

Kubeflow

Danish Amjad [danish.amjad@est.fib.upc.edu]

Pratham Solanki [pratham.solanki@est.fib.upc.edu]

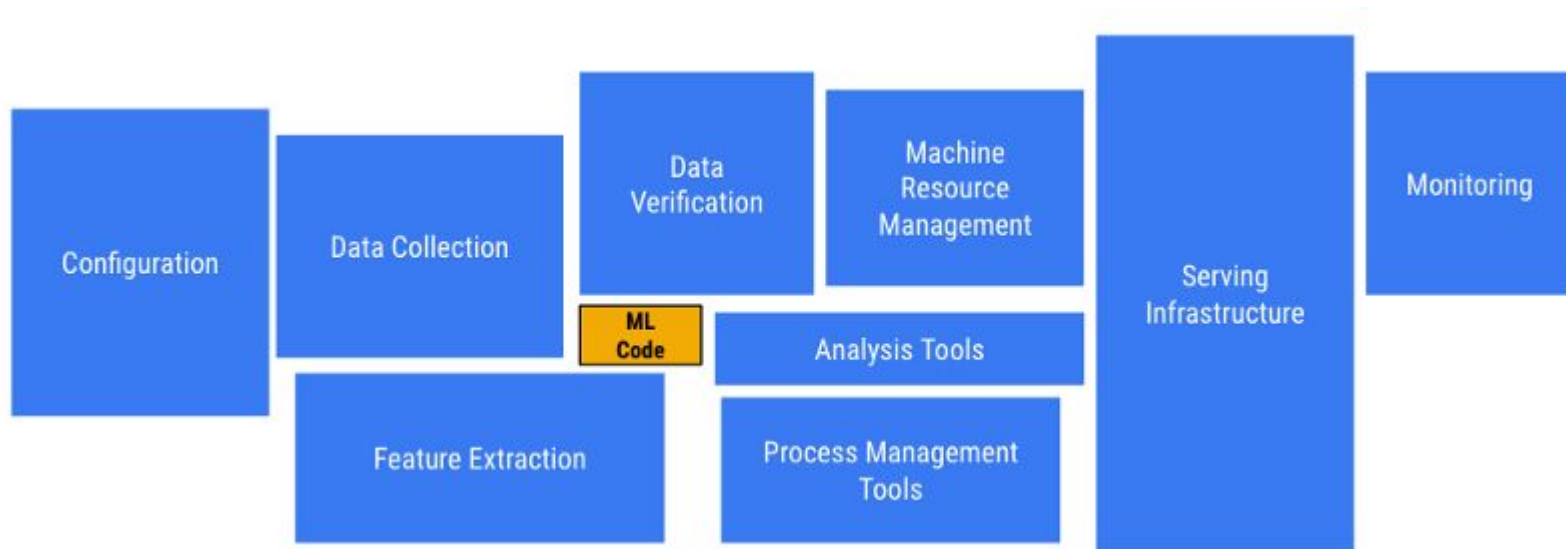
Machine Learning is hard

- Development and testing the models is quite hard in itself

HOWEVER,

- Making a production scale ML solution is harder

Technical Debt



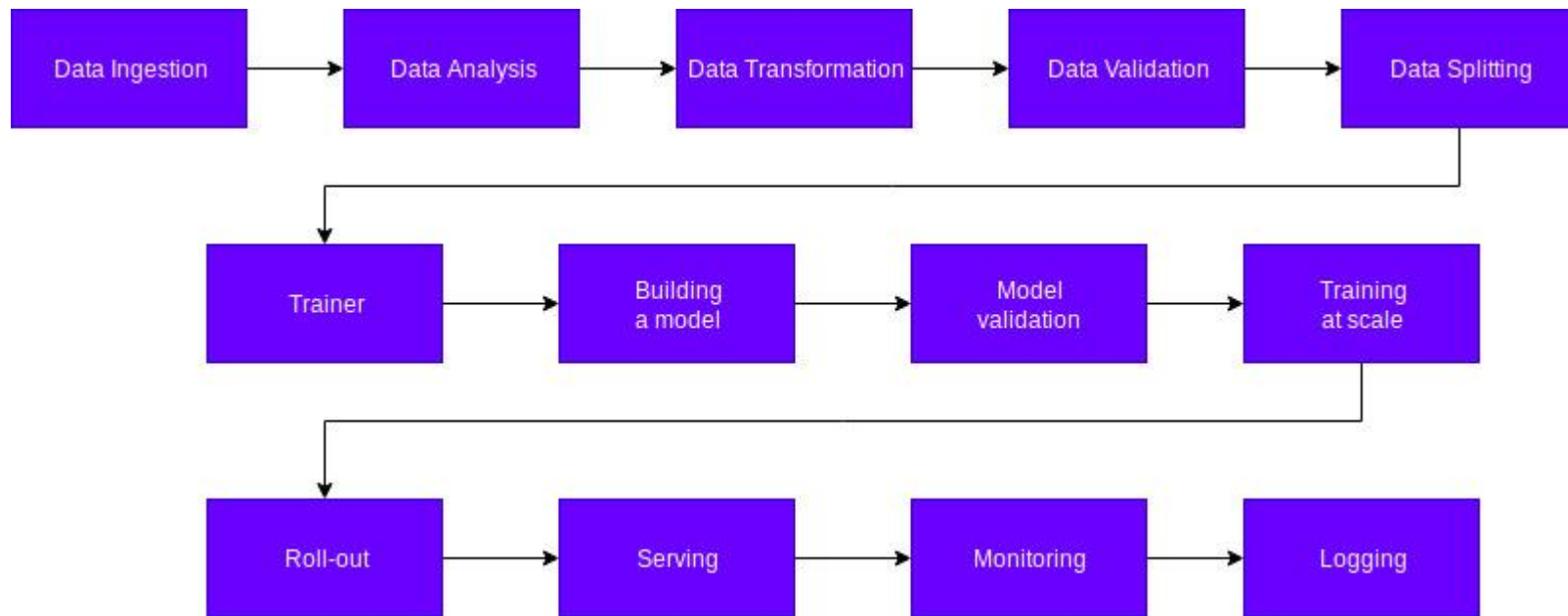
Credits: ["Hidden Technical Debt in Machine Learning Systems", D. Sculley et al. in "NIPS-2015"](#)

