Machine Learning Development Life Cycle

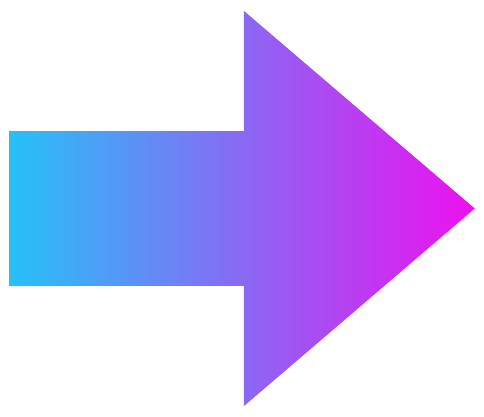
Model Training

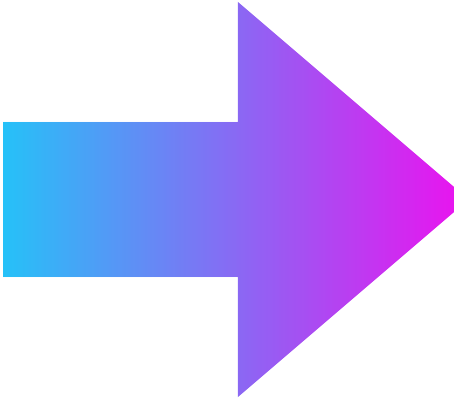
Data Preparation

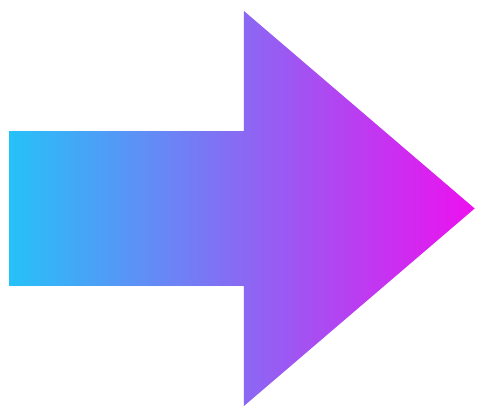
Evaluation

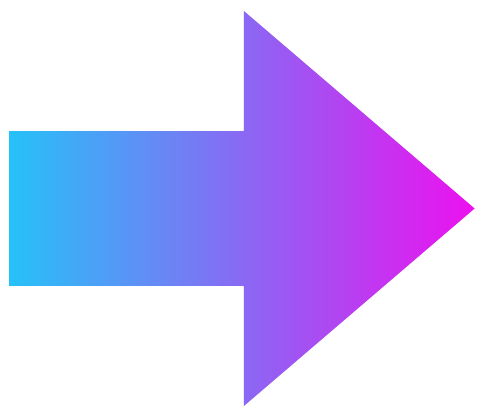
Deployment

