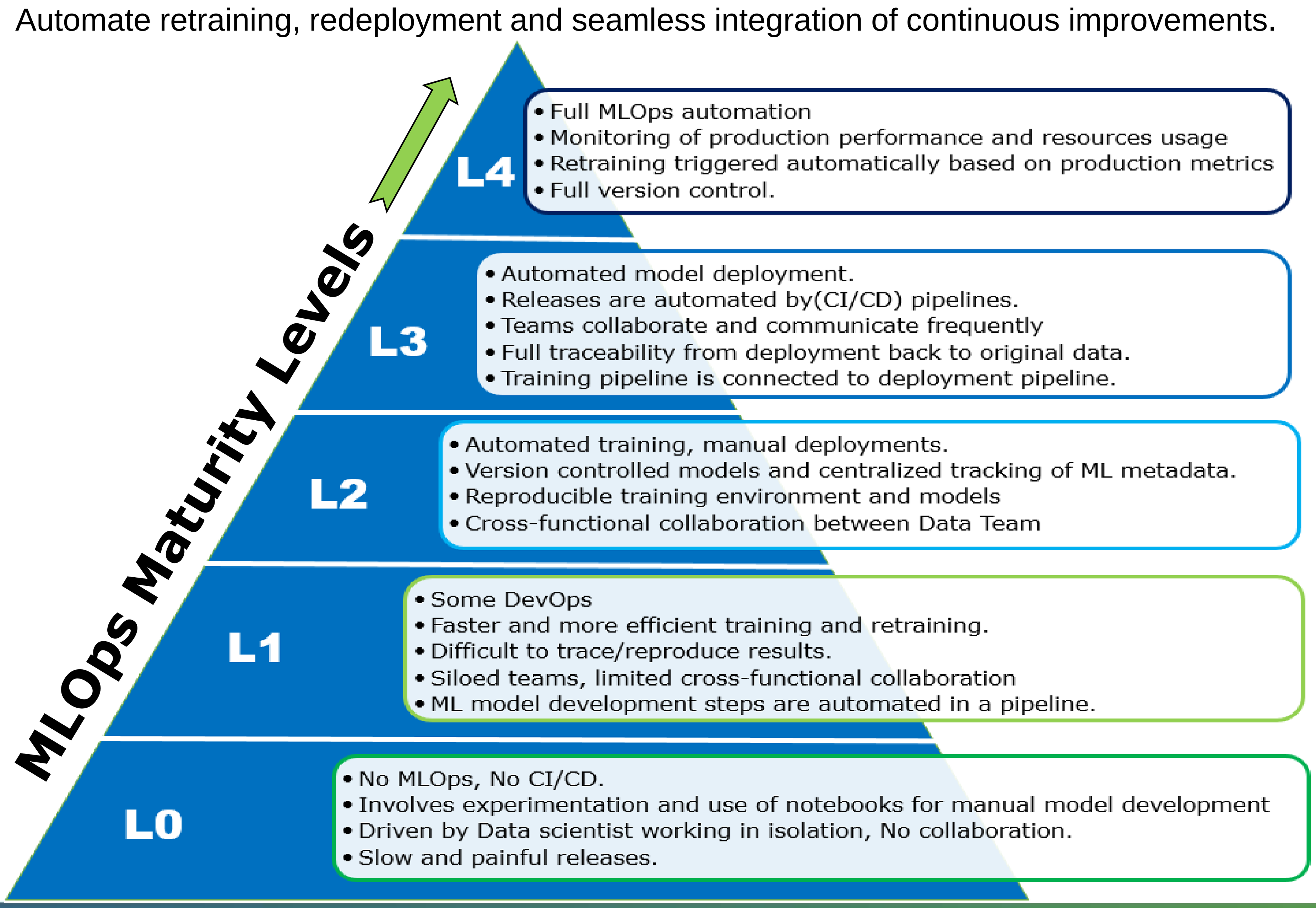
Infrastructure, expertise, reproducibility, scalability.

- **Governance Considerations**

Responsible AI solution, interpretability and accountability throughout the model's lifecycle.

