Introduction to Machine Learning in Production

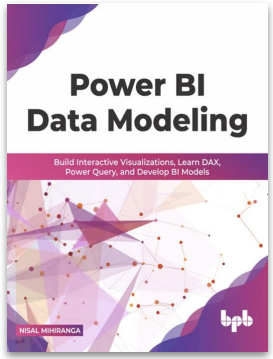
Nisal Mihiranga

PDU
PROFESSIONAL DEVELOPMENT UNIT

Microsoft
MCT

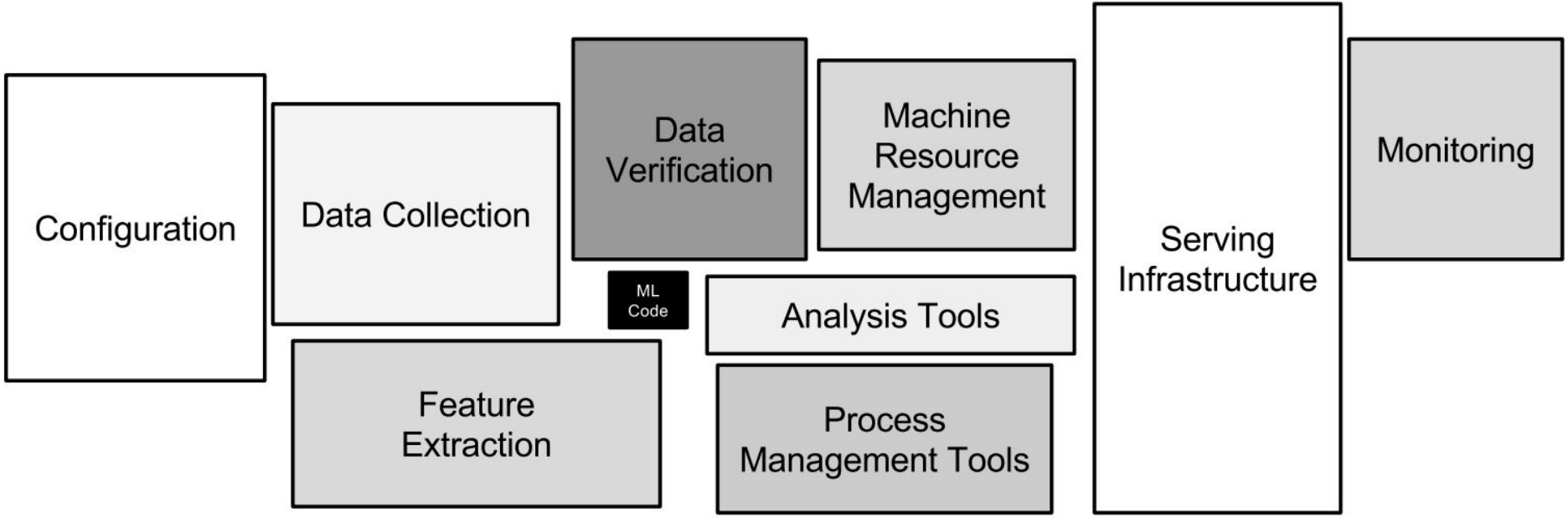
Nisal Mihiranga

- AI & Data Science Architect at Zone24x7
- Author and Trainer
- Author of the book “**Power BI Data Modeling**”



