

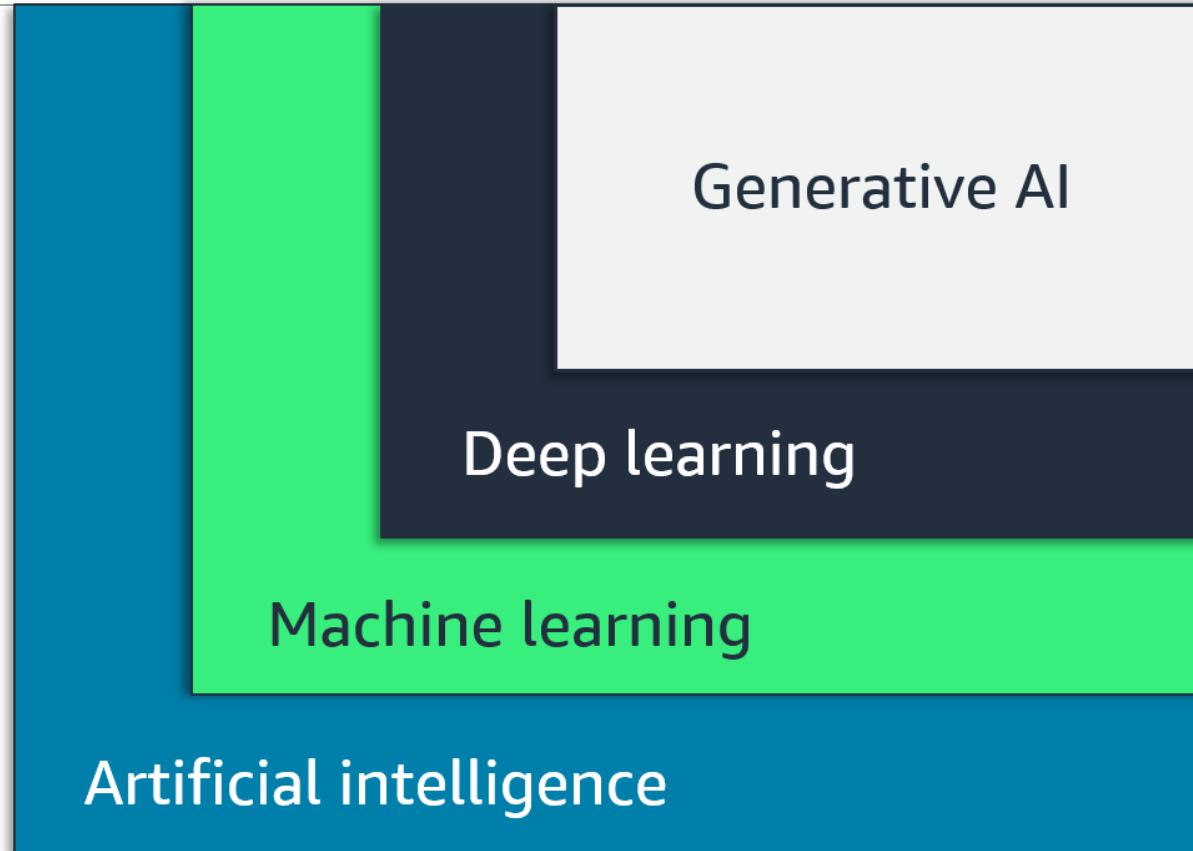
What is Machine Learning?

Outline

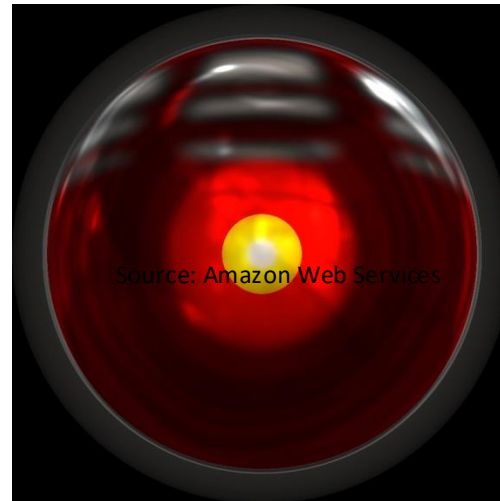
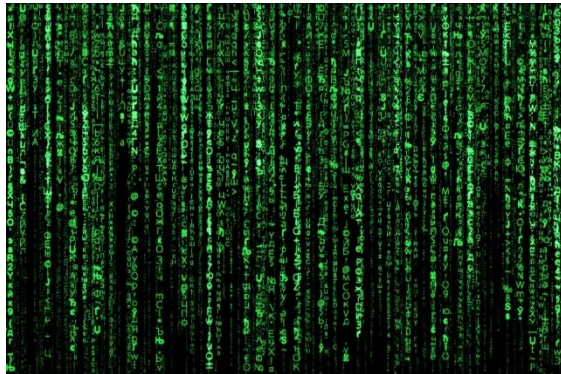
- Introducing machine learning
- Real-world problems solved with machine learning
- Machine learning process
- Machine learning tools

