ML Lifecycle and productionization

Milen Chechev 11/07/2021





- ~10 години опит като ML Engineer/ML/DS Lead
- ~10 години опит като Software Engineer

Образование:

- Докторантура "Препоръчващи системи", Софийски университет
- Специализация "Machine Learning" в Aalto University

Преподавателски опит в ФМИ, СУ(~10 години):

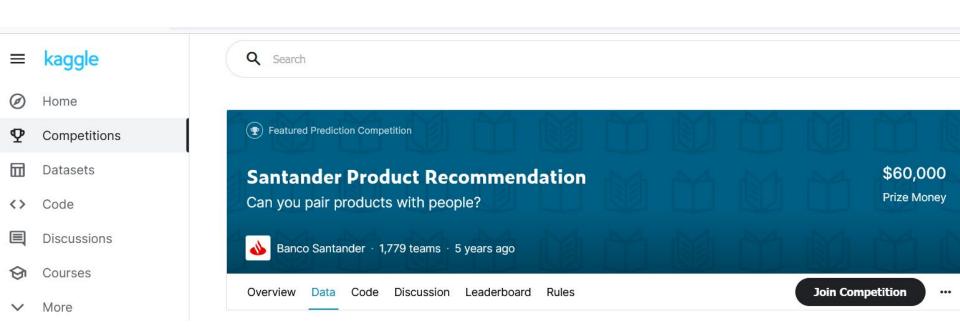
- Препоръчващи системи
- Машинно обучение
- Изкуствен Интелект
- Структури от данни и алгоритми

LinkedIn profile

Agenda

- ML Development and Evaluation
- AB testing
- Data shift and drift
- Data and Model versioning
- ML Ops
- From Jupyter to ML pipelines and production
- Summary

ML Development and Evaluation



Submissions are evaluated according to the Mean Average Precision @ 7 MAP@7:

$$MAP@7 = rac{1}{|U|} \sum_{u=1}^{|U|} rac{1}{min(m,7)} \sum_{k=1}^{min(n,7)} P(k)$$

Is this evaluation good?

Pros:

- It's giving a possibility to compare algorithms on the dataset.
- It's easy to calculate and use

Cons:

- It penalize an algorithm which return different items
- It's selecting recommender algorithm which is similar to the previous one used for collecting the dataset

How we could test it then?

Online tests - AB Tests

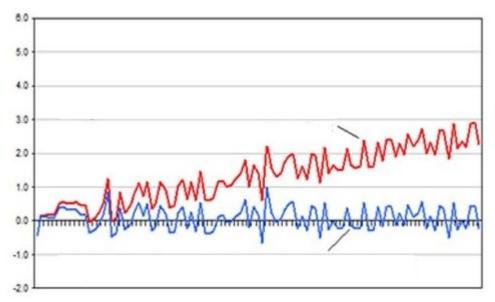
Pros:

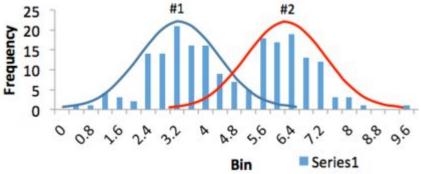
We are going to see the real performance of the model

Cons:

- Expensive
- Slow (when to stop to online test?)

Data Shift and Drift



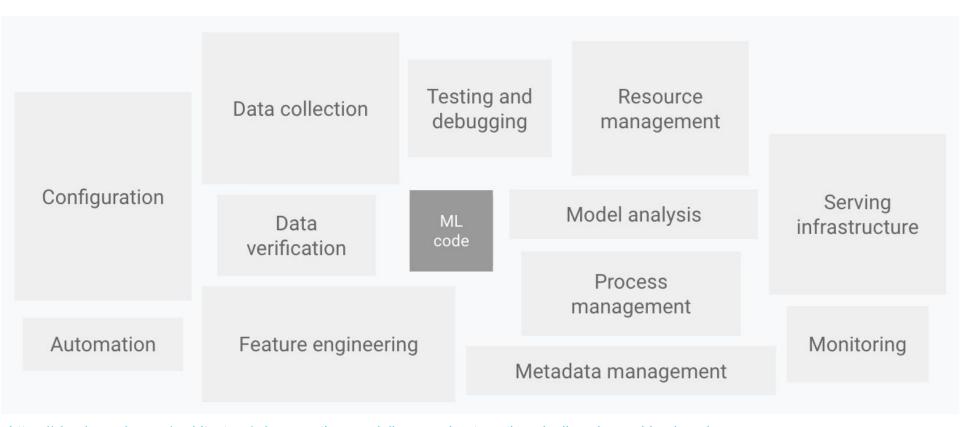


Reproducibility and Traceability

When working in constantly changing environment in order to be able to follow all the experiments and models and the conclusions from them to still be useful and valid we need:

