THE INTELLIGENT FACTORY

Building an Efficient MLOps Pipeline

HELLO!

I am Chi [Chaichana Thavornthaveekul]

Data Scientist at Data Wow

AGENDA

1. Introduction to Machine Learning
2. Definition of MLOps
3. MLOps Application
4. MLOps at Data Wow

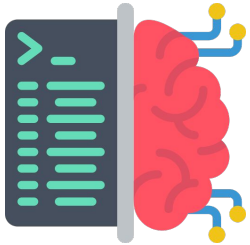
1

INTRODUCTION TO MACHINE LEARNING

WHAT IS MACHINE LEARNING?

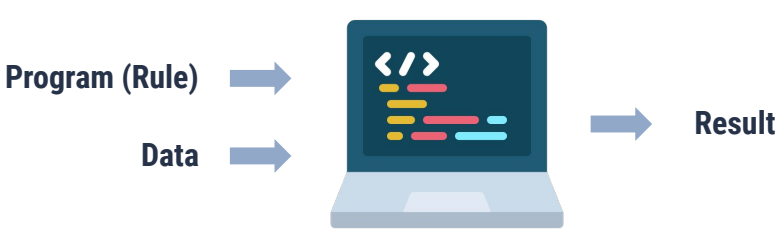
Traditional Programming

V/S



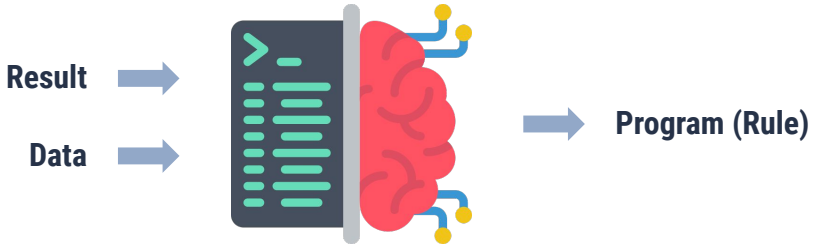
Machine Learning

WHAT IS MACHINE LEARNING?



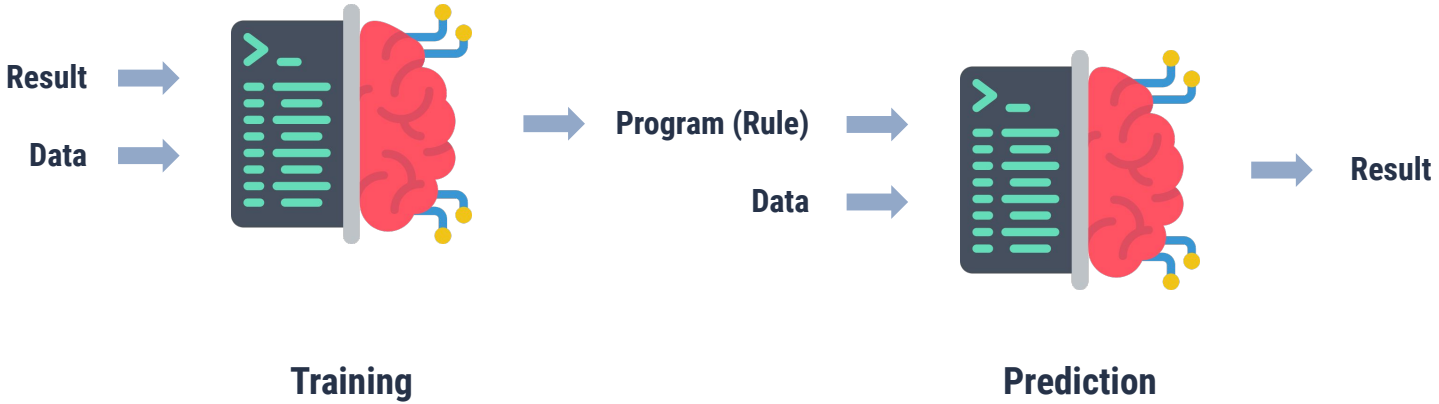
Traditional Programming

V/S



Machine Learning

WHAT IS MACHINE LEARNING?



