AWS Academy Machine Learning Foundations

Module 2: Introduction to Machine Learning

Module overview

Sections

1. What is machine learning?
2. Business problems solved with machine learning
3. Machine learning process
4. Machine learning tools overview
5. Machine learning challenges

Knowledge check

Demonstration

Introducing Amazon SageMaker

Module objectives

At the end of this module, you should be able to:

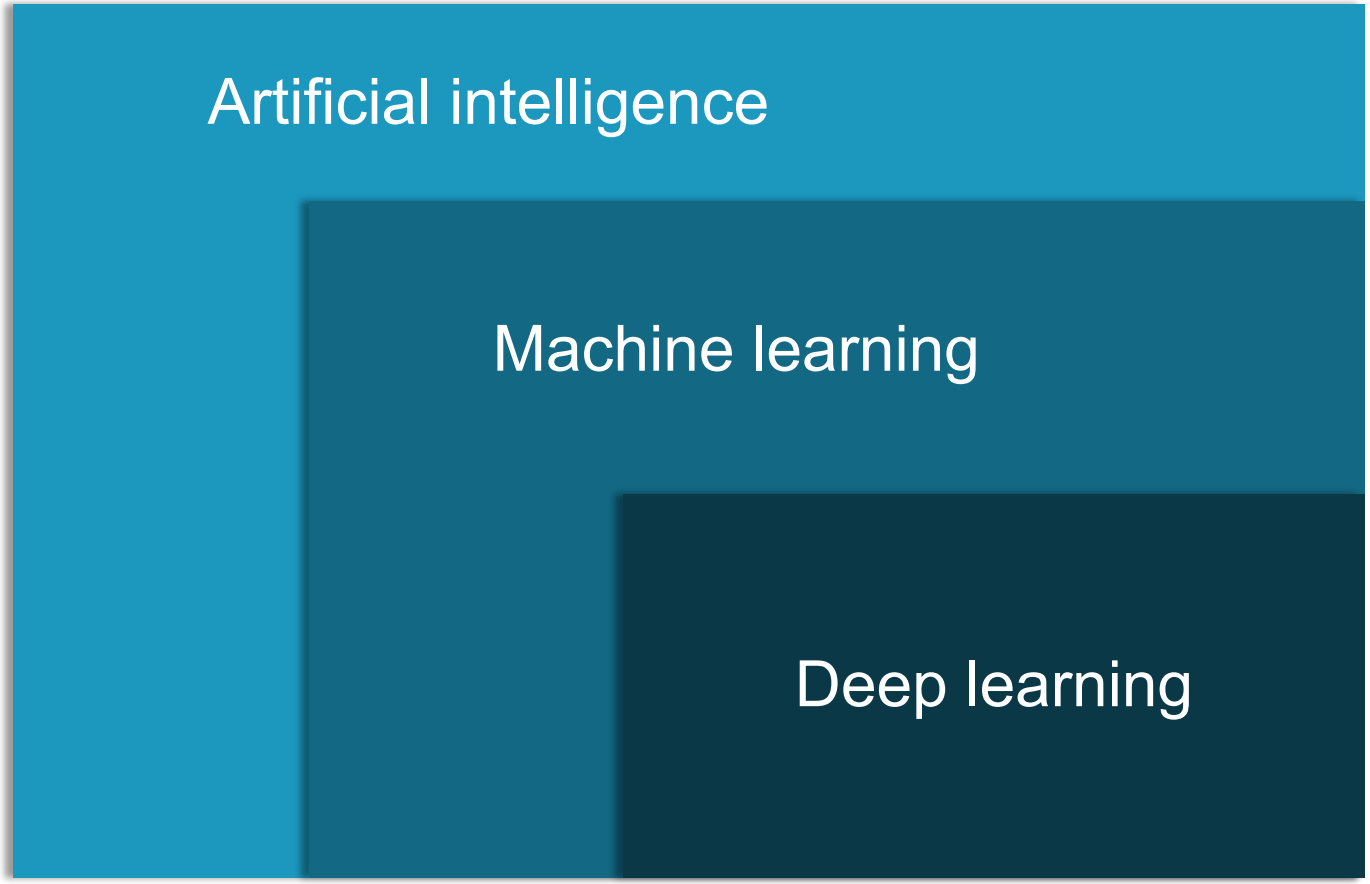
- Recognize how machine learning and deep learning are part of artificial intelligence
- Describe artificial intelligence and machine learning terminology
- Identify how machine learning can be used to solve a business problem
- Describe the machine learning process
- List the tools available to data scientists
- Identify when to use machine learning instead of traditional software development methods

Module 2: Introduction to Machine Learning

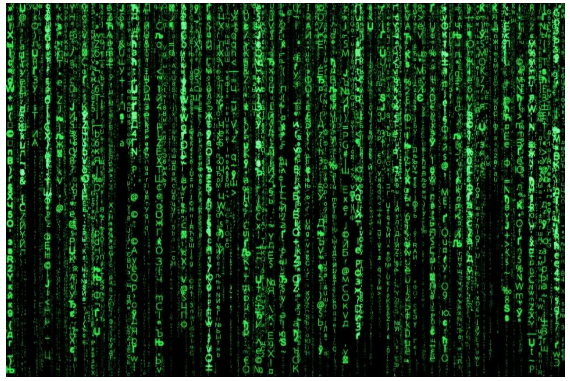
Section 1: What is machine learning ?

