Introduction to Machine Learning Models

04/01/2021

with

Burak Gozluklu, PhD
AWS Solutions Architect

aws training and certification

Foreword

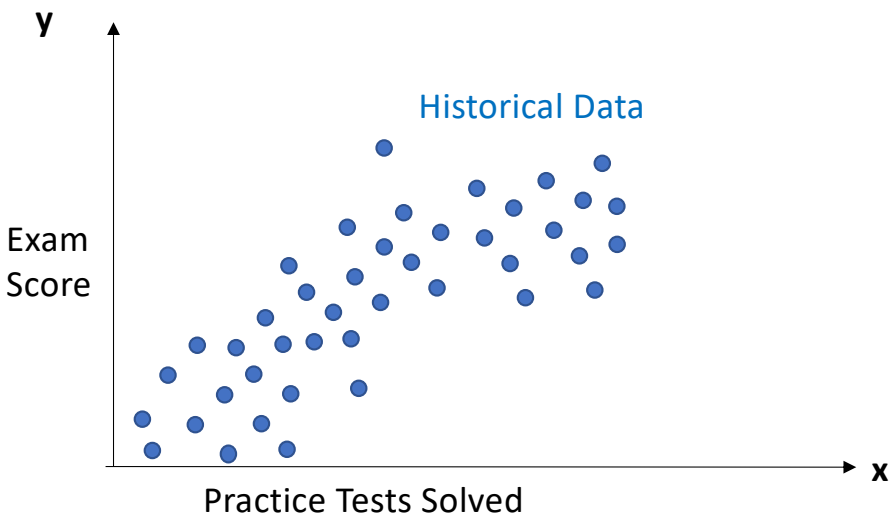
- It is a *crash course* on AI/ML modeling.
- It is an introduction and expected you to ask me questions.
- Do not worry if you cannot catch up with everything!

It may seem difficult at first, but
everything is difficult at first.
Miyamoto Musashi
quotefancy

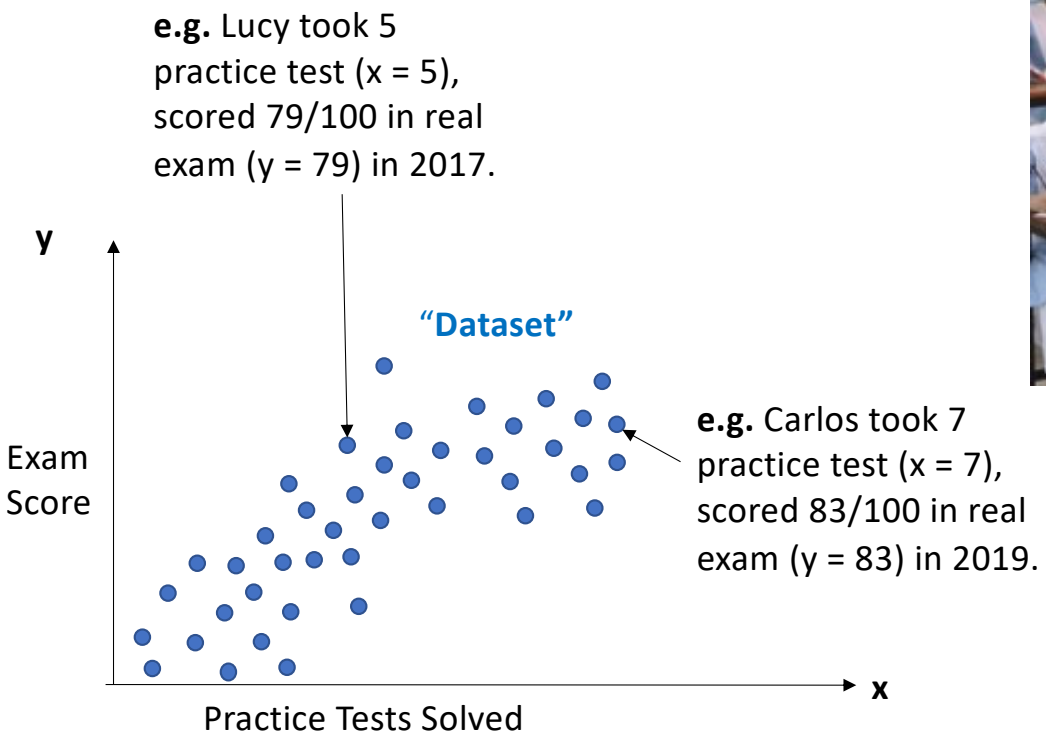
Example: Data Scientist in an Education Institute

Problem: You are expected to create a service that **predicts** the *exam score* based on the number of practice tests solved so that the student is advised about how many extra practice tests is recommended to reach the desired score before entering the actual exam.