TYPES OF MACHINE LEARNING

4 TYPES OF
MACHINE LEARNING

01 Supervised Learning
Classification
Regression

02 Unsupervised Learning
Clustering
Anomaly Detection

03 Semi-Supervised Learning
Self-training (Pseudo-label)
Image Generation

04 Reinforcement Learning
Robotic
Autonomous Driving

CURRENT STATE OF MACHINE LEARNING

- Focusing mostly on building ML models
- Operationalization was an afterthought

LAST DECADE

TODAY

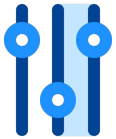
- Only **53%** of PoCs are built into production
- Taking an average of 9 months

BY END OF 2024

- **75%** of organisations will shift from piloting to operationalising AI
- Driving a **400%** increase in streaming data and analytics infrastructures

Ref: Gartner: accelerating AI deployments — paths of least resistance

WHY MOST ML PROJECTS ARE NOT MADE INTO PRODUCTION?



Production

Model Development

- Data collection
- Data exploration
- Model training
- Model evaluation

Packaging

- Dependencies
- Dockerization
- Productionisation
- Deployment

Scaling-out

- Load-balance
- Data partitions
- Model distribution

Tuning

- Parallelism
- GPU support
- Query tuning
- Catching

Instrument

- Monitoring
- Logging and Tracing
- Versioning
- Security

Automation

- CI/CD
- Workflows and Pipelines
- Rolling upgrades
- A/B testing

Weeks

with **one** data scientist

Months

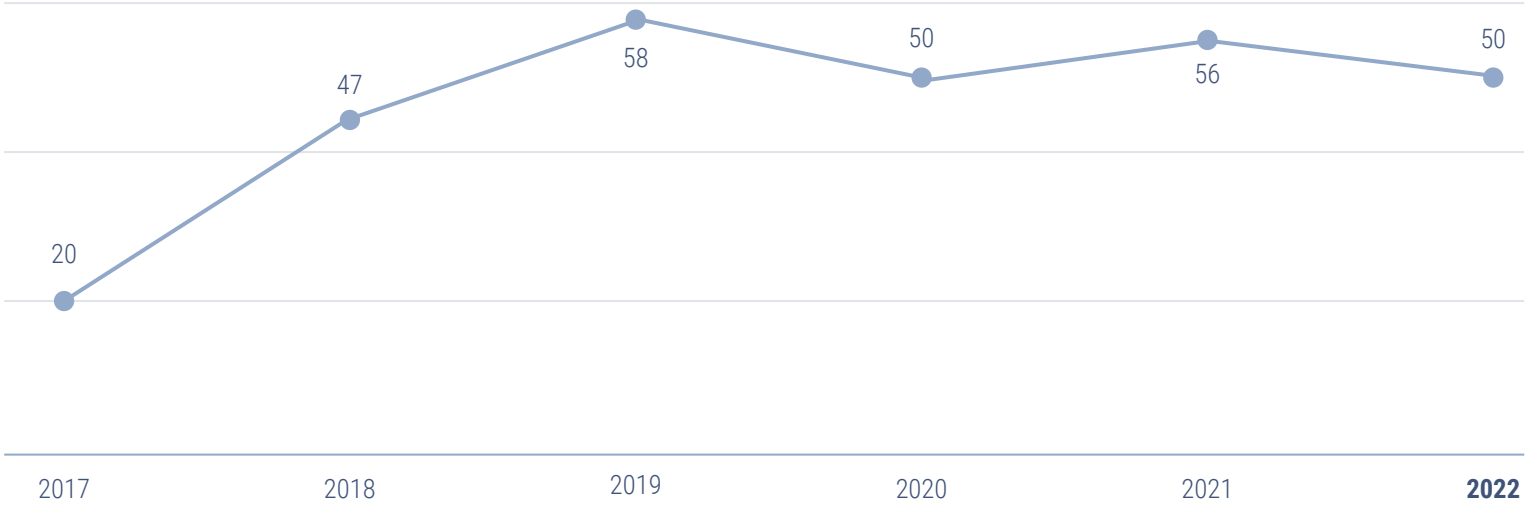
with **a large team** of data scientists, data engineer, ML engineer, software engineer, and DevOps