Part 1: Introducing machine learning

AI, ML, deep learning, and generative AI



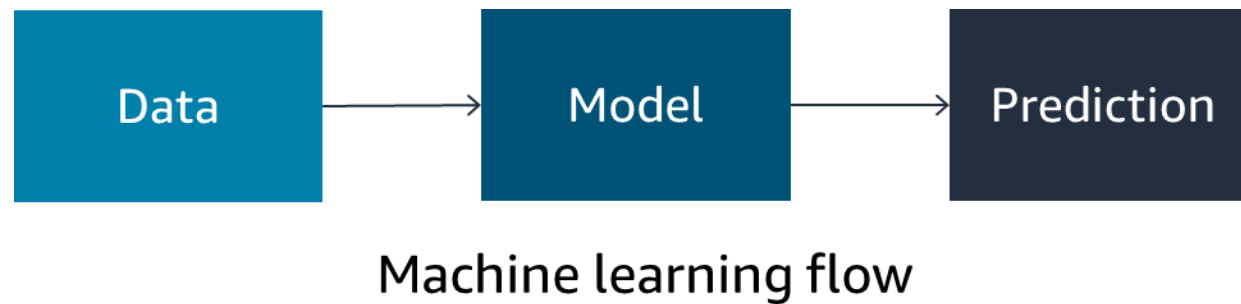
Artificial intelligence



Source: Amazon Web Services

Machine learning

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

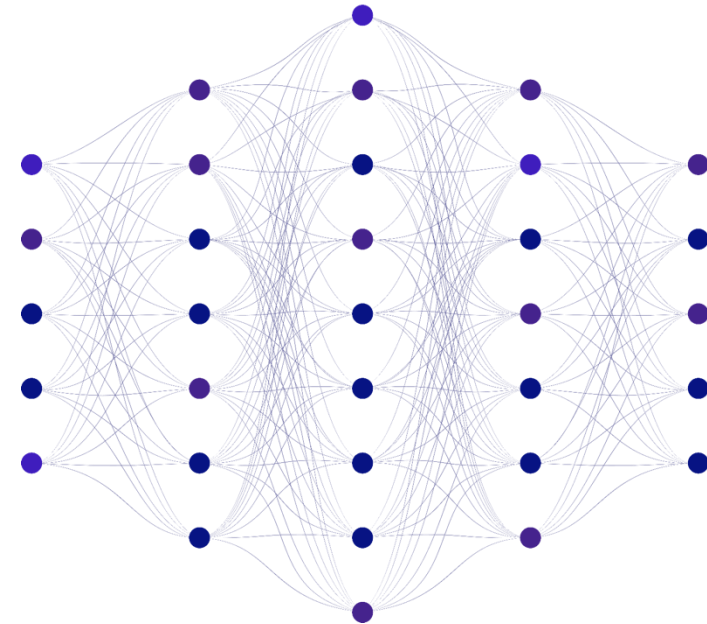


Source: Amazon Web Services

Deep learning

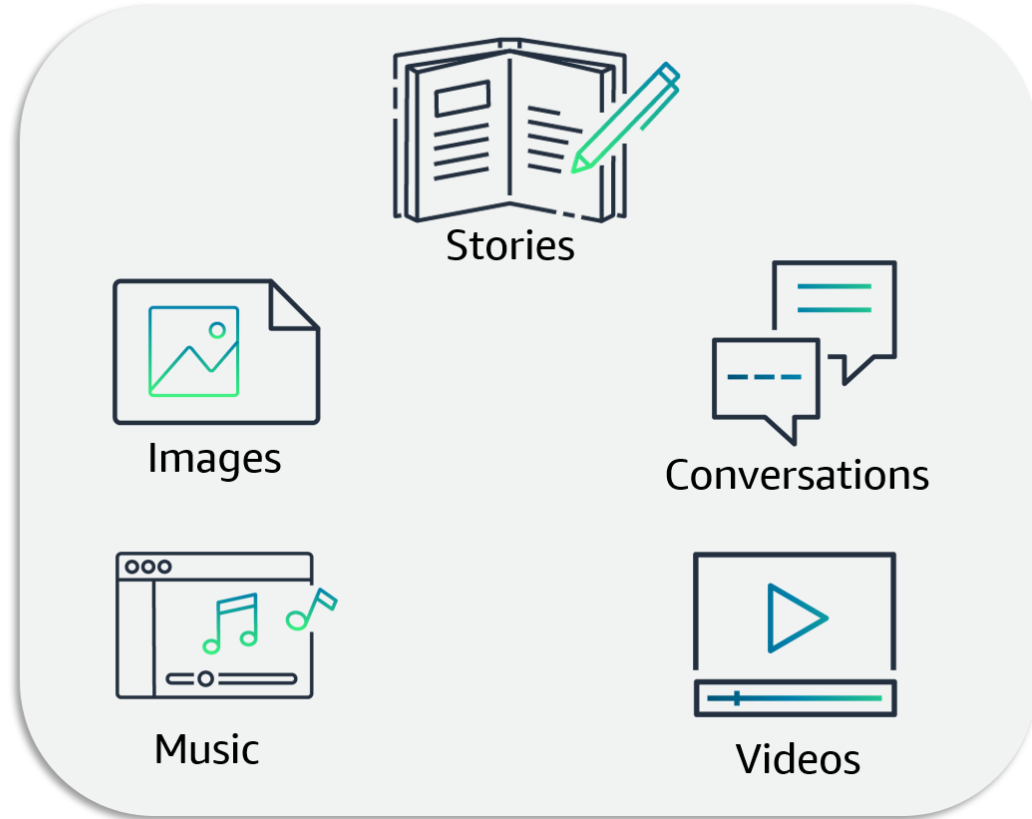
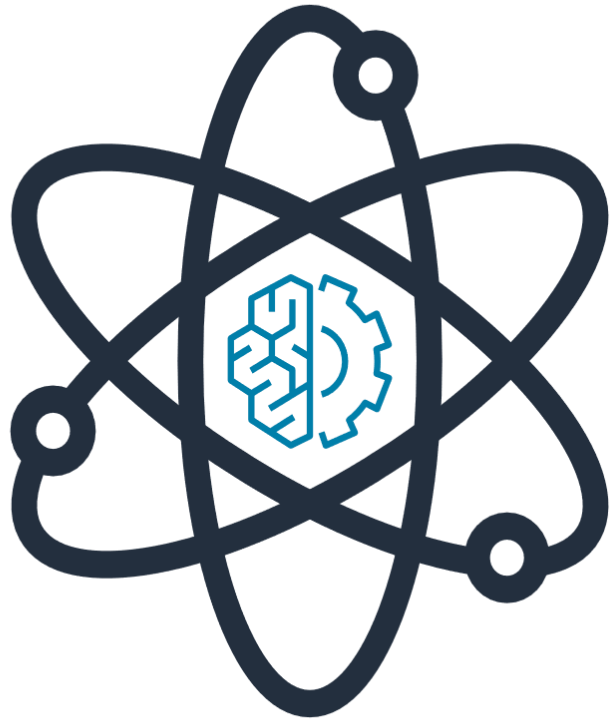


