MLOps

Agenda

1. MLOps
2. AutoML Hands-On

MLOps – Why?

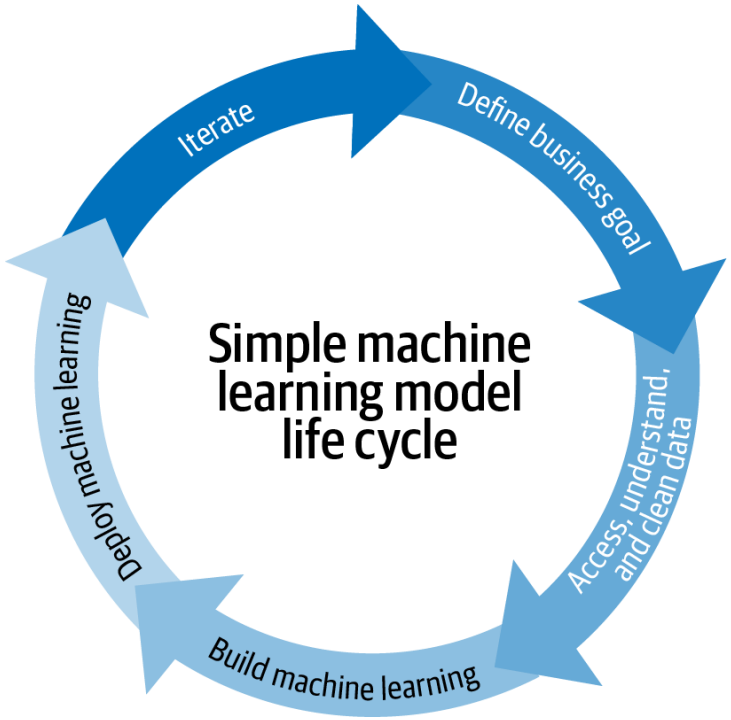
The average number of AI projects for a business is expected to increase to 35 by 2022 from four this year, according to a Gartner Inc. survey of about 100 organizations of various sizes, many of them with annual revenue of $1 billion to $3 billion.

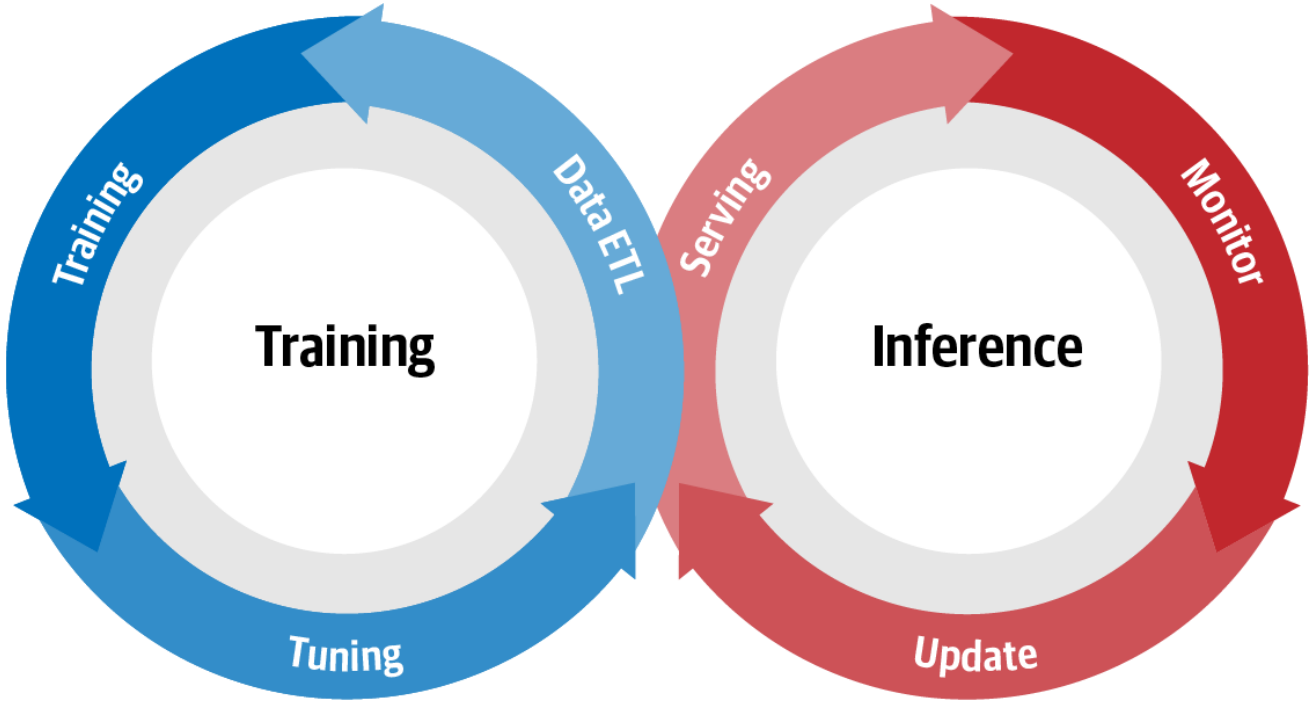
As per the report by NASSCOM and Blueocean, India is reigning big data analytics with a value of $1.2 billion placing it among the top 10 big data analytics markets in the world. They have also anticipated the growth becoming eight-fold by 2025, soaring to $16 billion. With this vision in mind, every sector is now looking forward to Data analytics for its evolution.