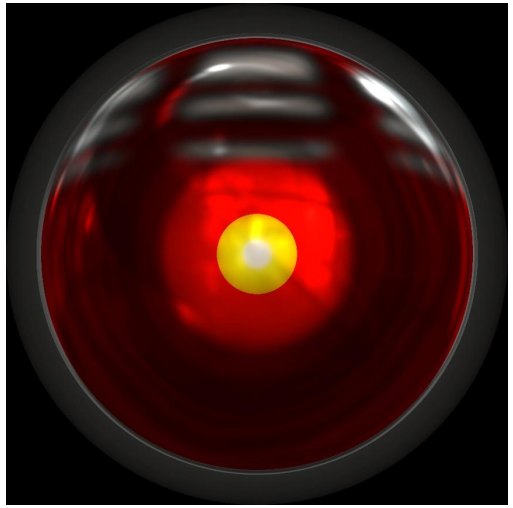
aws academy

Artificial intelligence, machine learning, and deep learning



Artificial intelligence





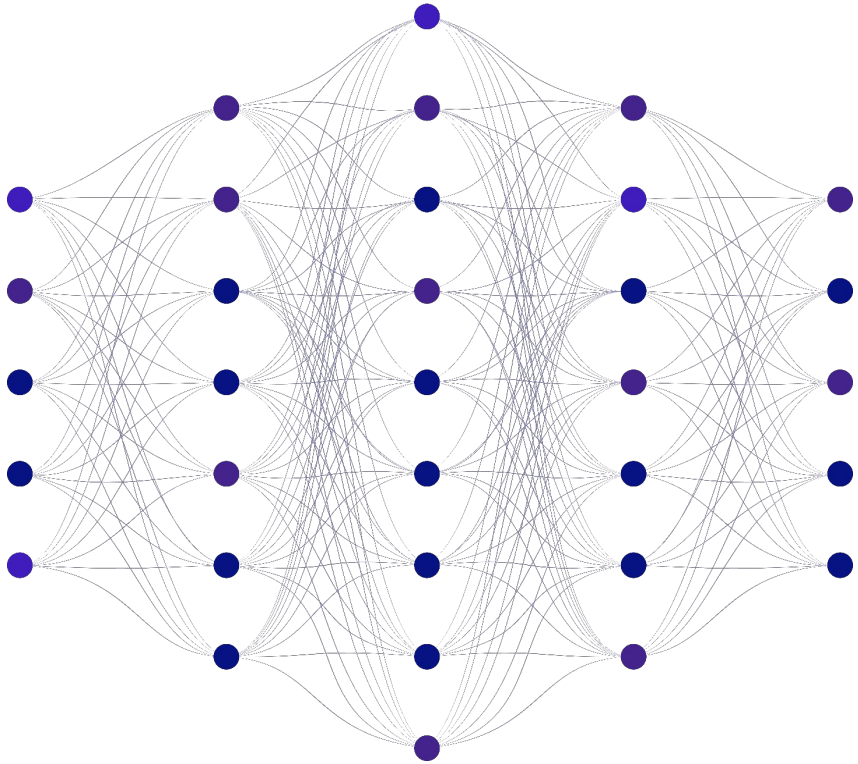
Machine learning

Machine learning is the scientific study of algorithms and statistical models to perform a task using inference instead of instructions.

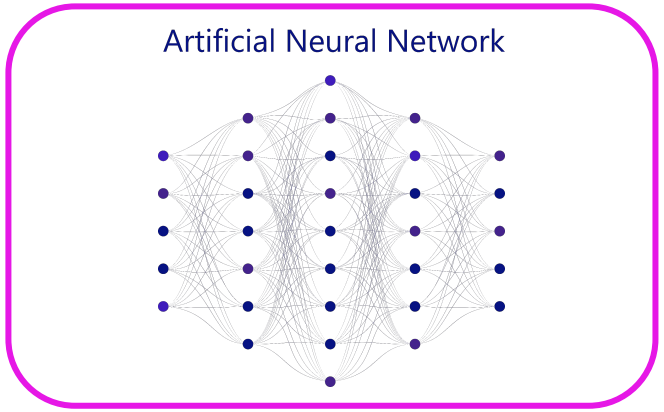
Machine learning flow

Deep learning

Artificial Neural Network



ML and technology advancements



Traditional computing

Cloud computing and Moore's law

Modern machine learning

Section 1 key takeaways

- Artificial intelligence
  - Machines performing human tasks
- Machine learning
  - Training models to make predictions
- Deep learning
  - Neural networks
- Technology and economic advancements have made machine learning more accessible to individuals and organizations