Artificial Neural Network



Source: Amazon Web Services

Generative AI



ML and technology advancements



Traditional
computing

Cloud computing and
Moore's law

Modern machine
learning

Source: Amazon Web Services

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

Part 2: Real-world problems solved with machine learning

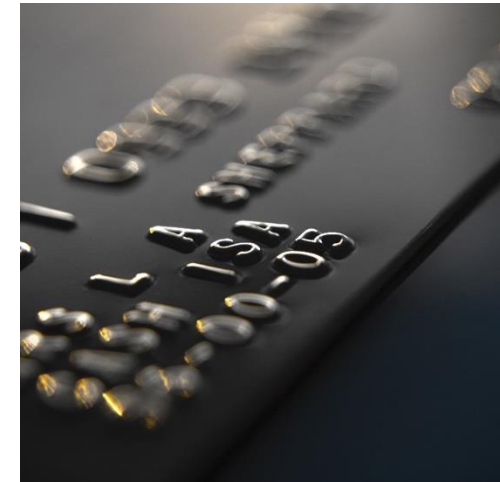
Common business use cases



Spam versus
regular email



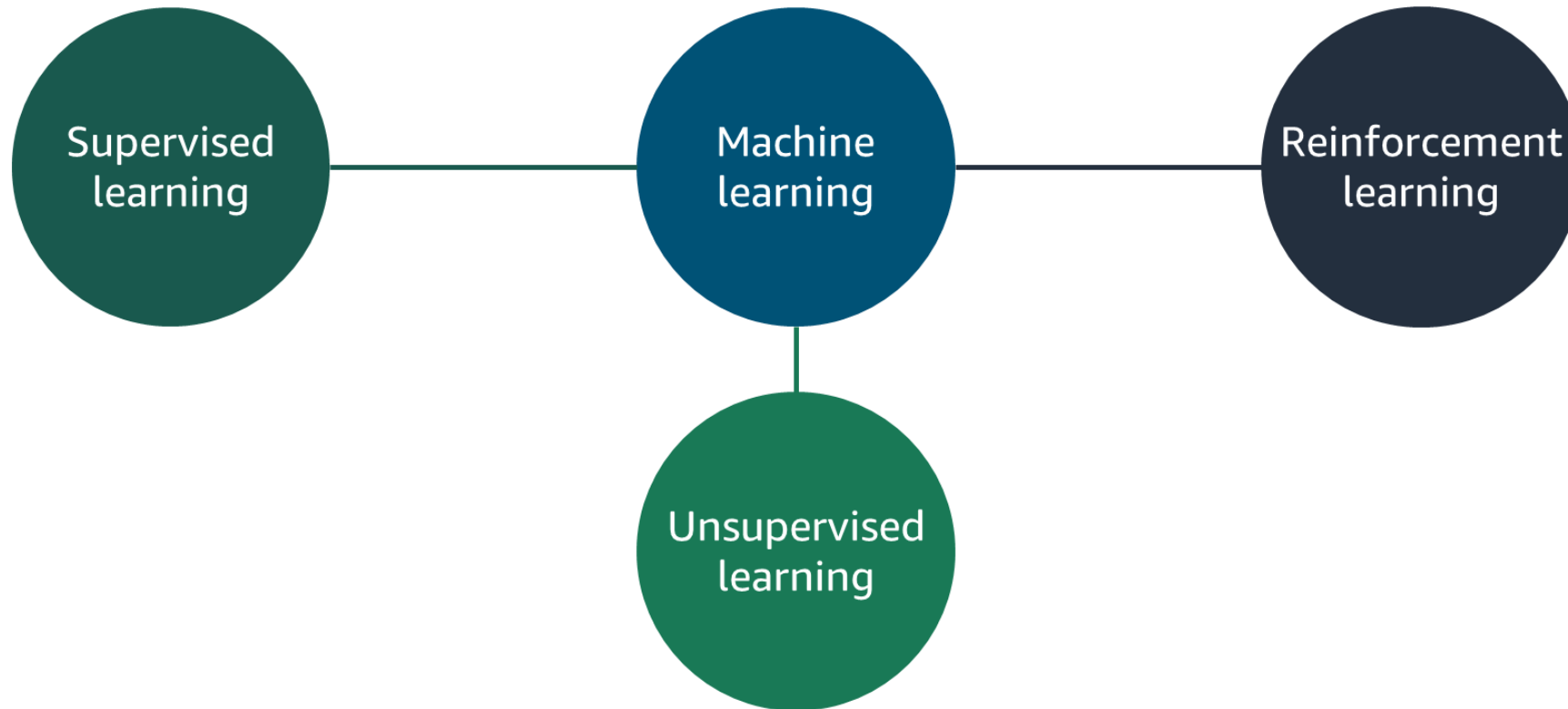
Recommendations



Fraud

Source: Amazon Web Services

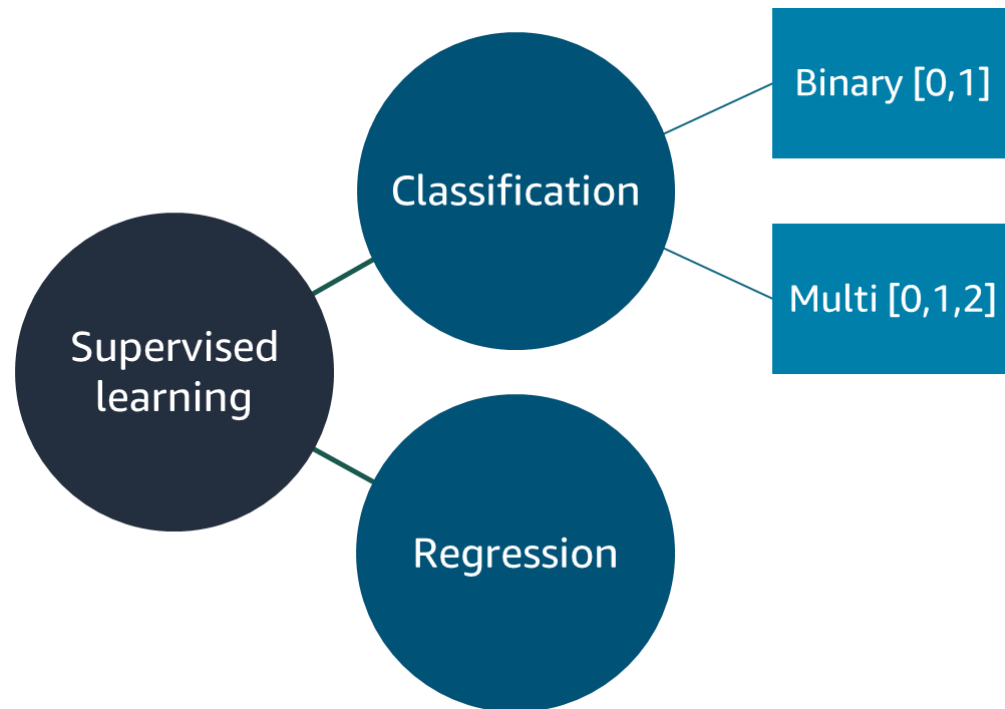
Types of machine learning