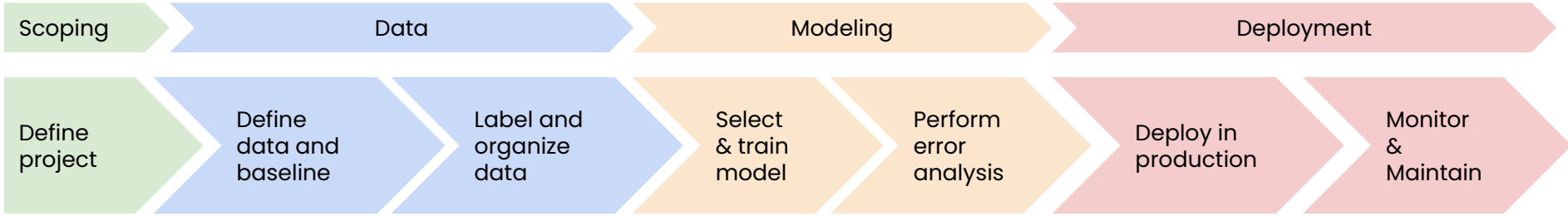
ML Project Lifecycle

ML Infrastructure

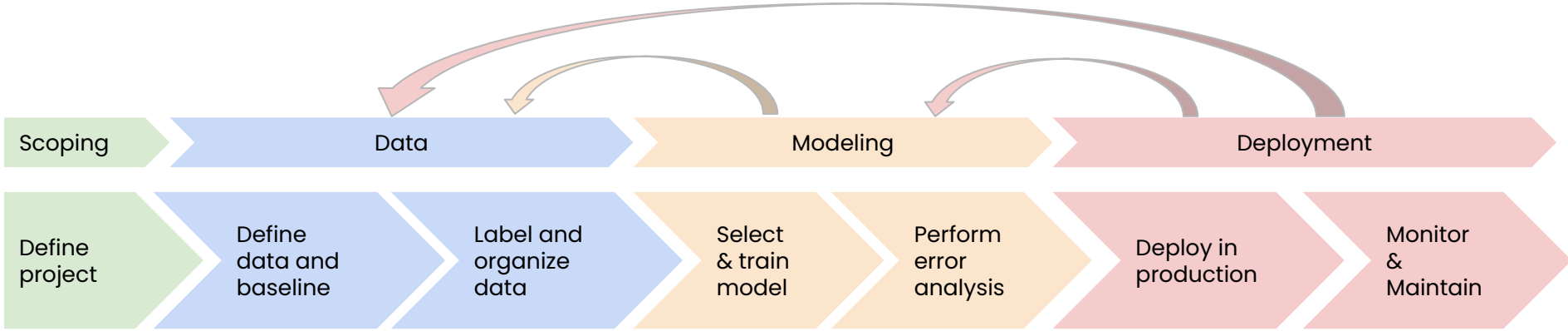


Sculley et. al. NIPS 2015: Hidden Technical Debt in Machine Learning Systems]