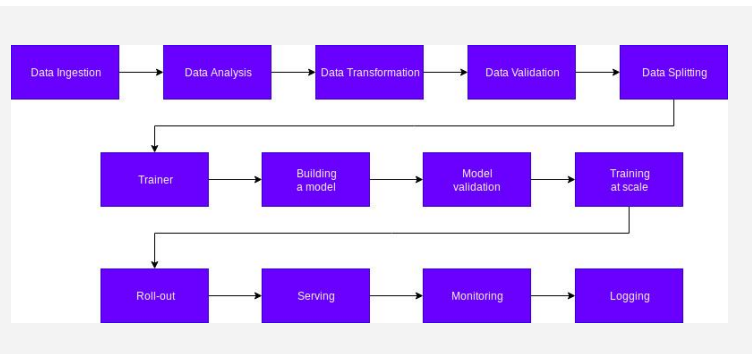
Current Systems

- Data processing pipelines
- Model training pipeline
- Ad-hoc experiments



Typical ML Task

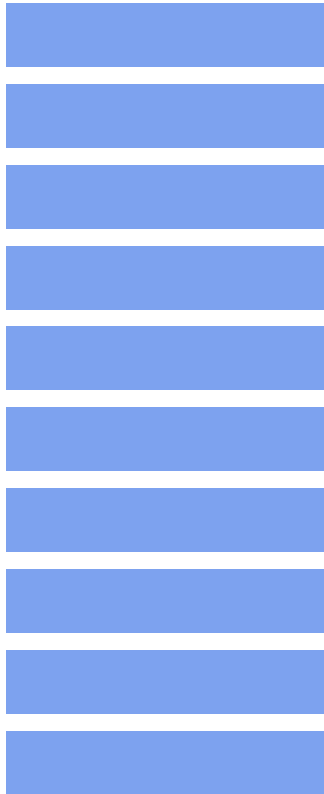




Experimentation



Training



Cloud



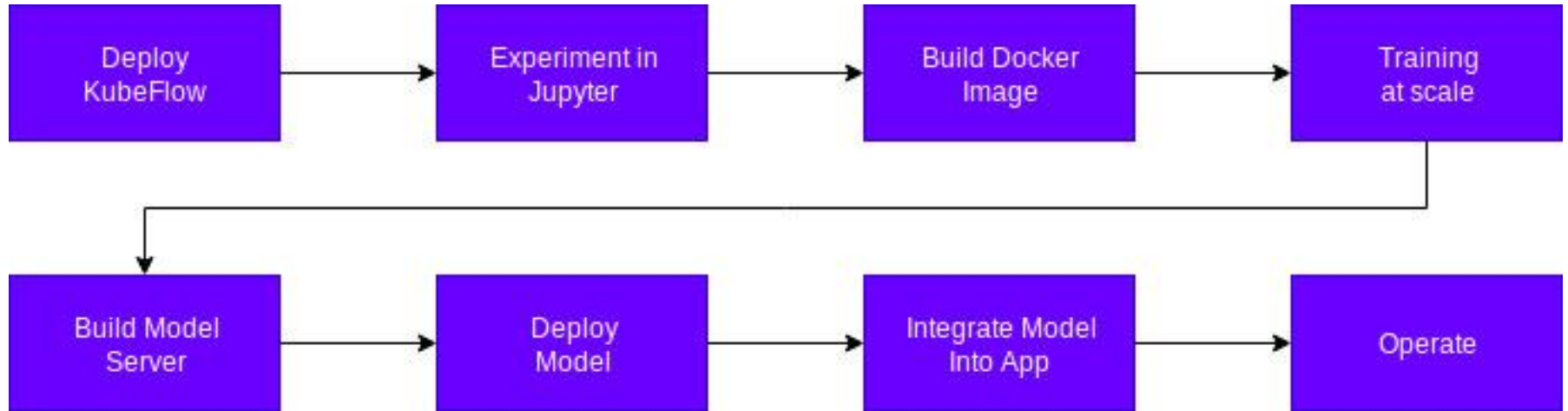
Experimentation

Training

Cloud



User Experience with Kubeflow

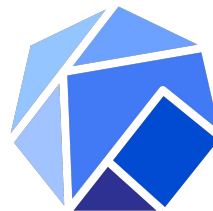