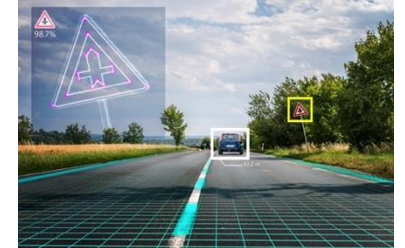
Source: Amazon Web Services

Supervised learning

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

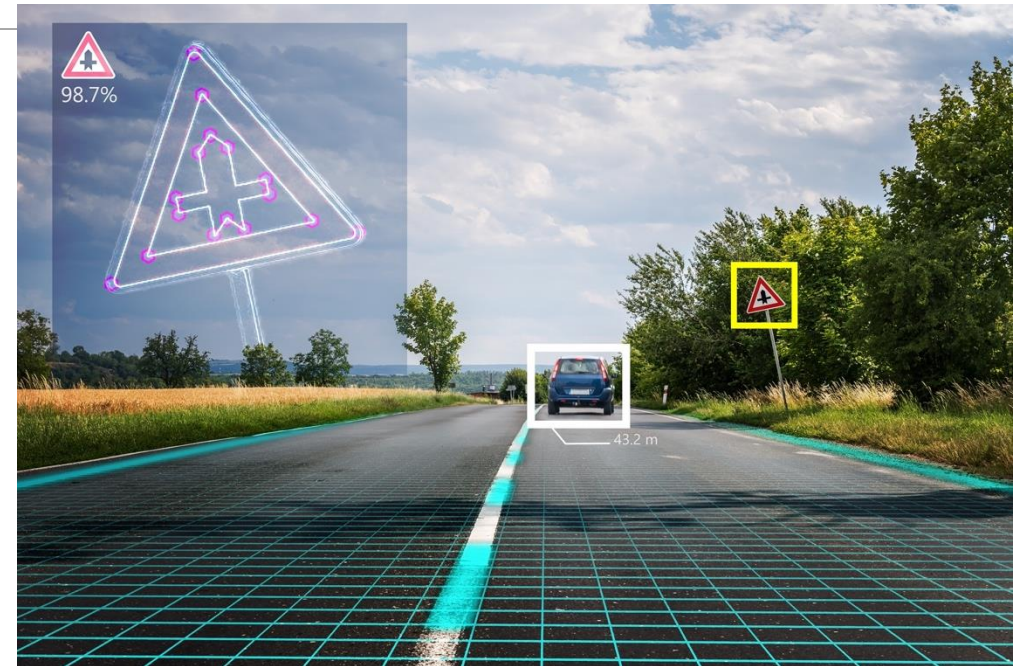


- Fraud detection
- Image recognition
- Personalized advertising
- Medical diagnostics
- Product sales prediction
- Weather forecasting
- Population growth prediction



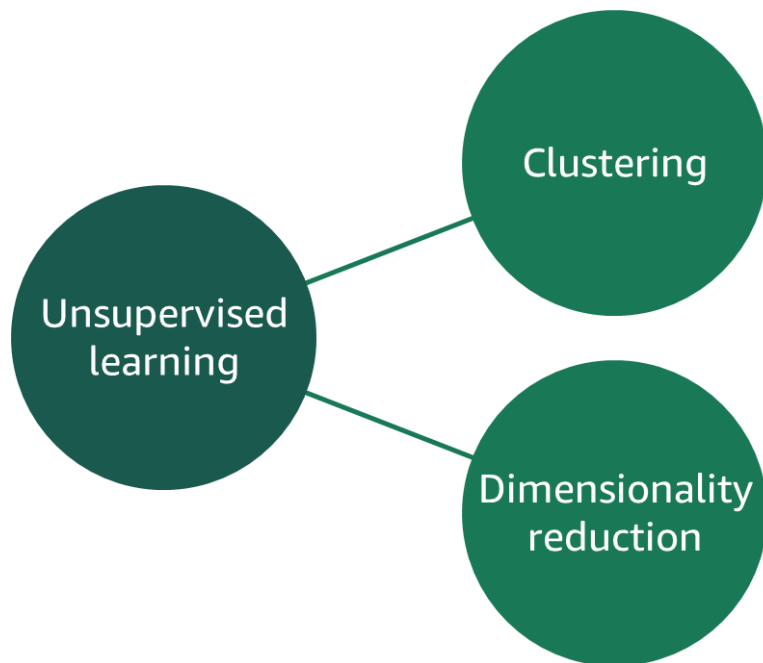
Source: Amazon Web Services

Computer vision



Unsupervised learning

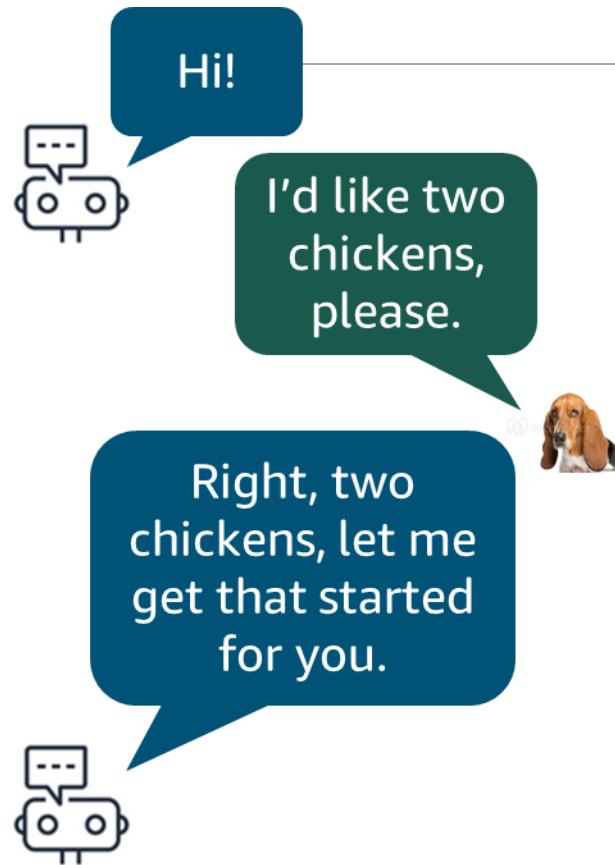
The machine must uncover and **create the labels** itself.



