Reaching for the MLOps level 1

Understanding and implementing MLOps into practice
The scientific community, as well as statistical community, focuses on the development of ML models, not in producing ready-made machine learning products.

That is the reason why so many ML projects fail (based on research).

Focus should be on building ML products, not only on developing and building ML models.
MLOps

- Collection of principles and components that implement those principles
  - Sometimes MLOps as a concept can seem a little “fuzzy” and everyone has a slightly different understanding and description of it.

- Why do we need it?
  - That we could implement necessary quality dimensions (or responsible ML?)
  - Kind of a “horror picture” is that statistics would be based on predictions produced by machine learning models, about which we know nothing

- MLOps includes also a standardized process (or workflow as they call it) and roles

- ML model is just one component among many other components
  - Actually, model is not a component in MLOps... (it’s one the main artifacts)
Principles

- Versioning
  - Data, code and model
- Automation
- Reproducibility
- Monitoring
- Deployment (CI/CD)
Reproducibility

• Every phase (data processing, model training, model deployment) will produce identical results with the same input
• In other words: you can repeatedly run your algorithm on certain datasets and obtain the same results
• Reproducibility is a major principle of the scientific method (predictions are produced by scientific methods)
• The lack of reproducibility (and transparency) undermines predictions scientific value
• Reproducibility creates trust and credibility with the ML product
• Versioning is one of the key components behind reproducibility
Automation

• The level of automation (data, model, code pipelines) determines the maturity of the ML process
• From manual process (all the steps are executed manually) to full automation (three levels, from manual data science process to full automated ML process)
• Three “official” levels:
  • Manual process (Google Cloud’s level 0). Every step in ML process is performed manually.
  • ML pipeline automation (Google’s Level 1). Continuous Training and data validation/model monitoring.
  • CI/CD pipeline automation (Google’s Level 2). Building, testing and deploying data and model automatically.
1. Can ML solve the business problem? (Tool: ML Canvas)

2. Understand the case and data Tool: What do you want to do? (from Azure)

3. Collecting training data

4. Analyzing data

5. Cleaning the data and splitting data

6. Training the model

7. Evaluating the model (creating the metrics)

8. Validating the model

9. Deploying the model into production environment

10. Monitoring the model
Roles

Subject matter expert with "data science –twist" (understands data science methods and R/Python -programming)

ML engineer

"Strategist"

Data scientist

Governance person

Coming soon: Citizen data scientist (understands data science)
...or it can look like this!

Figure 3. Roles and their intersections contributing to the MLOps paradigm
Monitoring

• Monitoring accuracy/quality of models prediction (also the quality of training data should be monitored)
• In fully automated MLOps workflow monitor triggers the CT (continuous training)
• But, when the process is not fully automated the monitor can give an alarm to a governance person
• There are open source –libraries and off-the-shelf –products (I also presume, that in some off-the-shelf –ML platforms, monitor components are implemented into the platform)
Components

- Technical components implement MLOps principles
  - Source control/source code repository
  - Feature store
  - CI/CD component
  - Model serving
  - Model registry
  - Monitoring component
  - Pipeline/workflow orchestrator
  - Metadata store
MLOps platform?

• Definition of MLOps platform is even “fuzzier” than the definition of MLOps
  • Sometimes difficult to describe the 1:1 connections between ML platforms and principles of MLOps –theory
• ML Platforms can be more like infrastructure platforms or data management or business intelligence platforms (example: Pyramid Analytics) or...
• Off-the-shelf machine learning platform versus in-house solution, using open-source tools and libraries on existing cloud service provides platform
• Some platforms go more to Citizen data scientist –direction!
An example: MLOps platform (or Stack!) from Valohai
MLOps Stack?

MLOps Stack Canvas, Online Whiteboard for Visual Collaboration (miro.com)
Figure 4. End-to-end MLOps architecture and workflow with functional components and roles
Responsibile ML (Statistics Canada) and MLOps

Framework for Responsible Machine Learning Processes at Statistics Canada

- Respect for People
  - Value to Canadians
  - Prevention of harm
  - Fairness
  - Accountability

- Respect for Data
  - Privacy
  - Security
  - Confidentiality

- Sound Application
  - Transparency
  - Reproducibility of process and results

- Sound Methods
  - Quality learning data
  - Valid inference
  - Rigorous modeling
  - Explainability

Assessed through self-evaluation and peer review, using a checklist and producing a report or dashboard
What we have done at Statistics Finland?

- We have built a ML platform (in-house)
- Deployed two ML models into production at that platform
- We have also studied the theory of MLOps...
- We are planning to use Model Cards as Model’s metadata description (improving the transparency and explainability)
- From MLOps (automation) maturity point of view, we are still at the level 0...

Model Card Toolkit  |  Responsible AI Toolkit  |  TensorFlow
GitHub - tensorflow/model-card-toolkit: a tool that leverages rich metadata and lineage information in MLMD to build a model card
We use Data Factory -instance to move data between cloud and on-prem environments and launch prediction runs.

Workspace for developing machine learning models

Serves as a data storage location

Our “in-house” platform
MLOps principles at Statistics Finland

- Versioning
  - models are versioned in model registry
  - code (no versioning)
  - data (not versioned although it’s possible with Delta Lake)
- Automation
  - now at level 0 (+)
- Reproducibility (missing parts...)
- Monitoring (we are looking for a solution...)
- Deployment (CI/CD) (when predictions are requested, the register is automatically searched, but nothing else...)
- Many “development lines“:
  - Implementing ML to statistical business processes (automating manual classification)
  - The more machine learning solutions there are in production, more principles must be implemented into our MLOps platform
Future (global development)

• Citizen data scientist
  • Is part of the low code/no-code/democratization – development
  • All the big consulting companies (Gartner, Deloitte etc., see it coming)

• Let’s be ready for this development as well, but first MLOps to implement features of Responsible ML.