Module 2: Introduction to Machine Learning

Section 2: Business problems solved with machine learning

aws academy

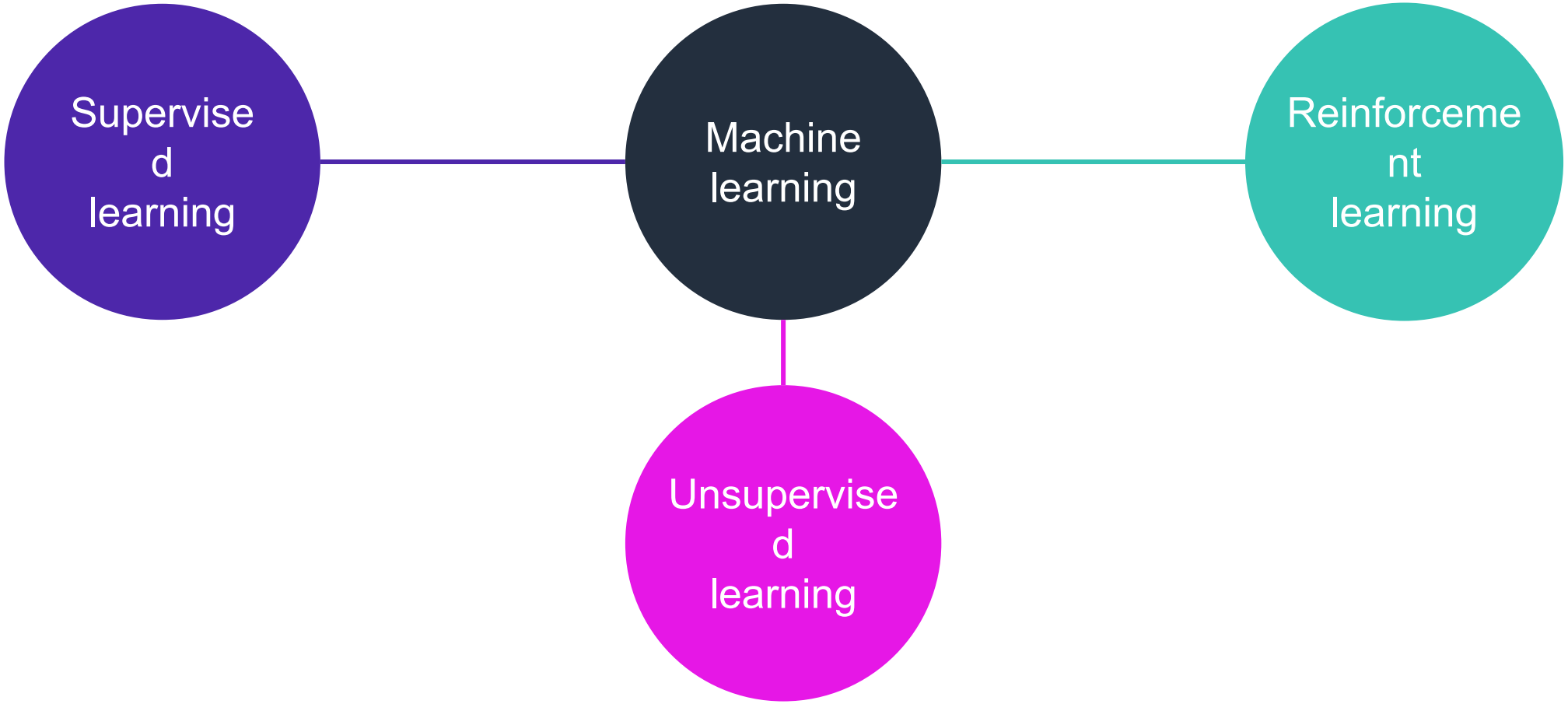
Common business use cases

Spam versus regular email

Recommendations

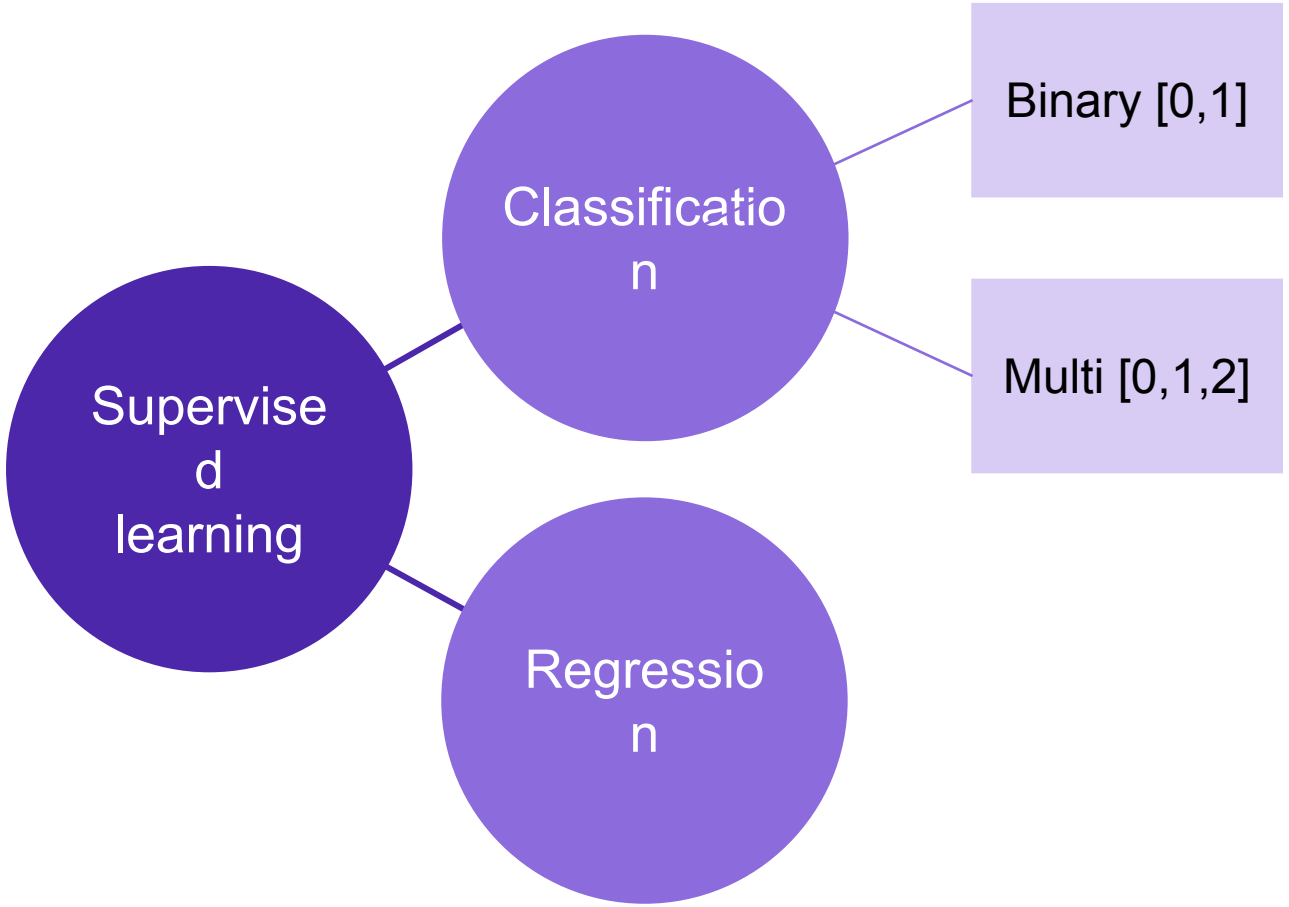
Fraud

Types of machine learning



Supervised learning

Learn by identifying patterns in data that is **already labeled**.

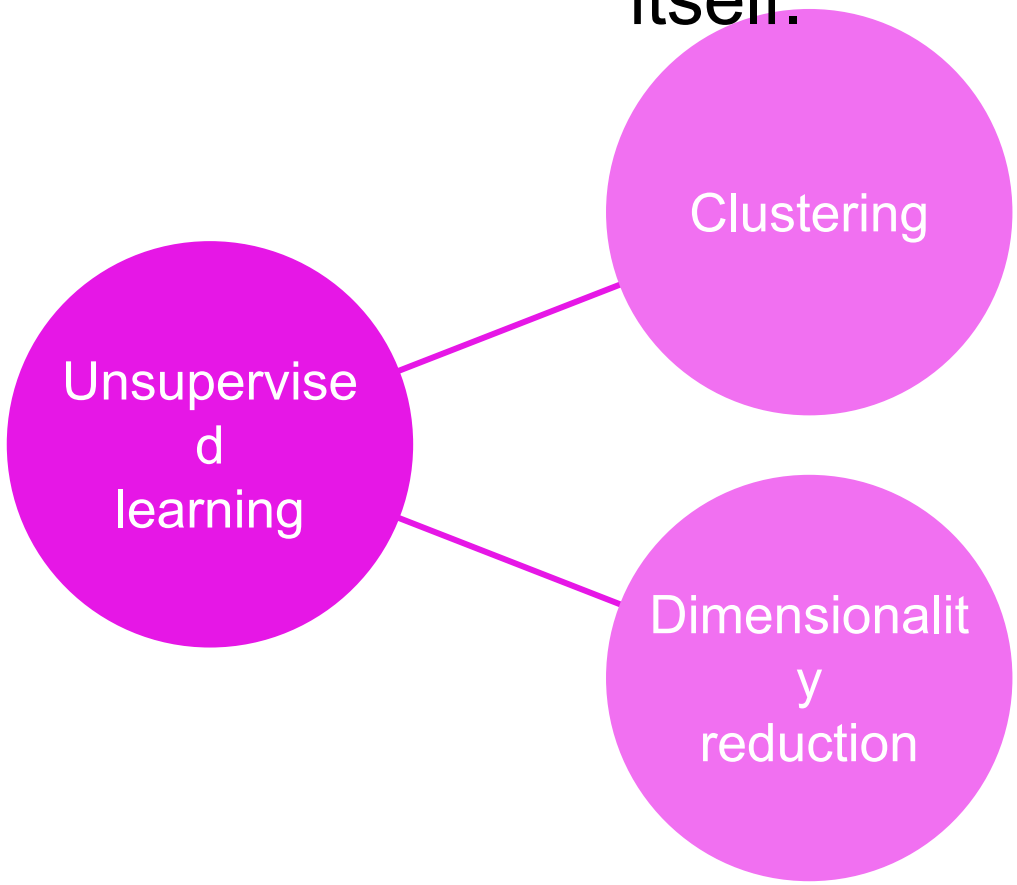


- Fraud detection
- Image recognition
- Customer retention
- Medical diagnostics
- Personalized advertising
- Product sales prediction
- Weather forecasting
- Market forecasting
- Population growth prediction

Computer vision

Unsupervised learning

The machine must uncover and create the labels itself.

