

Seminar

Foresight in Data Science



Lecturer:
Sajad Heydari
Aspiring Data Scientist
MSc. Industrial Engineering





❖ MY JOURNEY

MSc. Industrial Engineering, Tarbiat Modares University
Data Science Researcher especially in Intelligent Transportation Systems
Data Science Courses Instructor in Iranian top-notch Universities
Data Scientist at Snapp!Food
Data Science Consultant in Industries: Healthcare, Supply Chain, Transportation,...
Marketing Data Analyst in Insurance Co.



Importance of Data

180 zettabytes Data generated every minutes **NOW!**

ONLY 0.5% of generated Data are analyzed



64,140
Instagram stories
posted



336,480
Skype calls



567,360
Tweets sent



5,365,260
YouTube videos
watched



5,500,560
Google searches
conducted



181,331,340
Emails sent

Picture source: SAS

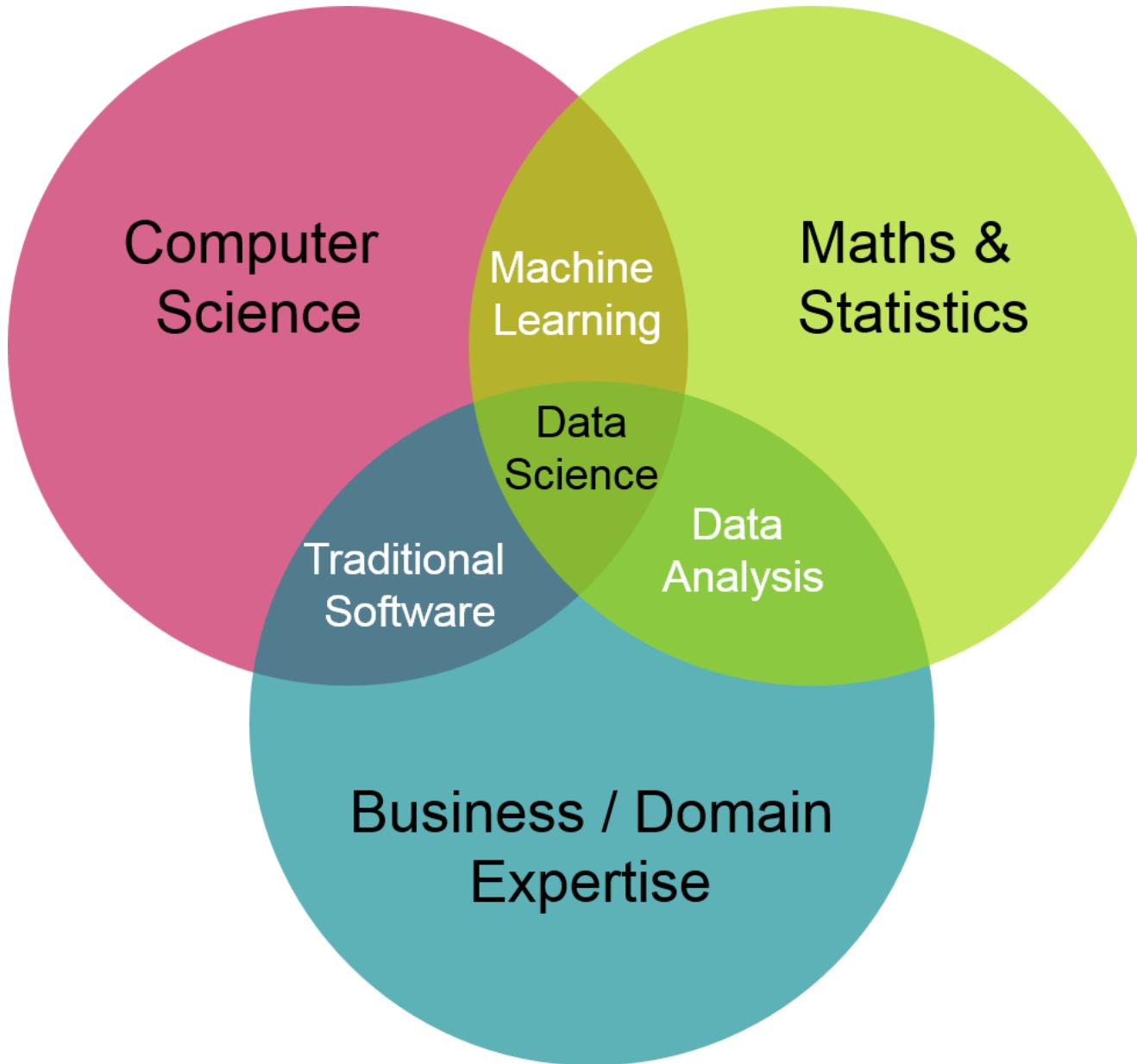
What is Data Science?



Data science is a
Multidisciplinary approach to
extracting **actionable insights**.



What is Data Science?



WHY LEARN **DATA SCIENCE**

Fuel of 21st Century

Problem of Demand & Supply

A Lucrative Career

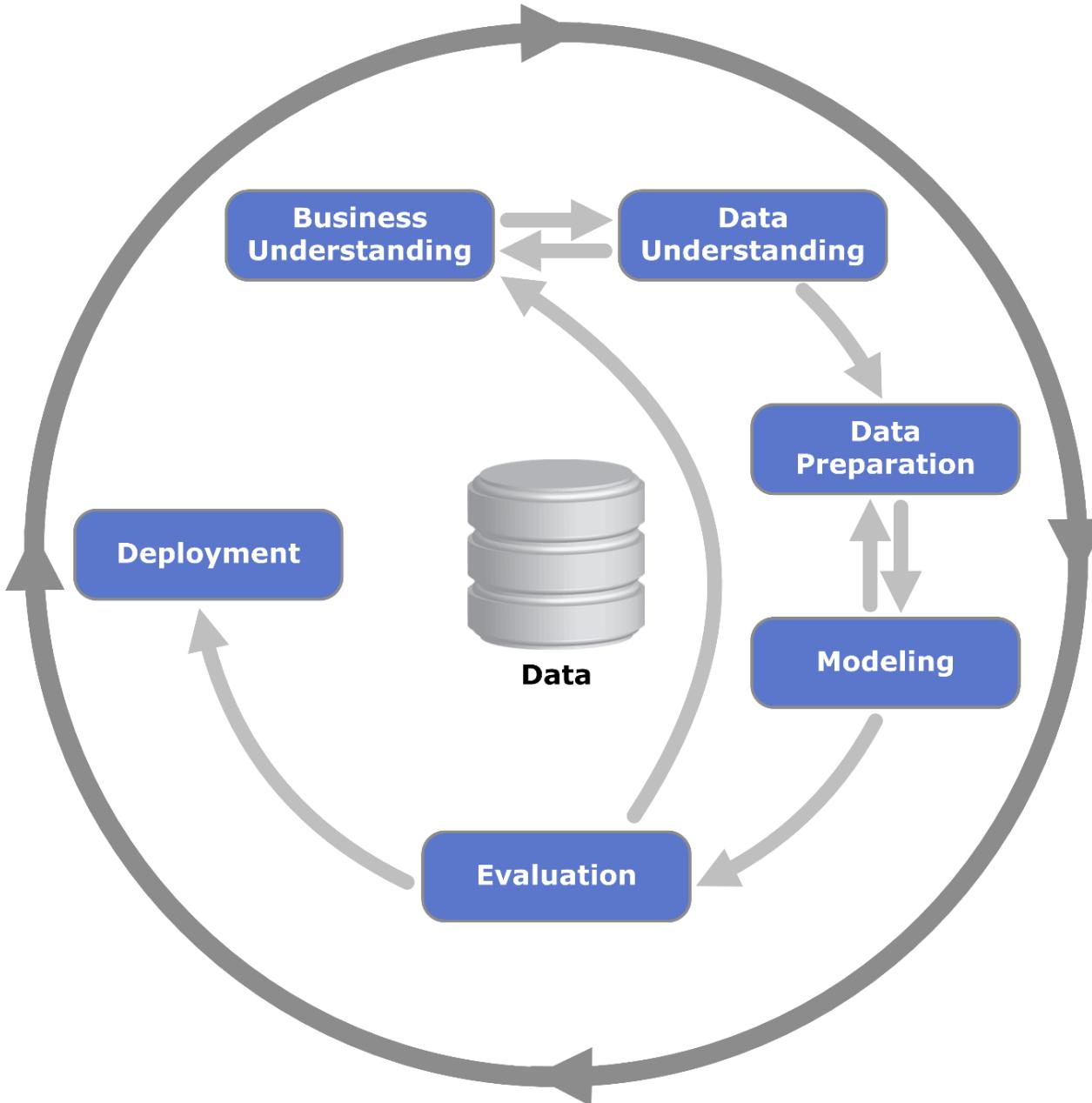
Data Science is Changing the world

Data Science Is The Future



Crisp-DM Methodology

Cross-Industry Standard
Process for Data Mining

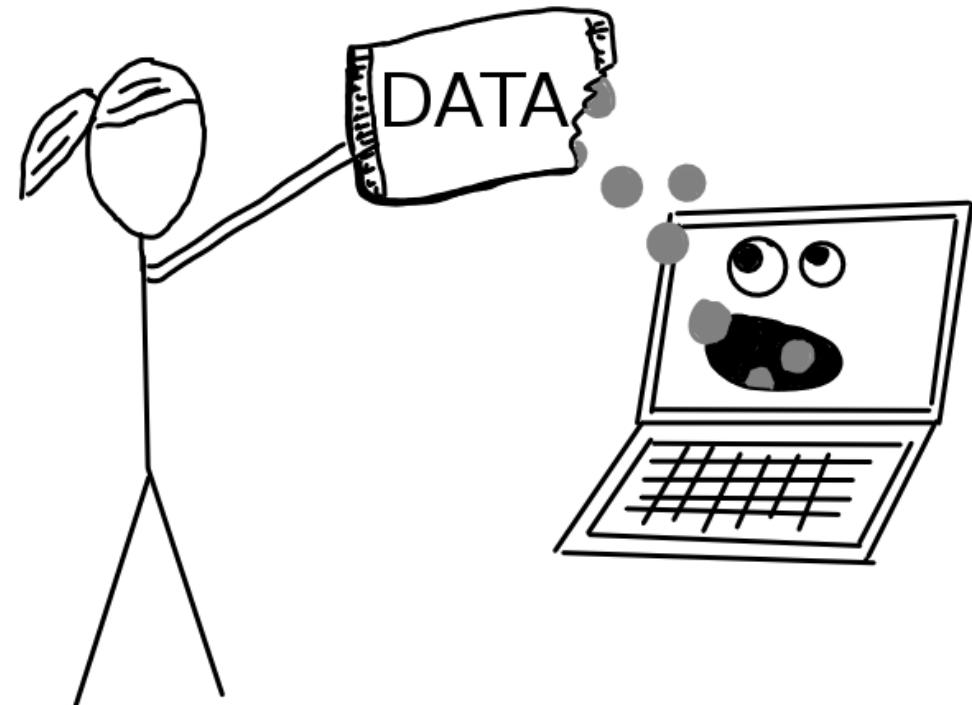


What is Machine Learning?

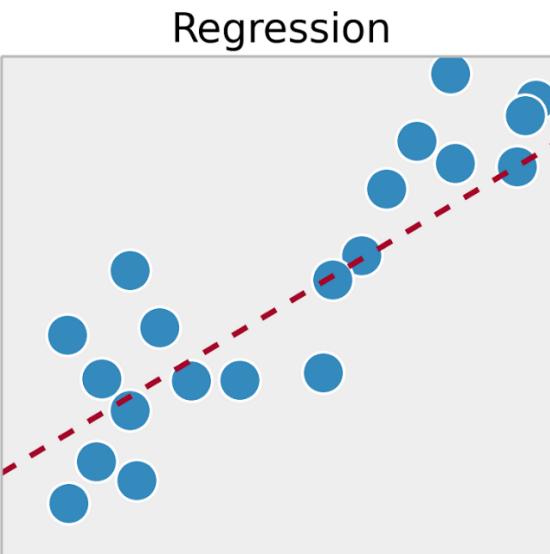
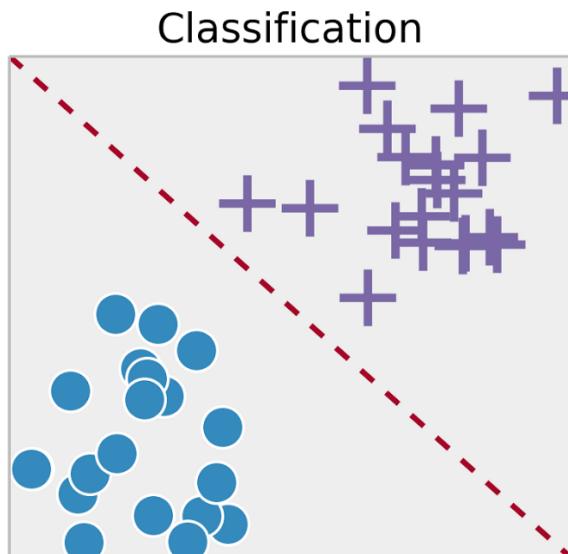
Without Machine Learning



With Machine Learning



What is Supervised Learning?