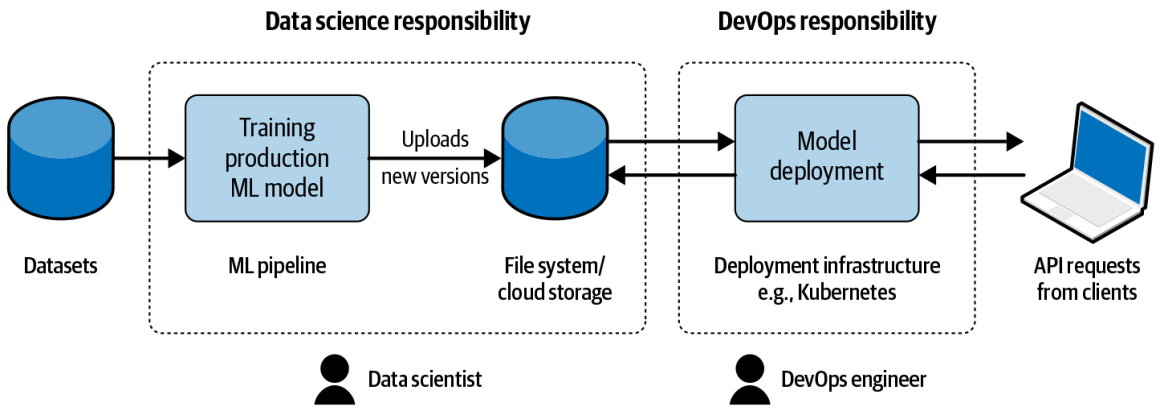
A recent report by PwC forecasts that AI will contribute as much as $15.7 trillion to the world economy by 2030—this number is a strong testament to its potential.

A study by Accenture further estimates that AI will add $957 billion or 15% of current gross value added to India's economy by 2035.

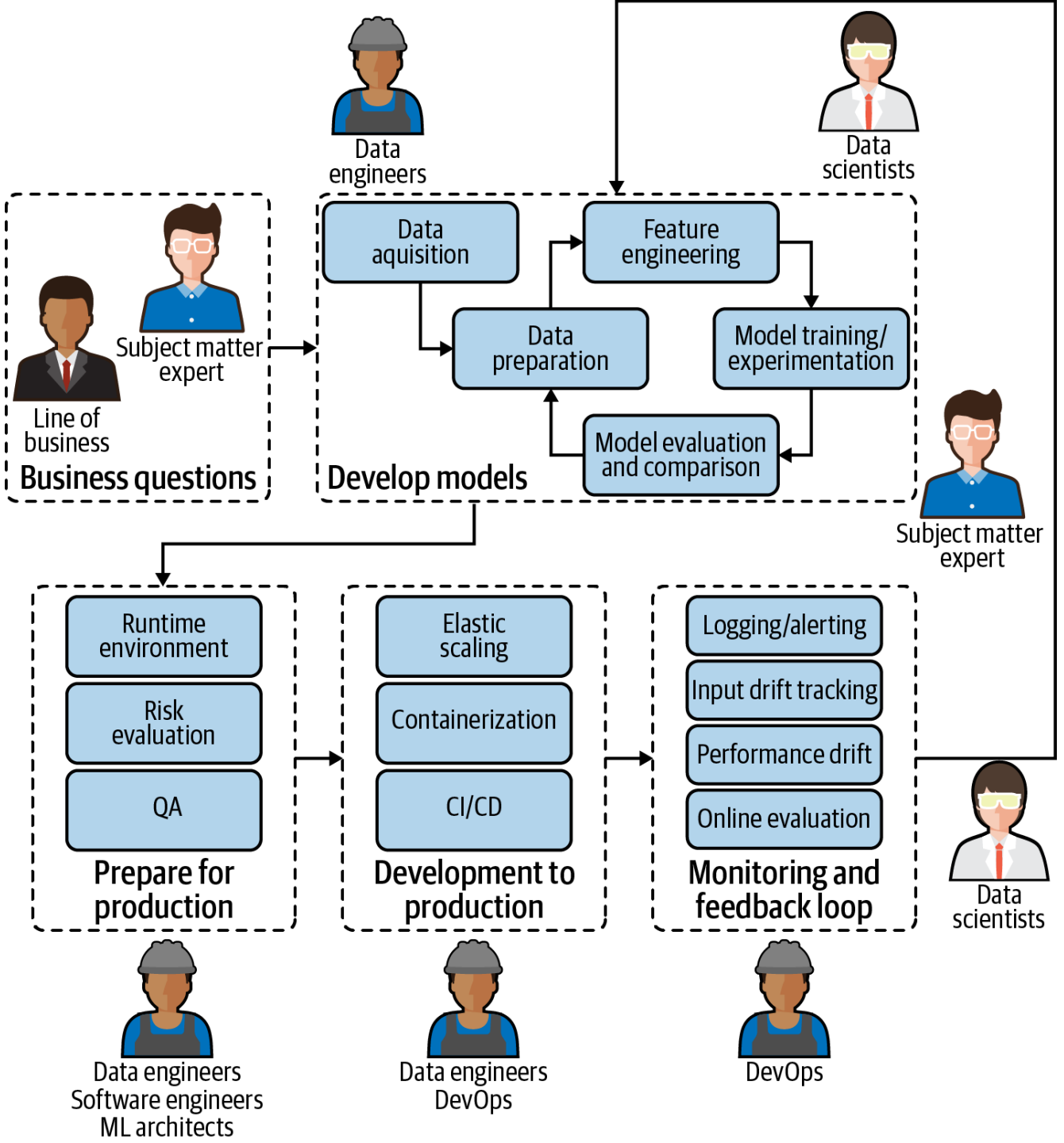
MLOps


Simple machine learning model life cycle
Define business goal
Access, understand, and clean data
Build machine learning
Deploy machine learning
Iterate