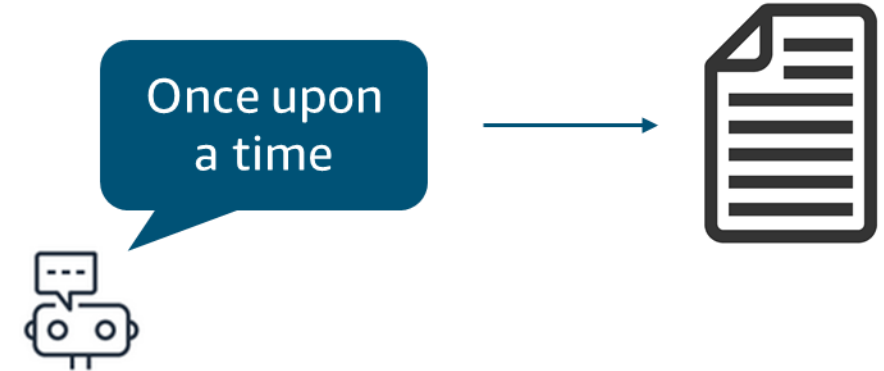
- Product recommendations
- Customer segmentation
- Targeted marketing
- Medical diagnostics
- Visualization
- Natural language processing
- Data structure discovery
- Gene sequencing

Source: Amazon Web Services

Natural language processing



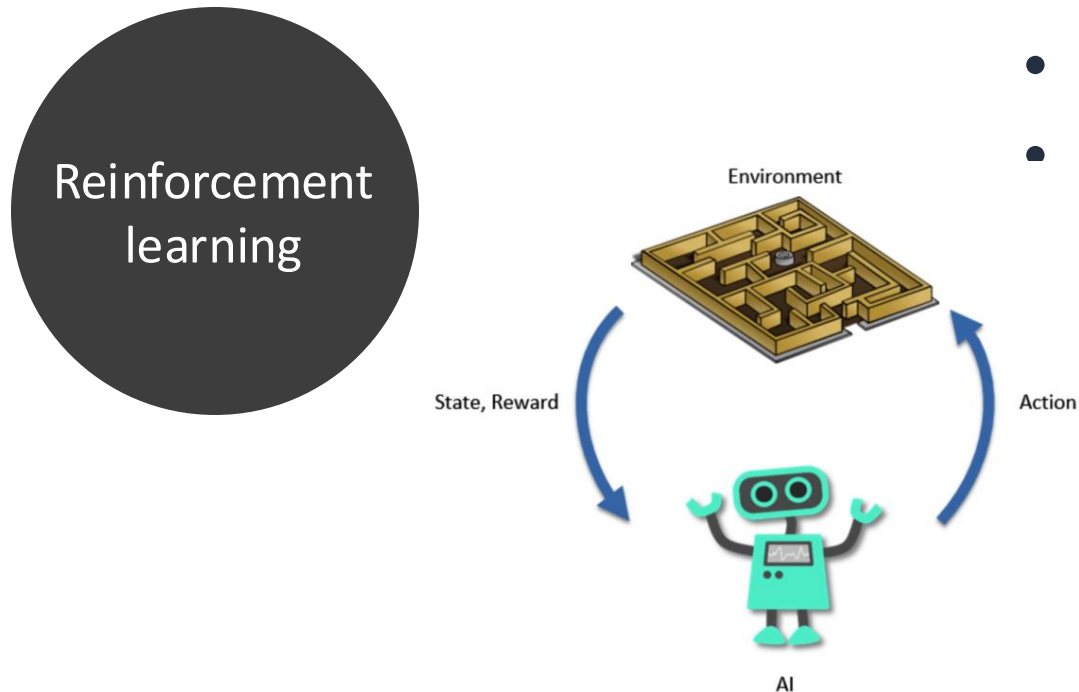
gögn eru lykilatriði \longleftrightarrow los datos son clave



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 let's not talk about the mechanics of building something so huge in such an inhospitable place...

Reinforcement learning

Learning through **trial and error**.



- Game AI
- Self-driving cars
- Robotics



Source: Amazon Web Services,
Image: <https://hub.packtpub.com/5-key-reinforcement-learning-principles-explained-by-ai-expert/>

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

Source: Amazon Web Services

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

Source: Amazon Web Services

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
- Machine learning usability & challenges