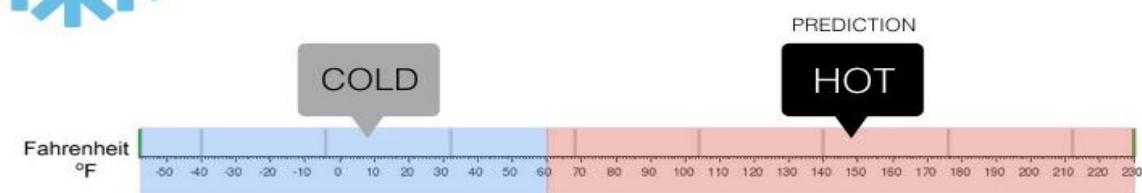
Regression

What is the temperature going to be tomorrow?



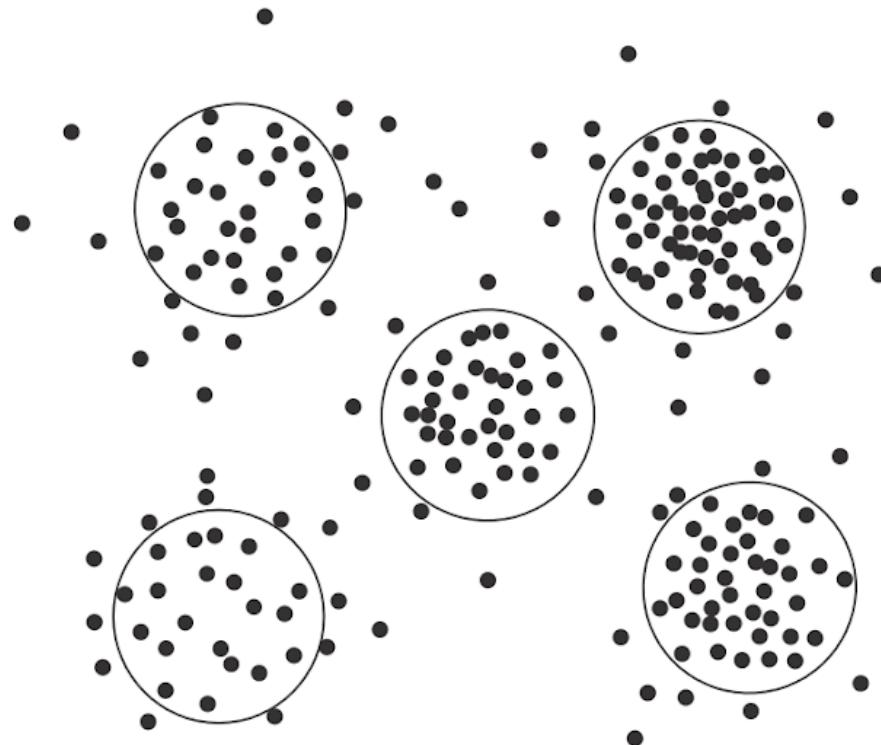
Classification

Will it be Cold or Hot tomorrow?

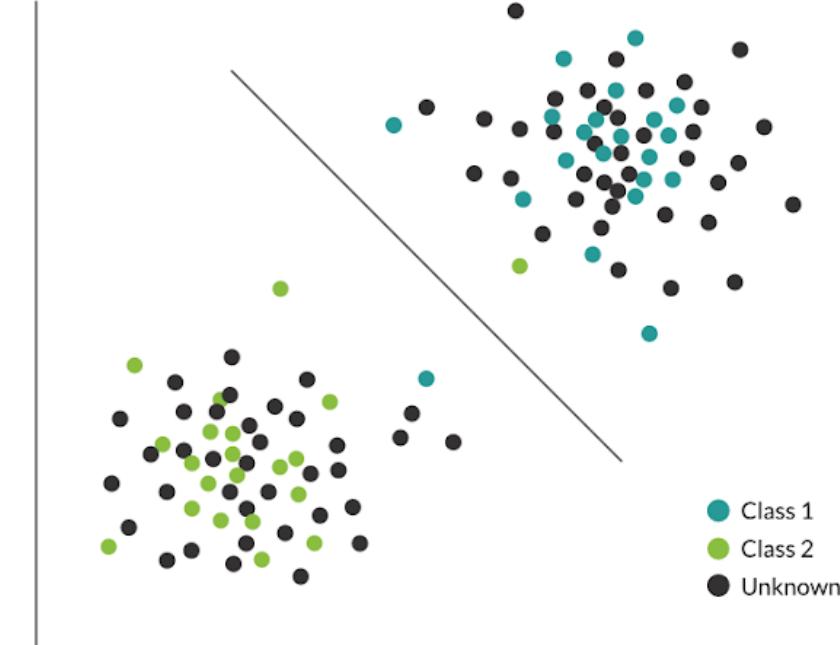


What is Unsupervised Learning?

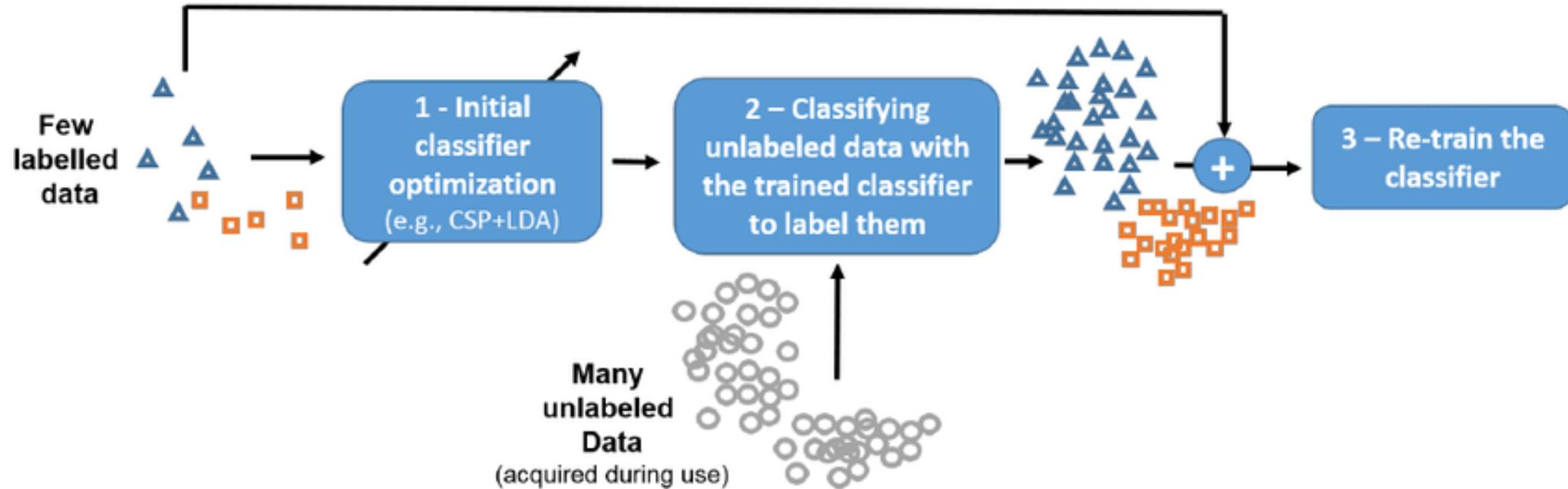
Unsupervised



Supervised

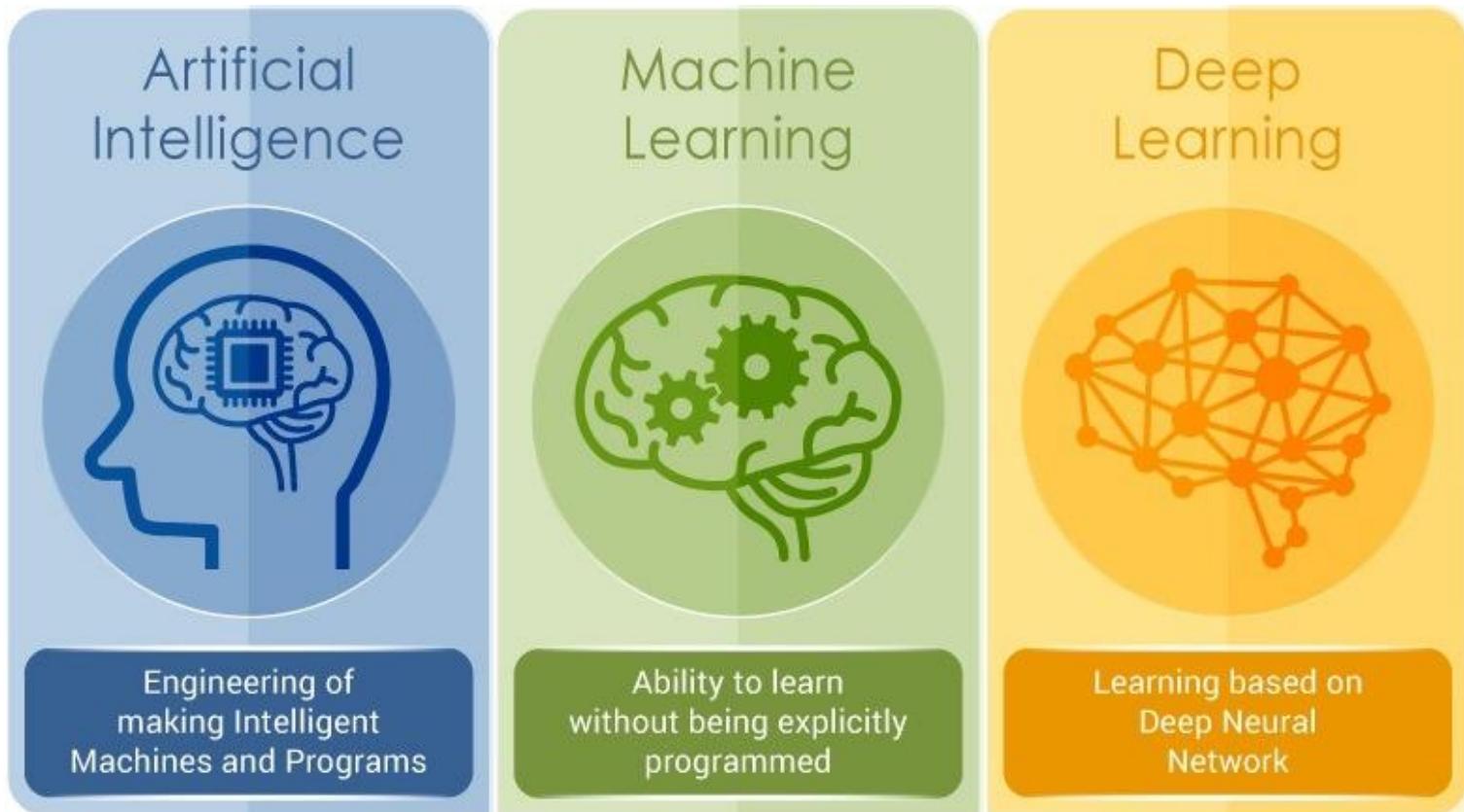
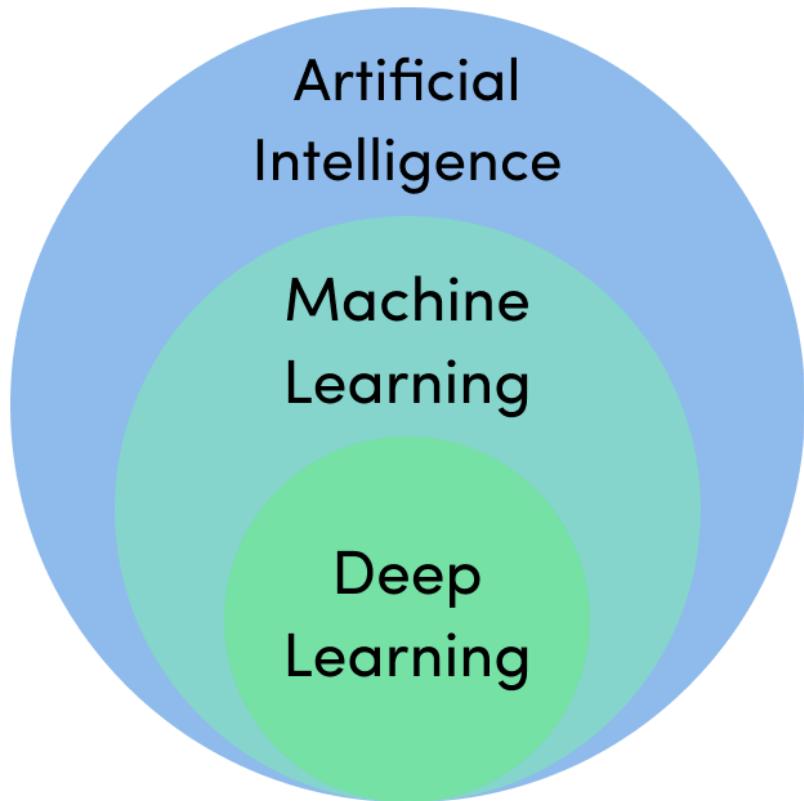


What is Semi-Supervised Learning?



