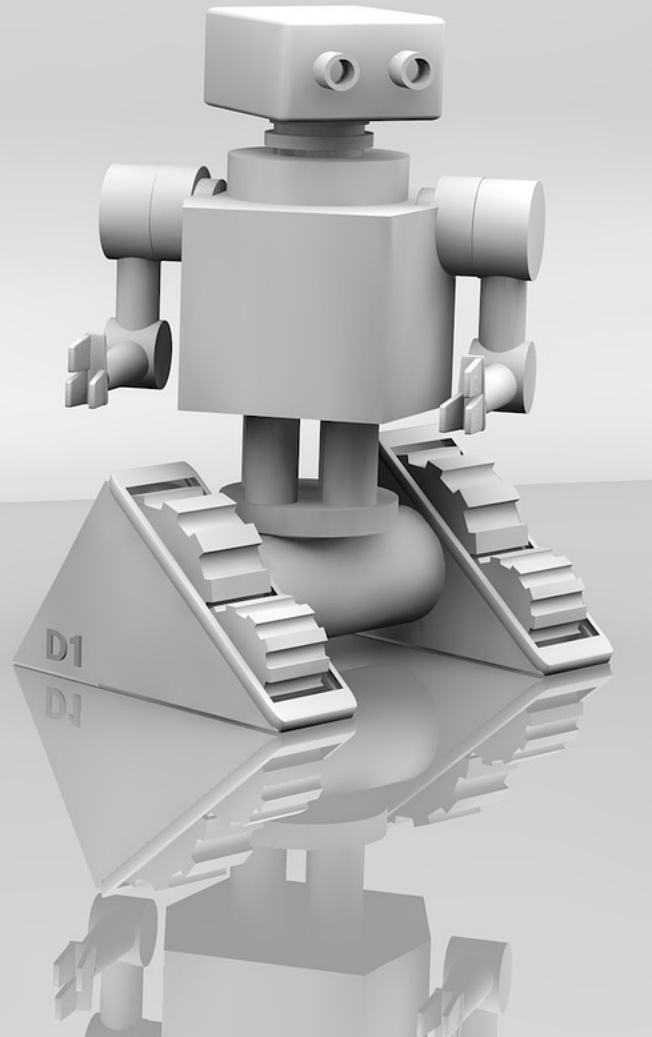
Machine Learning on AWS

Stefan Bergstein



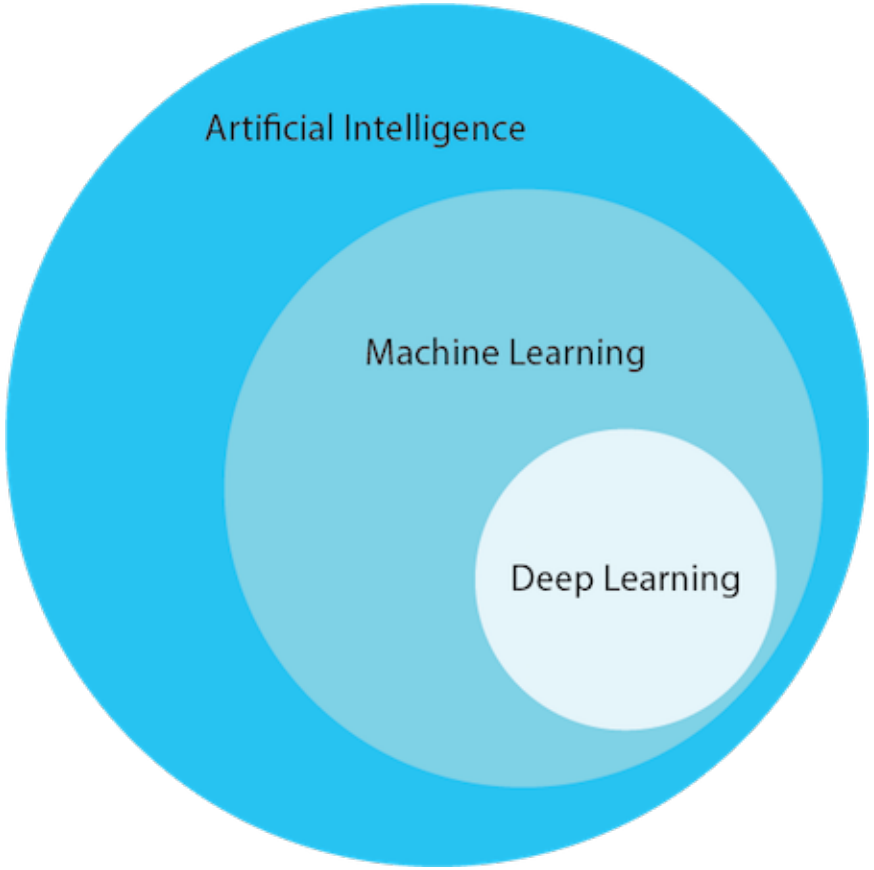
Intro

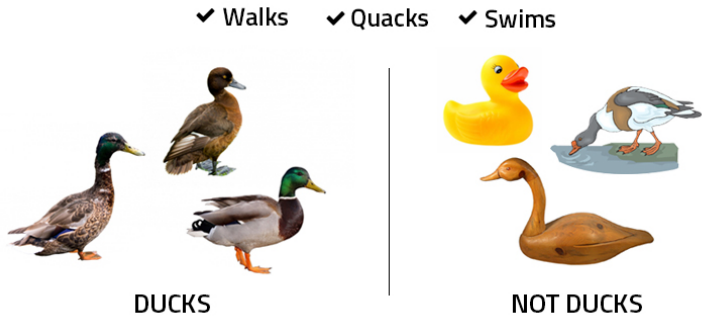
- ML Use cases
- Type of problems
- Learnings techniques
- Algorithms

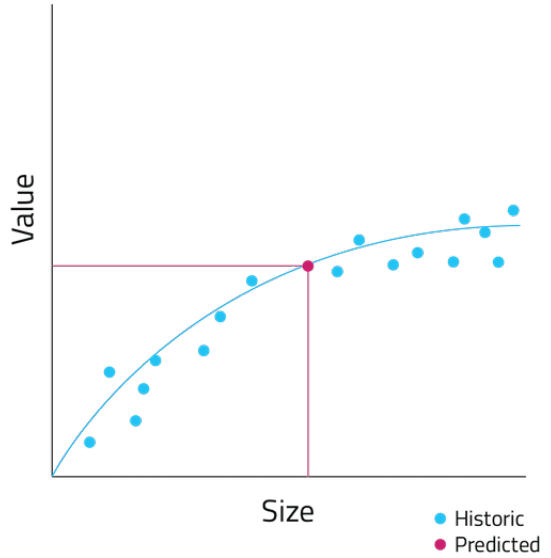
AWS

- Service overview
- AWS ML
- AWS SageMaker
- Deep Learning AMI

Ducks, not ducks, which size?