87%

of data science projects will never make it into production

Ref: VentureBeat

Share of respondents who say their organizations have adopted AI in at least one business unit, %



Ref:
https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review#review

2

DEFINITION OF MLOPS

> ***MLOps (Machine Learning Operations)*** *is a paradigm, including aspects like best practices, sets of concepts, as well as a development culture when it comes to the end-to-end conceptualization, implementation, monitoring, deployment, and scalability of* ***machine learning*** *products*

Source: Machine Learning Operations (MLOps): Overview, Definition, and Architecture (Kreuzberger et al., 2022)

MACHINE LEARNING PROCESS

Business Problem

Machine Learning Problem Framing

Data Collection

Data Cleansing and Preparation

Data Exploration

Data Visualization and Analysis

Feature Engineering

Model Training

Model Evaluation

Monitoring

Model Prediction

Model Deployment

Re-training

Researching

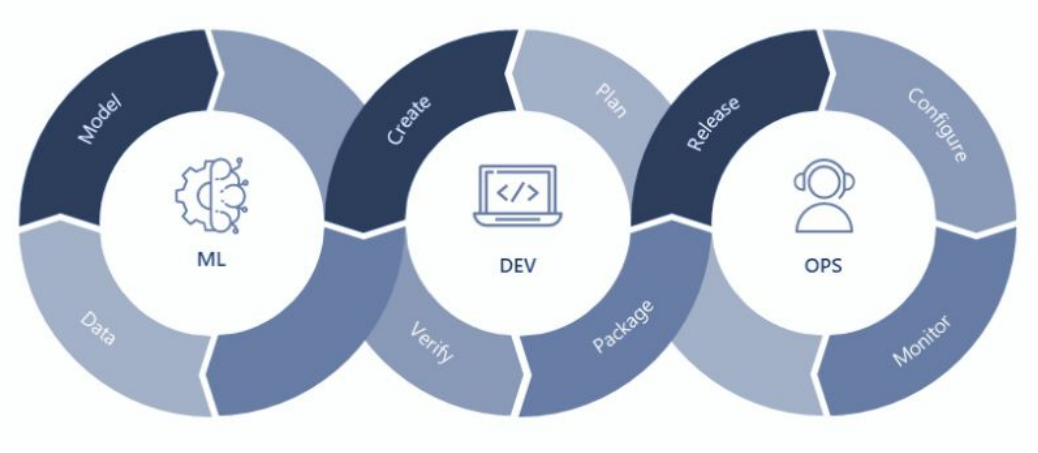
Feature Augmentation

NO

YES

Archive business goals?

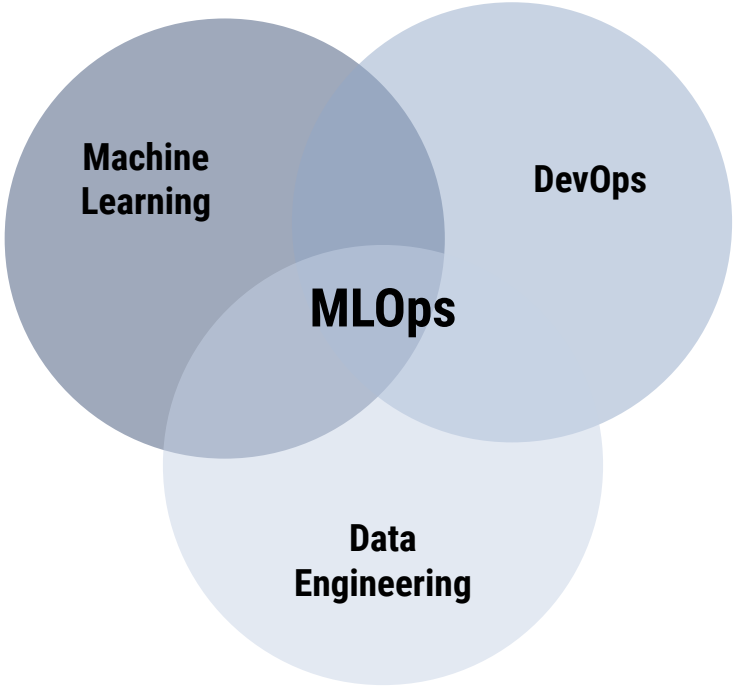
WHAT IS MLOPS?