ML Project Lifecycle



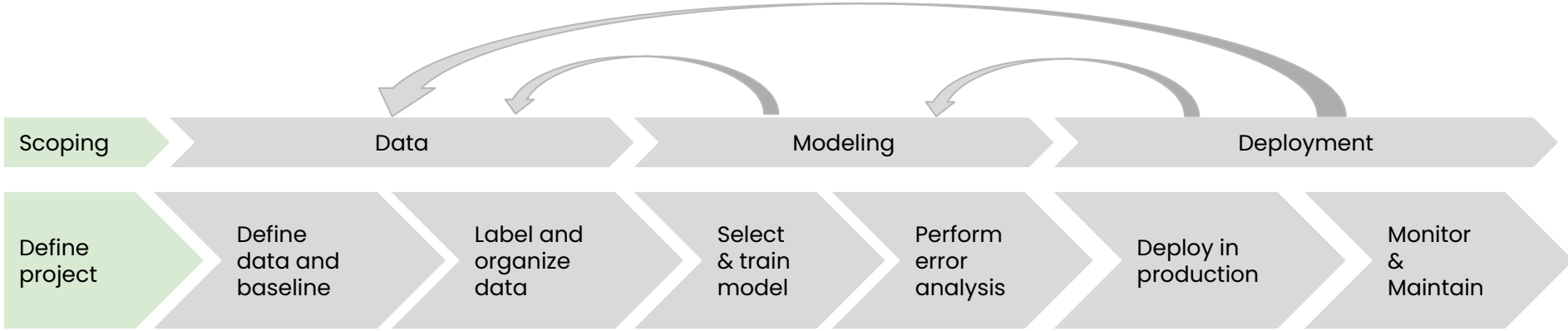
ML Project Lifecycle



ML Project Lifecycle

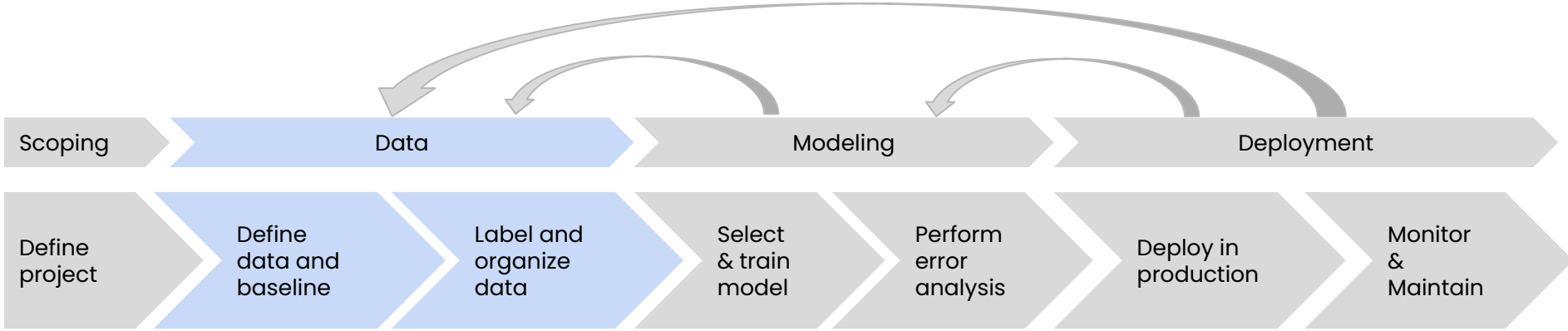
Case Study: Object detection

Image classification: Scoping



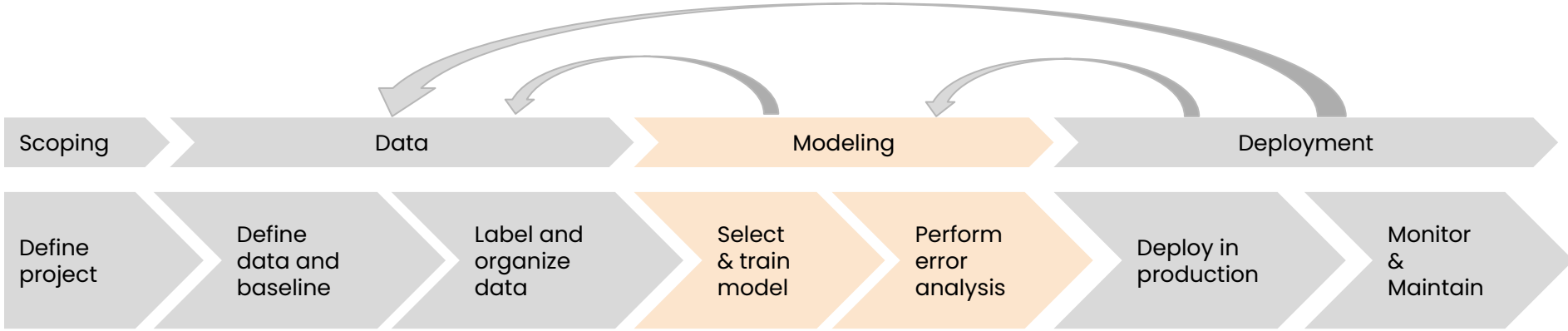
- Decide to work on Image recognition: object detection
- Decide on key metrics:
  - Accuracy, latency
- Estimate resources, project timeline

Image classification: Data Stage



- Define data
  - Is the data labeled consistently
  - How much data in each class

Image classification



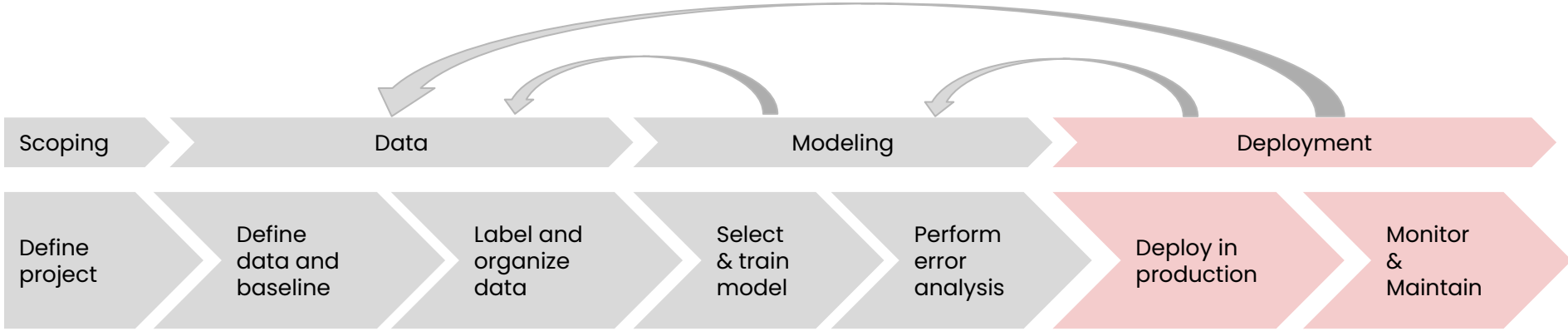
Code (algorithm/model), Research/Academic