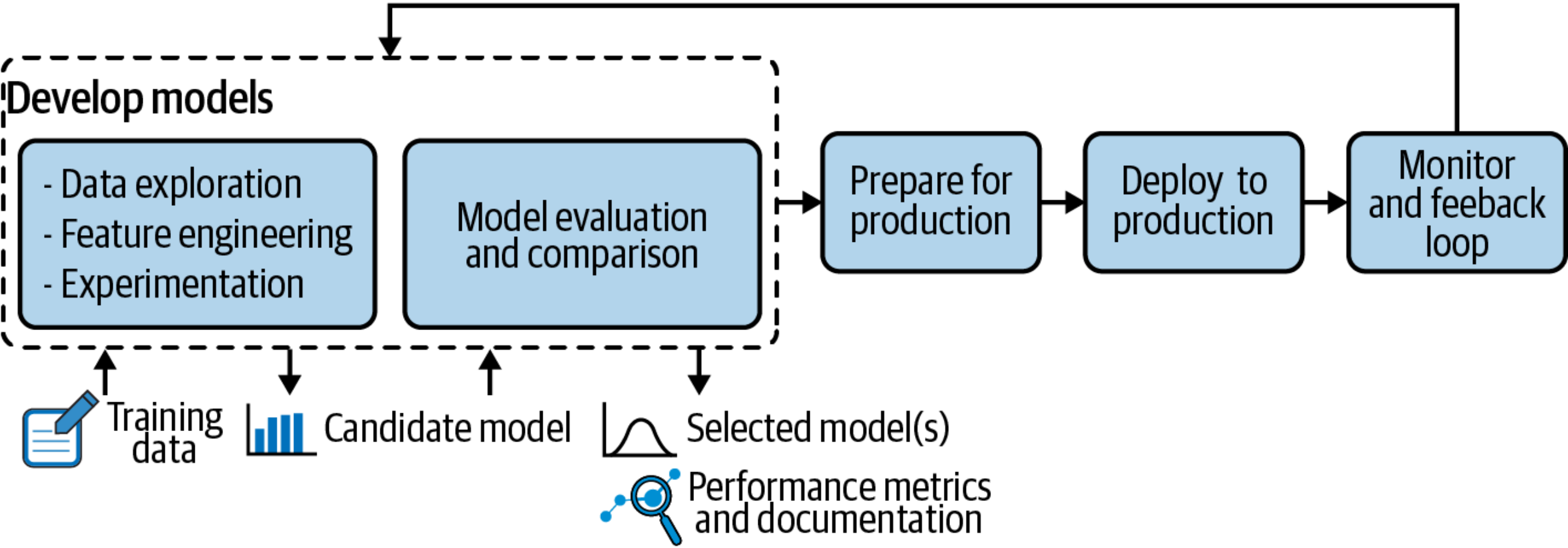

Training
Data ETL
Tuning
Training
Serving
Monitor
Update
Inference


Data science responsibility
DevOps responsibility
Training production ML model
Uploads new versions
Model deployment
Datasets
ML pipeline
File system/ cloud storage
Deployment infrastructure e.g., Kubernetes
API requests from clients
Data scientist
DevOps engineer

MLOps


Data engineers
Data scientists
Data aquisition
Feature engineering
Data preparation
Model training/ experimentation
Model evaluation and comparison
Subject matter expert
Line of business
Business questions
Develop models
Subject matter expert
Runtime environment
Risk evaluation
QA
Prepare for production
Elastic scaling
Containerization
CI/CD
Development to production
Logging/alerting
Input drift tracking
Performance drift
Online evaluation
Monitoring and feedback loop
Data scientists
Data engineers
Software engineers
ML architects
Data engineers
DevOps
DevOps

MLOps – Developing Models



Machine Learning

what society thinks I do

what my friends think I do

what my parents think I do

$$L_p = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{l} \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{w} + b) + \sum_{i=1}^{l} \alpha_i$$

$$\alpha_i \geq 0, \forall i$$

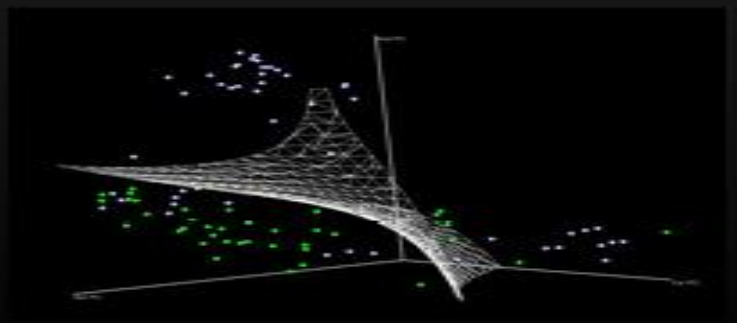
$$\mathbf{w} = \sum_{i=1}^{l} \alpha_i y_i \mathbf{x}_i, \sum_{i=1}^{l} \alpha_i y_i = 0$$

$$\nabla \hat{g}(\theta_t) = \frac{1}{n} \sum_{i=1}^{n} \nabla \ell(x_i, y_i; \theta_t) + \nabla r(\theta_t).$$

$$\theta_{t+1} = \theta_t - \eta_t \nabla \ell(x_{i(t)}, y_{i(t)}, \theta_t) - \eta_t \cdot \nabla r(\theta_t)$$

$$\mathbb{E}_{i(t)}[\ell(x_{i(t)}, y_{i(t)}; \theta_t)] = \frac{1}{n} \sum_i \ell(x_i, y_i; \theta_t).$$