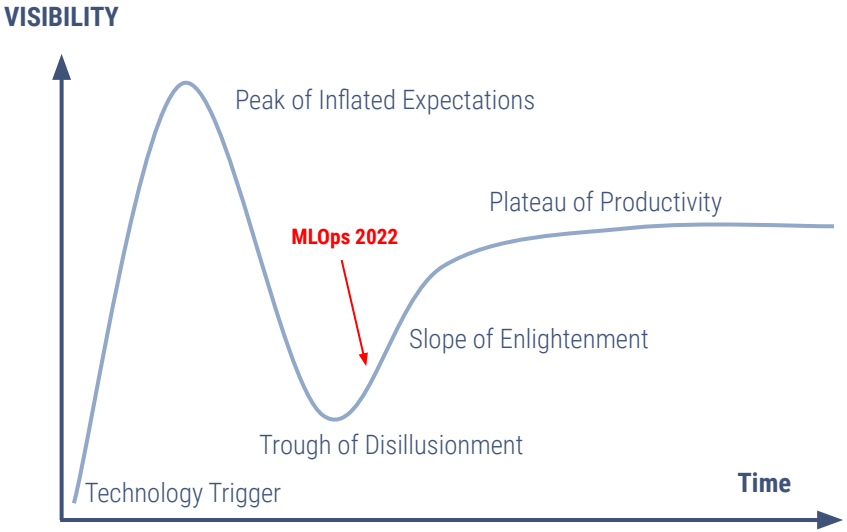
- **Collaborative and experimental in nature**
- **Automate as much as possible**
- **Continuous improvement of ML models**
- **Standardize and Scale**

Ref: https://blogs.nvidia.com/wp-content/uploads/2020/09/MLOps-Neal-Analytics-672x296.png.webp

COLLABORATION OF MLOPS



Gartner Hype Cycle



Ref: https://www.mihaileric.com/posts/mlops-is-a-mess/

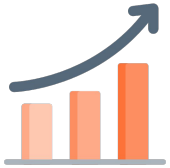
IMPORTANCE OF MLOPS

Agility

- Continuous and faster deliveries
- Faster modifications
- Faster bug-fixing

Experiments

- Faster and Controlled Experiments
- Faster Integration of successful experiments to other environments



Scalability

- Ease integration of new ML model
- Standardization of code
- Lower operational costs