Hyperparameters

Data

ML Project Lifecycle

PDU
PROFESSIONAL DEVELOPMENT UNIT

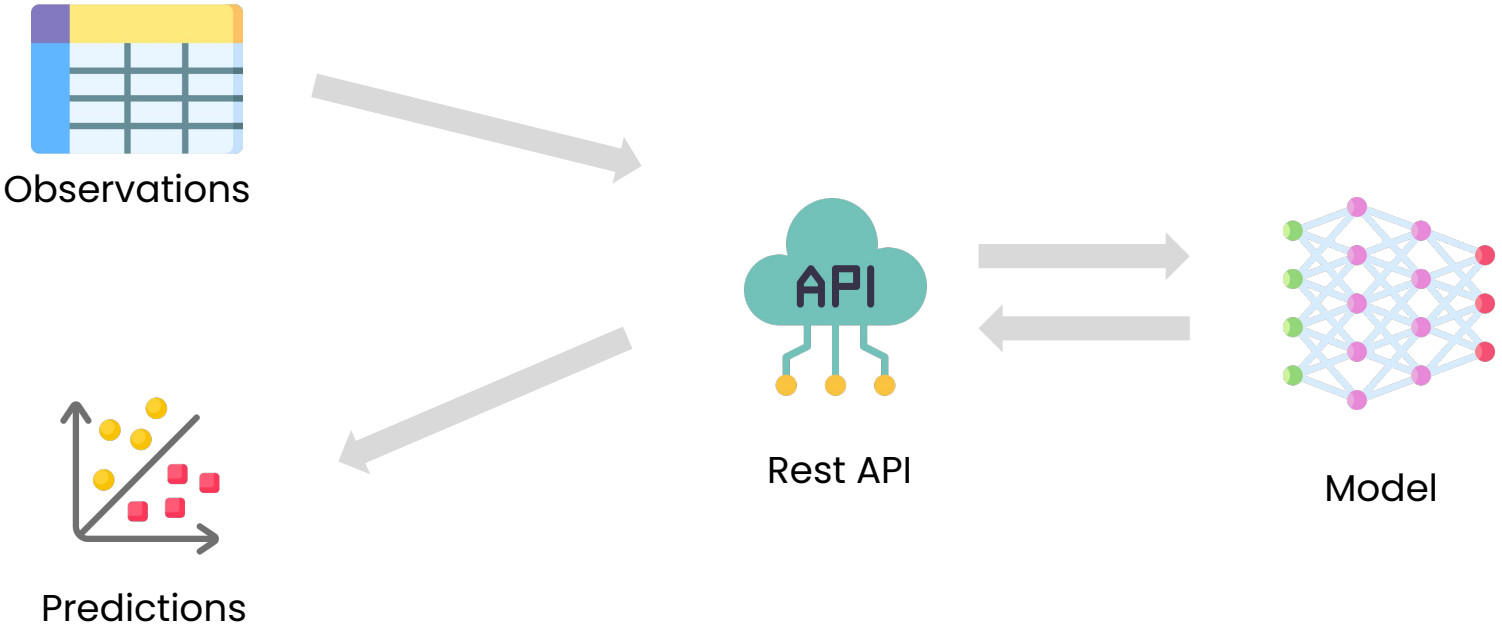


Production server

MLOps

- MLOps (Machine Learning Operations) is an emerging discipline, and comprises a set of tools, principals to support progress through ML lifecycle

Production System



Imagine You have deployed a predictive analytics application in production

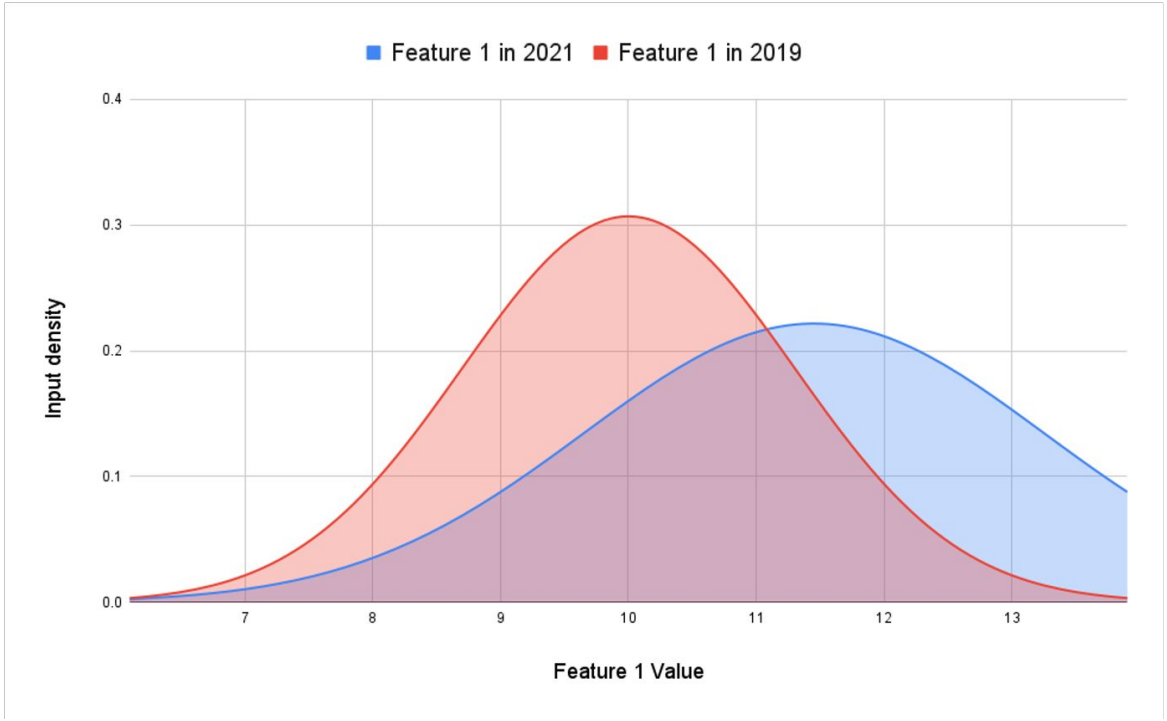
What could happen in Production after covid?