what other programmers think I do

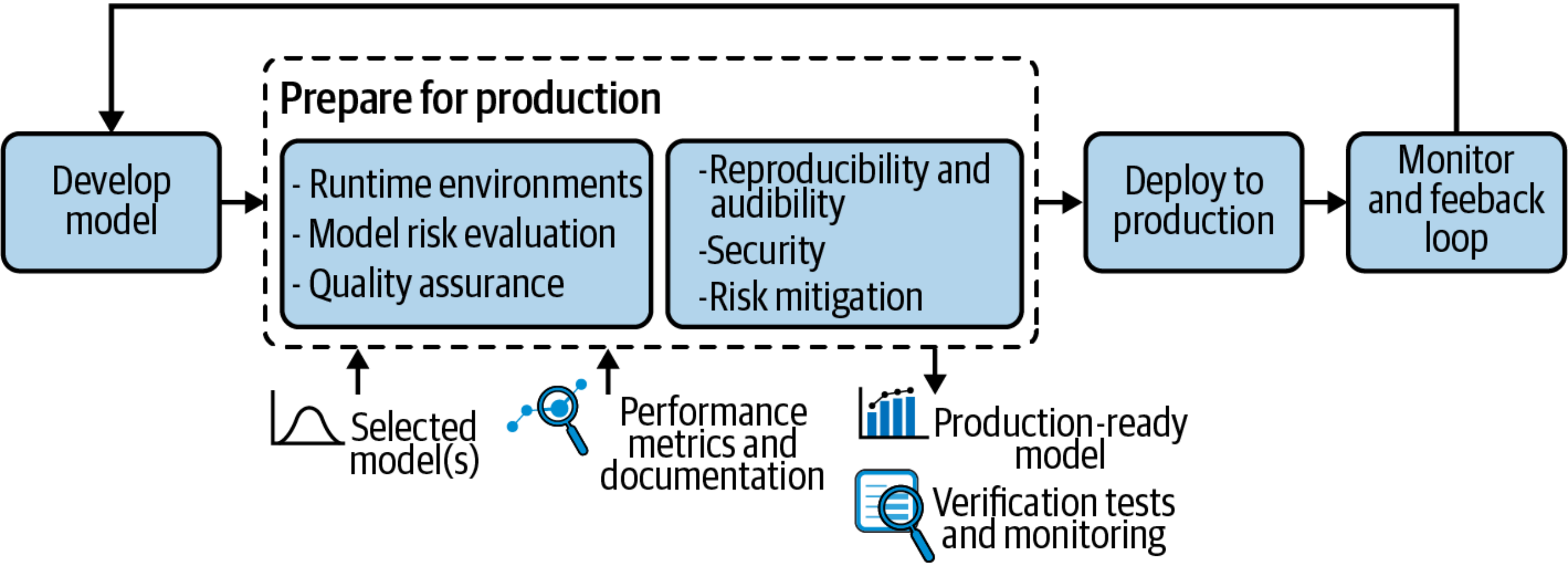


what I think I do

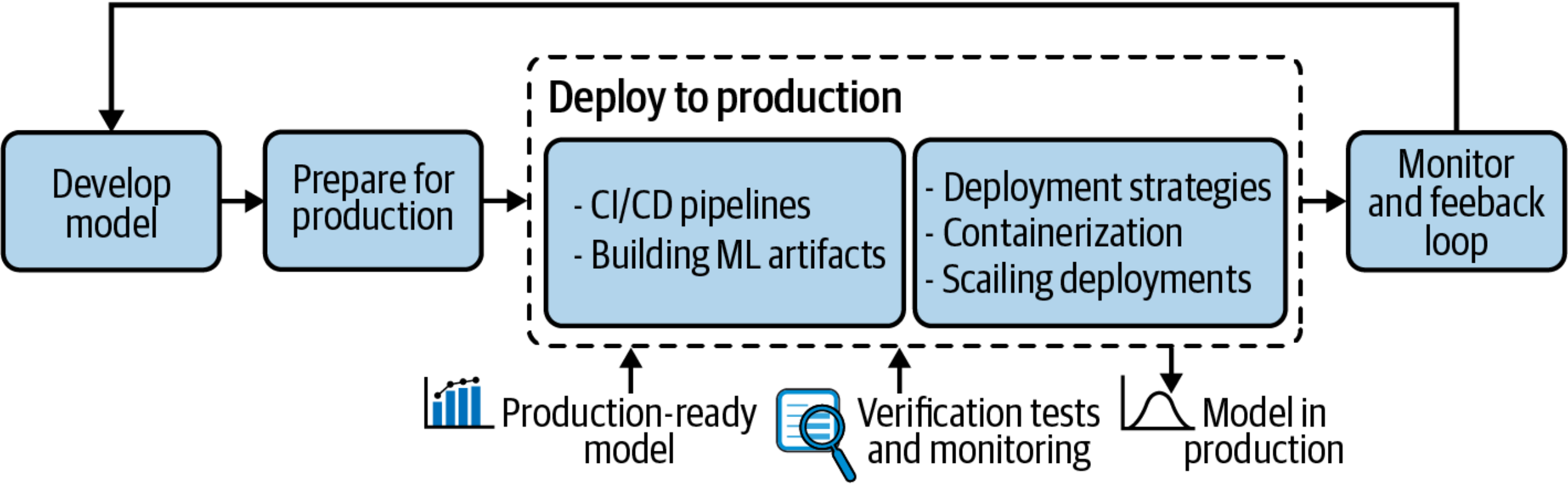
```
>>> from sklearn import svm