Data Collection









jupyter

SQL

APACHE
Spark

FEAST

TensorFlow

PyTorch

scikit
learn

dmlc
XGBoost

LightGBM

NumPy

APACHE
Spark

APACHE
Spark

TensorBoard

scikit
learn

Feature Engineering

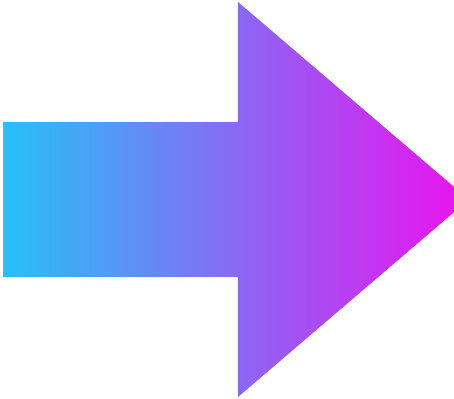
Containers at scale

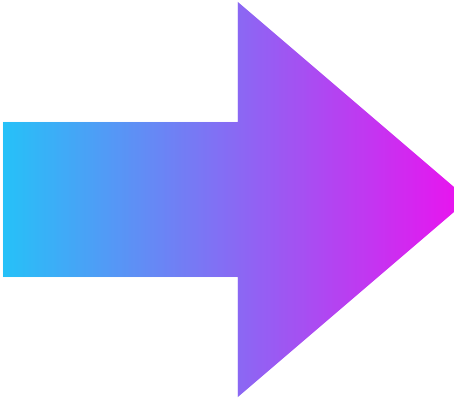
KServe

Data Engineer

Data Scientist

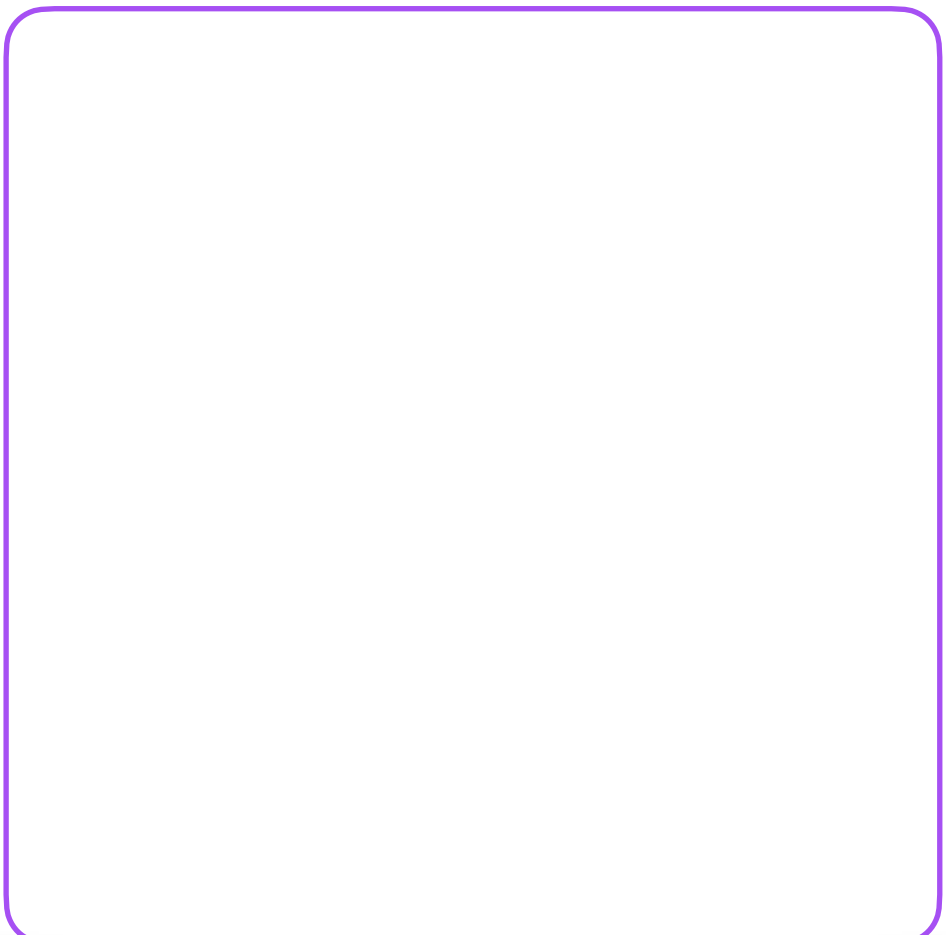
ML Engineer





4 common problems with scaling ML

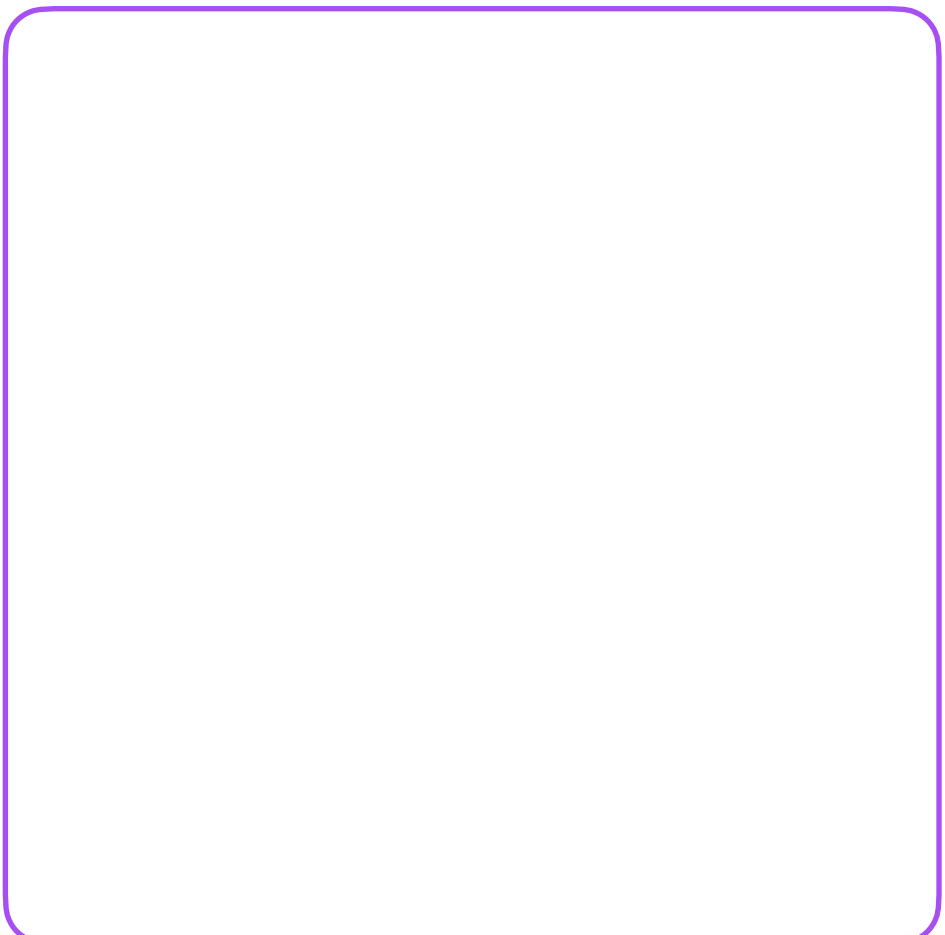
Why data science team needs MLOps?