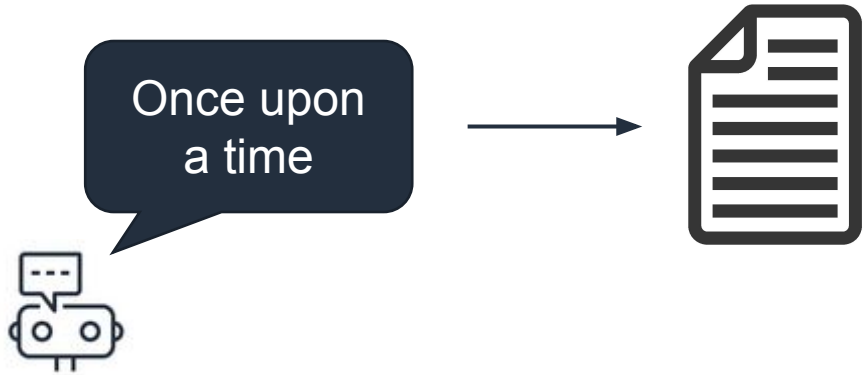


- Product recommendations
- Customer segmentation
- Targeted marketing
- Medical diagnostics

- Visualization
- Natural language processing
- Data structure discovery
- Gene sequencing

Natural language processing

gögn eru lykilatriði ⇄ los datos son clave



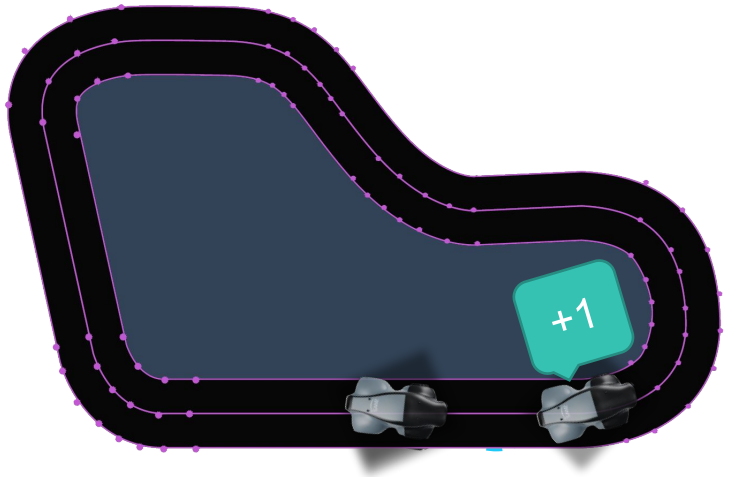
Poor story. Little character development. Jumps between scenes like you might get caught stealing. Unexplained bad guys appear with the thinnest of back story. Back to unlimited resources and lets not talk about the mechanics of building something so huge in such an inhospitable place...

Reinforcement learning

Learning through trial and error.

Reinforcement learning

- Game AI
- Self-driving cars
- Robotics
- Customer service routing



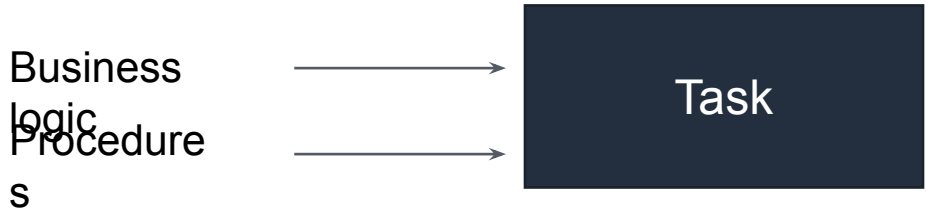
AWS DeepRacer

Best when the desired outcome is known but the exact path to achieving it is not known.

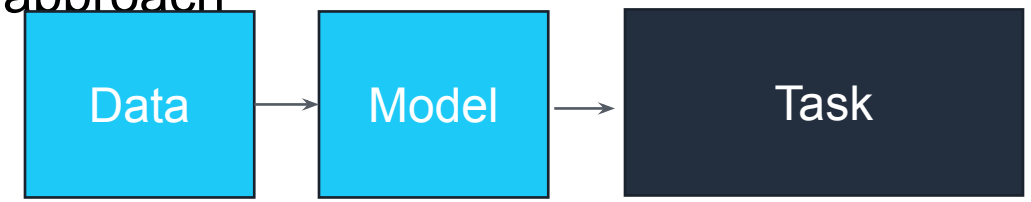
Self-driving vehicles

When to use machine learning?

Classical programming approach



Machine learning approach



Use machine learning when you have:

- ✔ Large datasets, large number of variables
- ✔ Lack of clear procedures to obtain the solution
- ✔ Existing machine learning expertise
- ✔ Infrastructure already in place to support ML
- ✔ Management support for ML

Section 2 key takeaways

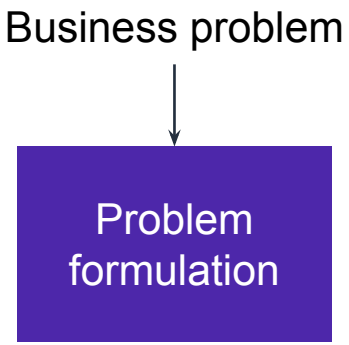
- Machine learning applications affect everyday life
- Machine learning can be grouped into –
  - Supervised learning
  - Unsupervised learning
  - Reinforcement learning
- Most problems are supervised learning