```

what I really do

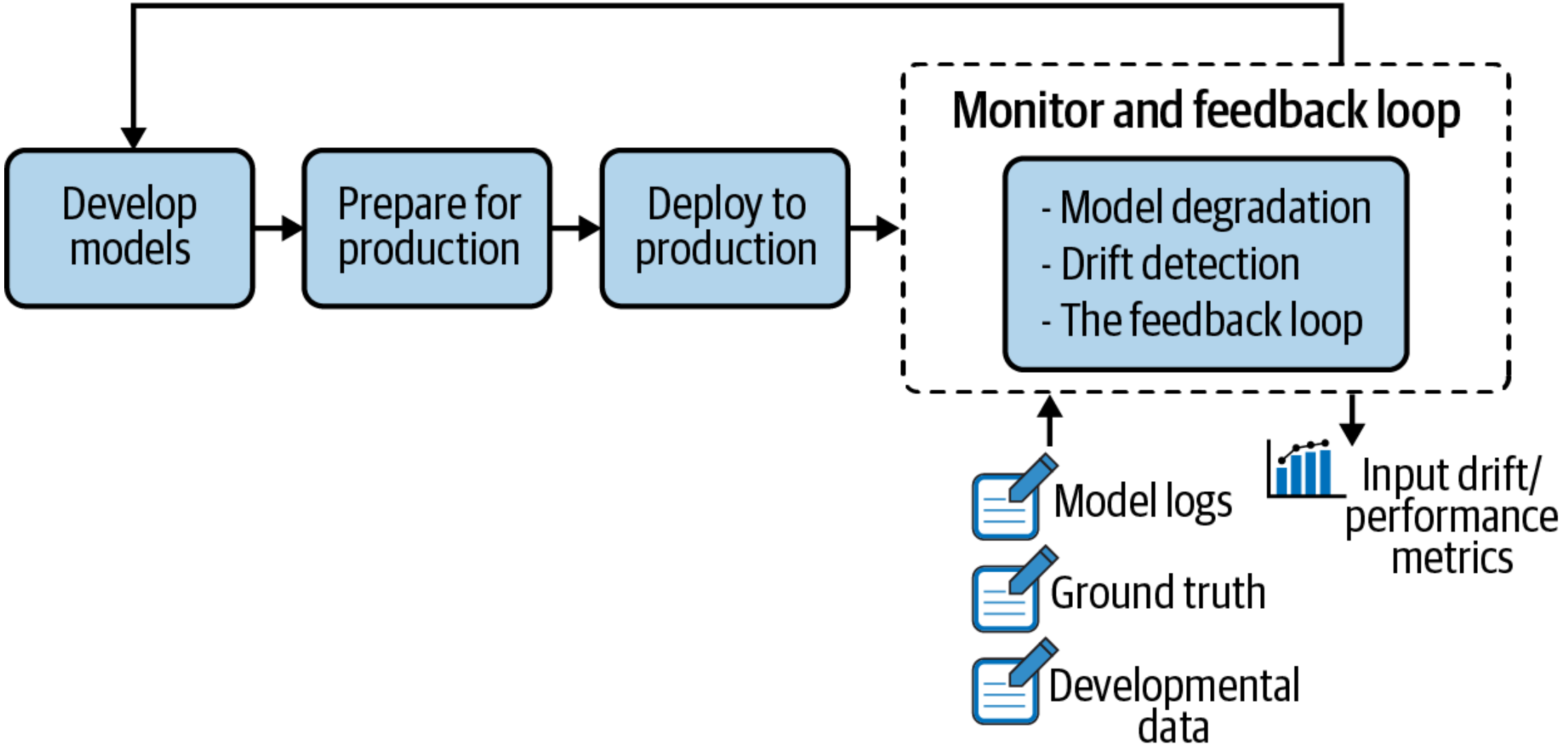
MLOps – Preparing for Production



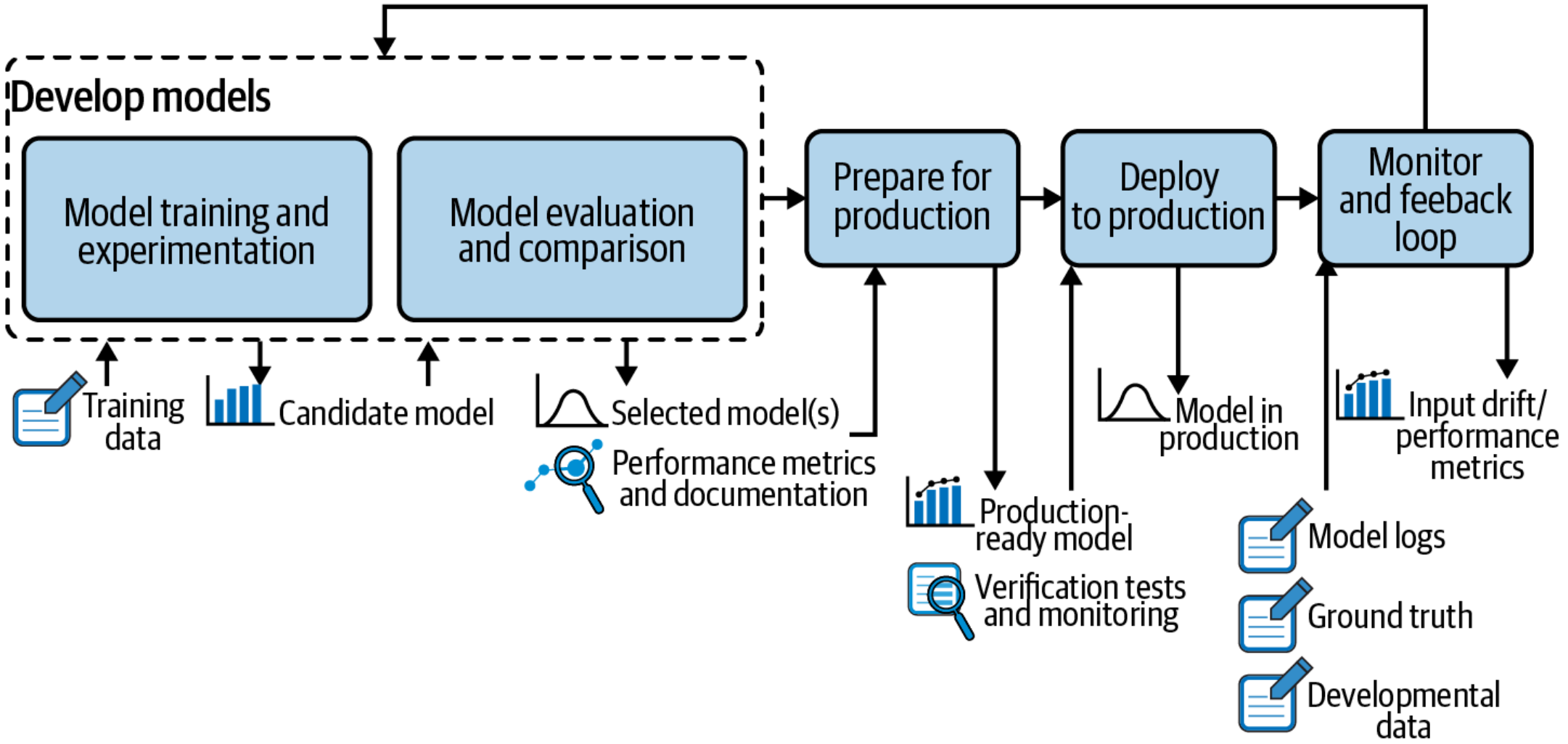