What could happen in Production after covid?

What could happen in Production with other major events?

“Data is mean to be changed. You know the world is changing and lot of things going on, pandemics, financial rescissions, inflation etc.Production data tends to be constantly changing over different dimension”

Model Drift


Feature 1 in 2021
Feature 1 in 2019
0.4
0.3
0.2
0.1
0.0
Input density
7
8
9
10
11
12
13
Feature 1 Value

Concept Drift

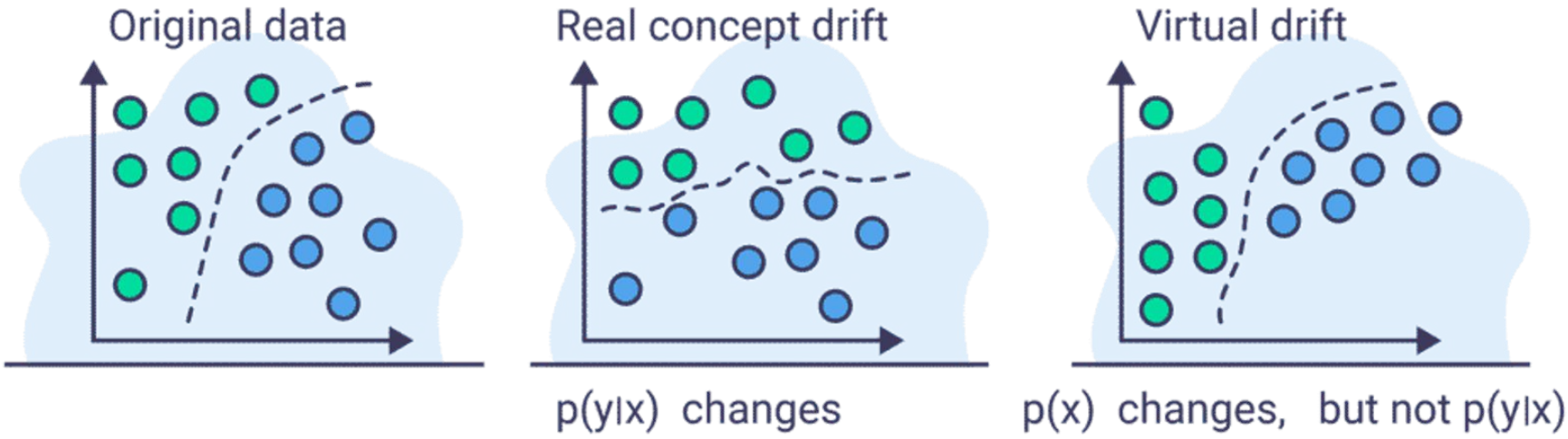
The distribution of the relationship between features and label P(Y) and P(y/x) changes over time.

Ex: income, credit score, age) more or less risky to loan money to.

Data Drift

distribution of features p(x) changes over time. Ex: if you train a model using specific demographic think about what happen when change the population over a period of time. Other examples for the data drift are changes due to Seasonality, Consumer preferences and new products arrivals

Model Drift (Model Decay)



How to Mitigate Model Decay

- Ensure the Data Validity by compare distribution changes
  - KL Divergence, KS Test
- Retrain the model with new data
- Refine the retraining policy

Goals of the MLOps Strategy

- Faster model development
- Faster production deployment of updated models
- Monitoring and quality assurance of production systems
- Experiments tracking and traceability

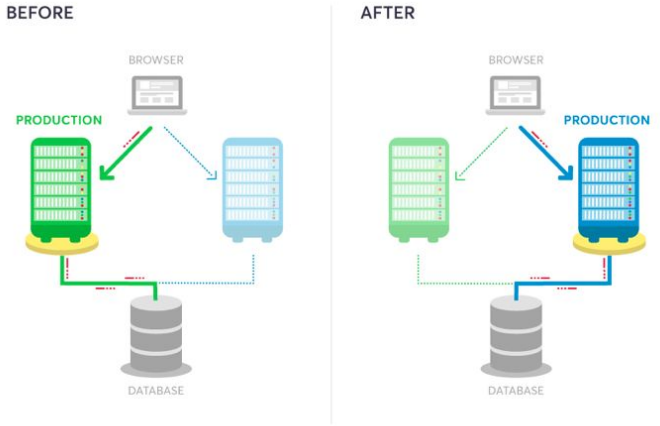
How to achieve MLOps goals

- Experimentation and tracking
- Source control the data and model artifacts
- Checkpoint the steps in the ML development lifecycle
- Automate the testing and deployment
- Monitor model performance, trigger and automatic retraining

Deployment Patterns: Blue green

Maintain two environments, one handle Live traffic and other idle.