Source: https://www.moogsoft.com/blog/aiops/understanding-machine-learning-aiops/

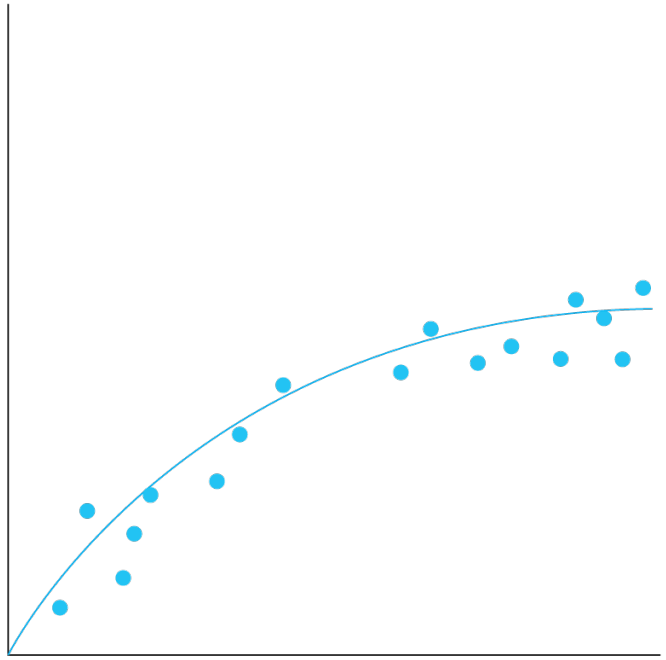
Top AI And Machine Learning Use Cases

- **Data Security -** predict malware
- **Personal Security -** spot things human screeners might miss
- **Financial Trading -** predict what the stock markets will do
- **Healthcare -** spotted cancers a year before were officially diagnosed
- **Marketing Personalization -** lead consumers reliably towards a sale
- **Fraud Detection -** PayPal is using ML to fight money laundering

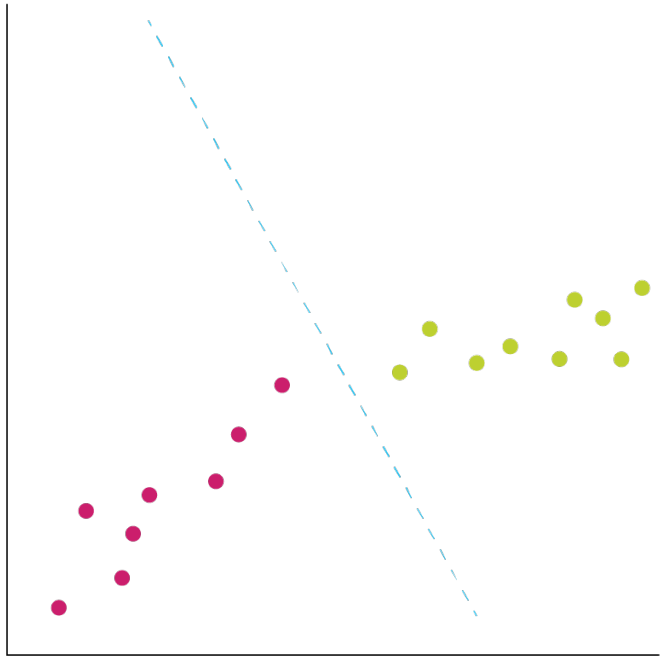
Source: https://www.forbes.com/sites/bernardmarr/2016/09/30/what-are-the-top-10-use-cases-for-machine-learning-and-ai/#424cb0594c90

