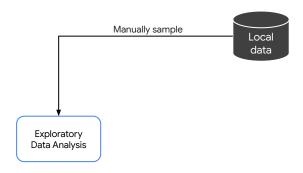
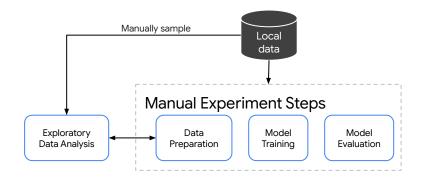
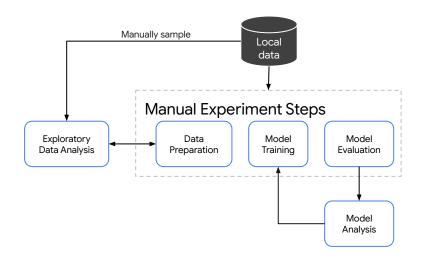
MLOps: A Use Case Overview

Building integrated ML pipelines









Explosion of things to worry about

Data preparation

Feature engineering

Architecture selection

Parameter selection

Model Training

Tuning strategy

Model evaluation

Properly handling:

- Imbalanced data
- Outliers
- Missing values
- High cardinality features
- Highly correlated features
- Target leakage
- Inconsistent feature definition
- Data that doesn't fit local memory

• ...

Selecting right preprocessing for:

- Numbers
- Classes
- Strings
- Dates
- Lists
- Nested fields
- .

Multiple options per column, 100s of columns in table

Selecting the best model architecture from dozens

available

- Linear
- Feed forward
- Random forest
- Decision tree
- Residual nets
- .

Keeping up with the onslaught of newest state of the art

For each architecture, selecting the right values for each hyperparameter

- Learning rate
- Regularization
- Layers
- Hidden nodes
- Activation fxn
- ..

Potentially more than a dozen values to set

Choose efficiently from O(1000s) of combinations.

Selecting right strategy for ensembling

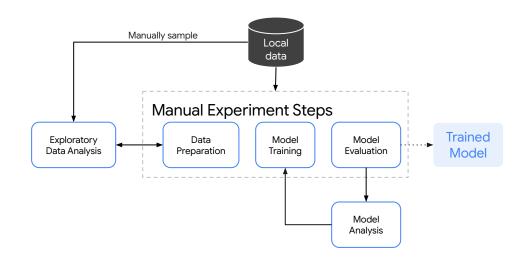
- Simple average
- Boosting
- Bagging
- ...

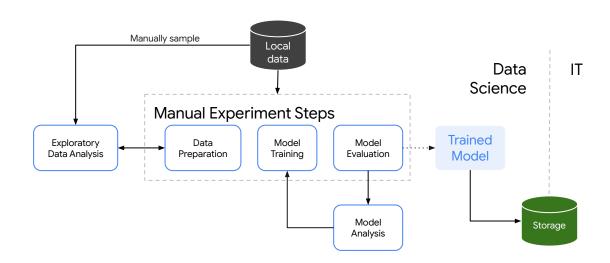
Evaluating model at

- Dataset-level
- Feature-level
- Prediction-level

Ensuring behavior is fully understood before deployment

Rinse & repeat up to 10s of times per use case





The never ending hand-overs

The MLOps **Personas**



ML Engineer



ML Researcher



Data Scientist



Data Engineer



Software Engineer



DevOps

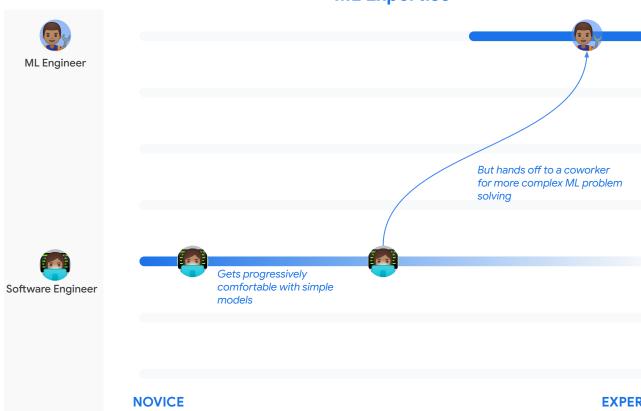


Business Analyst

ML Expertise

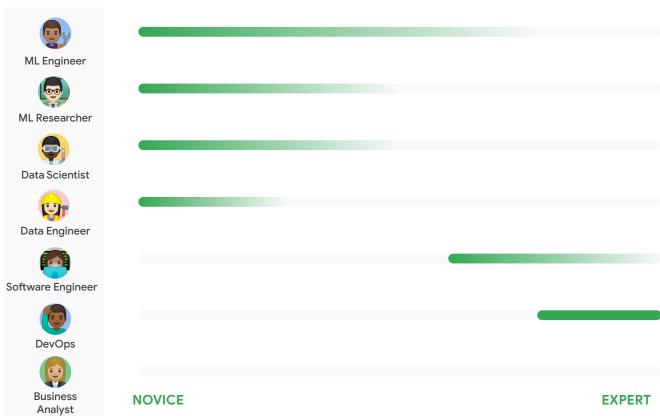


ML Expertise



EXPERT

Deployment Expertise



Deployment Expertise