Example: Data Scientist in an Education Institute

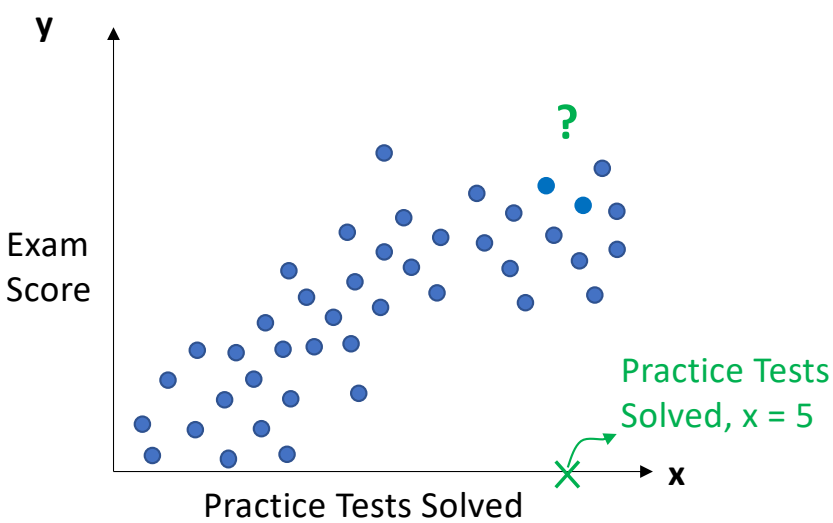


Example: Data Scientist in an Education Institute



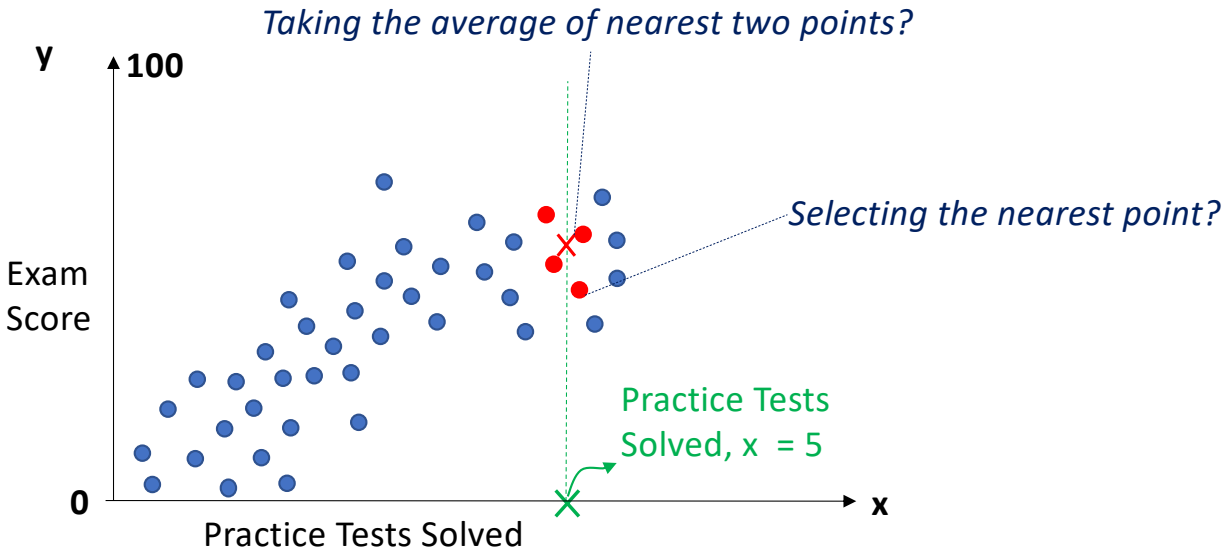
Example: Data Scientist in an Education Institute

"I solved 5 practice tests, what will be my predicted exam score?"