MLOps – Deploying to Production



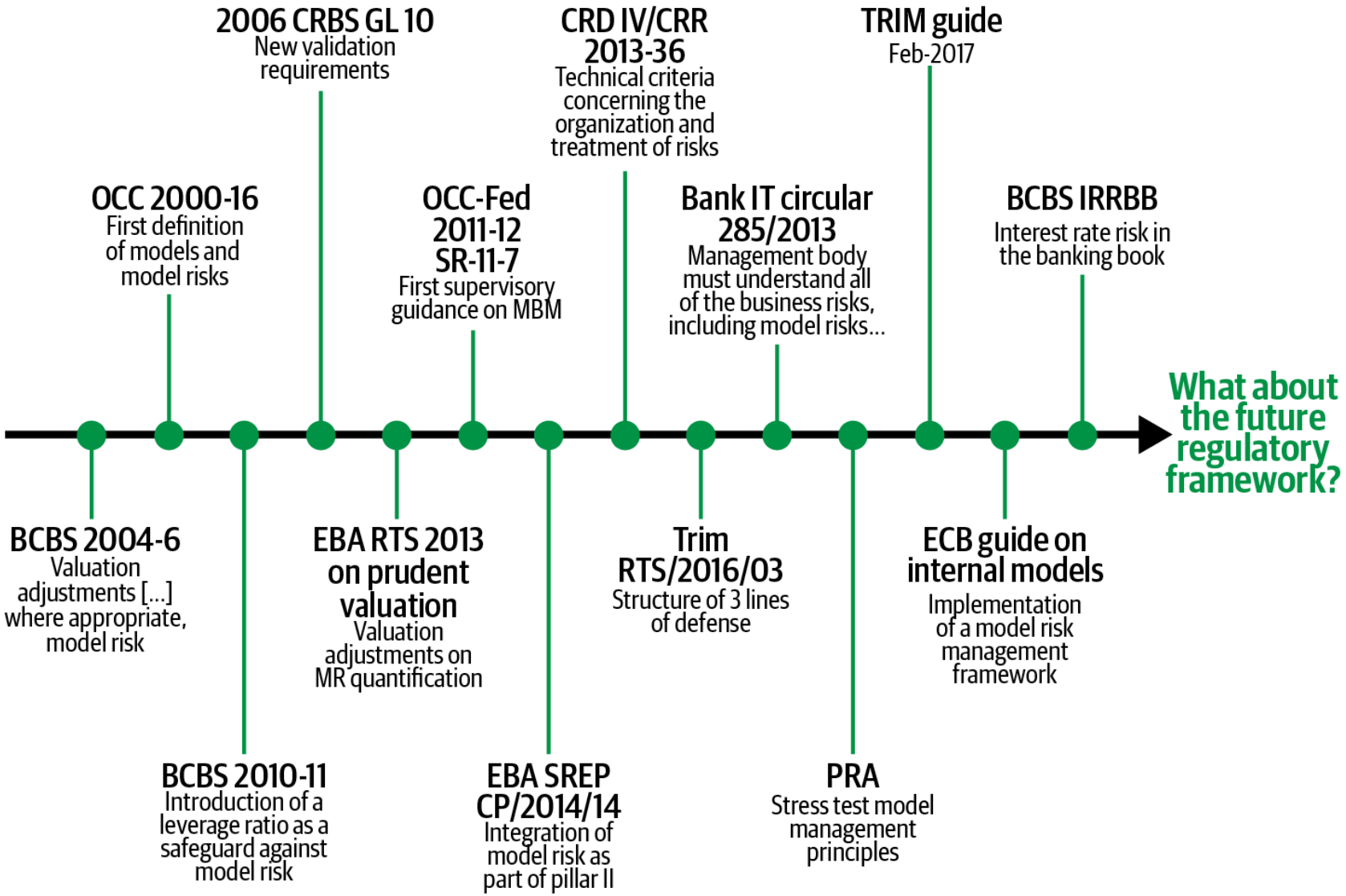
MLOps – Monitoring & Feedback Loop



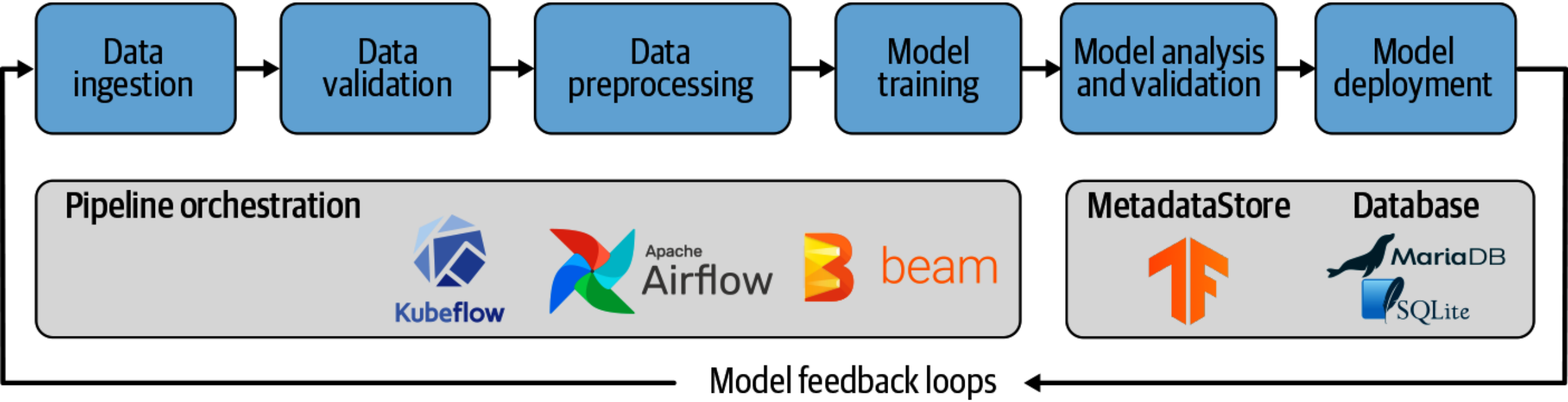