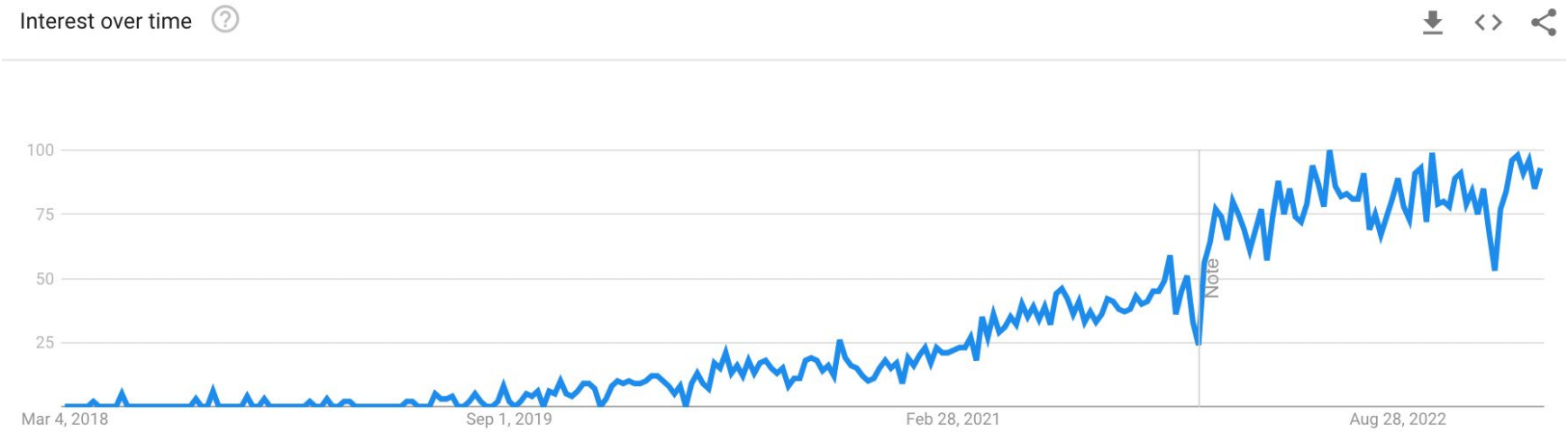
Time to Market

- Reduced time-to-market
- Faster planning and deliver expectations

Business Owners

- Strong collaboration
- Improve iterations

GOOGLE TRENDS (MLOps)