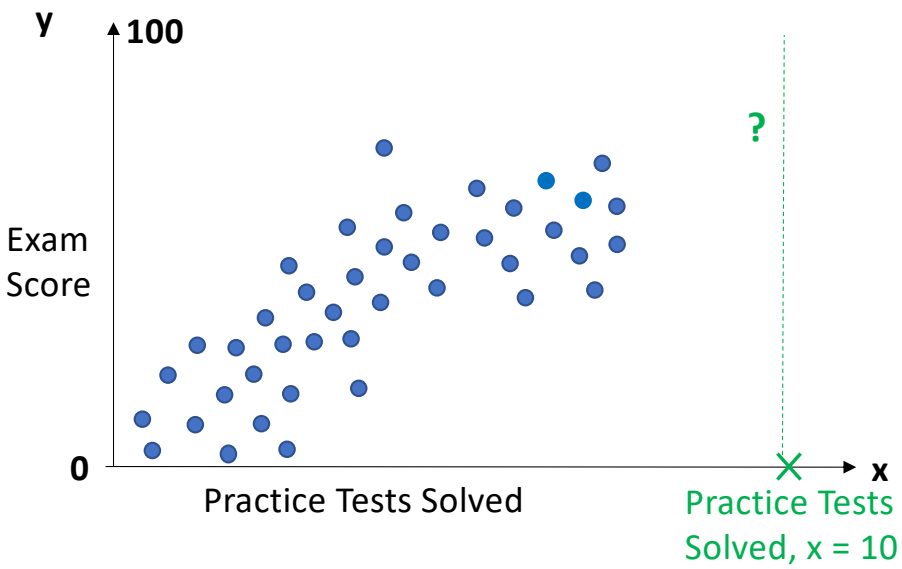
Example: Data Scientist in an Education Institute

"I solved 5 practice tests, what will be my predicted exam score?"



Example: Data Scientist in an Education Institute

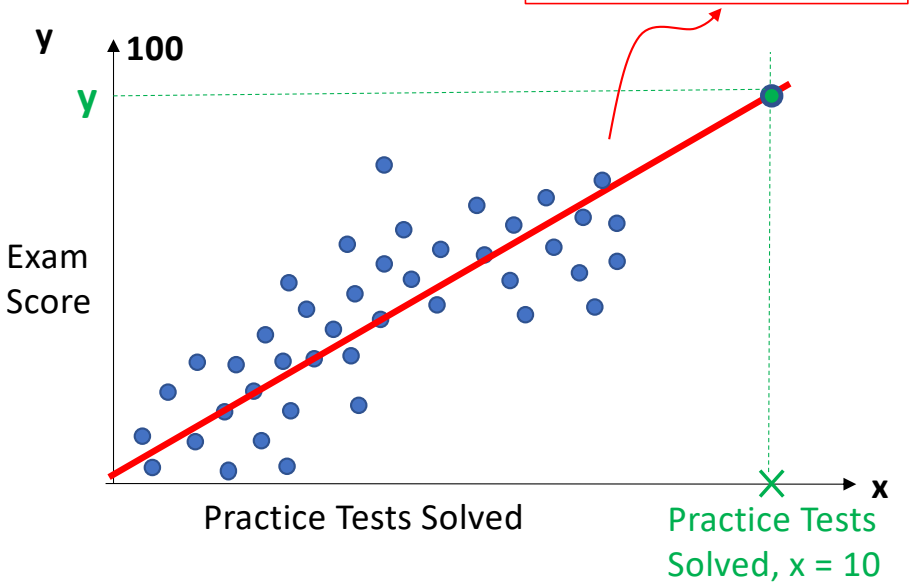
"I solved 10 practice tests, what will be my predicted exam score?"