Source: Lotte, F., (2015). Signal Processing Approaches to Minimize or Suppress Calibration Time in Oscillatory Activity-Based Brain–Computer Interfaces, Proceedings of the IEEE, 103(6), pp. 871-890.

Deep Learning

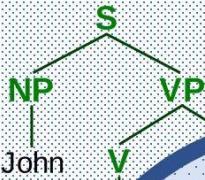
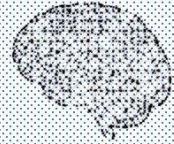


Application of Data Science

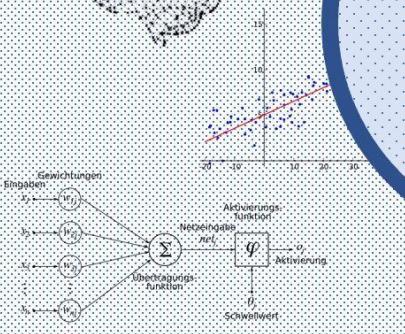




aisoma.de



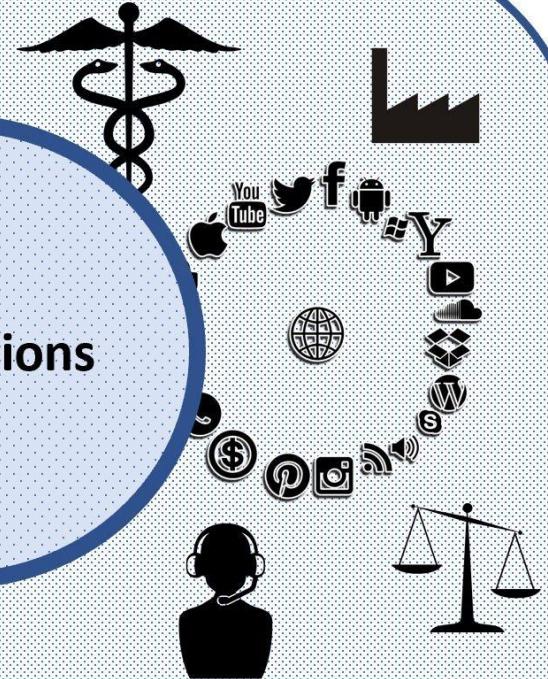
Algorithms



- Massive-scale Graph
- Geospatial Temporal Predictive Analytics
- Hyperfast Analytics
- Embedded Deep Learning
- Natural Language Generation
- Structured Database Generation
- Cognitive Machine & Federated Learning
- Fully automated Machine Learning



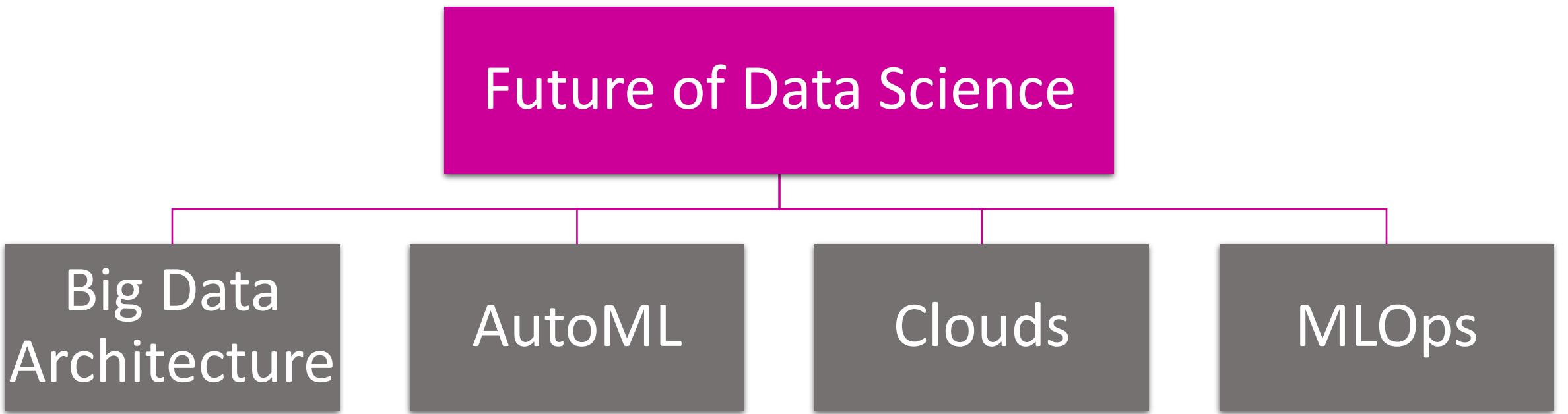
Applications



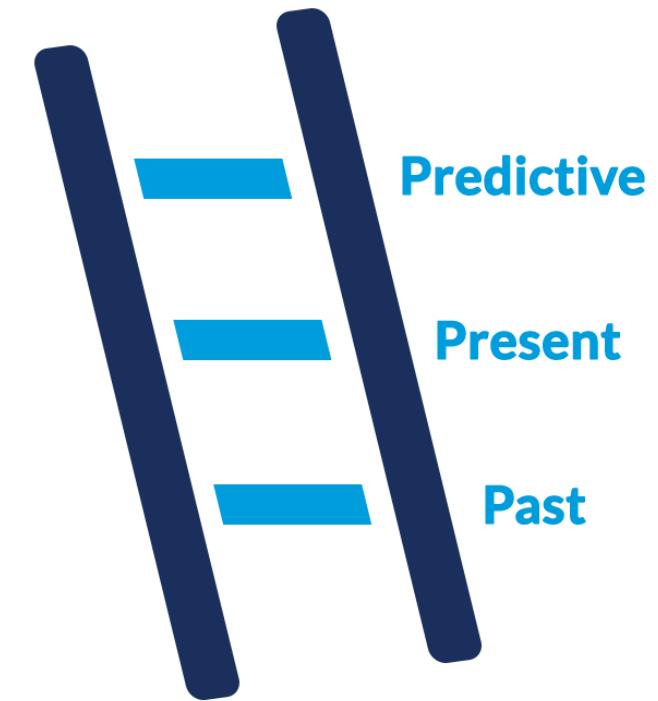
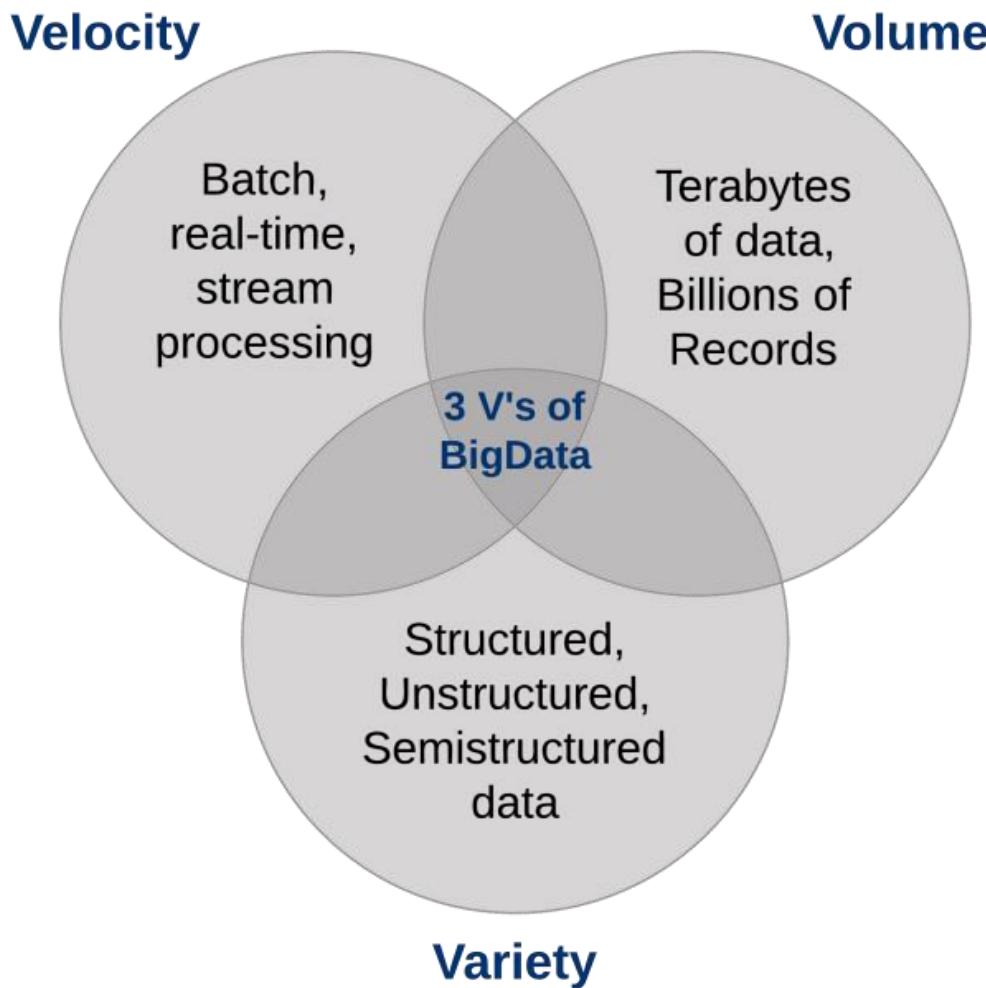
- Cybersecurity
- Internet of Things (Smart City, Industry 4.0)
- Healthcare
- Compliance
- Customer Engagement & Experience
- Human Capital
- Data for Social Good
- Smart Everything