CHALLENGES OF MLOPS

Defining the business requirements

Unrealistic expectations

Misleading success metrics

Machine Learning challenges

Data available

Current study

Data Drift

Productionisation challenges

Infrastructure

Monitoring

Logging & Tracing

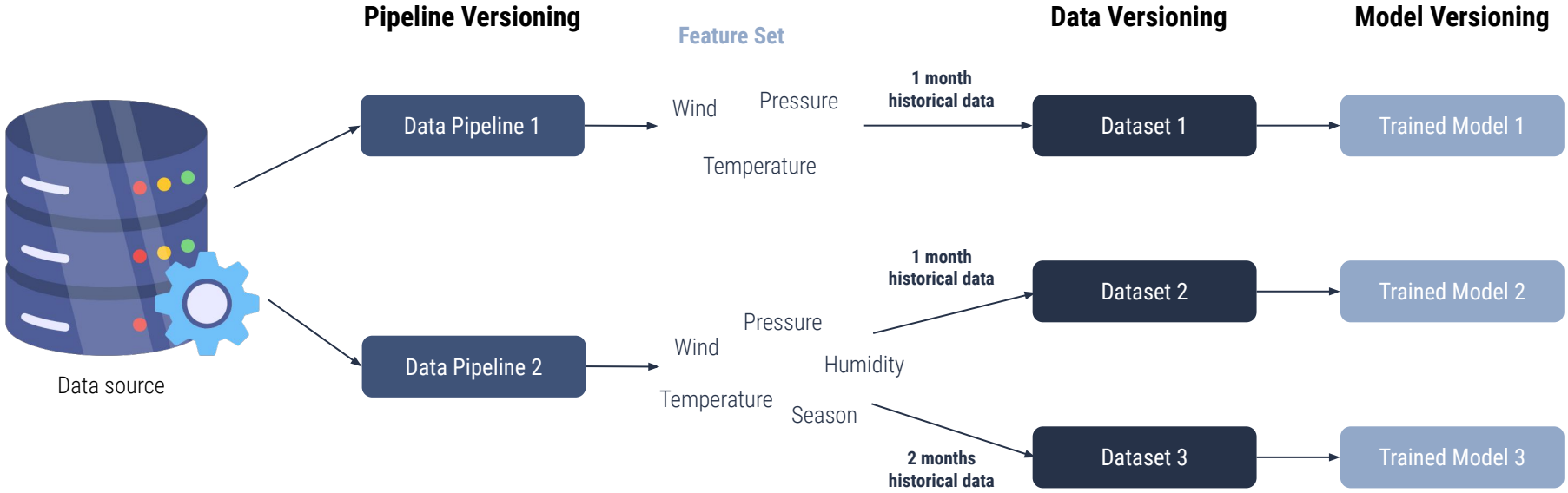
Many dependencies

Data consistency

Business need shift

Data/Pipeline/Model versioning

Dependencies



3

MLOPS APPLICATION

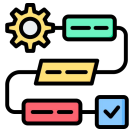
> ***Don’t be afraid to launch a product without machine learning***

1st rule of Machine Learning’s Google

START APPLYING ML TO CURRENT PROCESS

PROBLEM FINDING

- Current Process Exploration
- Process Bottleneck Finding



ML RESEARCH

- Possibility of Machine Learning Capability
- Current Machine Learning Study
- Data Availability
- Possibility of Data Gathering

BUSINESS VALUABLE

- ROI in Business Term
- Company Vision and Mission Alignment
- Business Impact

Start ML Project

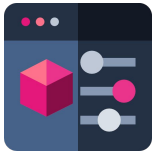
HOW DO WE KNOW ML PROJECT SHOULD START?

Ref: Skooldio

WHAT'S ABOUT MLOPS?