Project Goals

- Cloud Native Machine Learning
- Sharing, re-use and composability
- Rapid and reliable experimentation



AI Hub



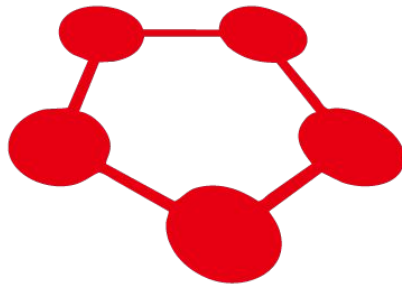
Kubeflow

Features

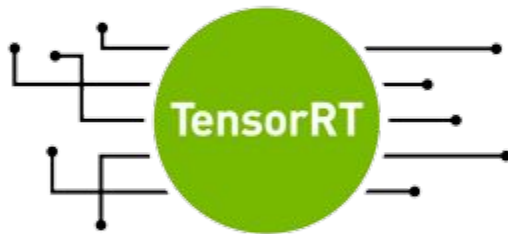
Development



Training/Testing

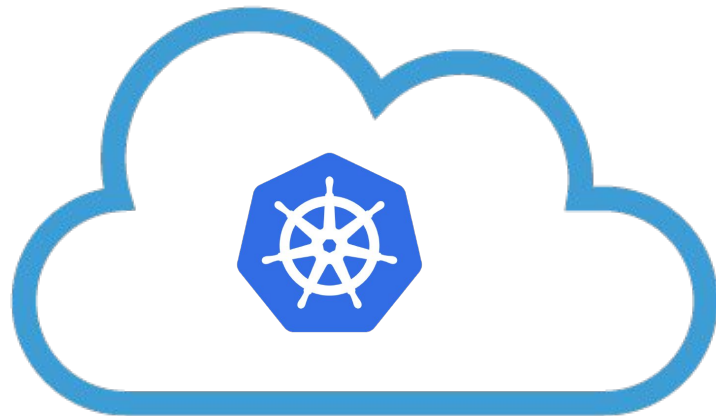
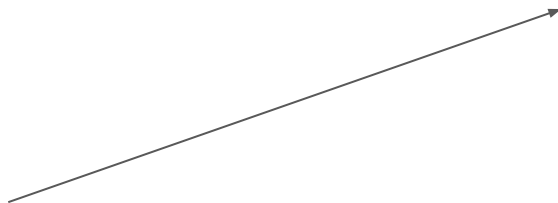