Module 2: Introduction to Machine Learning

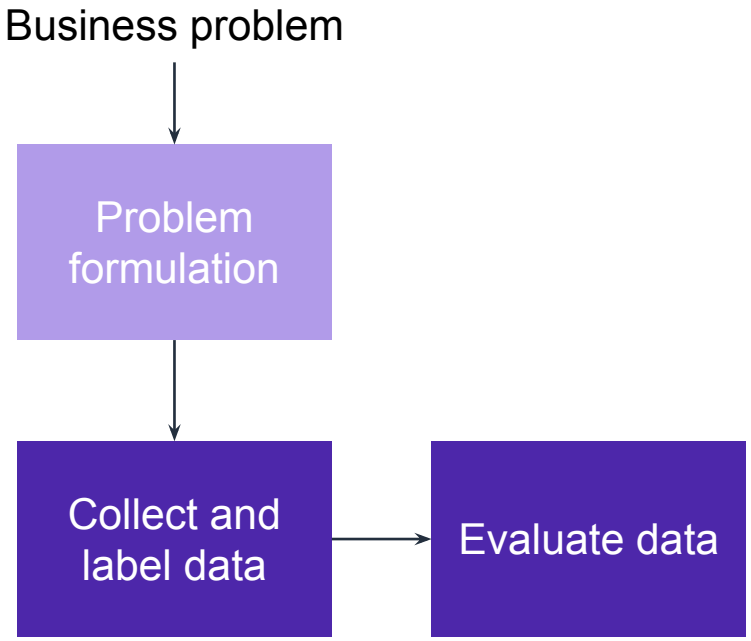
Section 3: Machine learning process

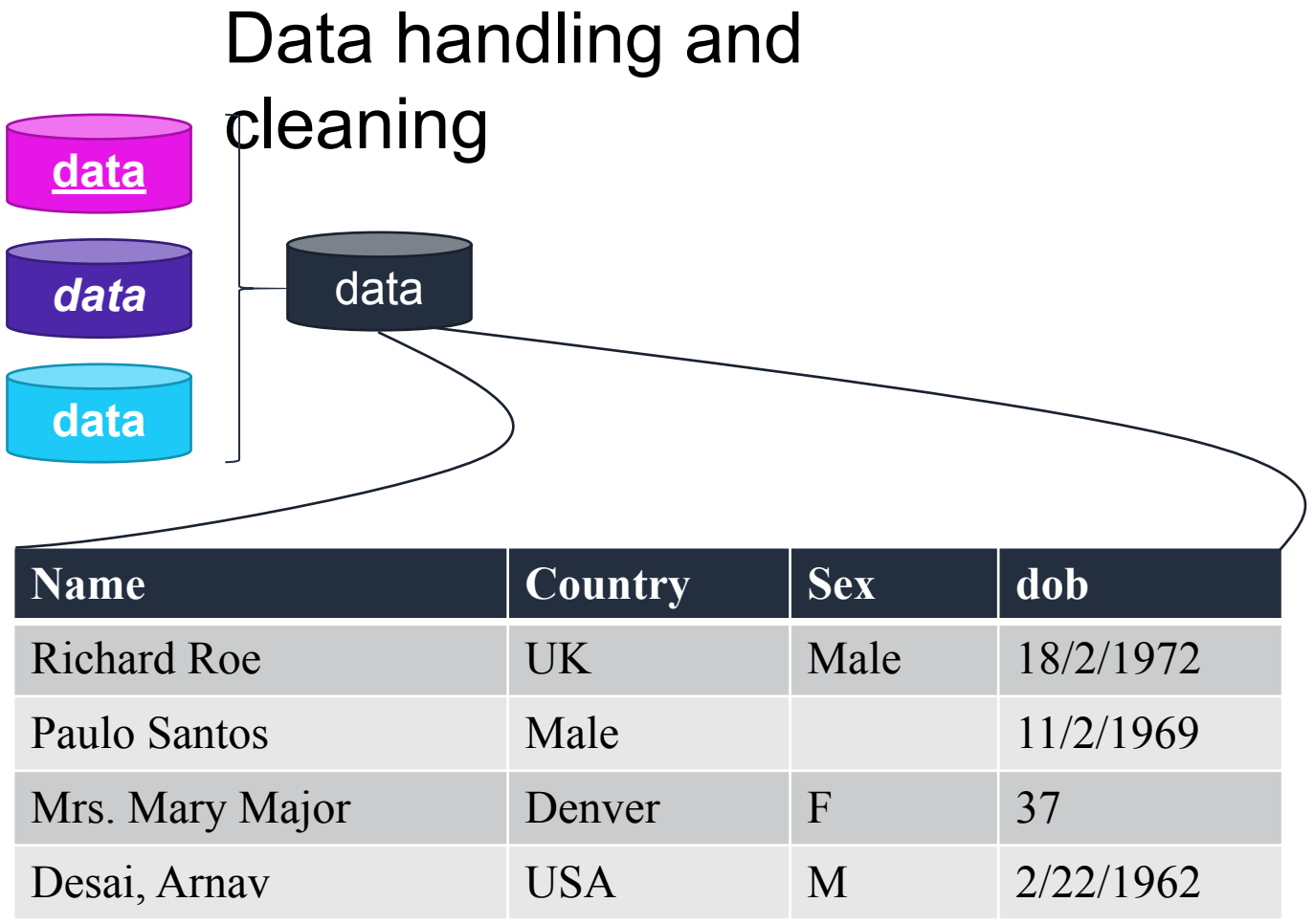
aws academy

ML pipeline: Business problem



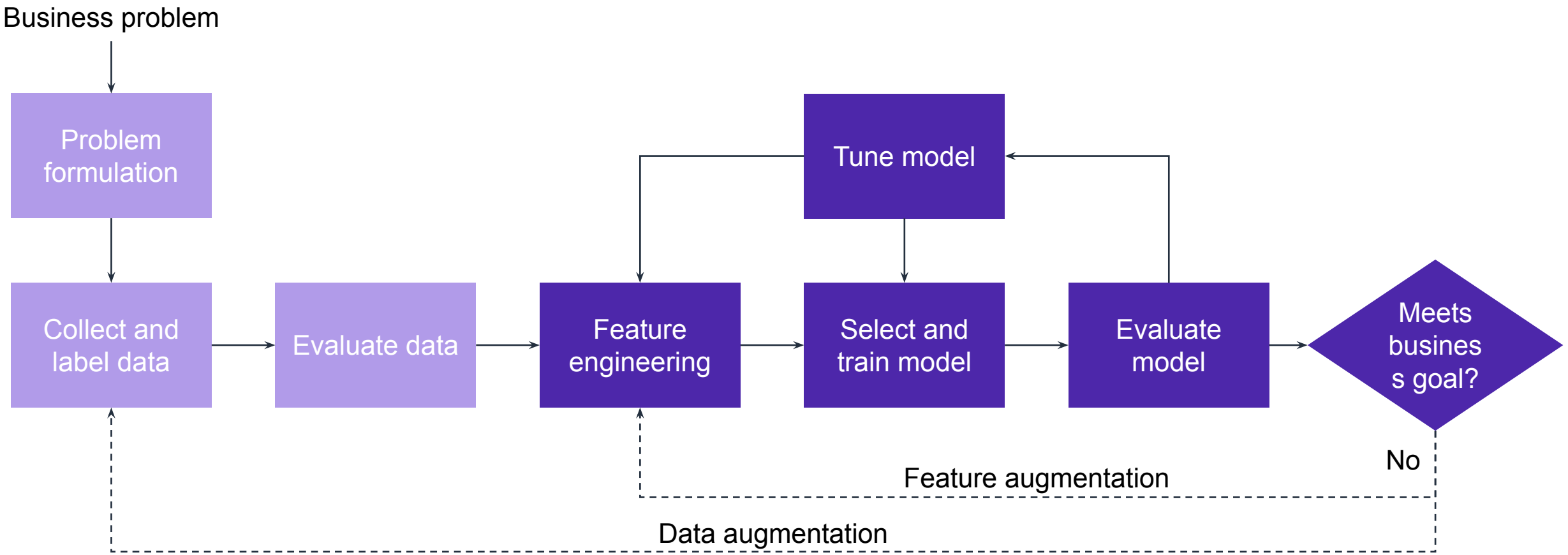
ML pipeline: Data preparation





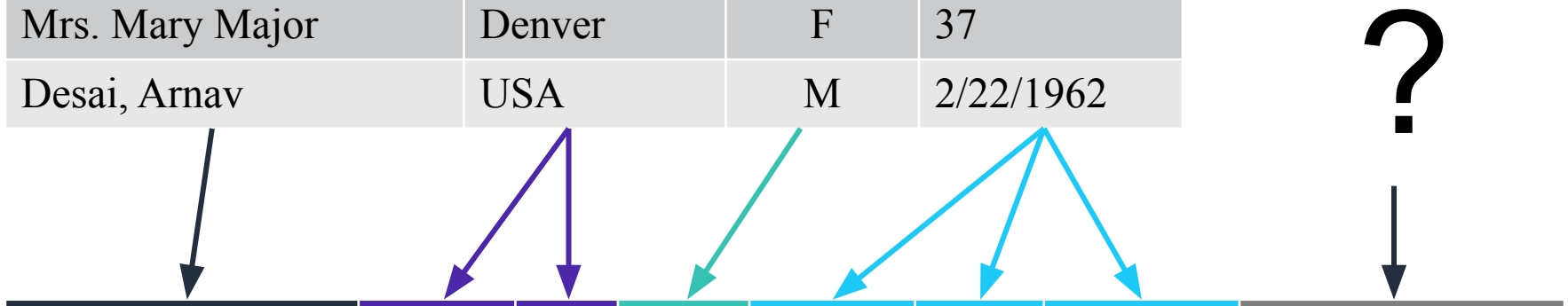
Name	Country	Sex	dob