based on Kirk Borne (Booz Allen Hamilton) Slide

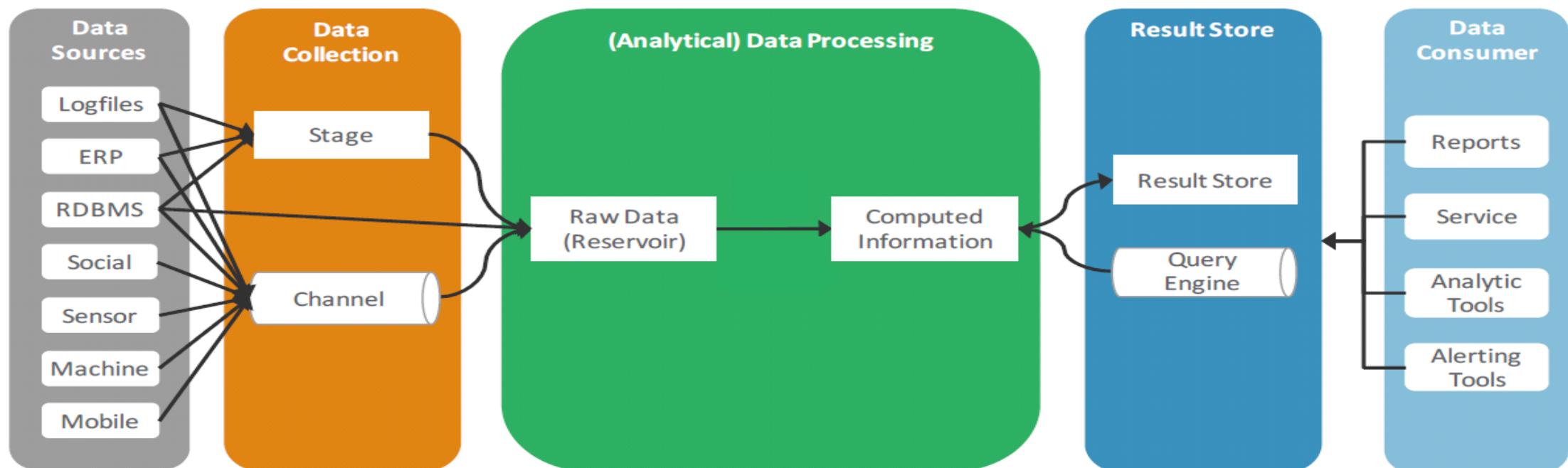
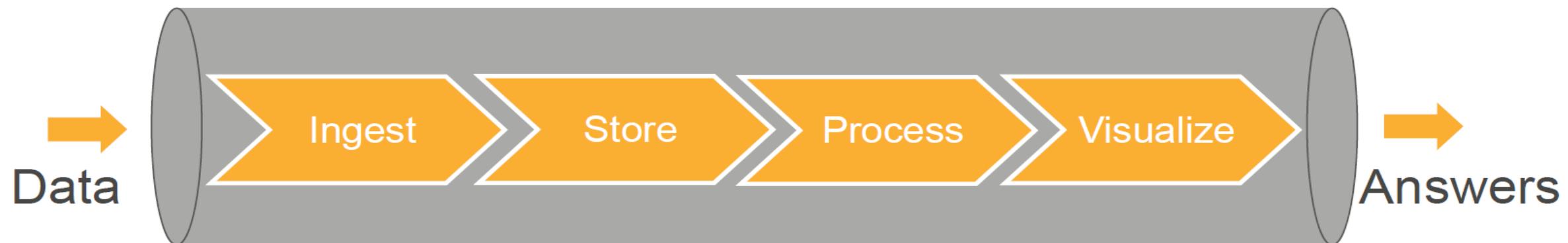
Future of Data Science



What is Big Data?



How Big Data Works?



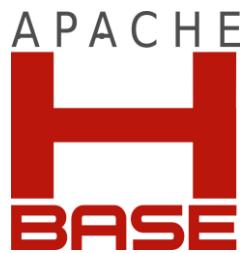
Big Data Layers: Data Ingestion

- Scalable, Extensible to capture streaming and batch data
- Provide capability to business logic, filters, validation, data quality, routing, etc. business requirements
- Technology Stack:
 - Apache Kafka
 - Apache Sqoop
 - Apache NiFi
 - Logstash
 - Fluentd
 - Facebook Scribe
 - Amazon Kinesis
 - ...



Big Data Layers: Data Storage

- Depending on the requirements data can placed into Distributed File System, Object Storage, NoSQL Databases, etc.
- Technology Stack:
 - HDFS, Hive
 - Redis, MongoDB, Hbase, Cassandra, ElasticSearch,...
 - RDBMS



Big Data Layers: Data Processing

- Processing is provided for batch, streaming and near real time cases
- Scale-Out Instead of Scale-Up
- Fault-Tolerant based methods
- Technology Stack:
 - MapReduce
 - Spark
 - Strom

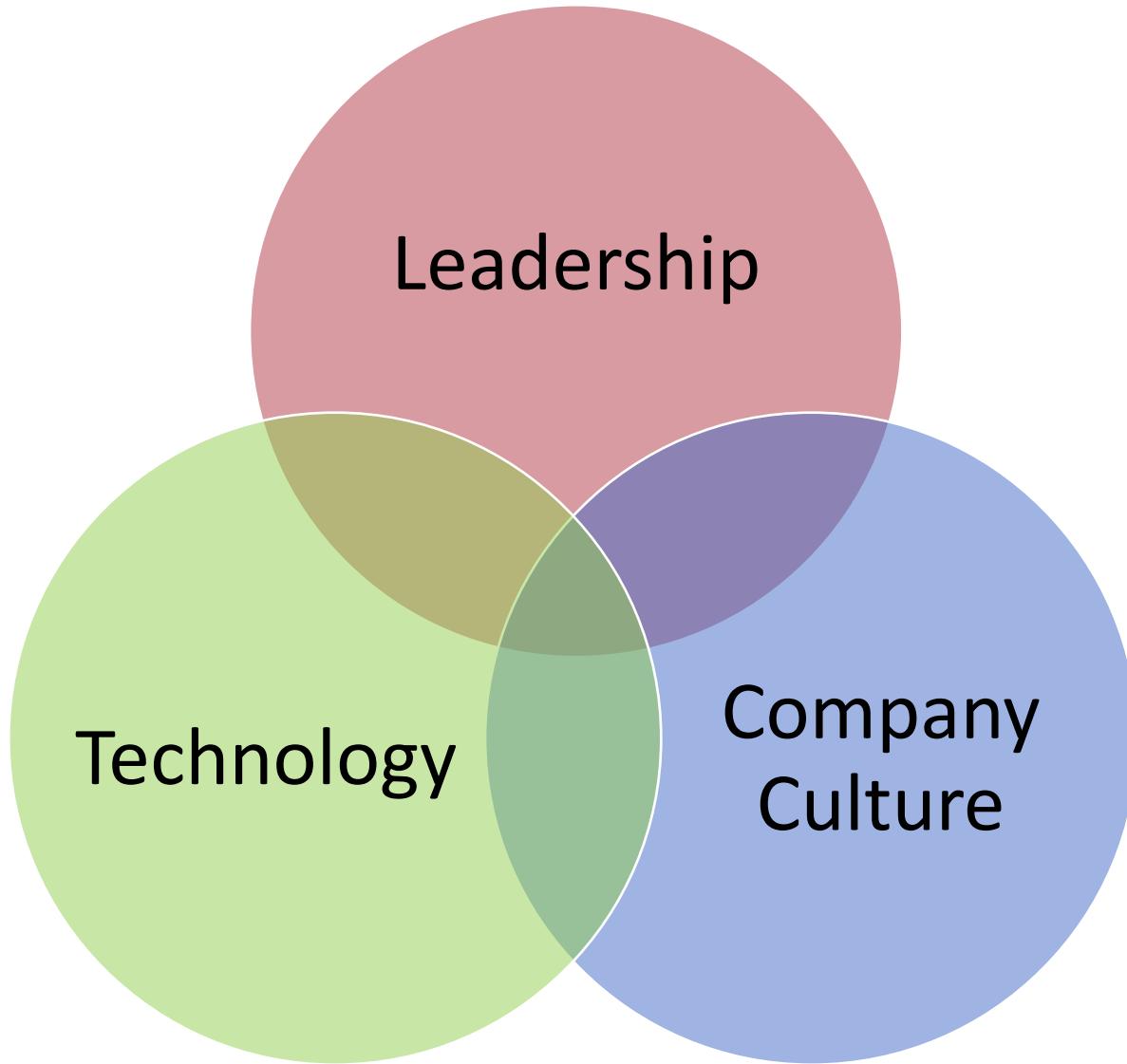


Big Data Layers: Data Visualization

- Dashboard and applications that provides valuable business insights
- Technology Stack:
 - Qlik
 - Tableau
 - Google Data Studio

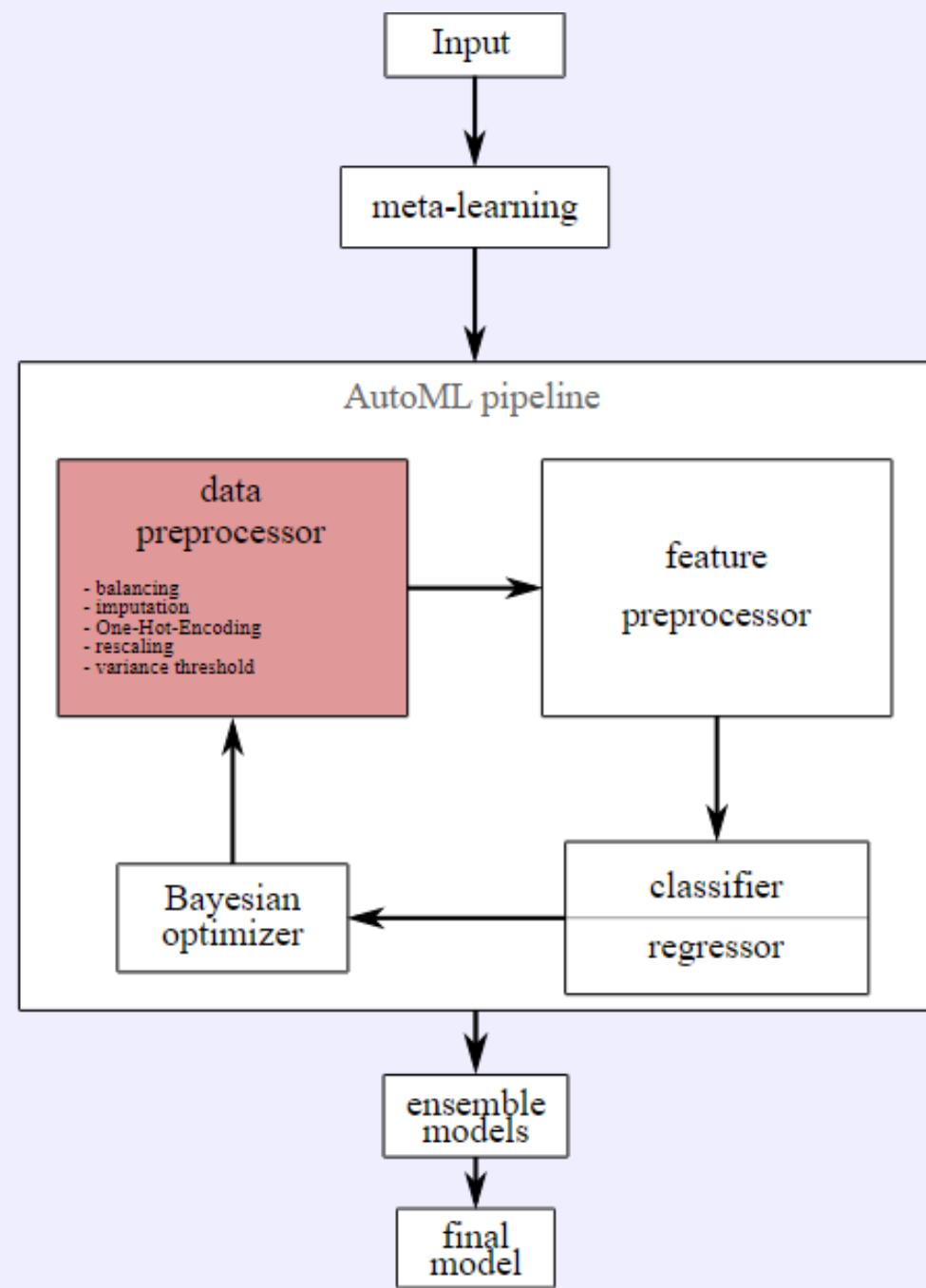


Business Challenges with Big Data

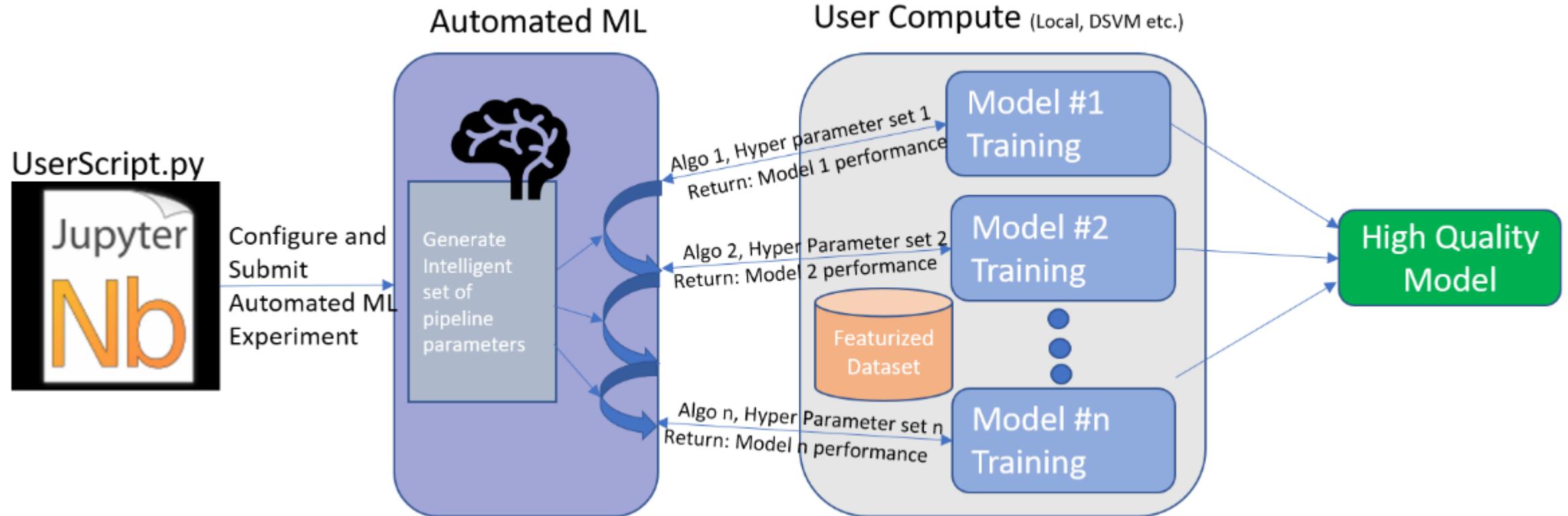


What is AutoML?