Serving the models

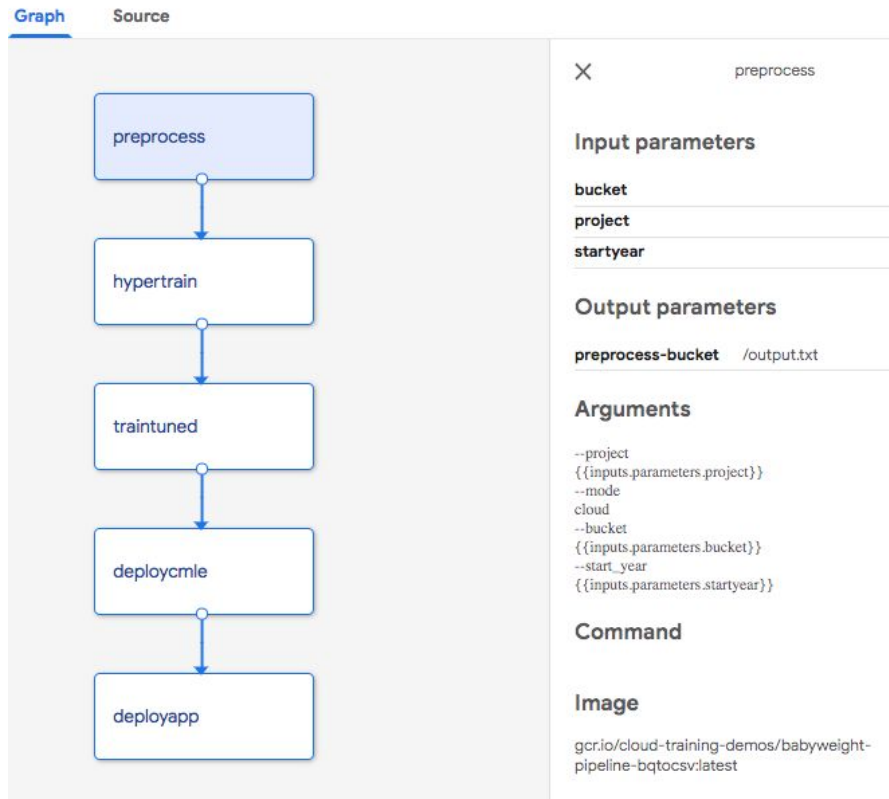


Additional Features

Fairing



Pipelines



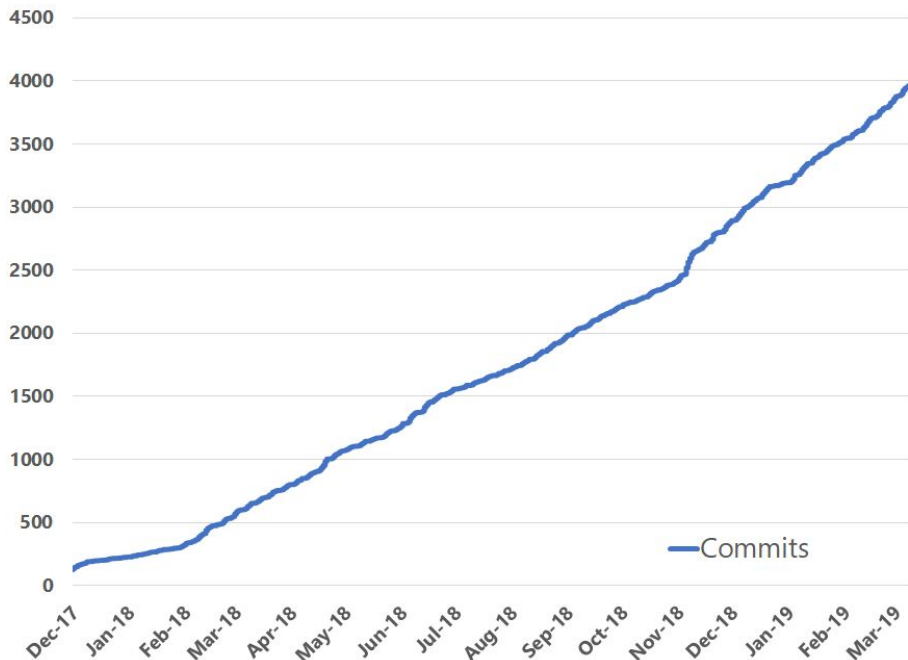
Hyper Parameter Tuning



K a t i b

MDB

Who is involved?



- ~4000 commits
- ~200 community contributors
- ~50 companies contributing, including:



Priority tasks for milestone 1.0

- Stabilized APIs for training (TFJob/PyTorch operators) and serving
- Integration with hyperparameter tuning with Katib
- Solve PyTorch Issues and TFJob Issues
- Enterprise Readiness
- Deployment and Development Experience
- Data Science UI
- Advanced ML Platform
- Test Release Infrastructure

Conclusion

Questions?