- Mundane steps with lots of

boilerplate

Getting started

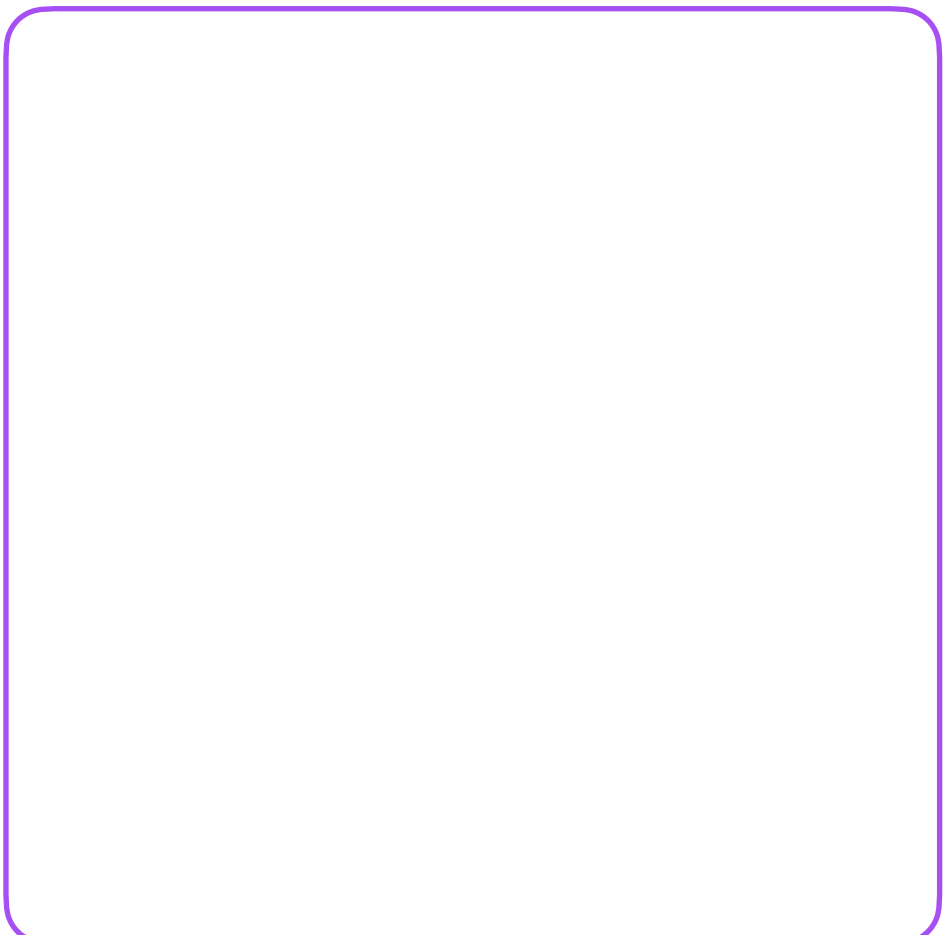


- Slow, redundant iterative

- Long time-to-market

Iteration speed

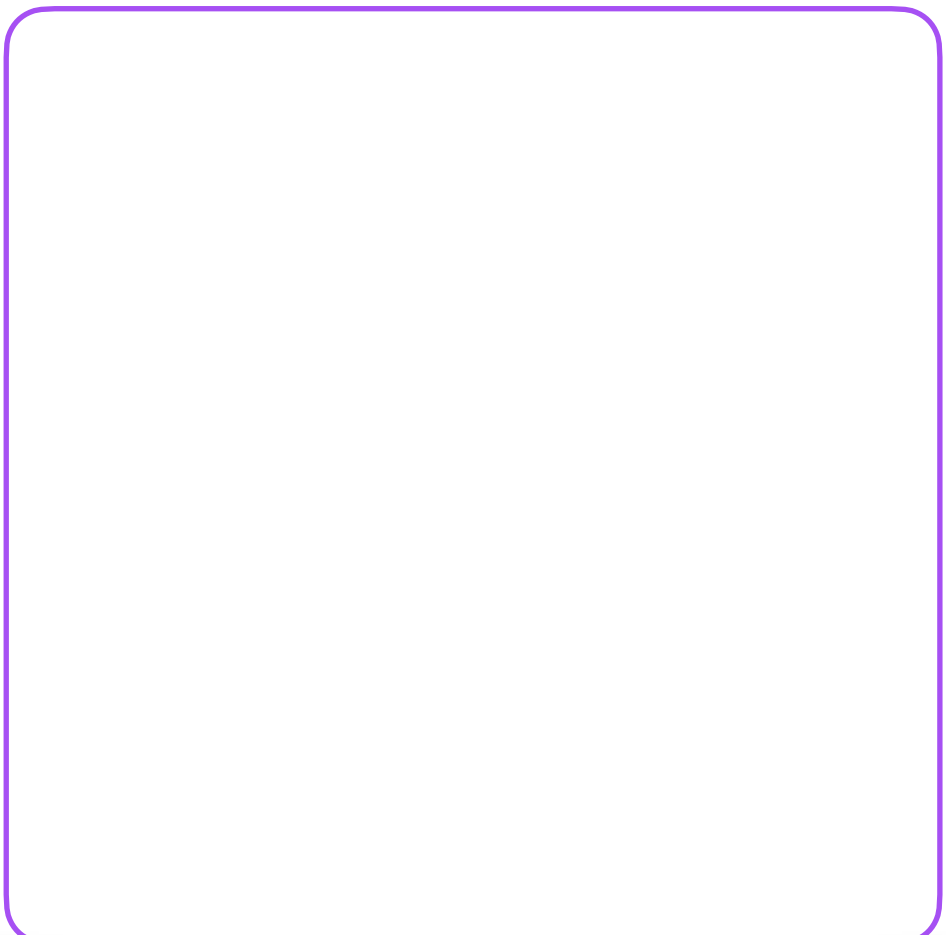
development



Collaboration

- Team growth

- Different skills and background



Productionization

- Difficulty to reproduce and

- Handoff to production is

control experiments

manual and bottlenecked

- Need to reduce operating costs

11

Amazon
S3

10
01

12