Types of machine learning problems

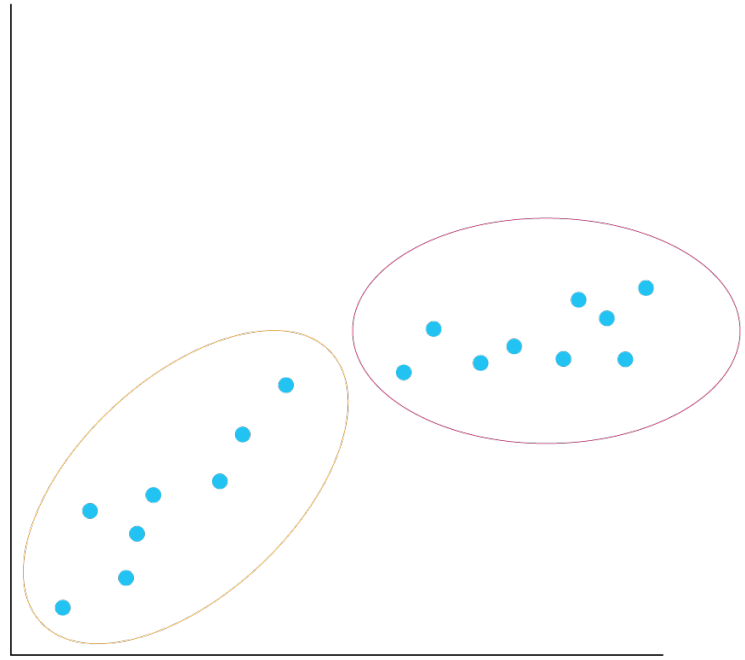
Regression



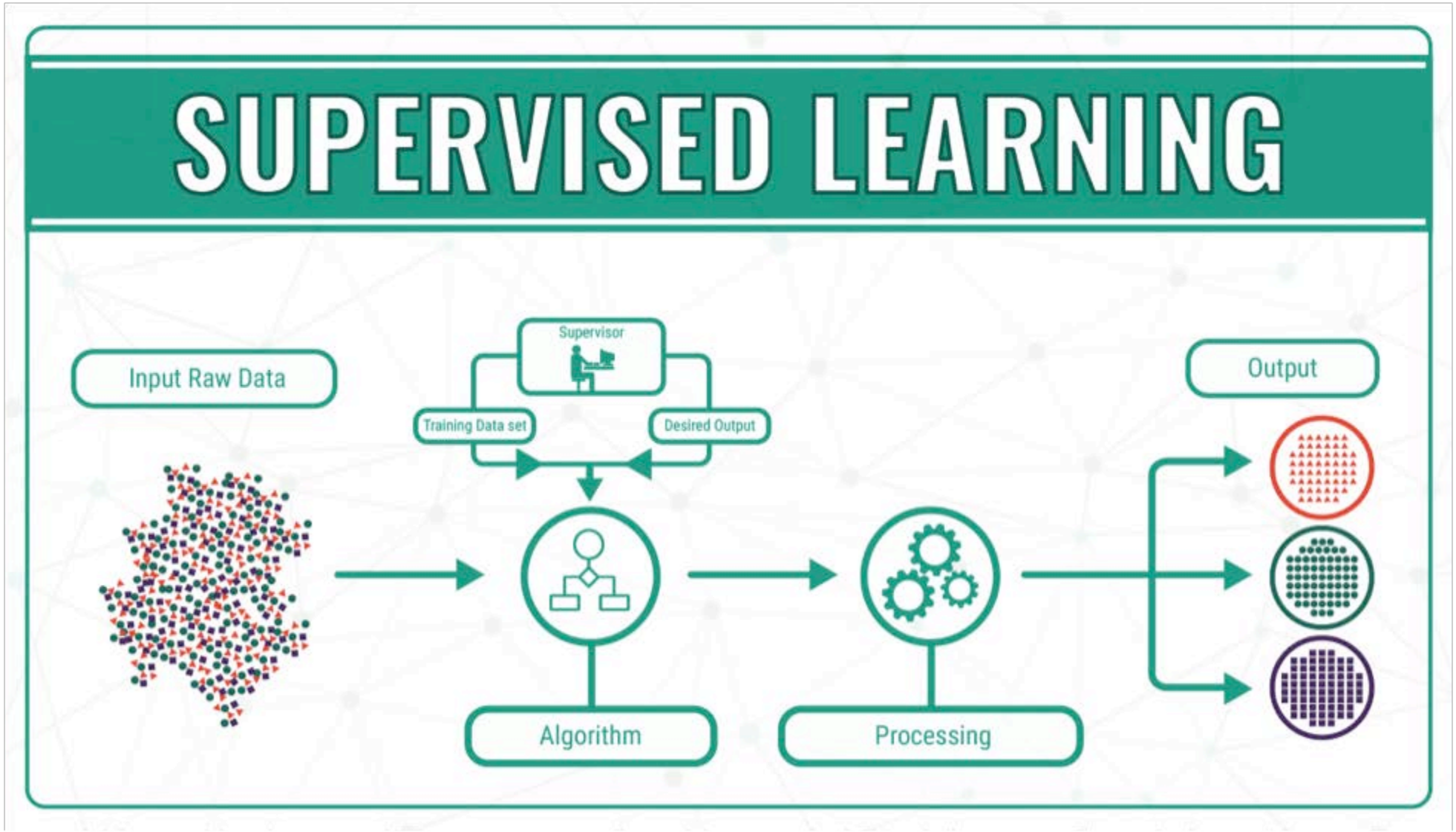
Classification



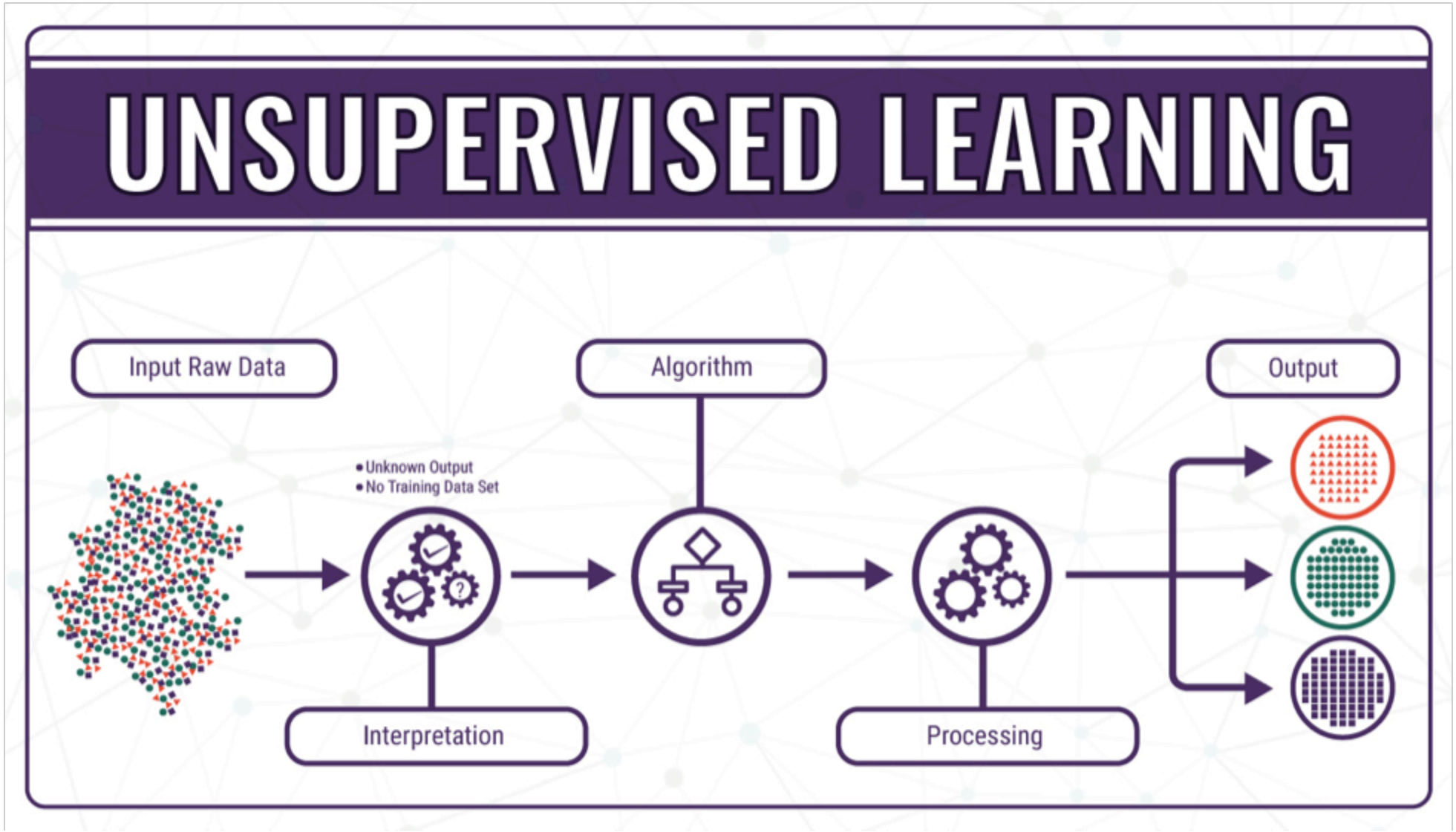
Clustering



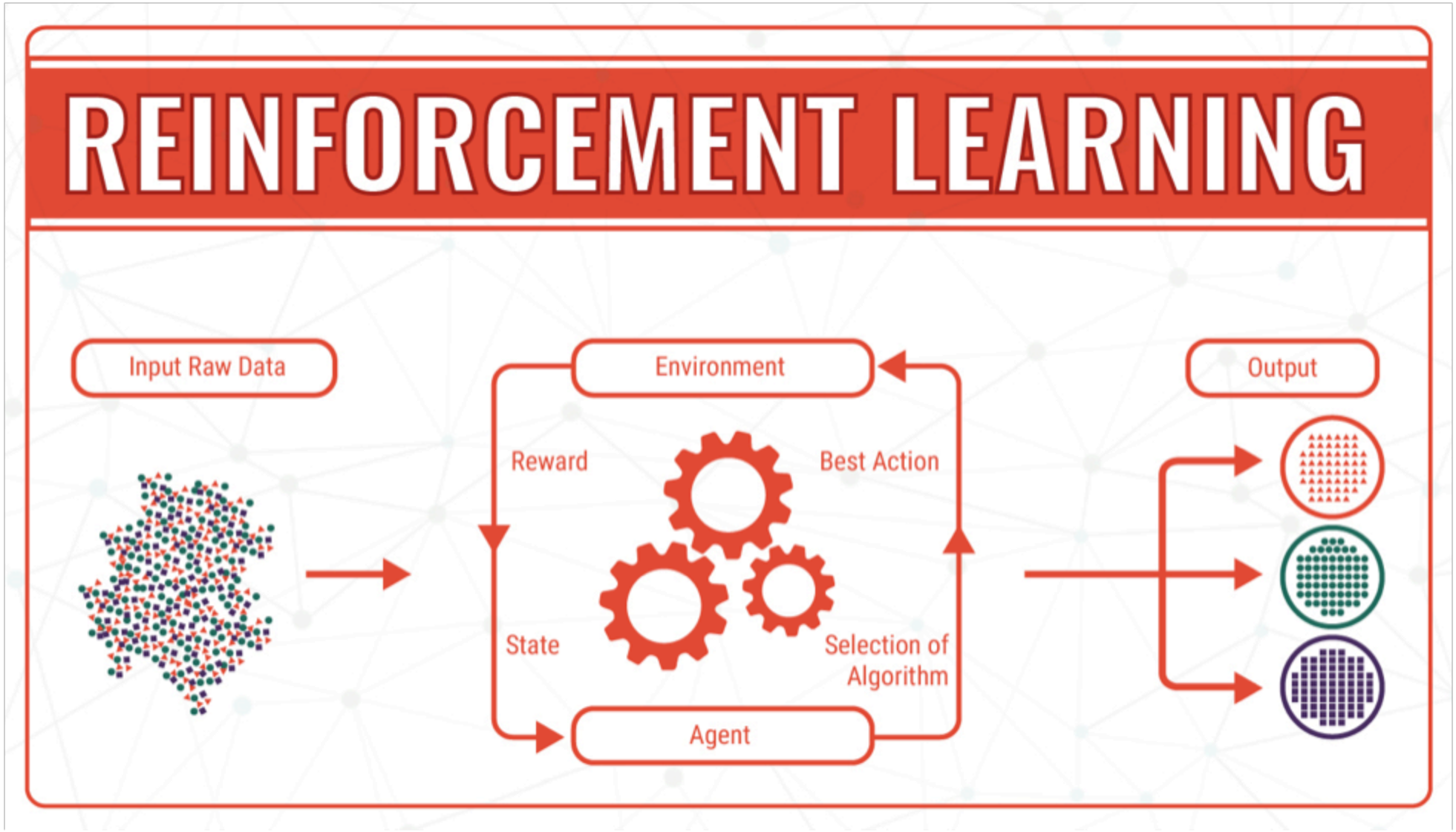
Source: https://www.moogsoft.com/blog/aiops/understanding-machine-learning-aiops/



Source: https://www.datasciencecentral.com/profiles/blogs/machine-learning-explained-understanding-supervised-unsupervised