Part 3: Machine learning process

ML pipeline: Formulating problem

Business problem

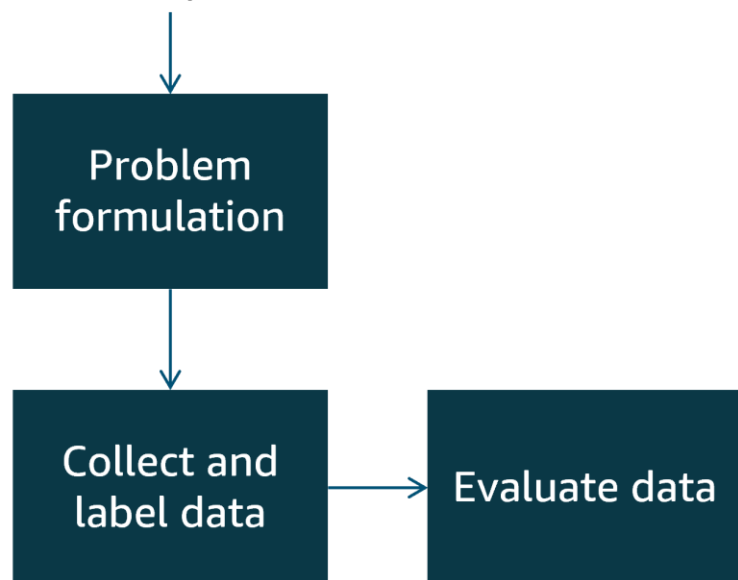


Problem
formulation

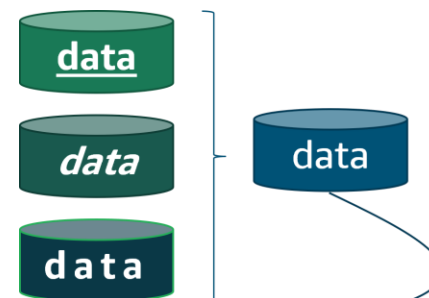
Source: Amazon Web Services

ML pipeline: Data preparation

Business problem



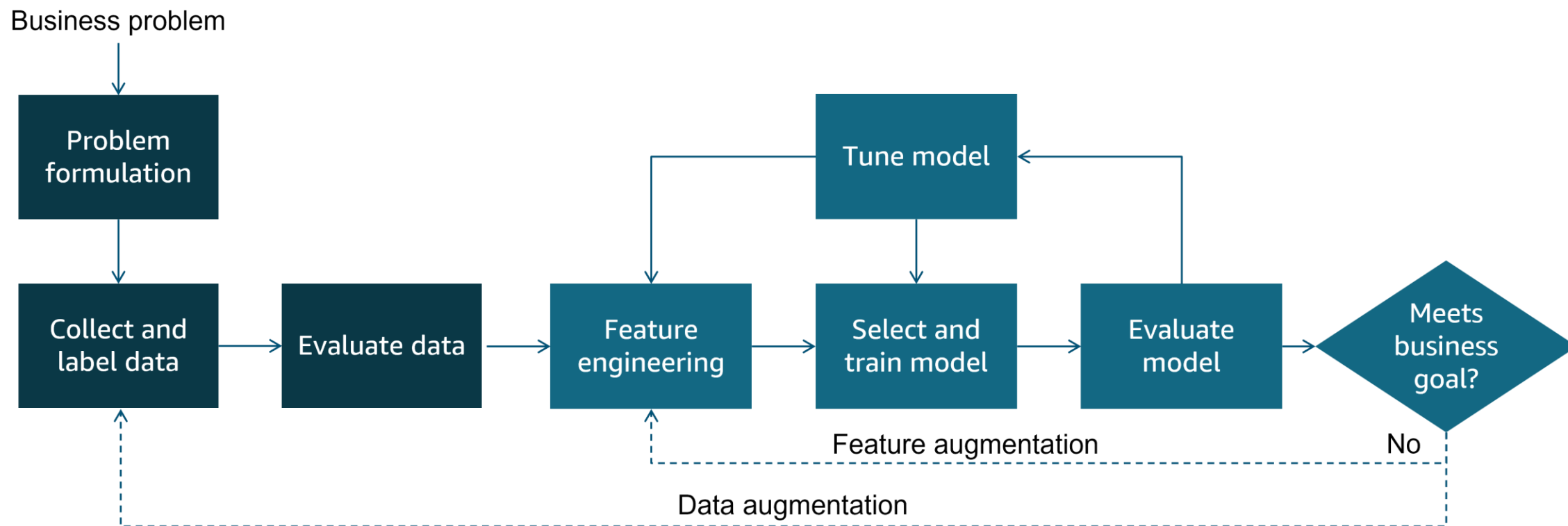
Data handling and cleaning



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

Source: Amazon Web Services

ML pipeline: Iterative model training



Source: Amazon Web Services

ML pipeline: Feature engineering

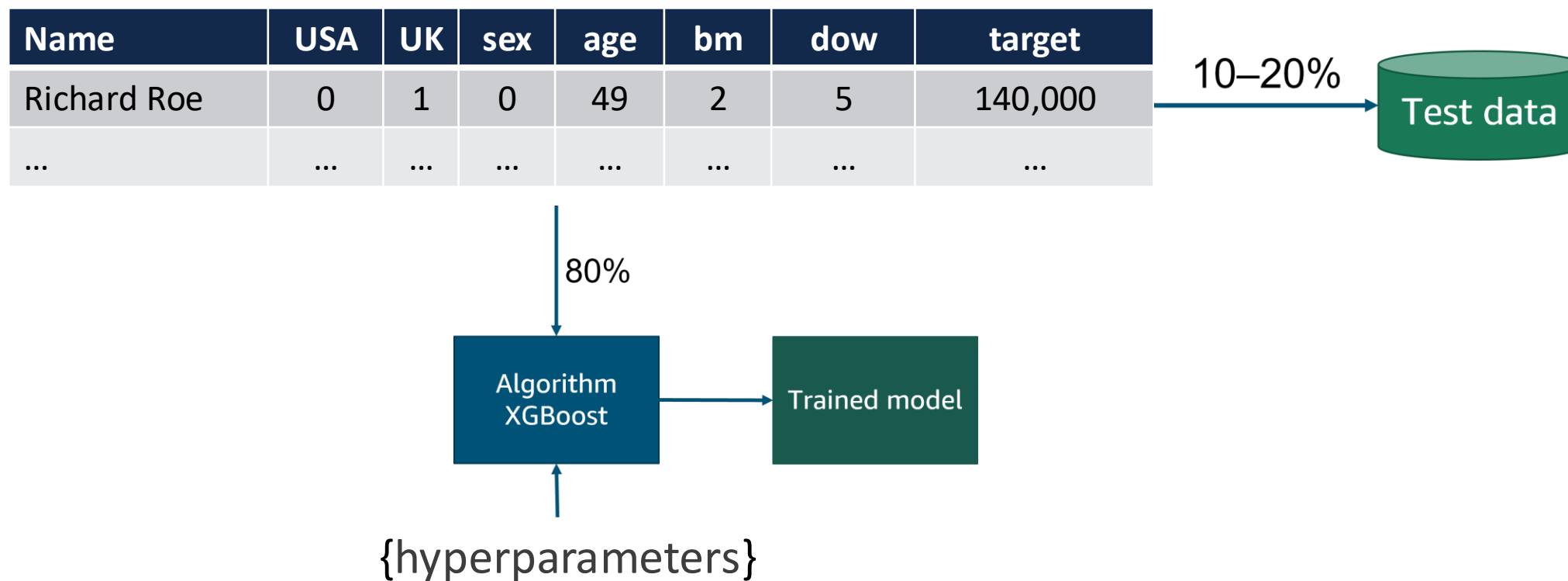
Name	Country	Sex	dob
Richard Roe	UK	Male	18/2/1972
Paulo Santos		Male	11/2/1969
Mrs. Mary Major	Denver	F	37
Desai, Arnav	USA	M	2/22/1962

?