Richard Roe	UK	Male	18/2/1972
Paulo Santos	Male		11/2/1969
Mrs. Mary Major	Denver	F	37
Desai, Arnav	USA	M	2/22/1962

ML pipeline: Iterative model training



ML pipeline: Feature engineering

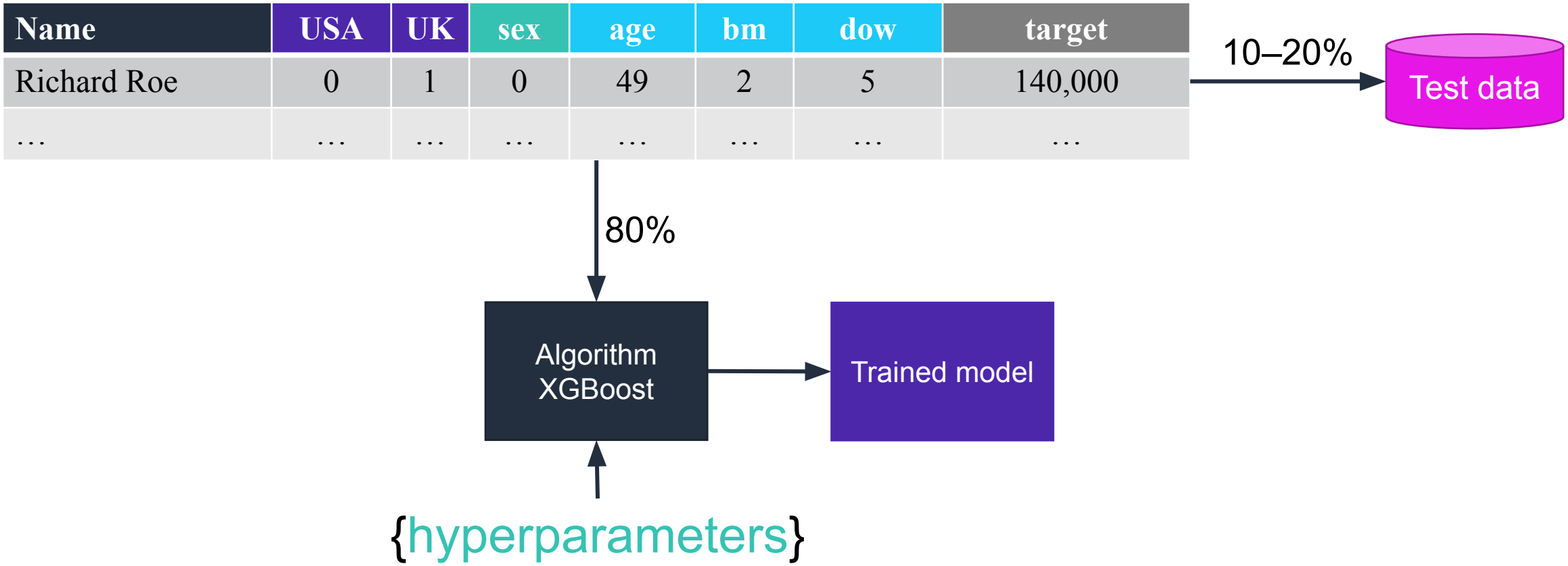
Name	Country	Sex	dob
Richard Roe	UK	Male	18/2/1972
Paulo Santos	Male		11/2/1969
Mrs. Mary Major	Denver	F	37
Desai, Arnav	USA	M	2/22/1962