Source: https://www.datasciencecentral.com/profiles/blogs/machine-learning-explained-understanding-supervised-unsupervised

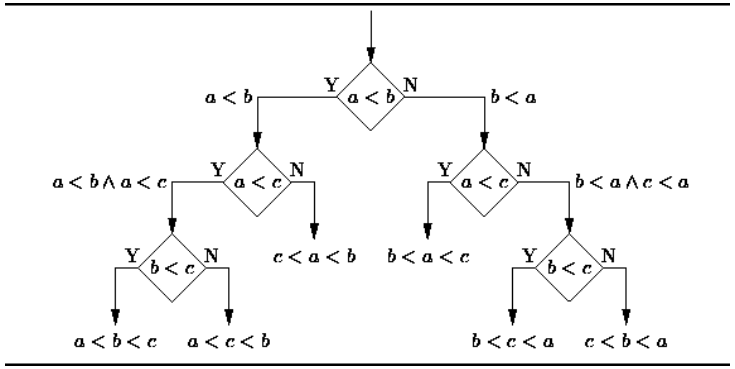


Source: https://www.datasciencecentral.com/profiles/blogs/machine-learning-explained-understanding-supervised-unsupervised

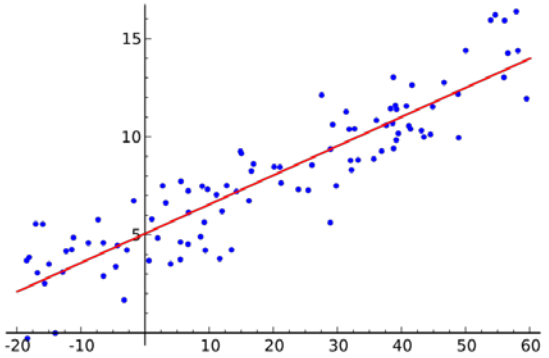
Source: https://s3.amazonaws.com/MLMastery/MachineLearningAlgorithms.png

Algorithms ML Engineers Need to Know (4/10)