MLOps – Continuous Delivery



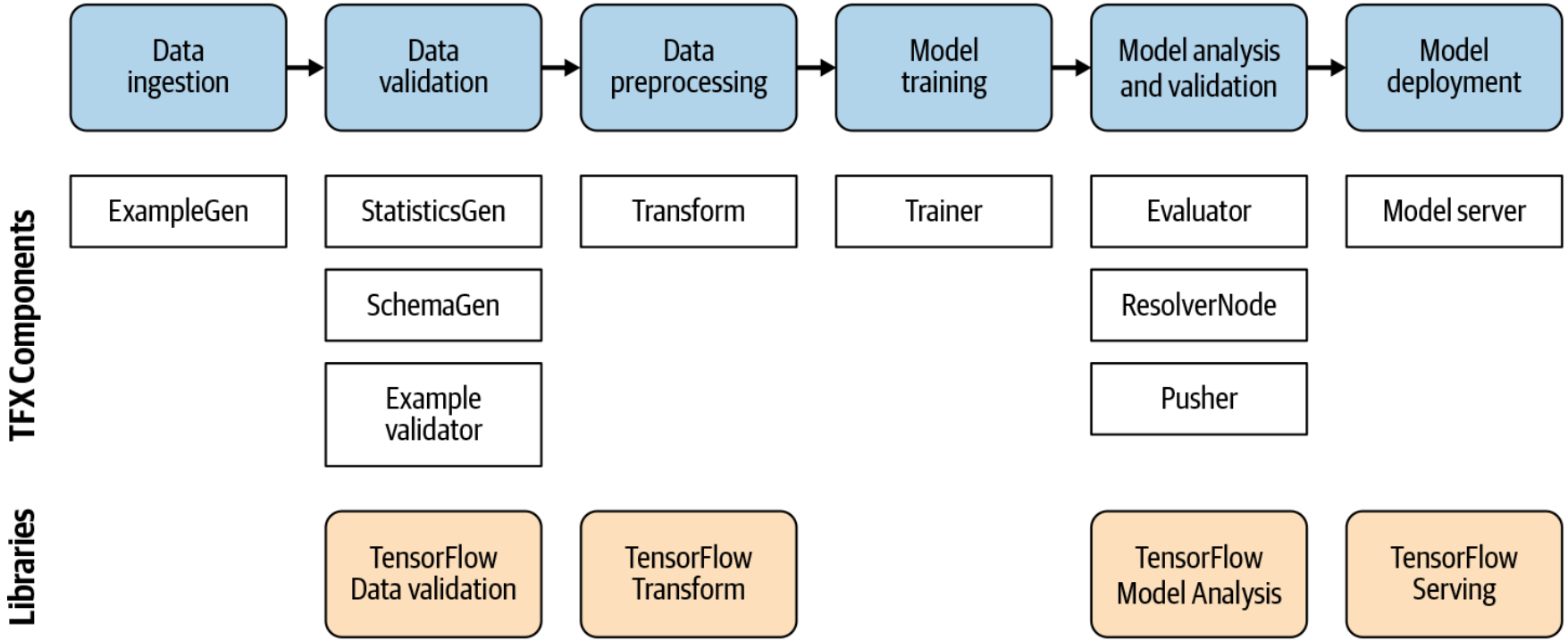
MLOps – Model Governance

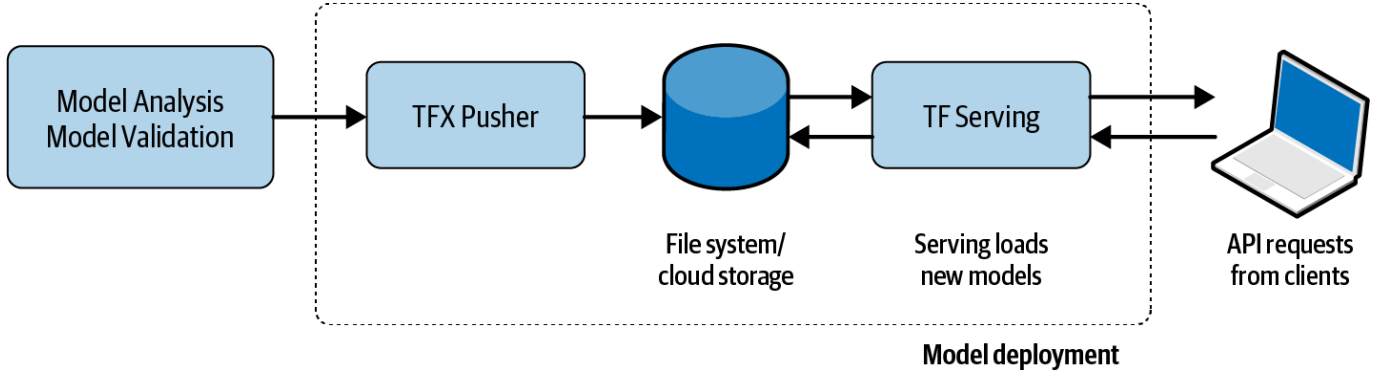


MLOps – Example of Tools