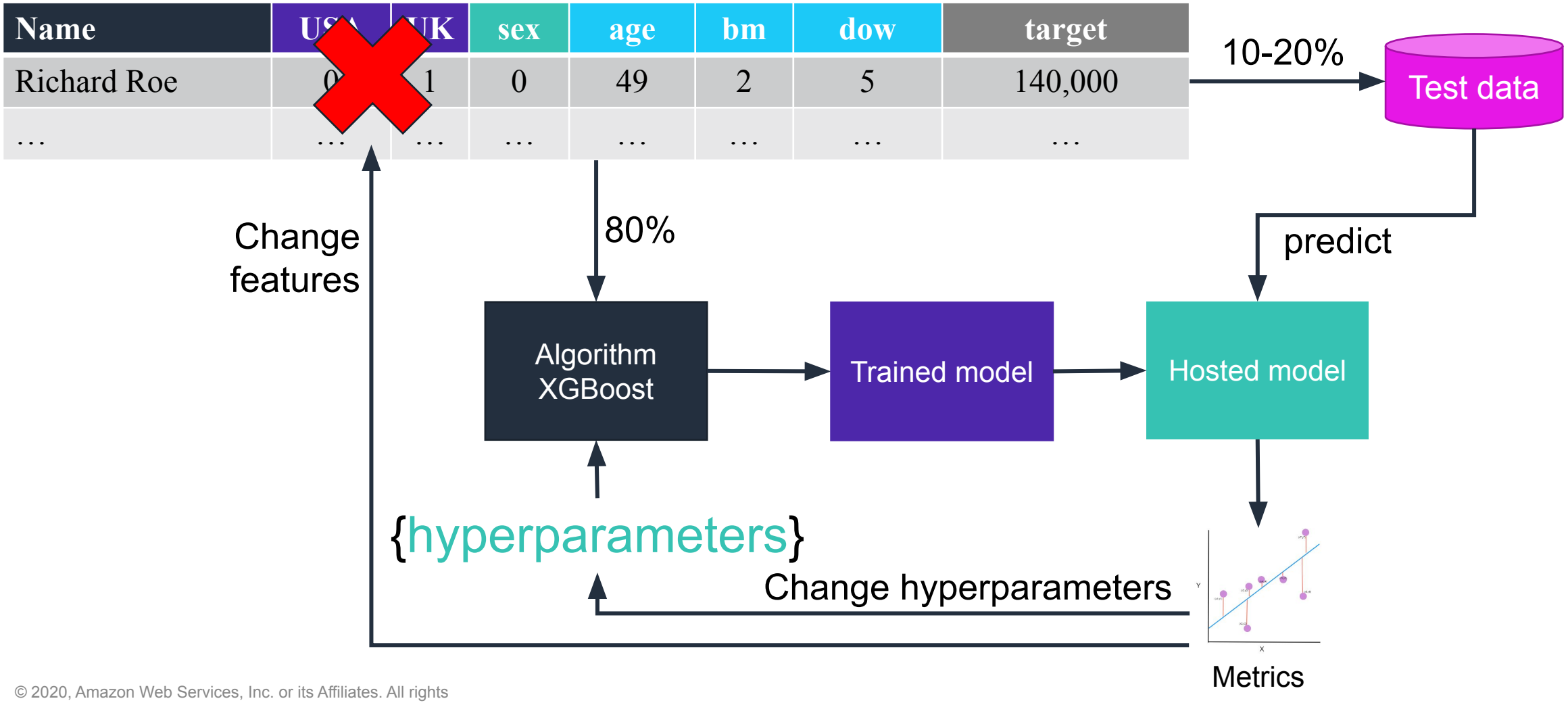
Name	USA	UK	sex	age	bm	dow	target
Richard Roe	0	1	0	49	2	5	140,000
Paulo Santos	1	0	0	51	11	7	78,000
Mary Major	1	0	1	37	NAN	0	167,000
Arnav Desai	1	0	0	58	2	4	100,000

ML pipeline: Model training

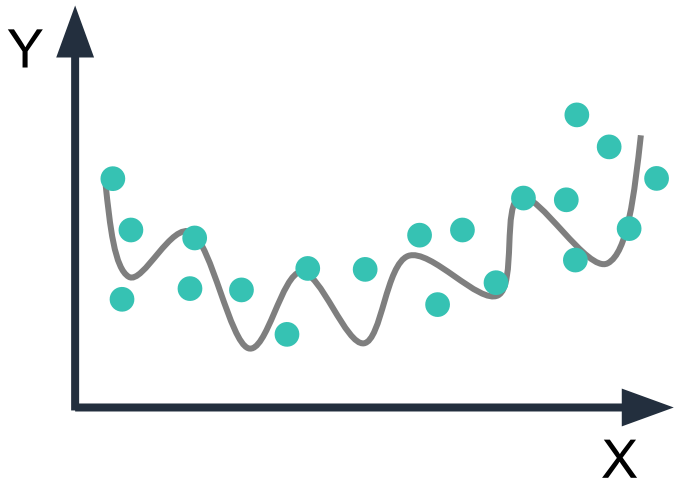
Name	USA	UK	sex	age	bm	dow	target
Richard Roe	0	1	0	49	2	5	140,000
…	…	…	…	…	…	…	…



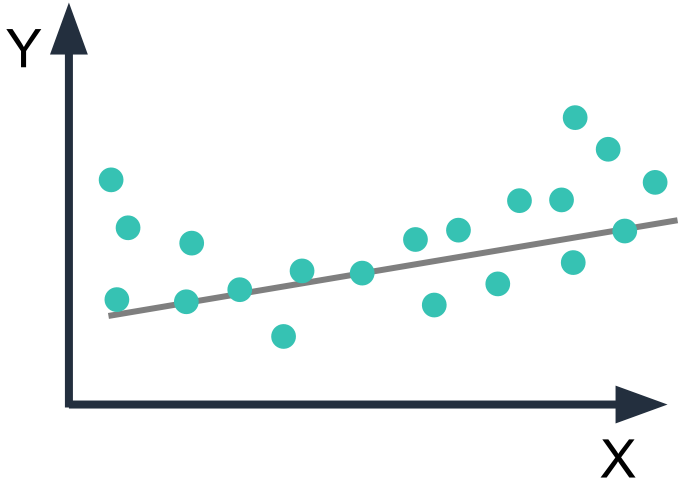
ML pipeline: Evaluating and tuning the model