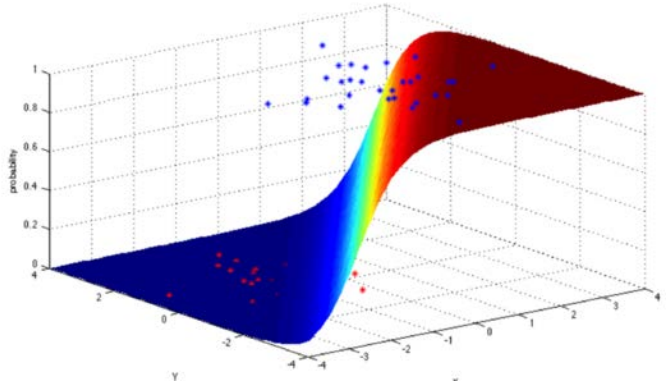
Decision Trees



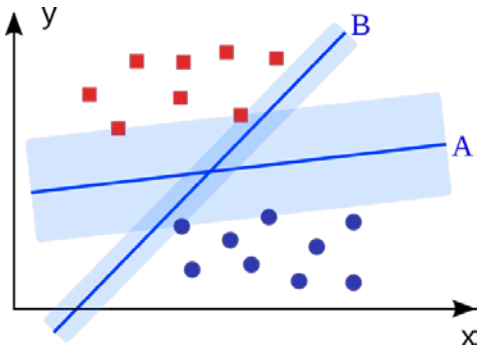
Least Squares Regression



Logistic Regression

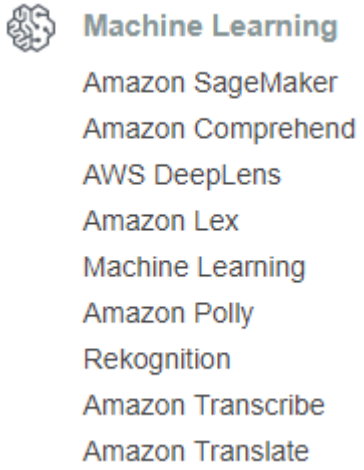


Support Vector Machines



Source: https://www.kdnuggets.com/2016/08/10-algorithms-machine-learning-engineers.html

AWS ML Services



Today's discussion:

Amazon Machine Learning
Amazon SageMaker
AWS Deep Learning AMI

Solving Business Problems with Amazon ML

Examples of binary classification problems

Examples of multiclass classification problems

Examples of regression classification problems

Source: https://docs.aws.amazon.com/machine-learning/latest/dg/machine-learning-problems-in-amazon-machine-learning.html

Solving Business Problems with Amazon ML

- Examples of binary classification problems:
  - Will the customer buy this product or not buy this product?
  - Is this email spam or not spam?
  - Is this product a book or a farm animal?
  - Is this review written by a customer or a robot?

Source: https://docs.aws.amazon.com/machine-learning/latest/dg/machine-learning-problems-in-amazon-machine-learning.html

Solving Business Problems with Amazon ML

- Examples of multiclass classification problems:
  - Is this product a book, movie, or clothing?
  - Is this movie a romantic comedy, documentary, or thriller?
  - Which category of products is most interesting to this customer?

Source: https://docs.aws.amazon.com/machine-learning/latest/dg/machine-learning-problems-in-amazon-machine-learning.html

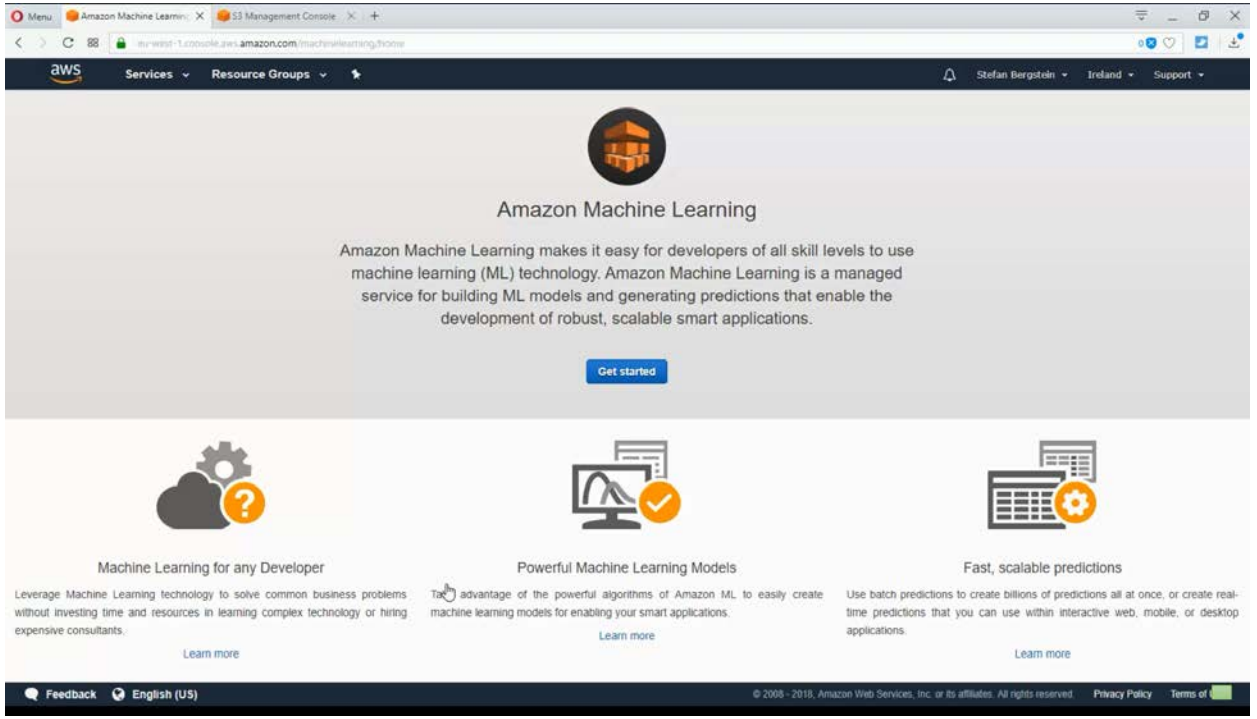
Solving Business Problems with Amazon ML

- Examples of regression classification problems:
  - What will the temperature be in Seattle tomorrow?
  - For this product, how many units will sell?
  - How many days before this customer stops using the application?
  - What price will this house sell for?