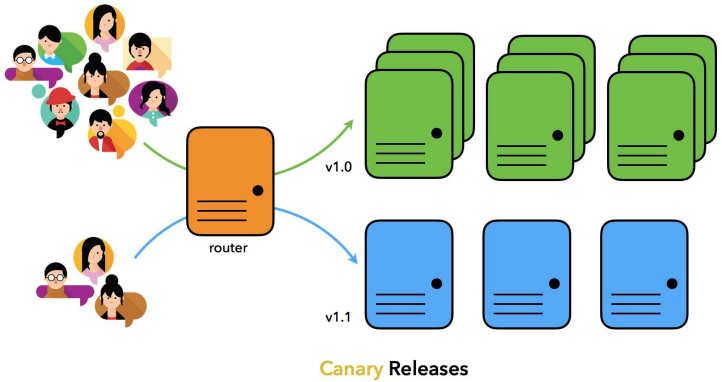
When new version deployed to idle after testing is done, switch the traffic



Deployment Patterns: Canary

Gradually rolling out a new version of the application to a subset of users, while leaving the rest on the current version.

This allows to test the new version in a live production environment.

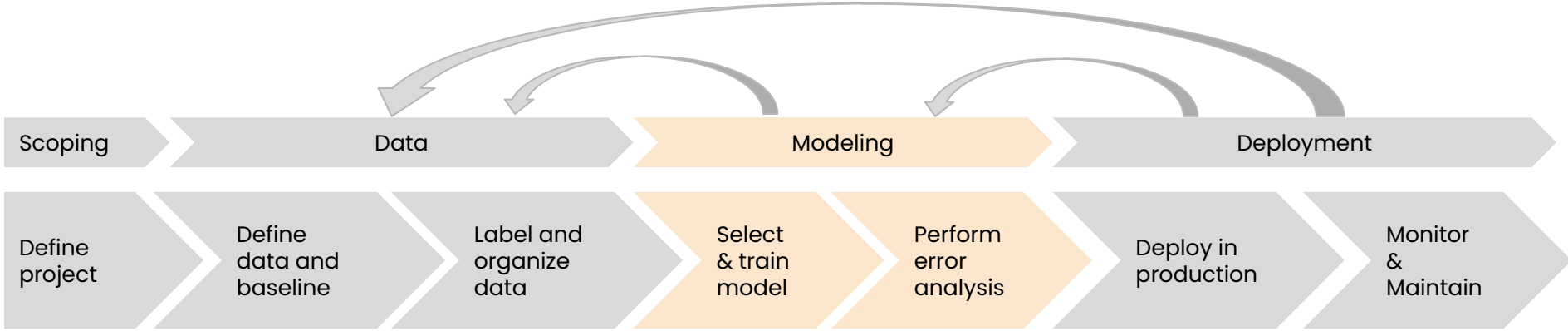


Monitoring dashboard

- Brainstorm on
  - the things that could go wrong
  - Few statistics/metrics that will detect the problem