- Data Versioning
- Model Versioning

From research to production



ML Lifecycle

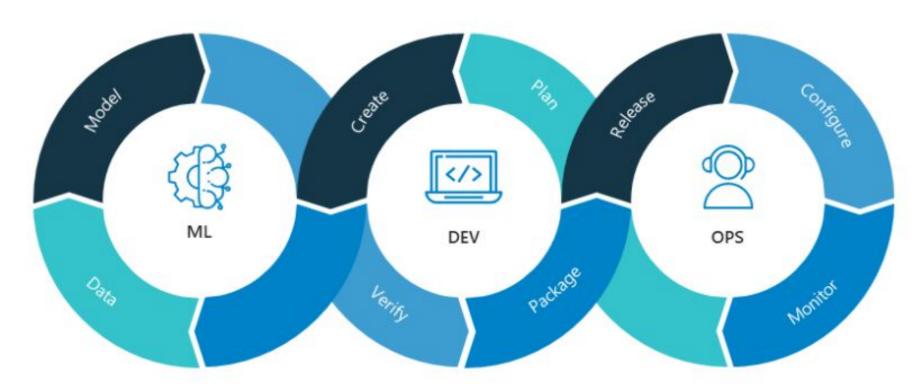
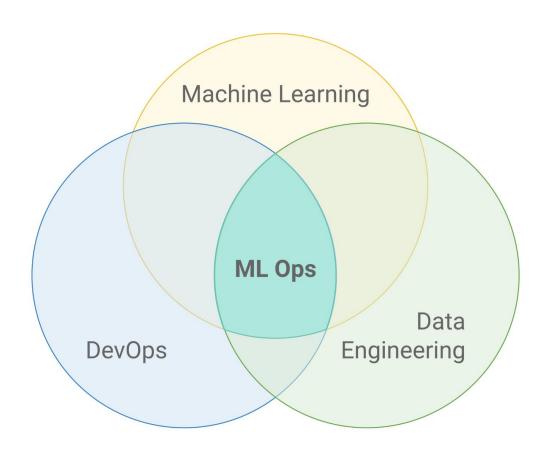
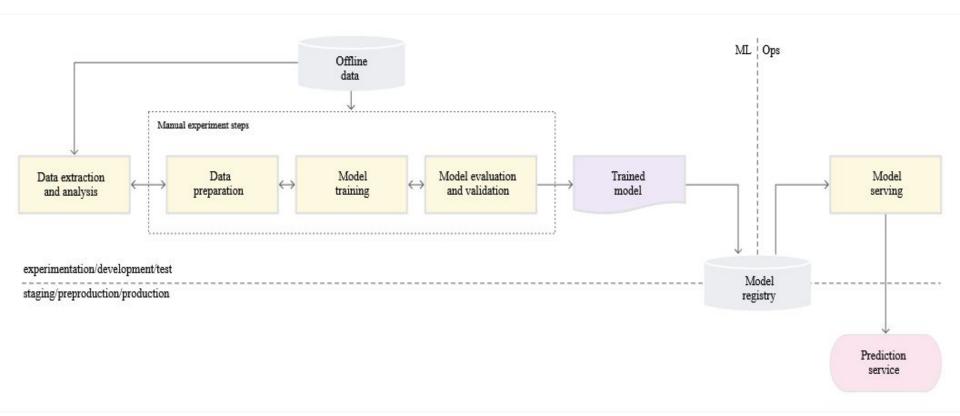


Image from: https://blogs.nvidia.com/blog/2020/09/03/what-is-mlops/

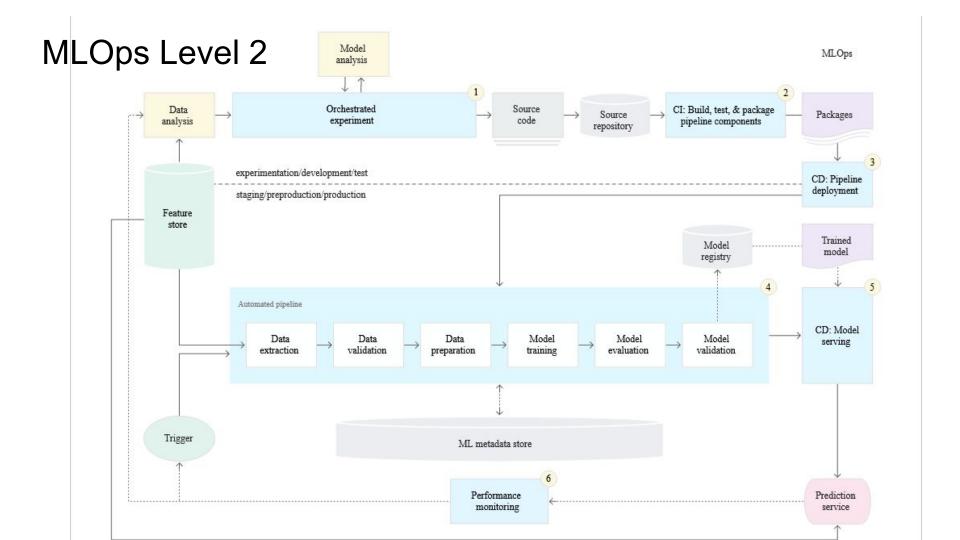
ML Ops



MLOps level 0: Manual process



MLOps - Level 1 Model analysis Orchestrated experiment Pipeline deployment Model Model Model Data Data Data Source Source validation analysis preparation training evaluation validation code repository Offline extract experimentation/development/test staging/preproduction/production Feature store Trained Model model registry Batch fetching Automated pipeline CD: Model Data Data Data Model Model Model serving validation preparation evaluation validation extraction training Trigger ML metadata store Performance Prediction monitoring service



Demo

Azure ML