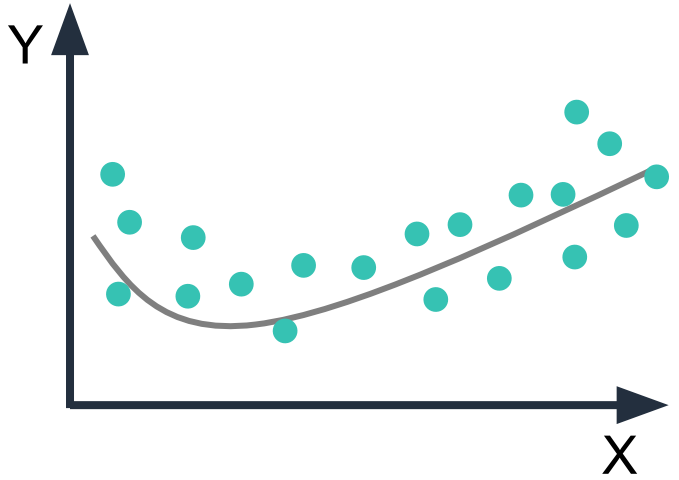
Overfitting and underfitting



Overfitting

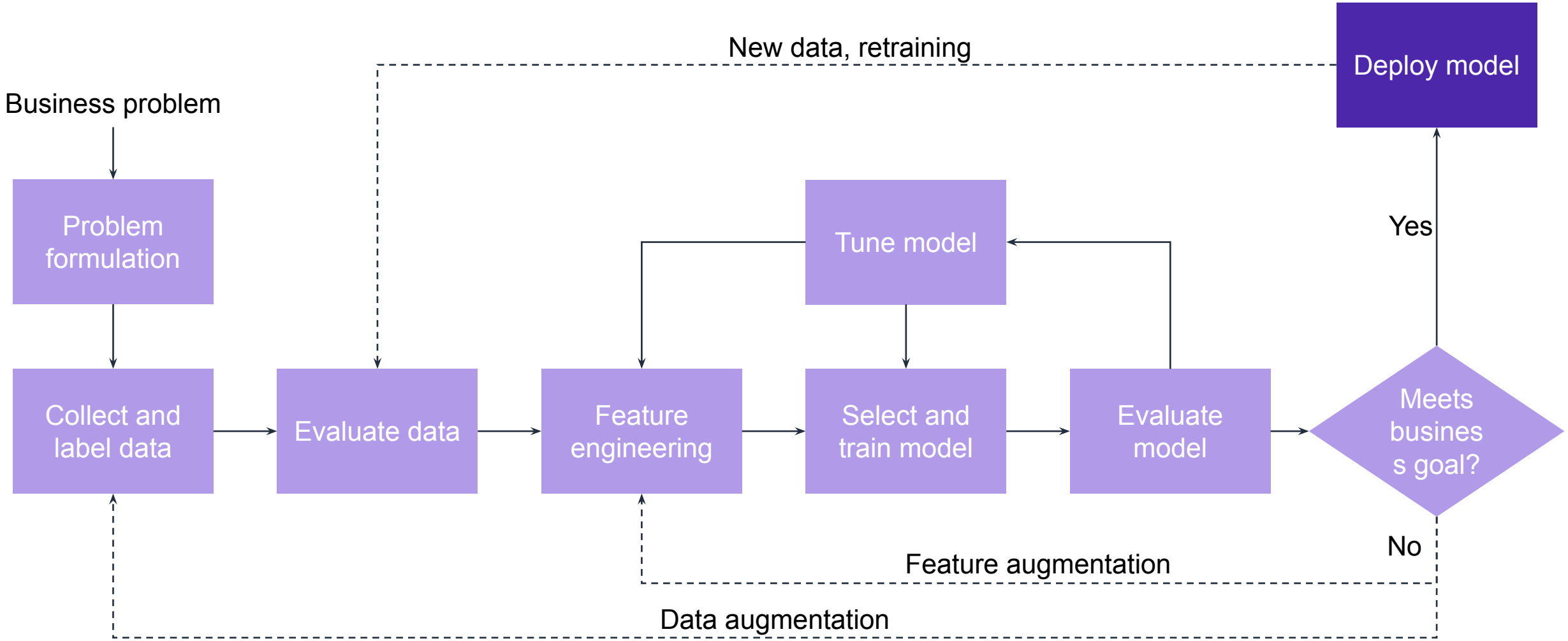


Underfitting



Balanced

ML pipeline: Deployment



Section 3 key takeaways

- Machine learning pipeline guides you through the process of evaluating and training a model
- Iterative process of –
  - Data processing
  - Training
  - Evaluation

Module 2: Introduction to Machine Learning

Section 4: Machine learning tools overview

aws academy

Python tools and libraries

- Jupyter Notebook
- JupyterLab

- pandas
- Matplotlib
- Seaborn
- NumPy
- scikit-learn

Machine learning frameworks and infrastructure

Machine learning frameworks provide tools and code libraries:

- Customized scripting
- Integration with AWS services
- Community of developers

PyTorch	Caffe2	Torch
TensorFlow	Gluon	Chainer
Keras	CNTK	Apache MXNet

Amazon instances that are designed for machine learning applications:

- AWS IoT Greengrass provides an infrastructure for building machine learning for IoT devices
- Amazon Elastic Inference reduces costs for running machine learning applications

EC2 P3 instances

EC2 C5 and C5n instances

AWS IoT Greengrass

Amazon Elastic Inference

Amazon SageMaker

Ground Truth

Set up and manage labeling jobs for highly accurate training datasets by using active learning and human labeling.

Notebook

Provide AWS and SageMaker SDKs and sample notebooks to create training jobs and deploy models.

Training

Train and tune models at any scale. Use high-performance AWS algorithms, or bring your own.

Inference

Create models from training jobs, or import external models for hosting so you can run inferences on new data.

AWS Marketplace

Find, buy, and deploy ready-to-use model packages, algorithms, and data products in AWS Marketplace.

Demonstration: Introducing Amazon SageMaker

Machine learning managed services

These managed services don’t require ML experience.

Computer vision

Amazon Rekognition

Amazon Textract

Chatbots

Amazon Lex

Speech

Amazon Polly

Amazon Transcribe

Forecasting

Amazon Forecast

Language

Amazon Comprehend

Amazon Translate

Recommendations

Amazon Personalize

Section 4 key takeaways

- Python is the most popular ML language
- Jupyter Notebooks
- Many open-source tools
- Frameworks and services for all requirements
  - Low-level frameworks
  - Amazon SageMaker
  - Managed ML services

Module 2: Introduction to Machine Learning

Section 5: Machine learning challenges

aws academy

Machine learning challenges

Data

- Poor quality
- Non-representative
- Insufficient
- Overfitting and underfitting

Users

- Lack of data science expertise
- Cost of staffing with data scientists
- Lack of management support

Business

- Complexity in formulating questions
- Explaining models to the business
- Cost of building systems

Technology

- Data privacy issues
- Tool selection can be complicated
- Integration with other systems

Using existing models and services

Amazon ML managed services

- Amazon ML managed services
- No ML experience needed

You Only Look Once (YOLO)

AWS Marketplace

- Use existing trained and tuned models
- Enhance with domain-specific instances
- Over 250 ML model packages and algorithms
- Over 14 industry segments

Section 5 key takeaways

- Machine learning challenges
  - Data
  - People
  - Business
  - Technology
- Managed services simplify machine learning

Module 2: Introduction to Machine Learning

Module wrap-up

aws academy

Module takeaways

- Machine learning is a subset of artificial intelligence
  - Machine learning applies learning algorithms to develop models from large datasets
- The machine learning pipeline describes the different stages for developing a machine learning application
- The Amazon Machine Learning stack has three key layers
  - Managed services, machine learning services, machine learning frameworks
- Machine learning development is different from traditional development
  - Training algorithm is applied to data to create a model for making predictions

Module summary

In summary, in this module, you learned how to:

- Recognize how machine learning and deep learning are part of artificial intelligence
- Describe artificial intelligence and machine learning terminology
- Identify how machine learning can be used to solve a business problem
- Describe the machine learning process
- List the tools available to data scientists
- Identify when to use machine learning instead of traditional software development methods

Complete the knowledge check

Additional resources

- What is Machine Learning?
- Machine Learning on AWS

Thank you

aws academy