- **Sustain Optimal Model Performance**

Automate retraining, redeployment and seamless integration of continuous improvements.



Methodology

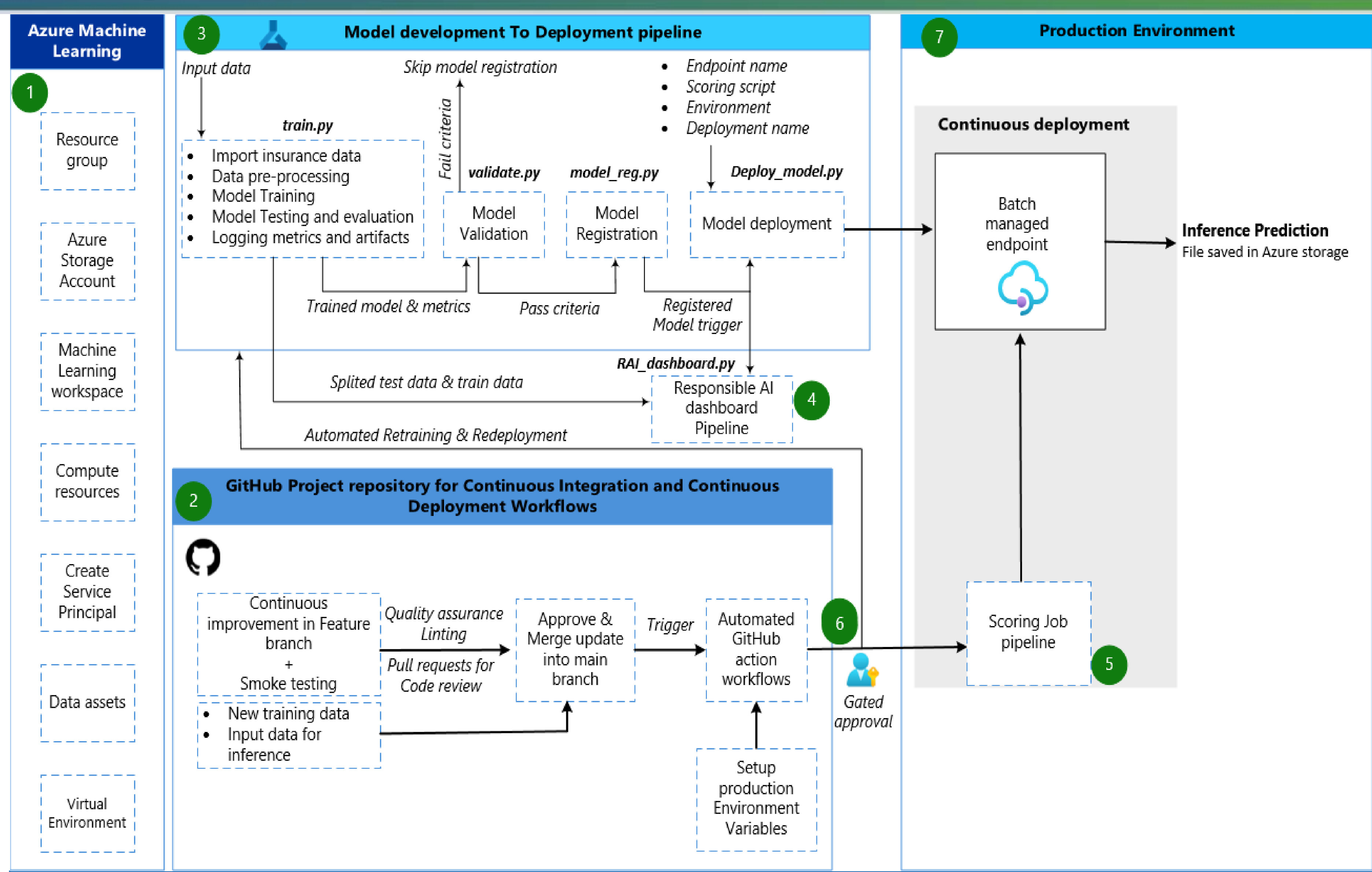
The project was done in the four stages below from experimentation to end-to-end automation. Much of the coding utilized Azure Machine Learning(AML) Python SDK v2. For terminal command execution, AML command line interface (CLI) was effective and the AML Studio served as user interface central hub for viewing outputs, logs, and jobs.