Source: https://docs.aws.amazon.com/machine-learning/latest/dg/machine-learning-problems-in-amazon-machine-learning.html

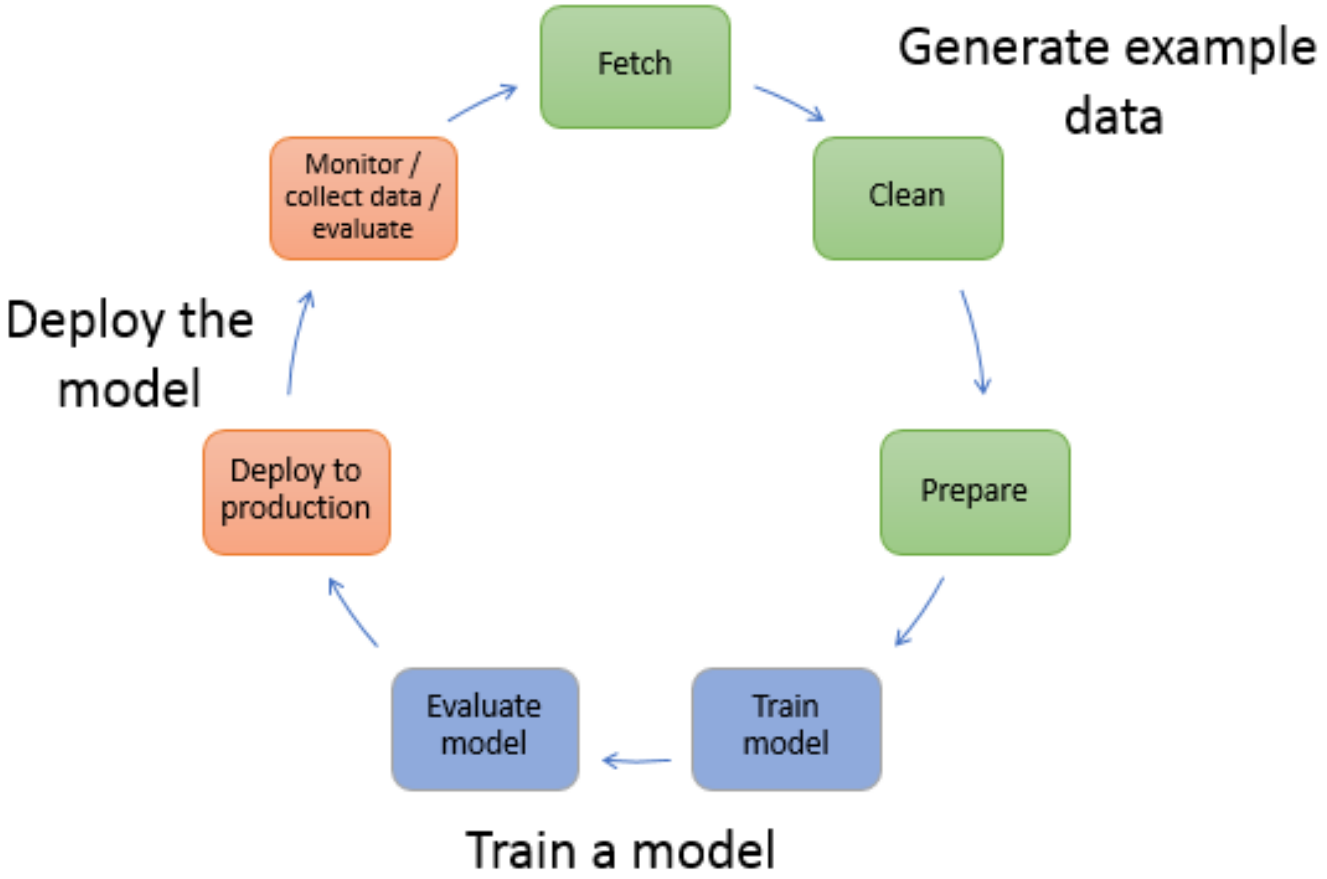
Amazon Machine Learning Demo

- Live Demo:
- AWS Tutorial:
  Using Amazon ML to Predict Responses to a Marketing Offer



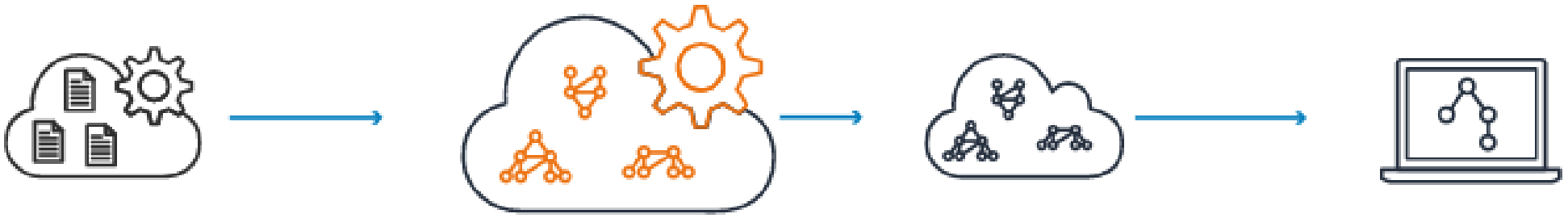
DIY:
https://docs.aws.amazon.com/machine-learning/latest/dg/tutorial.html?icmpid=docs_machinelearning_console

Typical workflow for creating a ML model



Source: https://docs.aws.amazon.com/sagemaker/latest/dg/how-it-works-mlconcepts.html

SageMaker - How it works



Build

Connect to other AWS services and transform data in Amazon SageMaker notebooks

Train

Use Amazon SageMaker's algorithms and frameworks, or bring your own, for distributed training

Tune

Amazon SageMaker automatically tunes your model by adjusting multiple combinations of algorithm parameters

Deploy

Once training is completed, models can be deployed to Amazon SageMaker endpoints, for real-time predictions

Source: https://eu-west-1.console.aws.amazon.com/sagemaker/home?region=eu-west-1#/landing

SageMaker - Benefits and features

Fully-managed notebook instances

For training data exploration and preprocessing, Amazon SageMaker provides fully managed instances running Jupyter notebooks that include example code for common model training and hosting exercises.

Highly-optimized machine learning algorithms

Amazon SageMaker installs high-performance, scalable machine learning algorithms optimized for speed, scale, and accuracy, to run on extremely large training datasets.