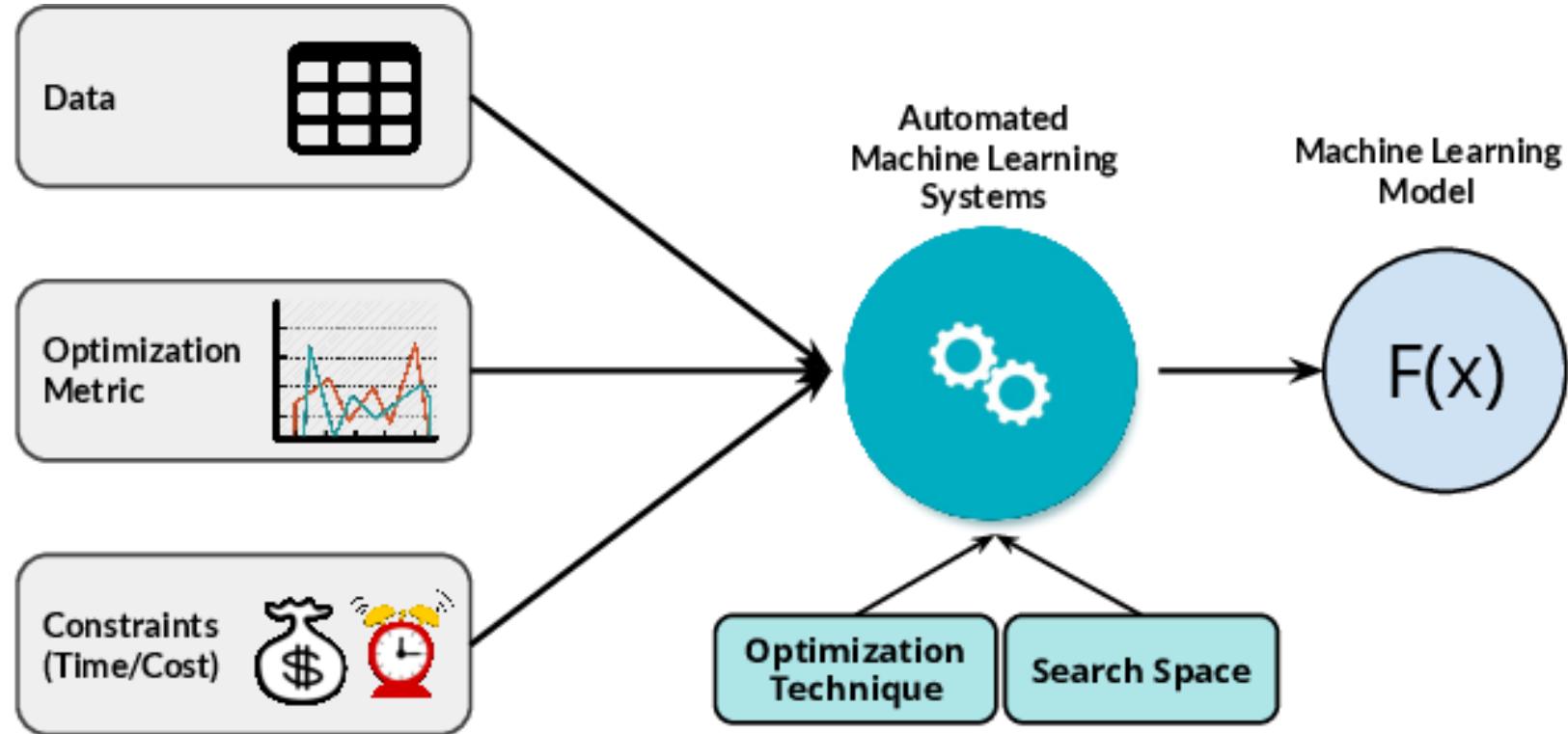
AutoML (automated machine learning) refers to the automated end-to-end process of applying machine learning in real and practical scenarios.



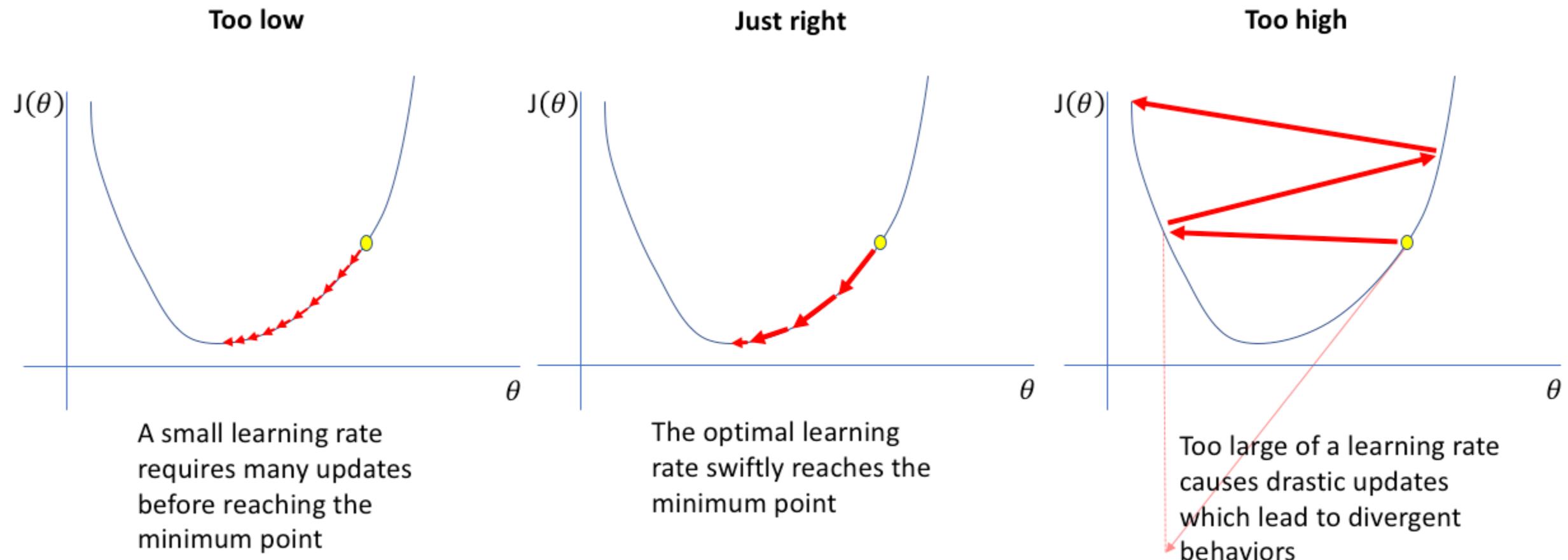
AutoML Process



AutoML Process



Optimization of Algorithms



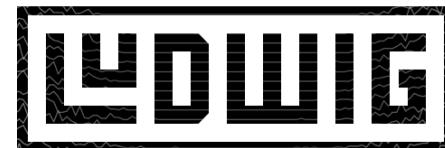
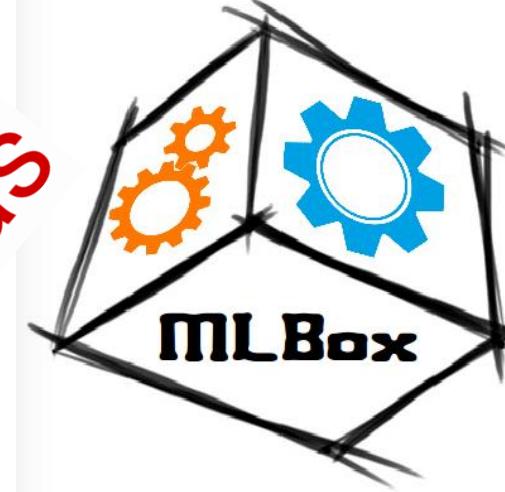
AutoML Frameworks

ML has been subject to the "black box" criticism

AutoML helps to make machine learning less of a black box by making it more accessible.