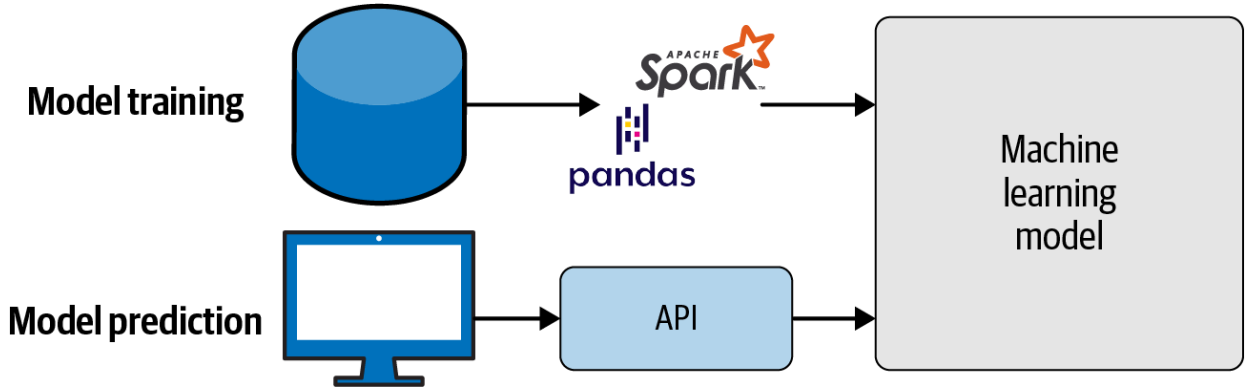


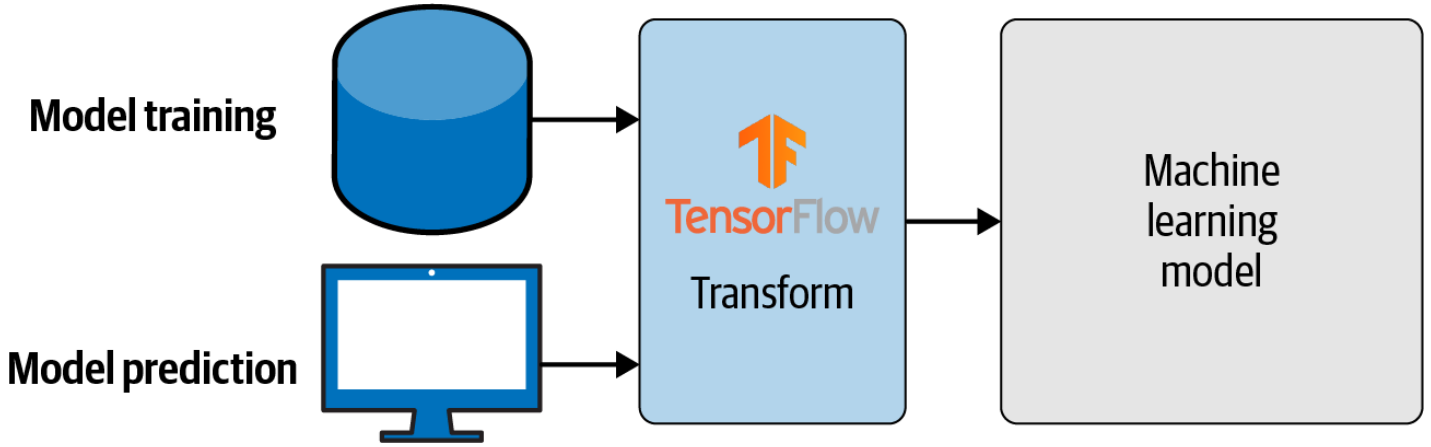
MLOps – Example of Tools





MLOps – Training-Serving Skew





THANK YOU