Modeling Overview: Select & Train Model

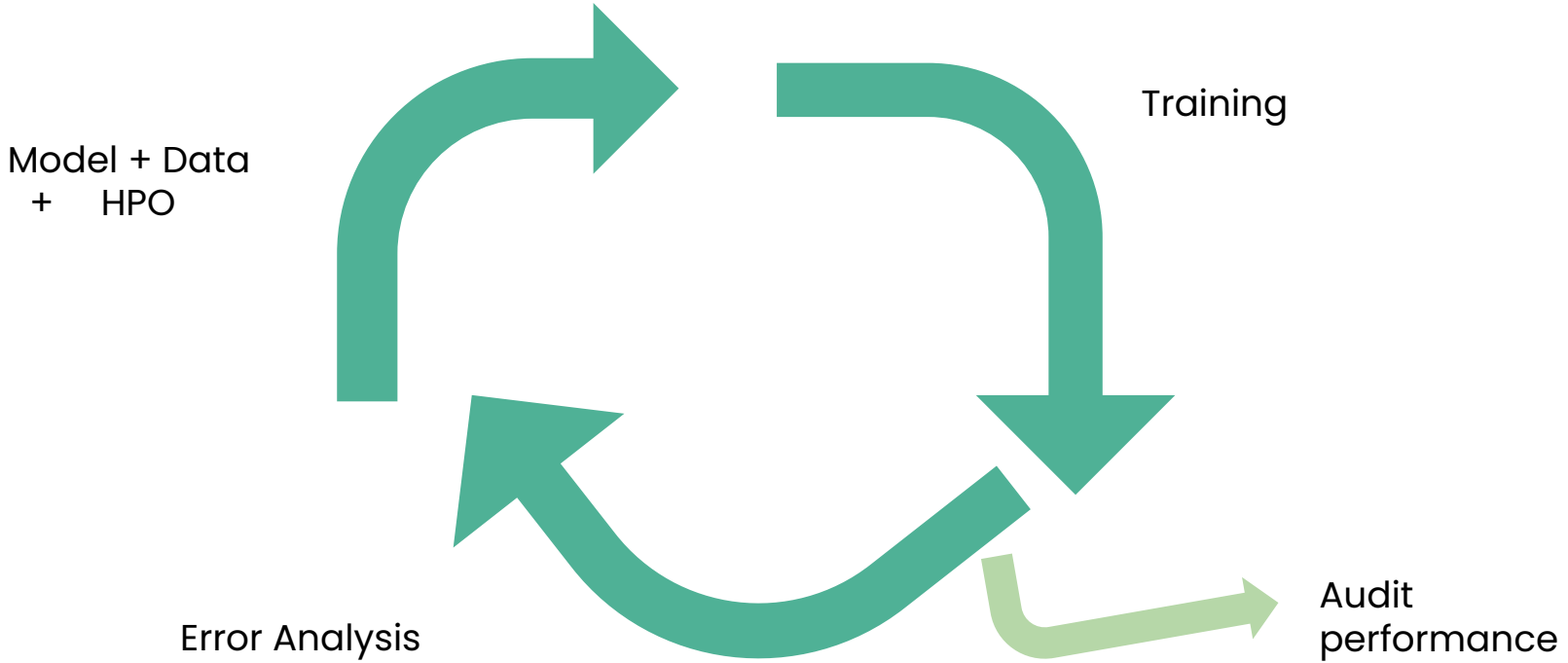
ML Modeling


Scoping
Data
Modeling
Deployment
Define project
Define data and baseline
Label and organize data
Select & train model
Perform error analysis
Deploy in production
Monitor & Maintain

AI System = Code + Data

(algorithm/model)

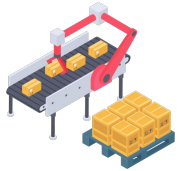
Model development is an iterative process



Error analysis and performance auditing

Examples of skewed data

- Manufacturing example
  - 99.7% no defect
  - 0.3% defect



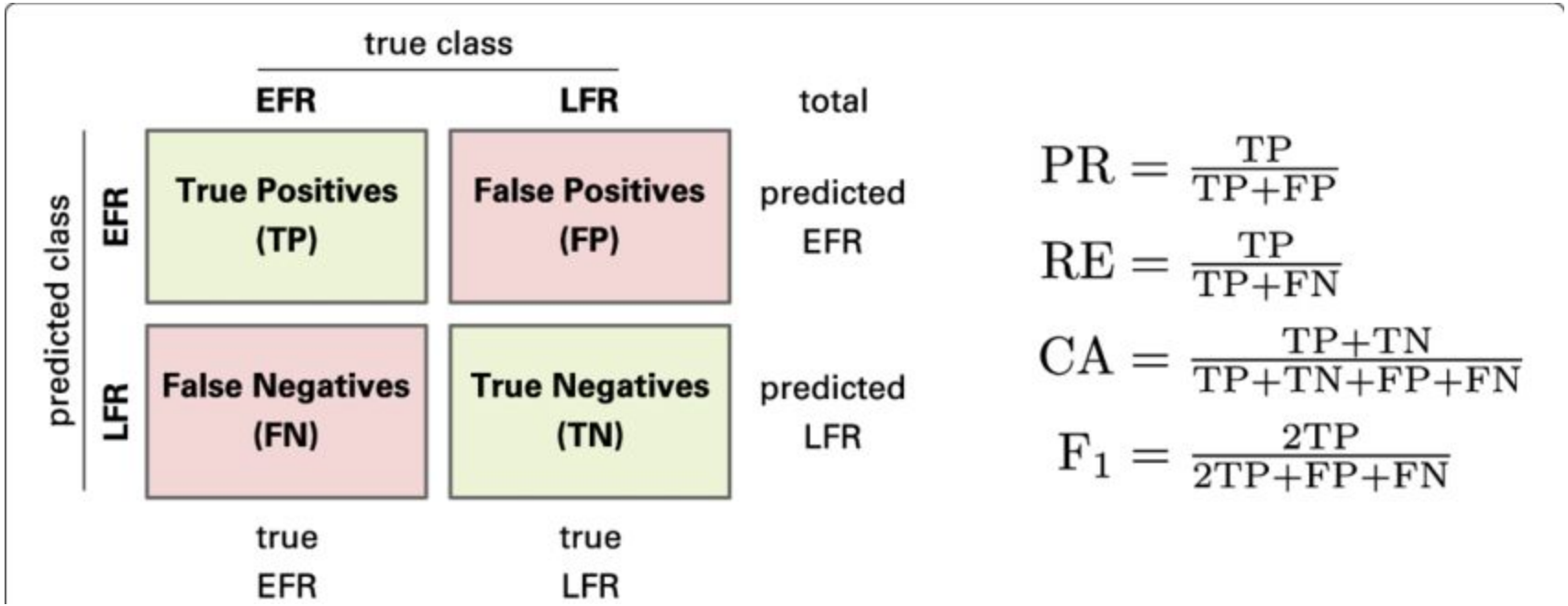
- Medical diagnosis: 98% of patients don't have a disease