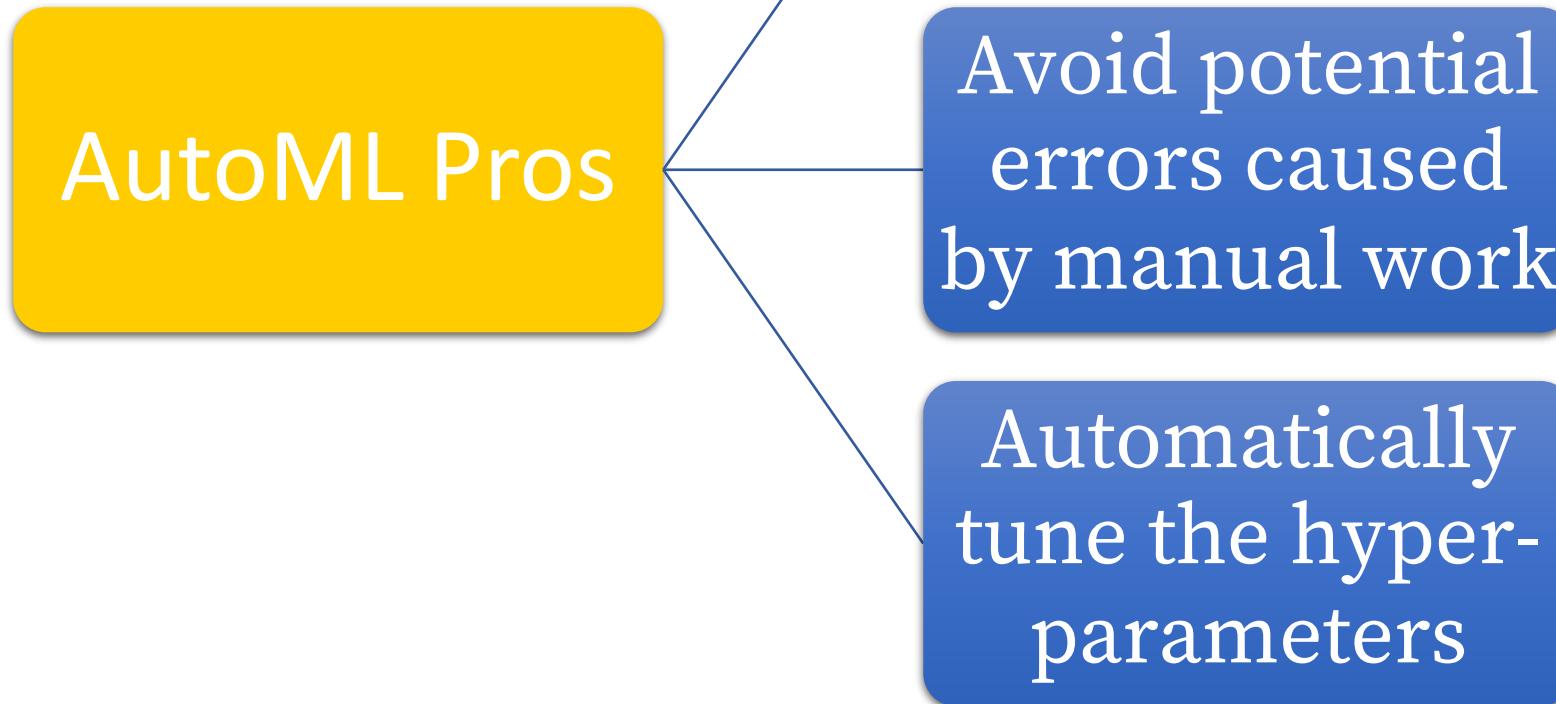
FLAML



AutoML Modern Frameworks

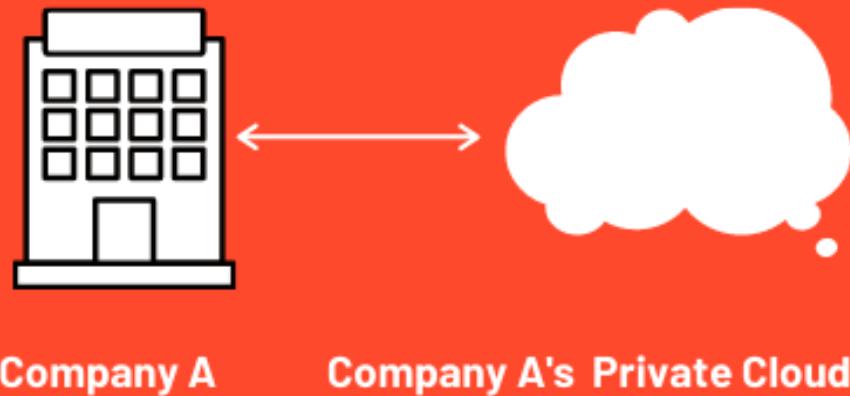


AutoML Advantages

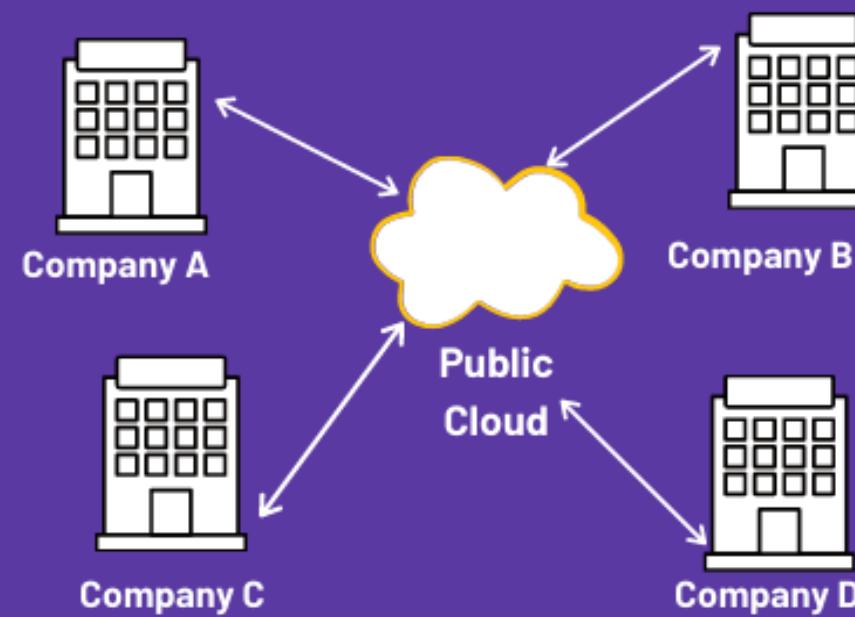


Cloud Systems

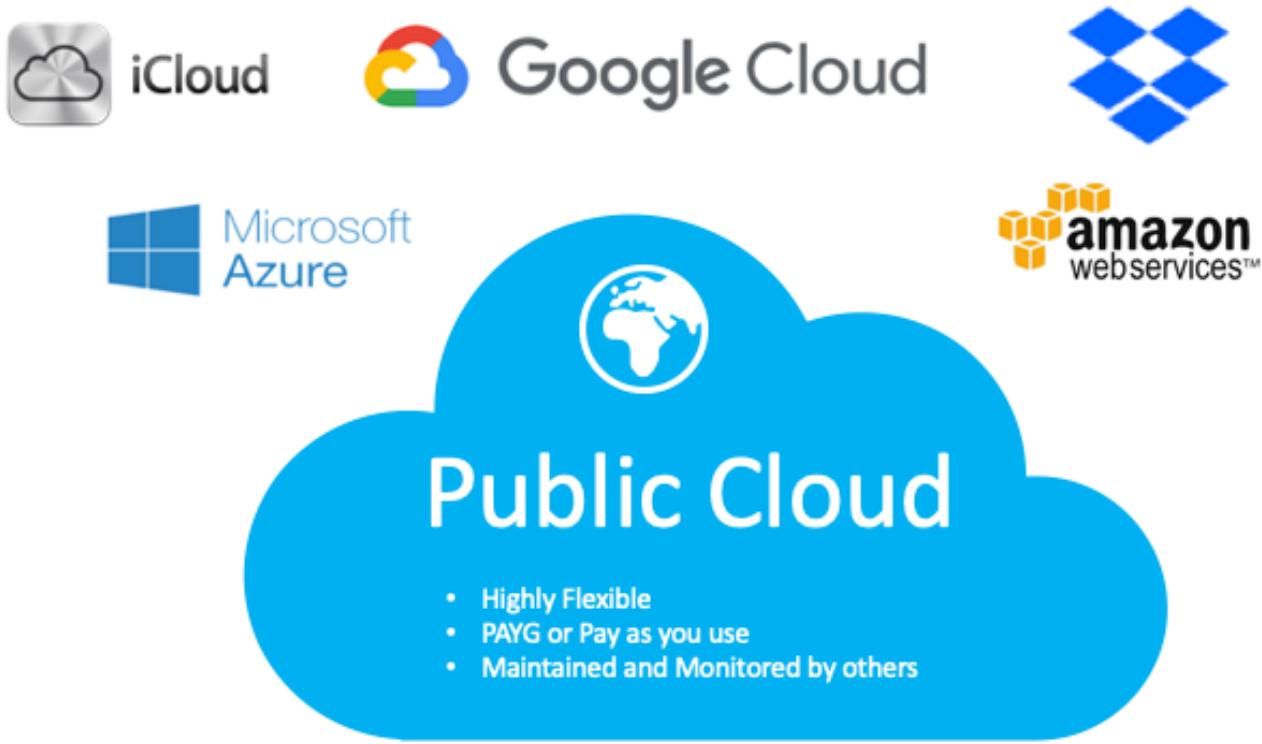
Private Cloud



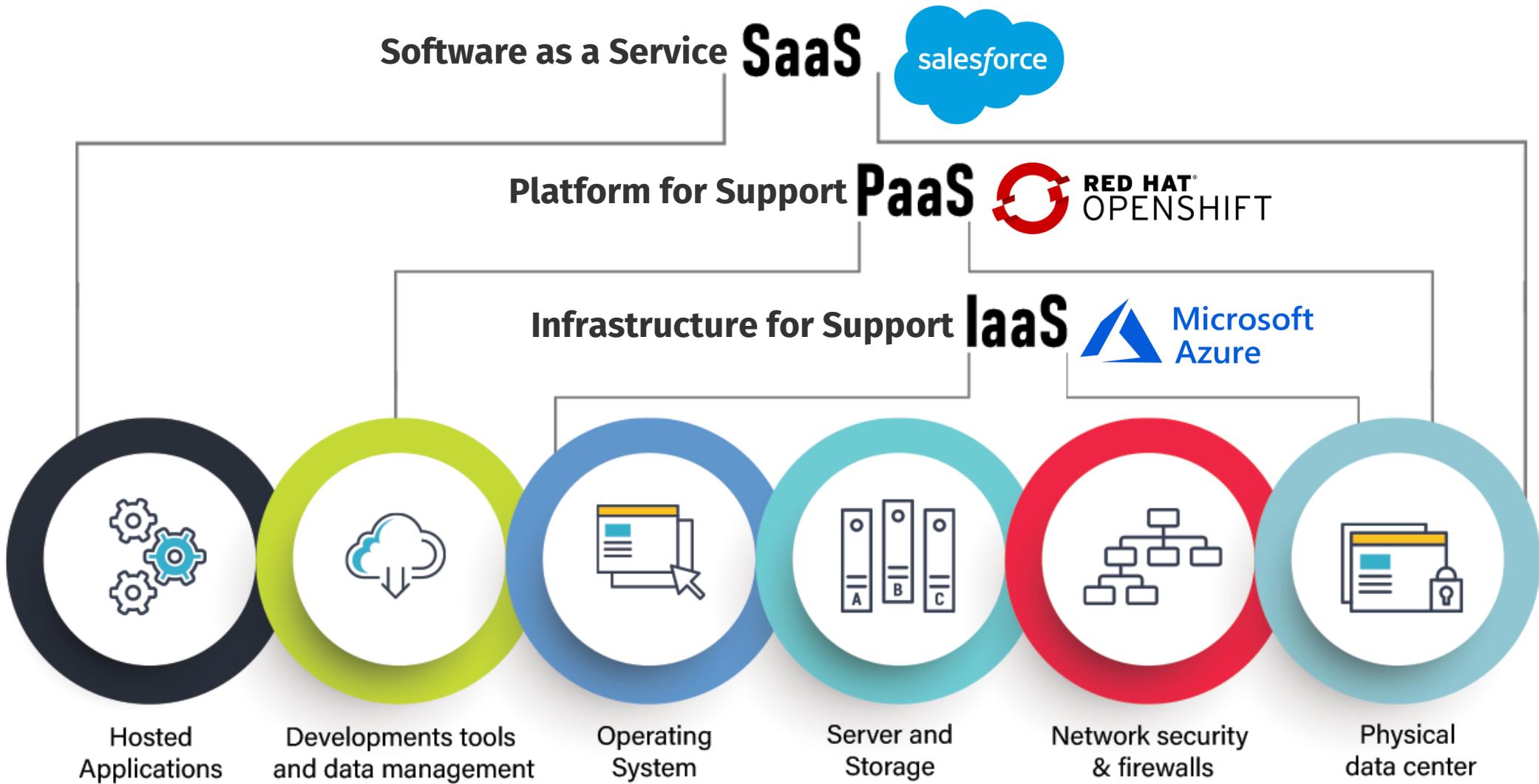