Encourage collaboration among teams

Automate as much as possible as early as possible



Make sure teams have **the right tools**

Monitor and retrain your models continuously

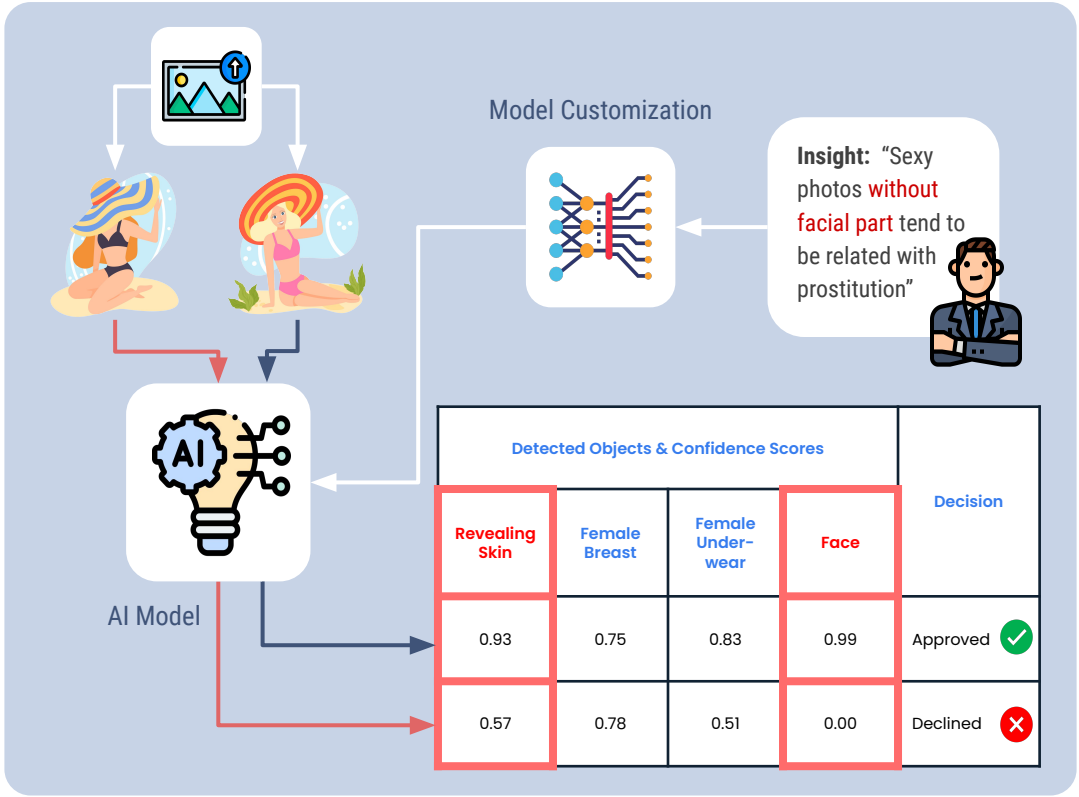
4

MLOPS AT DATA WOW

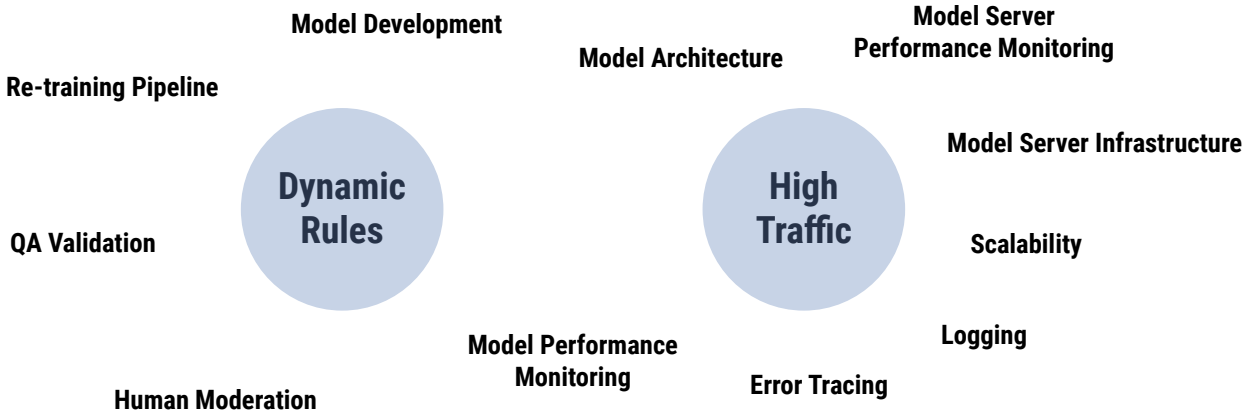
Posmoni Content Moderator

Business Requirements

- **Dynamic set of rules**
- More than **35 million accounts** across all platforms and more than **4 million images** every month!



REQUIREMENT UNDERSTANDING



Data Preparation

Data Gathering: Application, Google Image Search, Public Dataset
Data Cleansing
Data Labeling: Accurately, RoboFlow, Label Studio

accurately

Model Development

Model Training: TensorFlow, PyTorch, Pre-trained Model, On-Premise Server (On-Cloud)
Model Evaluation
Model Versioning: Database, Data Table Format
Model Registry: DVC, Local Disk, AWS S3, Cloud Storage
Experiment Tracking: TensorBoard, Streamlit, WandB, MLflow, Kubeflow

DVC mlflow W&B Kubeflow

Model Deployment

Model Prediction Development: GitHub (Code Version Control)
CI/CD: CircleCI, Jenkins
Server: On-Cloud (AWS)
Message Queue: SQS, KafKa
Deployment: Kubernetes, Docker, Canary, Scaling, Load-balancing

aws

Maintaining

Re-training: Airflow (Pipeline)
Model Performance Monitoring: Airflow (Weekly Data Extraction), Streamlit
Model Server Monitoring: New relic
Logging & Tracing: New relic, Django + PostgreSQL

THANKS!

Any questions?

You can find me at

chaichana.t@datawow.io

Chaichana Thavornthaveekul

datawow

Scan me

MLOps Made Simple เรื่องต้องรู้ เมื่อต้องจัดการ ML กับ คุณชัยชนะ ถาวรทวีกุล | Tech Monday

Scan me