One-click training

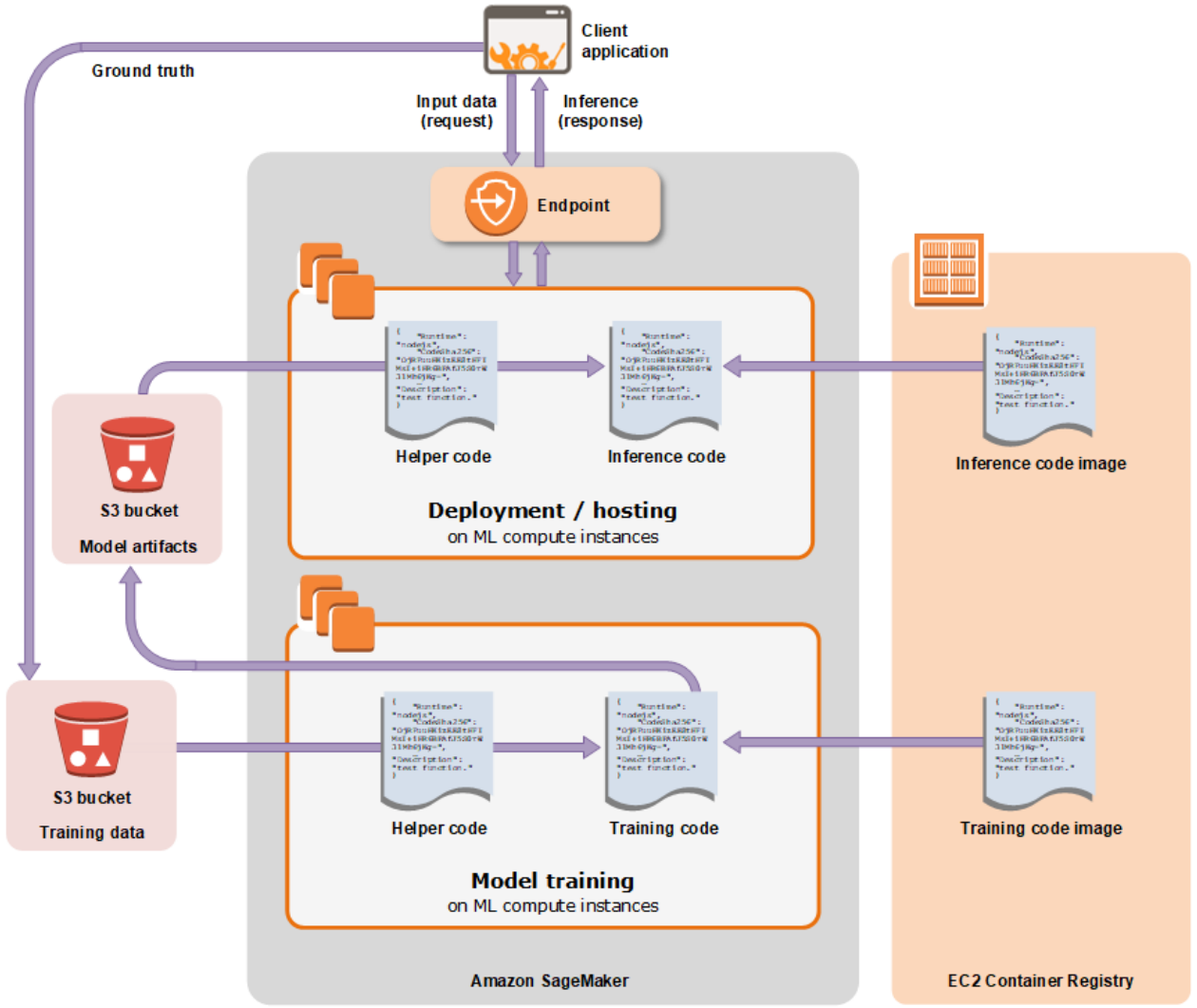
When you're ready to train in Amazon SageMaker, simply indicate the type and quantity of instances you need and initiate training with a single click.

Deployment without engineering effort

After training, SageMaker provides the model artifacts and scoring images to you for deployment to Amazon EC2 or anywhere else.

Source: https://eu-west-1.console.aws.amazon.com/sagemaker/home?region=eu-west-1#/landing

Training a Model with Amazon SageMaker



Options for a training algorithm:

A. Use a built-in algorithm.
(see Using Built-in Algorithms with Amazon SageMaker)

B. Use SageMaker's Apache Spark MLLib

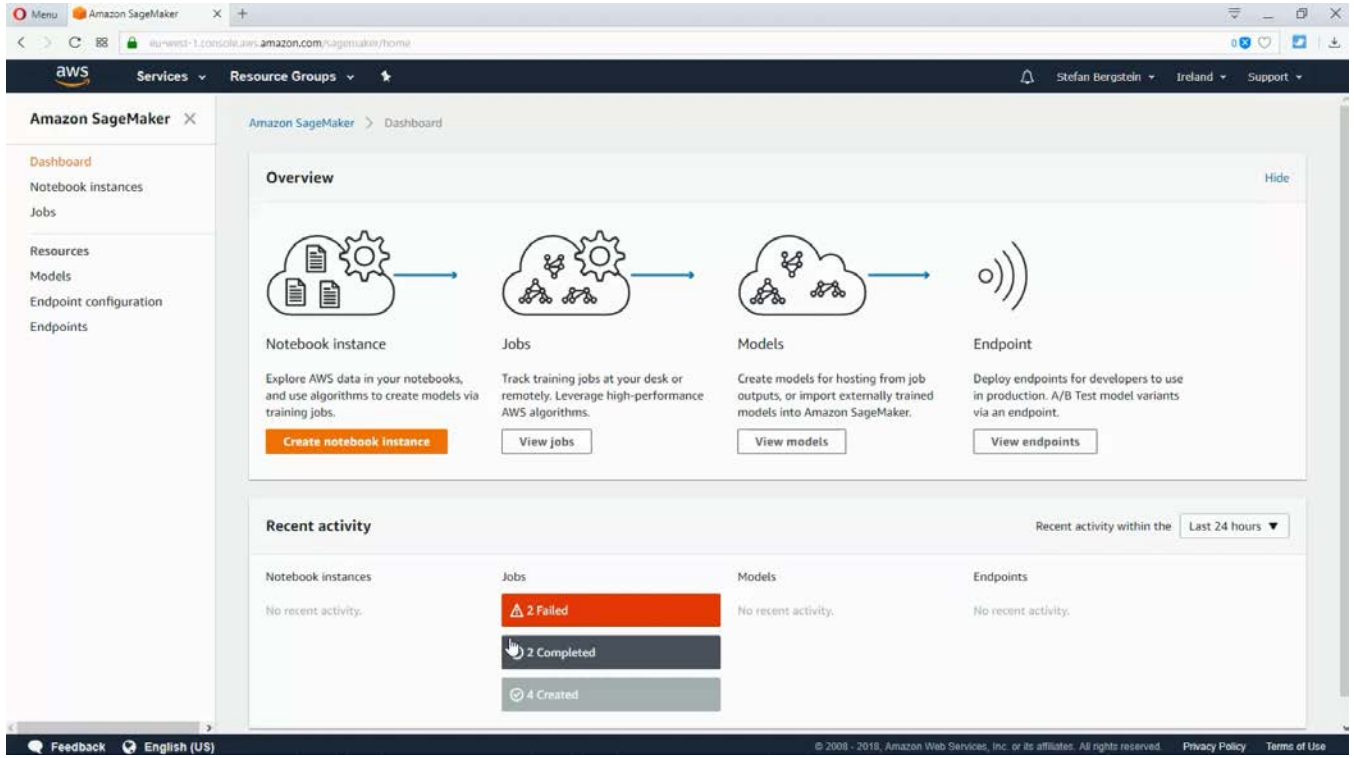
C. Submit custom code to train with deep learning frameworks (TensorFlow or Apache MXNet)