Public Cloud



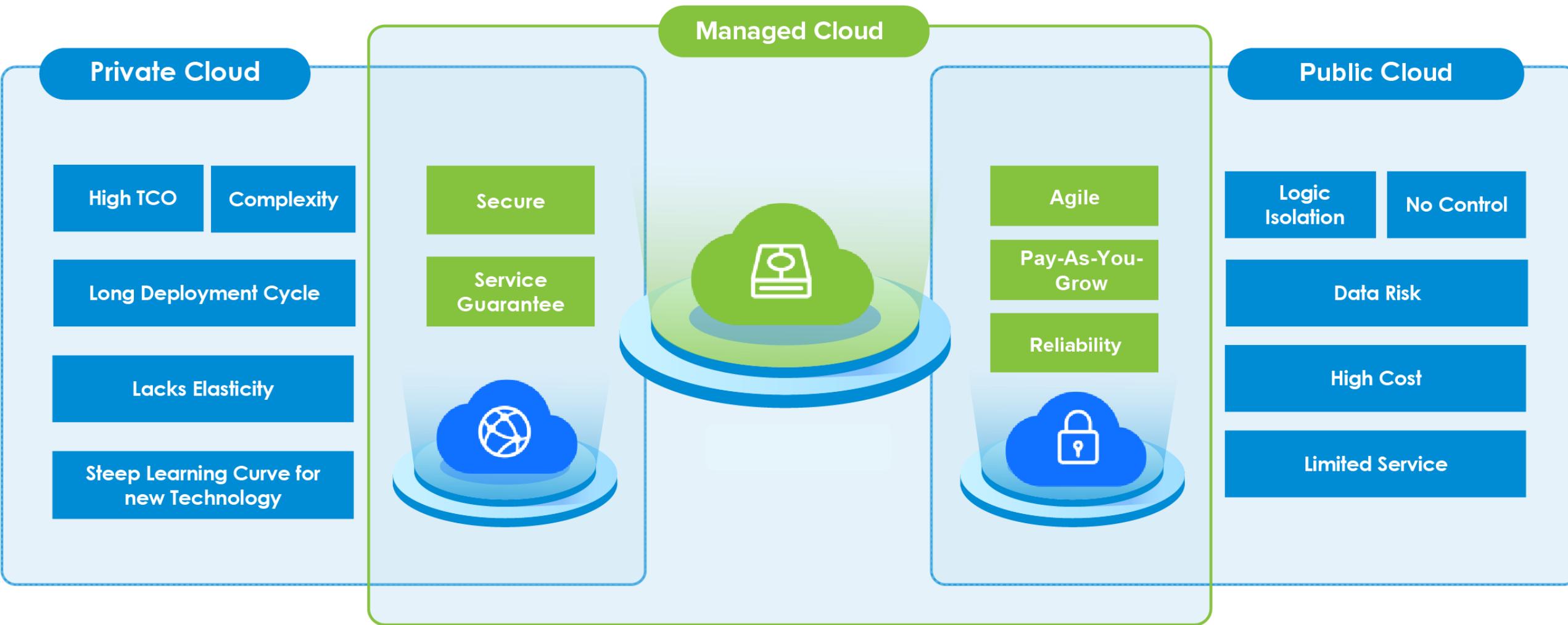
Public Cloud



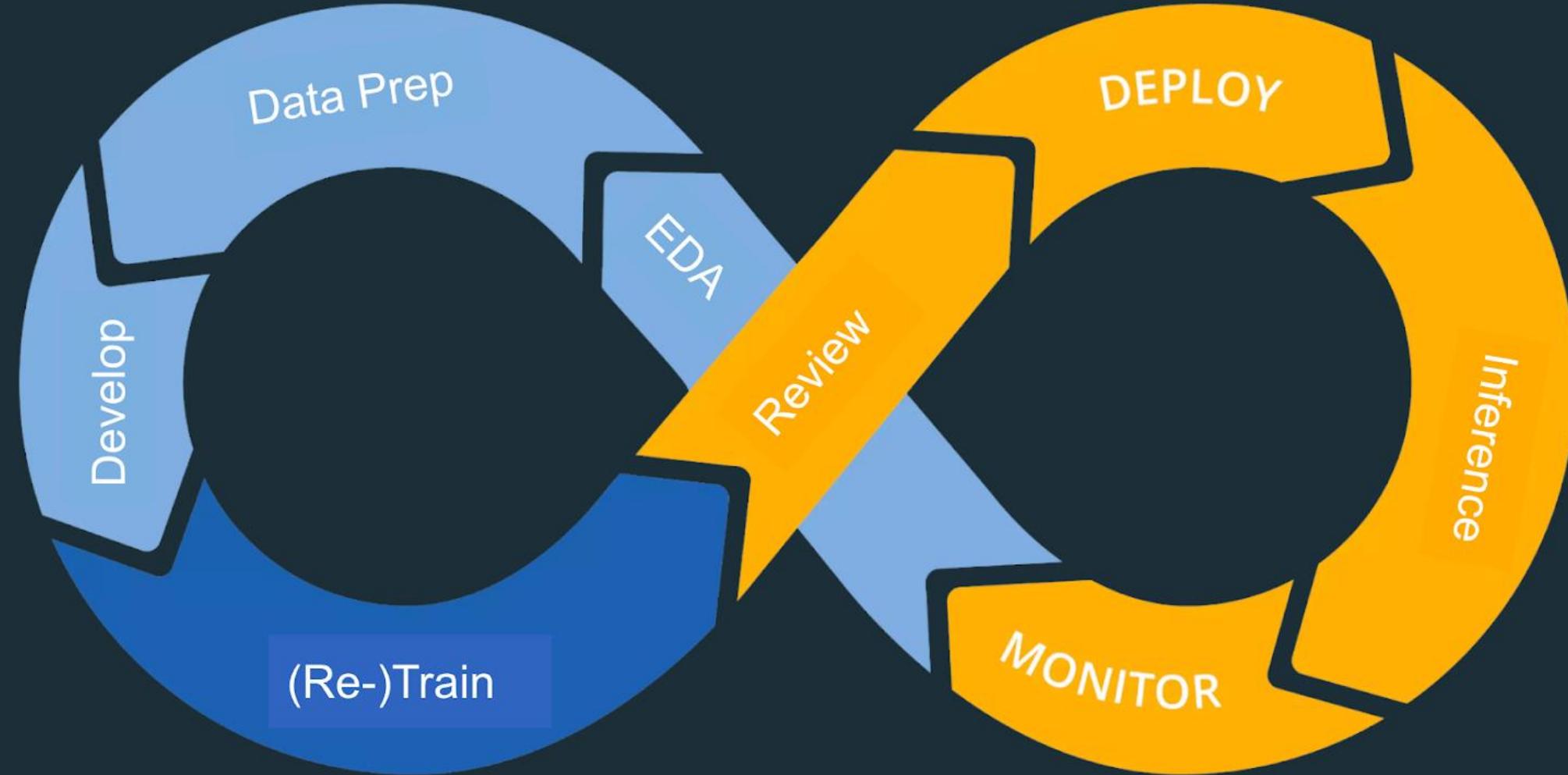
Private Cloud



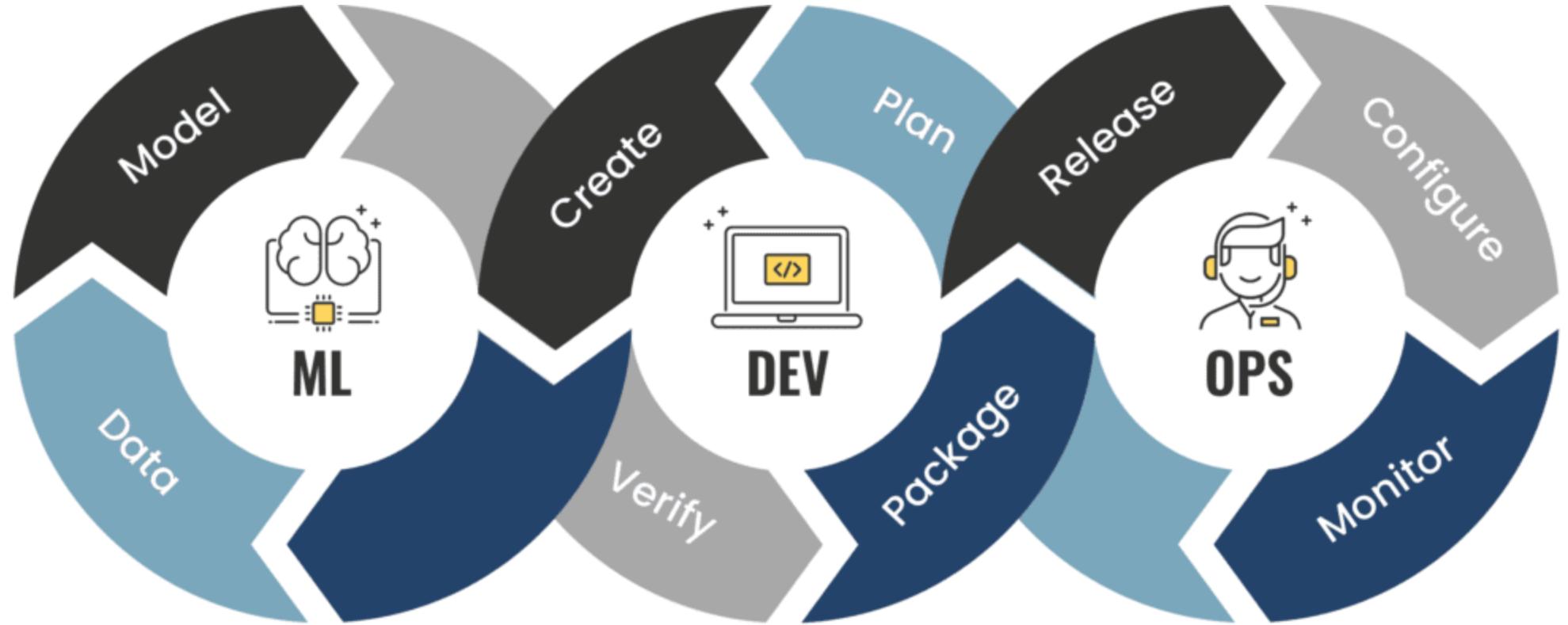
Private Cloud vs Public Cloud



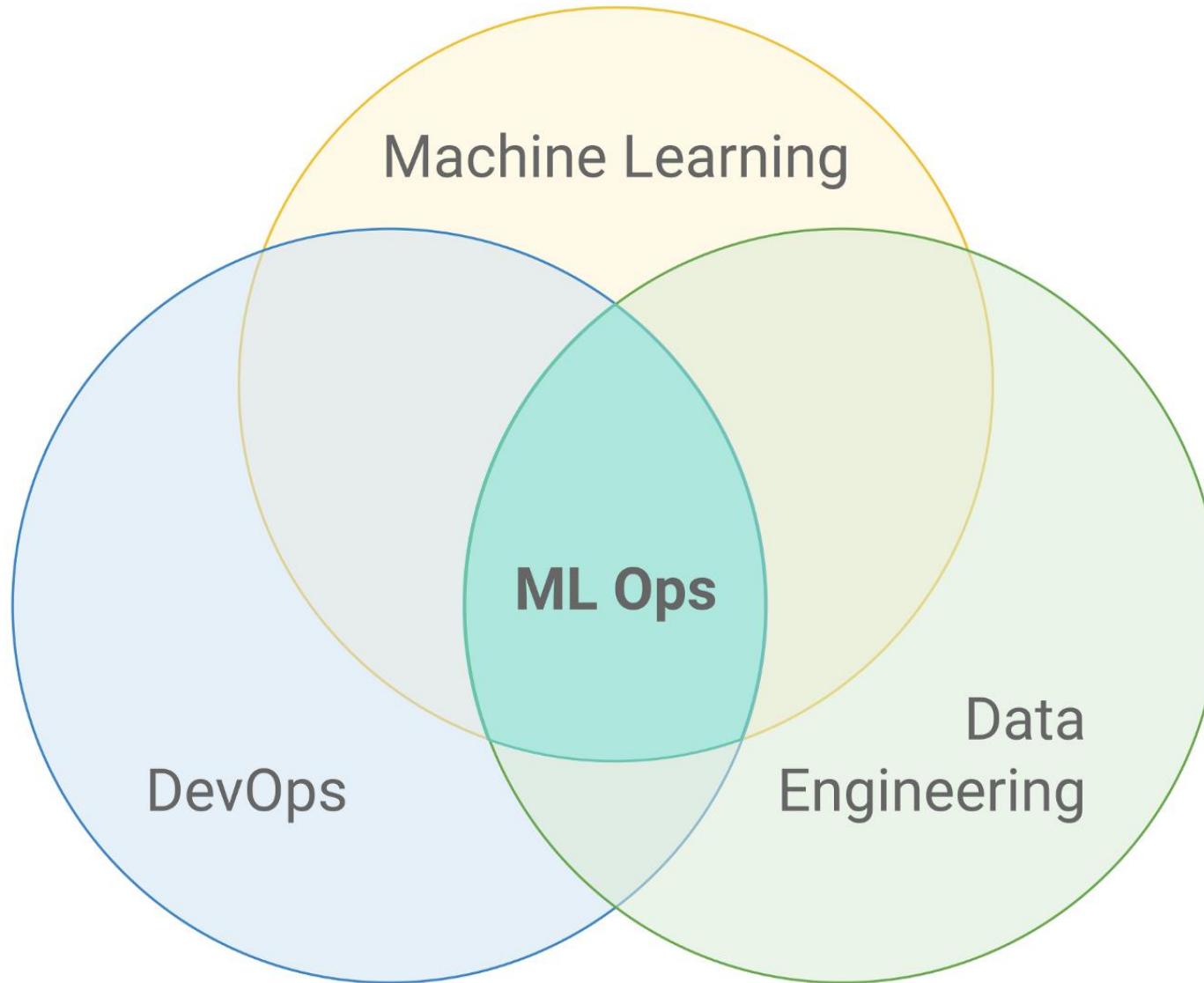
What is MLOps?