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

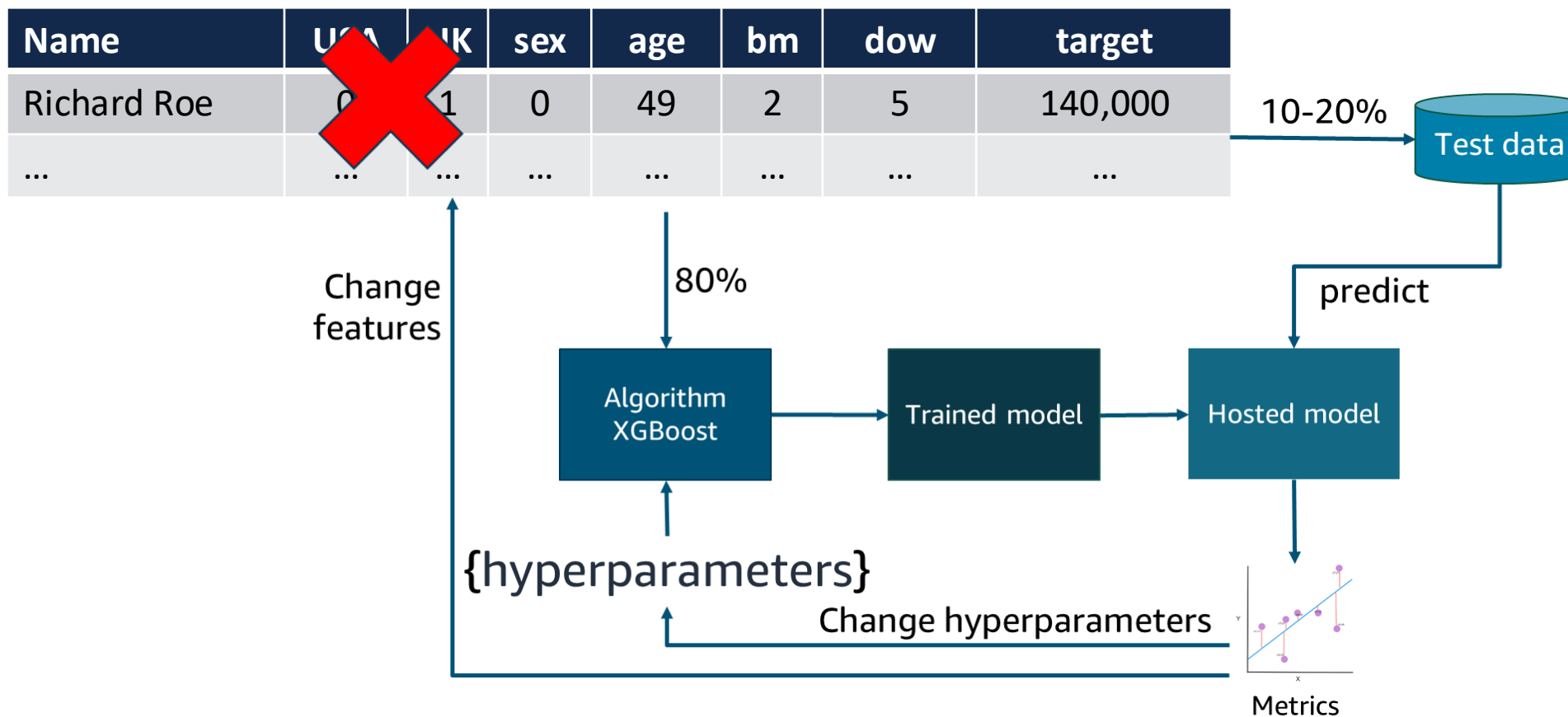
Source: Amazon Web Services

ML pipeline: Model training



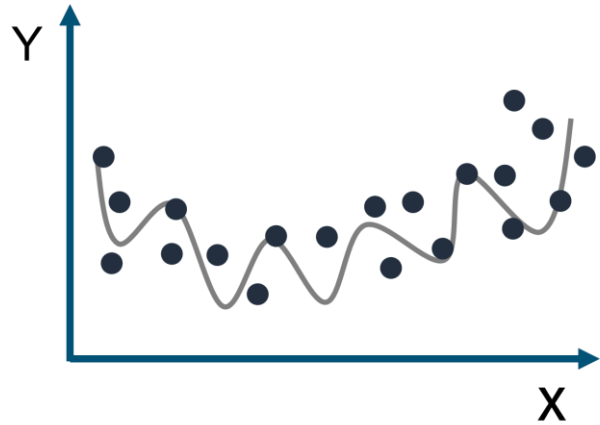
Source: Amazon Web Services

ML pipeline: Evaluating and tuning the model

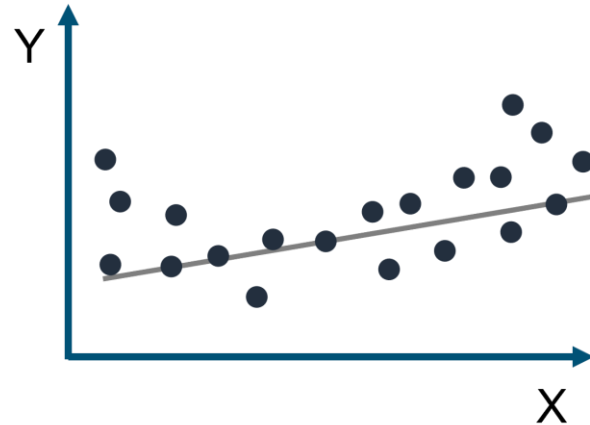


Source: Amazon Web Services

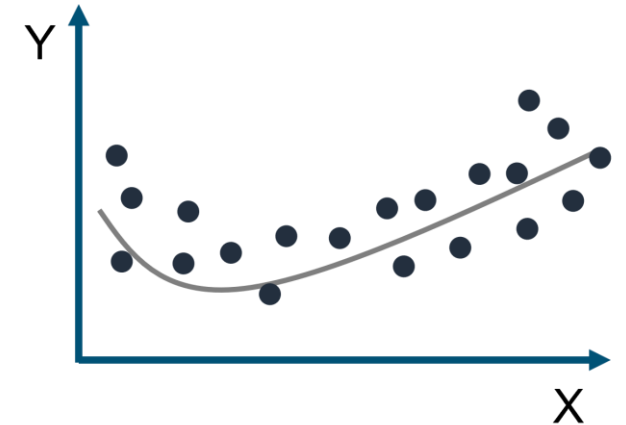
Overfitting and underfitting



Overfitting

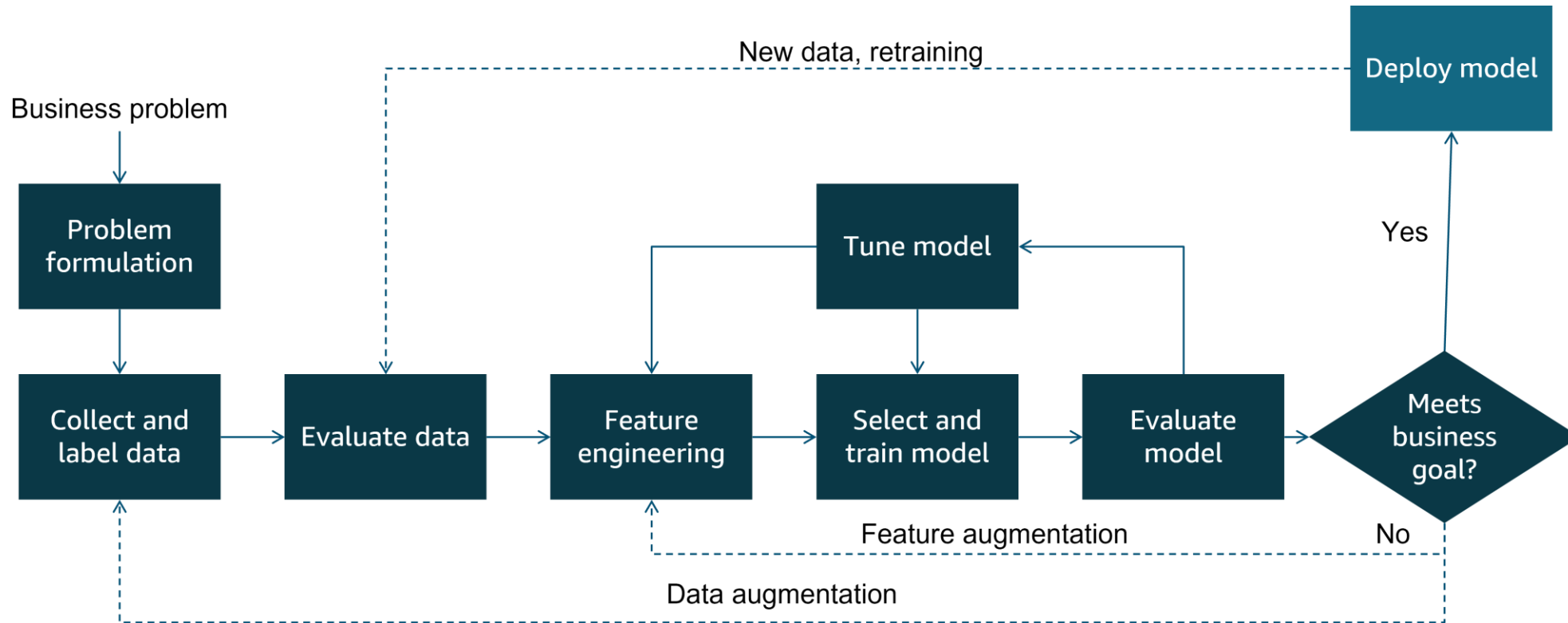


Underfitting



Balanced

ML pipeline: Deployment



Source: Amazon Web Services

Key takeaways



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

Part 4: Machine learning tools

Python tools and libraries

- NumPy
- Pandas
- Matplotlib
- Seaborn
- scikit-learn
- NLTK
- Jupyter Notebook
- Cloud-based ML platform
 - Amazon SageMaker
 - AWS Academy Learner Lab

Source: Amazon Web Services

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.

Source: Amazon Web Services

AWS Academy Learner Lab

- Home
- Announcements 
- Modules
- Discussions
- Grades
- Pages 
- Assignments 
- People 
- Rubrics 
- Quizzes 
- Syllabus 
- Files 
- Outcomes 
- BigBlueButton 
- Collaborations 
- Settings

AWS Academy Learner Lab [67885]



AWS Academy Learner Lab provides a long-running sandbox environment for ad hoc exploration of AWS services. Within this class, students will have access to a **subset of AWS services**. Not all AWS documentation walk-through or sample labs that operate in an AWS Production account will work in the Learner Lab environment. Access to the AWS resources set up in this environment for the duration of this course. We limit your budget (\$100USD), so you should exercise caution to prevent that will deplete your budget too quickly. If you exceed your budget, you will lose access to your environment and lose all of your work.

Each session lasts for 4 hours by default, although you can extend a session to run longer by pressing the start button to reset your session timer. At the end of a session, any resources you created will persist. However, we automatically shut EC2 instances down. Other resources, such as RDS instances, keep running. Keep in mind not to stop some AWS features, so they can still incur charges between sessions. For example, an Elastic Load Balancer or a NAT. You may wish to delete those types and recreate them as needed to test your work during a session. You will have access to this environment for the duration of the class that you are enrolled in. When the class ends, your access to the learner lab will also end.

- Sandbox Environment for exploring AWS services
- Restricted access to services
- \$100 budget for each student
- Lab session lasts for 4 hours by default
- Amazon SageMaker

Source: Amazon Web Services

Key takeaways



- Python is the most popular ML language
- Many open-source tools
- Jupyter Notebooks
- Cloud-based platform
 - Amazon SageMaker
 - AWS Academy Learner Lab

Summary

In this session, you learned how to:

- Describe types of machine learning
- Identify how machine learning can be used to solve real-world problems
- Identify when to use machine learning and what challenges data scientists encounter
- Describe the machine learning process
- List the tools available to train and evaluate machine learning models

Source: Amazon Web Services