D. Provide your own custom algorithms as a Docker container image

Source: https://docs.aws.amazon.com/sagemaker/latest/dg/how-it-works-mlconcepts.html

Amazon SageMaker Demo

- Live Demo:
- SageMaker sample notebook
- Multiclass classification with Amazon SageMaker XGBoost algorithm



AWS Deep Learning AMIs

Support of popular deep learning frameworks:

Accelerate model training:

Three types of AMIs:

CONDA AMI

For developers who want pre-installed pip packages of deep learning frameworks in separate virtual environments, the Deep Learning Conda-based AMI is available in in **Ubuntu** and **Amazon Linux** versions.

BASE AMI

For developers who want a clean slate to set up private deep learning engine repositories or custom builds of deep learning engines, the Deep Learning Base AMI is available in **Ubuntu** and **Amazon Linux** versions.

AMI WITH SOURCE CODE

For developers who want pre-installed deep learning frameworks and their source code in a shared Python environment, this Deep Learning AMI is available for P3 instances in **CUDA 9** **Ubuntu** and **Amazon Linux** versions as well as for P2 instances in **CUDA 8** **Ubuntu** and **Amazon Linux** versions.

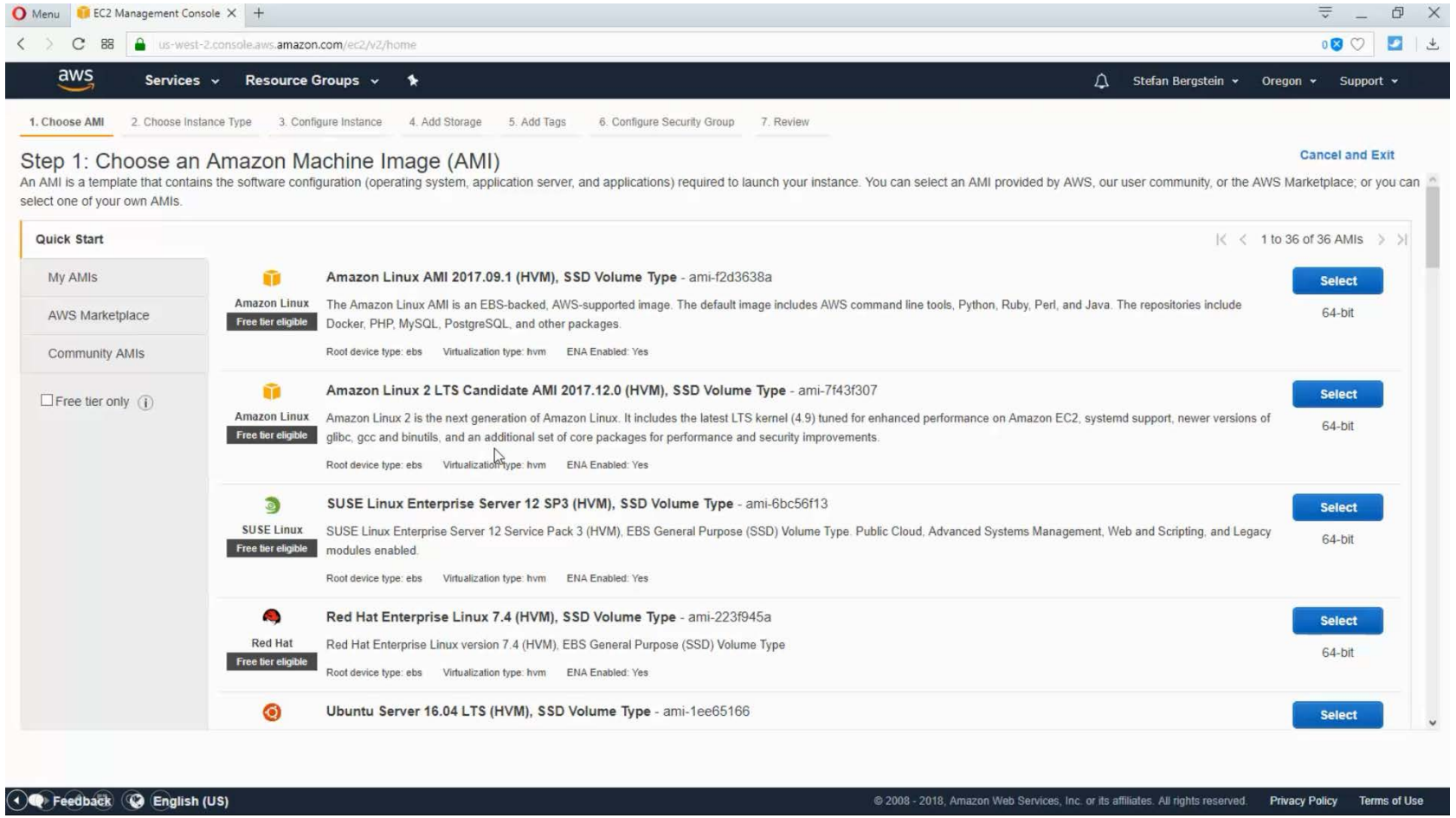
Source: https://docs.aws.amazon.com/dlami/latest/devguide/conda.html

DL AMIs available in different regions

Region	Code
US East (Ohio)	ec2-us-east-2
US East (N. Virginia)	ec2-us-east-1
US West (Oregon)	ec2-us-west-2
Beijing (China)	cn-north-1
Asia Pacific (Mumbai)	ec2-ap-south-1
Asia Pacific (Seoul)	ec2-ap-northeast-2
Asia Pacific (Singapore)	ec2-ap-southeast-1
Asia Pacific (Sydney)	ec2-ap-southeast-2
Asia Pacific (Tokyo)	ec2-ap-northeast-1
EU (Frankfurt)	ec2-eu-central-1
EU (Ireland)	ec2-eu-west-1

Source: https://docs.aws.amazon.com/dlami/latest/devguide/conda.html

AWS Deep Learning AMI Demo



Thank you!