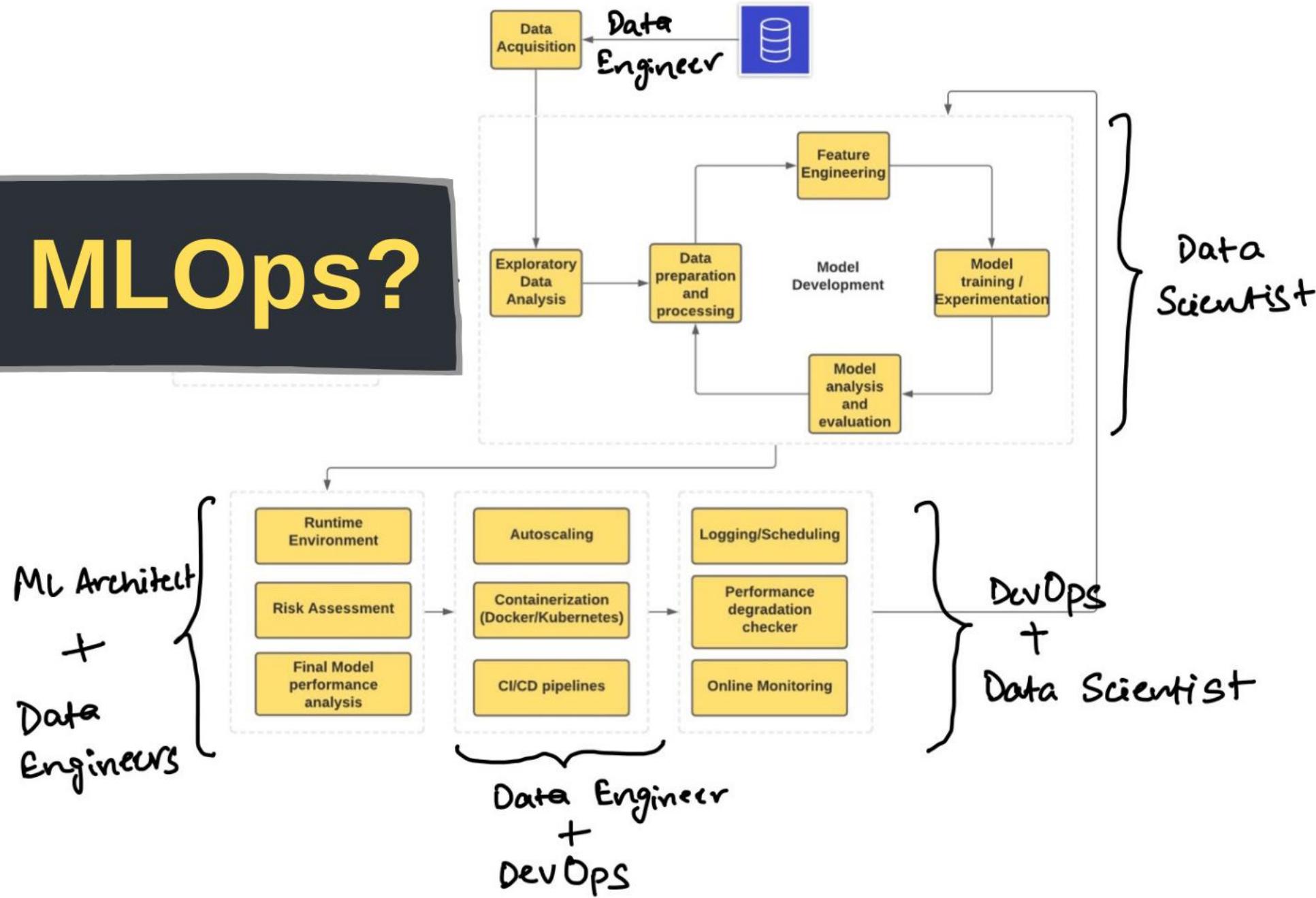
What is MLOps?



What is MLOps?



What is MLOps?



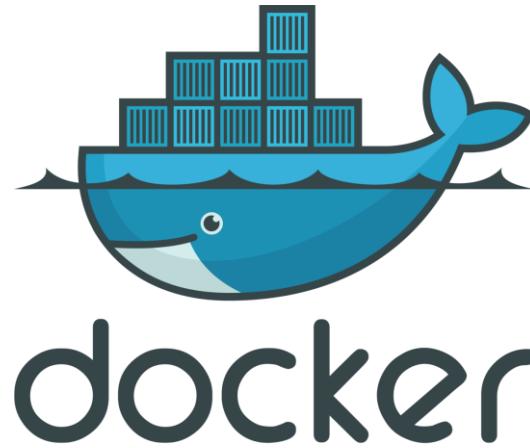
MLOps Technology Stacks



Google
Kubernetes Engine



Apache
Airflow



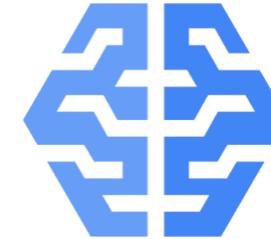
Cloud
DataFlow



Amazon
SageMaker



TensorFlow Extended



Google
AI Platform

Successful Career in Data Science

Completing High-Quality Courses



Newest courses in Data Science



100 Days Data Science Bootcamp: Build 100 Real Life...

Pianalytix .

4.7 ★★★★★ (14)

\$19.99

New



Build 75 Powerful Data Science & Machine Learning Projects

Pianalytix .

4.9 ★★★★★ (16)

\$19.99

New



Data Science Mega-Course: #Build {120-Projects In 120-...

Pianalytix .

4.4 ★★★★★ (18)

\$19.99

New



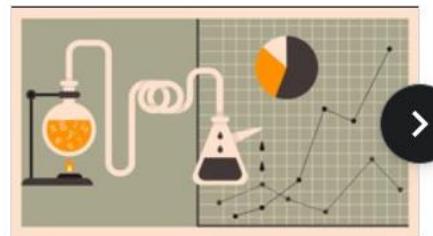
Mining and Analyzing LinkedIn Data

Jones Granatyr, IA Expert Academy

5.0 ★★★★★ (5)

\$19.99

Bestseller



Be Aware of Data Science

Robert Barcik, Jana Gecelovska, Patrik Zatko

4.8 ★★★★★ (39)

\$49.99

New

Topics recommended for you



Machine Learning



Data Science



Python



Deep Learning



Artificial Intelligence



Microsoft Power BI



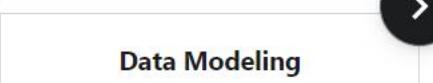
Business Intelligence



DAX



Data Visualization



Data Modeling

Successful Career in Data Science

Completing High-Quality Courses



coursera

Explore ▾

What do you want to learn?



Online Degrees

Find your New Career

For Enterprise

For Universities

Log In

Join for Free

Courses and Specializations

Courses You Can Complete in a Day

Over 4,000 courses in topics like business analytics, graphic design, Python, and more.



Foundations of
Mindfulness
Rice University

COURSE



Innovation Through
Design: Think, Make,
Break, Repeat
The University of Sydney

COURSE



Cloud Computing Basics
(Cloud 101)
LearnQuest

COURSE



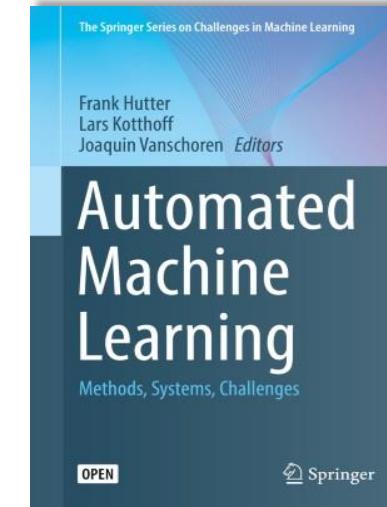
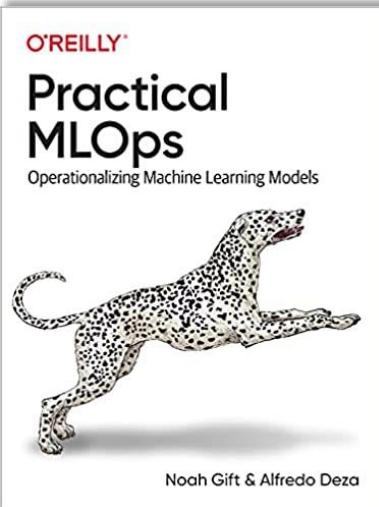
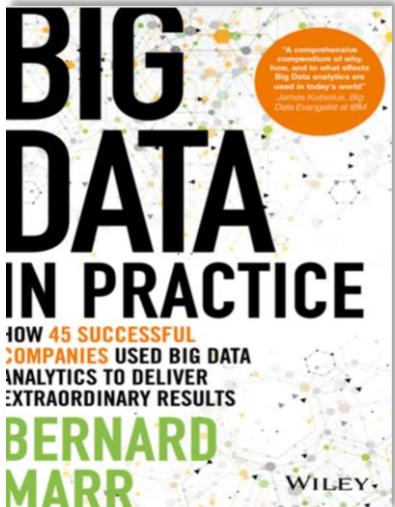
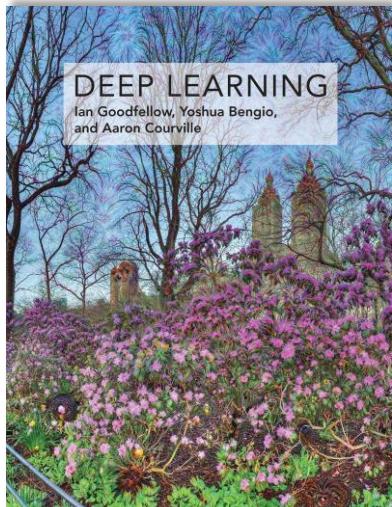
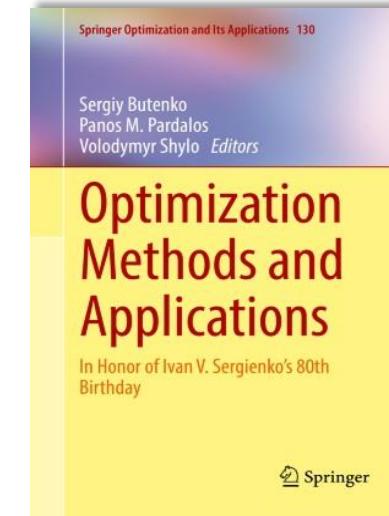
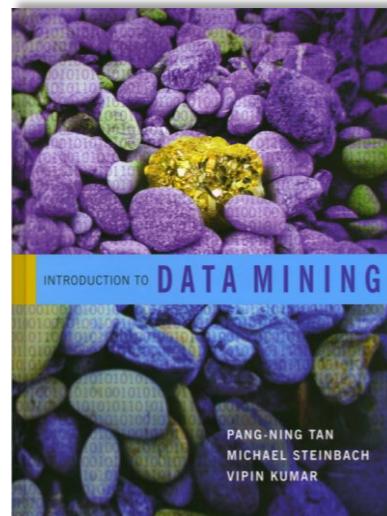
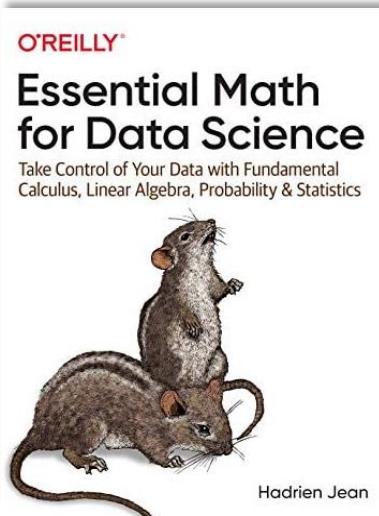
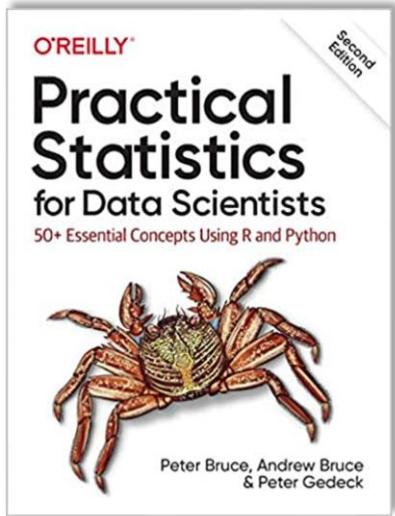
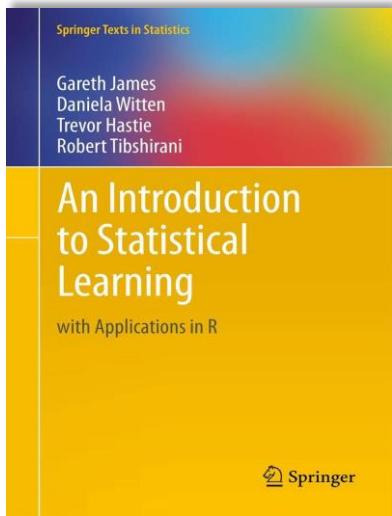
Statistics with SAS
SAS

COURSE



Successful Career in Data Science

Completing High-Quality Books



Successful Career in Data Science

Completing High-Quality Articles



ScienceDirect

ELSEVIER



Springer



Google Scholar



Conclusion and Questions?