**Note: This can lead to a poor performance of the model on the minority class*

Solutions for Skewed Data

- Collect more data
- Resampling techniques, over sampling the minority class or undersampling the majority
- Use different evaluation metrics
- Use different algorithms
- Adjust class weights

Confusion Matrix



$$\mathrm{PR} = \frac{\mathrm{TP}}{\mathrm{TP+FP}}$$

$$\mathrm{RE} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$$

$$\mathrm{CA} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}}$$

$$\mathrm{F}_1 = \frac{2\mathrm{TP}}{2\mathrm{TP+FP+FN}}$$


QUICK ICEBREAKER
ICEBREAKERIDEAS.COM

ML As a Service

MLAAS

On-premise Vs Cloud Base | Pros and Cons

On-Premises

Pros

- Full Data Control
- Full Hard- and Software Control
- Full Access Reliability
- No Operating Software Cost
- Performance
- Customizations

Cons

- Maintenance and Acquisition Costs
- Internal Knowledge
- Full Responsibility
- Long-term Reliability

Cloud-Base

Pros

- Easy Set-up
- Low Acquisition and Maintenance Costs
- Flexibility and Scalability
- Accessibility and Integration
- Updates and Security
- Back-up and Data Restore
- Disaster Recovery

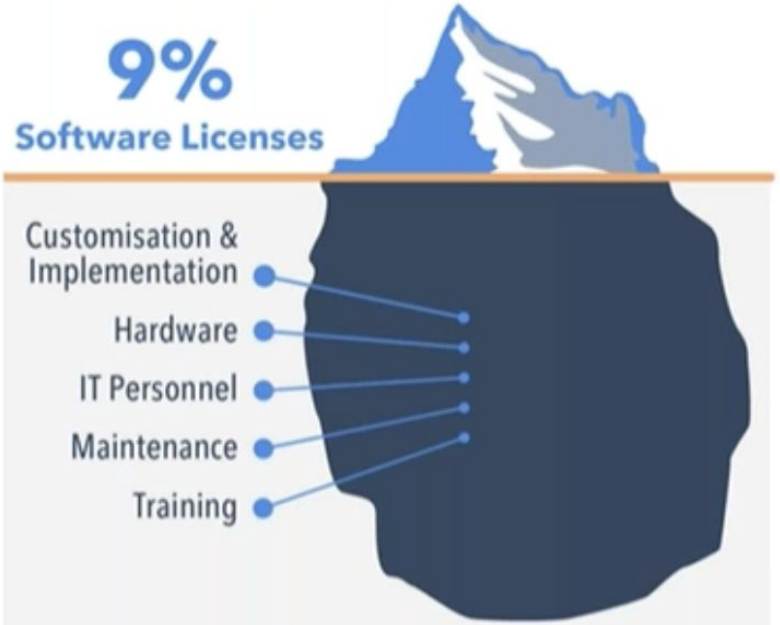
Cons

- On-going Software Costs
- Performance Limitations
- Internet Access

On-Premises

9%

Software Licenses



Ongoing Costs

- Apply Fixes, Patches, Upgrade
- Downtime
- Performance tuning
- Rewrite customizations
- Rewrite integrations
- Upgrade dependent applications
- Ongoing burden on IT
- Maintain/upgrade hardware
- Maintain/upgrade network
- Maintain/upgrade security
- Maintain/upgrade database

Cloud Computing

68%

Subscription Fee

Ongoing Costs

- Subscription fee

Google Cloud Platform (GCP)

offers a comprehensive suite of AI and machine learning services, including natural language processing, image and video analysis, predictive analytics, and machine learning platforms that enable developers to build and deploy intelligent applications quickly and easily.

Google Cloud

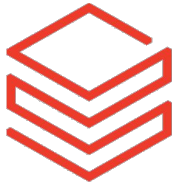
vertex.ai

Azure Machine Learning

Cloud-based service offered by Microsoft that provides a comprehensive set of tools and services to build, train, and deploy machine learning models at scale.

Databricks

Databricks is a unified analytics platform that enables **Data Engineering**, **Machine Learning**, and collaborative data science at scale using Apache Spark.



databricks

THE END.