Experimentation	Exploratory data analysis(EDA) Data preparation	Training and evaluation of ML classification models Select model with highest recall
Infrastructure Configuration	Setup AML workspace and AML Resources	Setup GitHub repository and Continuous Integration pipeline
Prepare scripts & build pipelines	Refactoring codes in scripts Test scripts	Build pipelines to connect ML cycle steps Test pipelines
Automating the MLOps	Set up production environment with Azure service principal credentials	Automation with GitHub action workflows for pipeline runs in production & CI/CD

The High level diagram shows the setup and flow of the end-to-end MLOps designed:

1. Available resources and managed infrastructures in AML that were configured.
2. CI/CD setup in GitHub repository to facilitate collaboration and update integration.
3. 1st pipeline automates training to deployment of the cross-selling model.
4. 2nd pipeline generates responsible AI insight dashboard for newly registered models.
5. 3rd pipeline is the inference pipeline which submits scoring job to the deployed model.
6. GitHub action workflows automatically triggers the pipelines to run in production.
7. Production environment where deployed model(s) make inference on new data.

Implementation

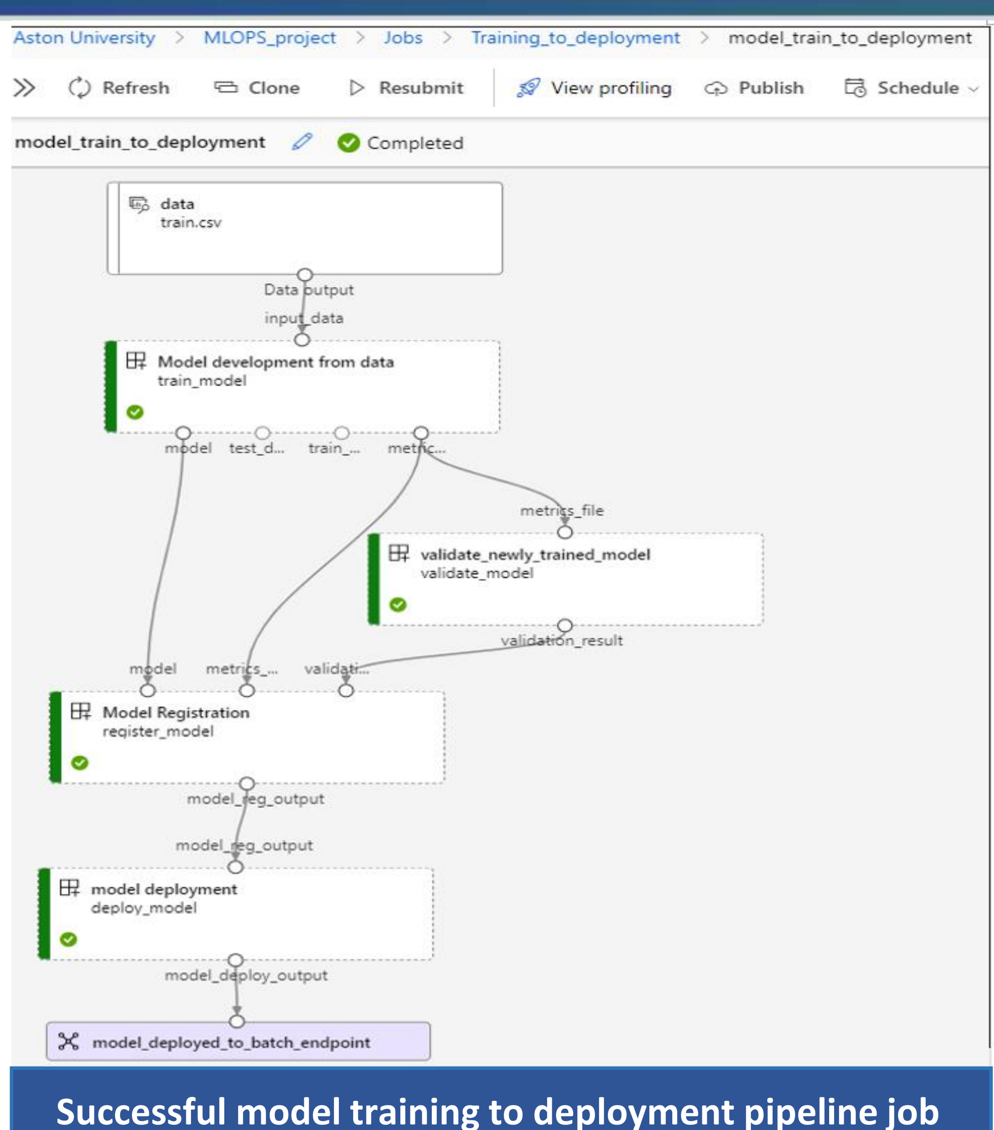


High level Flow diagram of MLOps implemented in Azure for Insurance cross-selling model

Results

The primary evaluation was recall, with the initial model deployed achieving a 85.9% recall. Executing MLOps with Azure ensures easy of reproducibility, scalability, and deployment of continuous improvement in unattended situation. Valuable features in the MLOps orchestrated with AML include:

- Automation through CT/CI/CD.
- Managed infrastructures, services, and scalable resources.
- Experimental tracking, logging and traceability.
- AML easily Integrates with MLFlow, GitHub, and other Azure services.
- User-friendly AML Workspace.
- Reliable and secure via credential authentication and RBAC.
- Interactive RAI insight dashboard.
- Versioning of models, data & codes.
- Supports reusable components.



Successful model training to deployment pipeline job

Conclusions

Considering that the primary focus of this project was not to create a high-performing model but to design and implement MLOps using Azure Machine Learning services, the research successfully achieved its goals. Comparing the MLOps implemented in this project with the MLOps maturity level illustrated earlier, shows that a 3rd maturity level was attained. Despite this success, it is important to note that this is not a fully matured MLOps. It is yet to be continuously exposed to real-life data and ongoing monitoring of production workloads to detect issues like data and model drifts which eventually provides a feedback for retraining. Automated retraining now occurs when new data or code improvements are pushed to the main GitHub branch

Leveraging this cross-selling model which is kept up-to date in production and automatically predicts the interest of costumers for new inference data, Insurers are empowered to make sustainable decisions on sales strategy which enhances competitive advantage, maximise customer relationship, cost efficiency, and customer retention. This confirms that DevOps is key for accelerating delivery of business value in data science projects and early deployment of ML models. Finally maintaining system is crucial to keep it up-to-date with Azure latest roll-outs .

Recommendations for future works

- Explore Azure DevOps and Azure event Grid for automation.
- Develop a self-orchestrated Bring Your Own (BYOD) container.
- Use of Serverless compute in Azure
- Write more comprehensive automated tests to increase robustness of the system.

References

[1] https://www.datarobot.com/wp-content/uploads/2023/08/DataRobot_MLOps_101-_The_Foundation_for_Your_AI_Strategy_eBook.pdf

[2] Raj, E. (2021): Engineering MLOps: Rapidly build, test, and manage production-ready machine learning life cycles at scale, Packt Publishing Ltd, ISBN 978-1-80056-288-2 pg 12-13, 18-243.

[3] Zeitgeist (2023): AI Readiness Report. Available at: https://go.scale.com/hubfs/Scale-Zeitgeist-AI-Readiness-Report-2023.pdf

[4] Microsoft (2023): MLOps 101 - Microsoft Solutions Playbook. Available at:https://playbook.microsoft.com/code-with-mlops/technology-guidance/mlops/#standardize-on-repeatable-architectural-patterns

Patricia.ekpat@yahoo.com

@Patricia-Inyang

Scan the QR code to access the project GitHub repository containing the project files and full report.