Example: Data Scientist in an Education Institute

Machine Learning Model: *Linear Regression (line fitting)*

General Model: $Y = f(\mathbf{x}) = a_1\mathbf{x} + a_0$

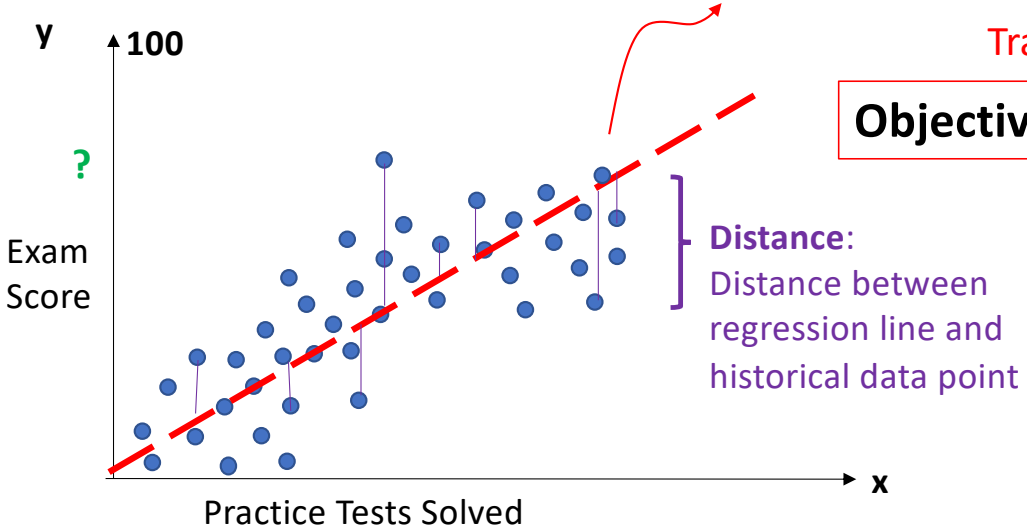


If we can find a_0, a_1, we can estimate Exam Scores from Practice Tests

Example: Data Scientist in an Education Institute

Machine Learning Model: *Linear Regression (line fitting)*

General Model: $Y = f(\mathbf{x}) = a_1\mathbf{x} + a_0$



Training:

Objective = minimize(distance) using historical data

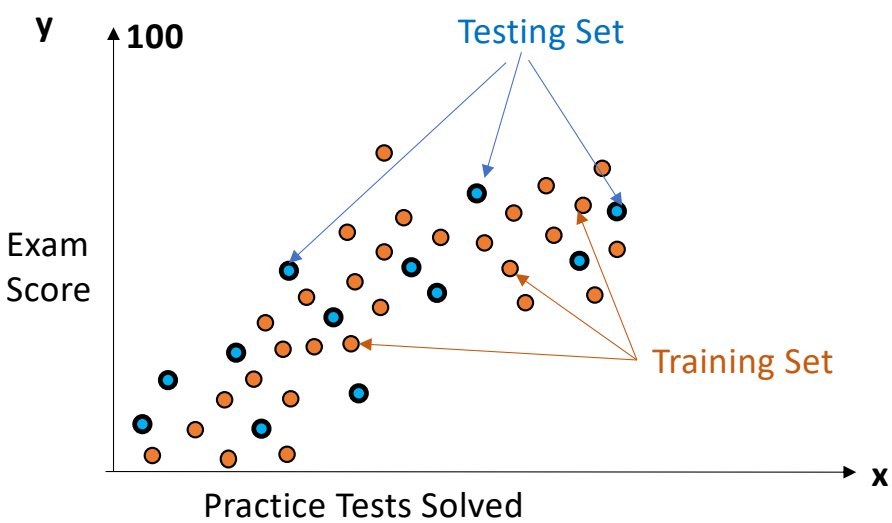
Finding the **optimum** fit

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Randomly split the data set into *Training* and *Testing*

80/20 split



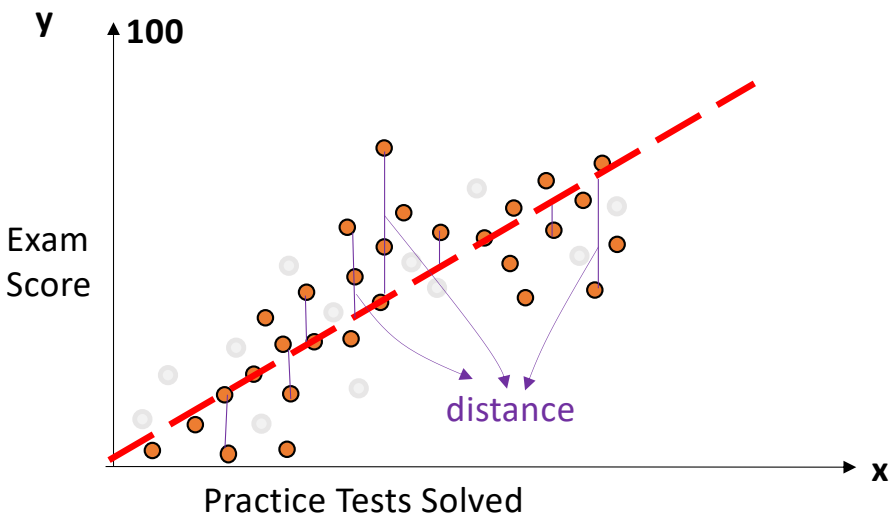
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Training:

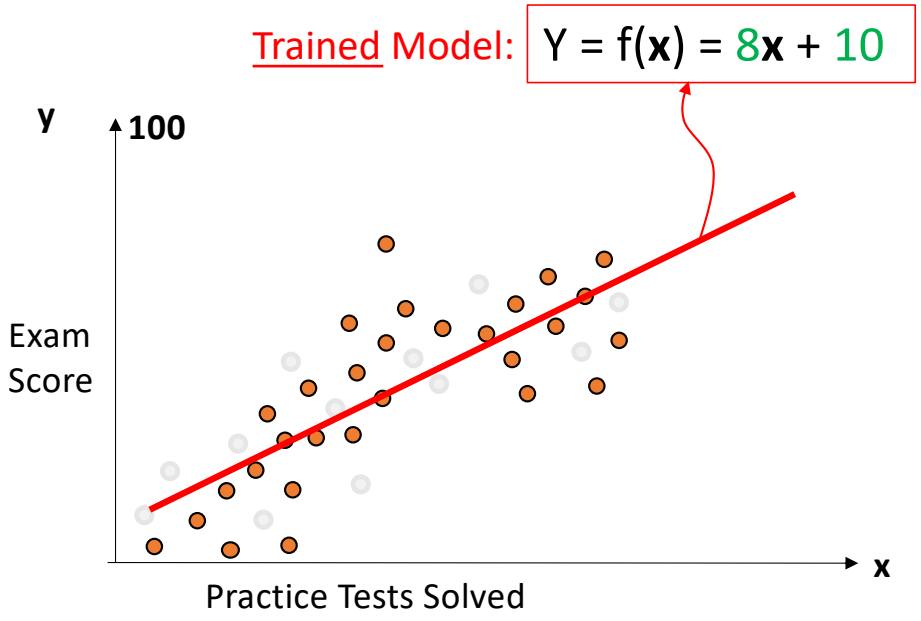
Training **data set is used to** minimize(distance) for training dataset **to find** a_1 **and** a_0

General Model: $Y = f(\mathbf{x}) = a_1\mathbf{x} + a_0$



© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute



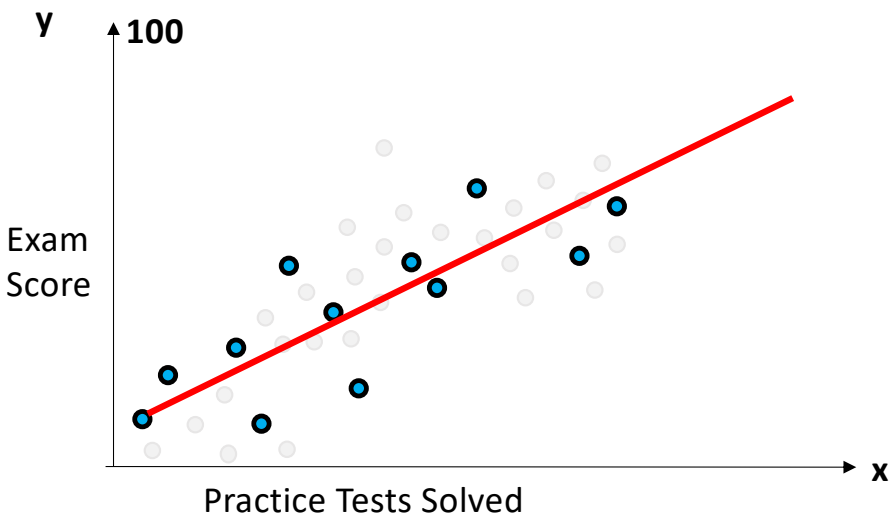
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Test:

Testing **data set is used to** Test model *error*

Trained Model: $Y = f(\mathbf{x}) = 8\mathbf{x} + 10$



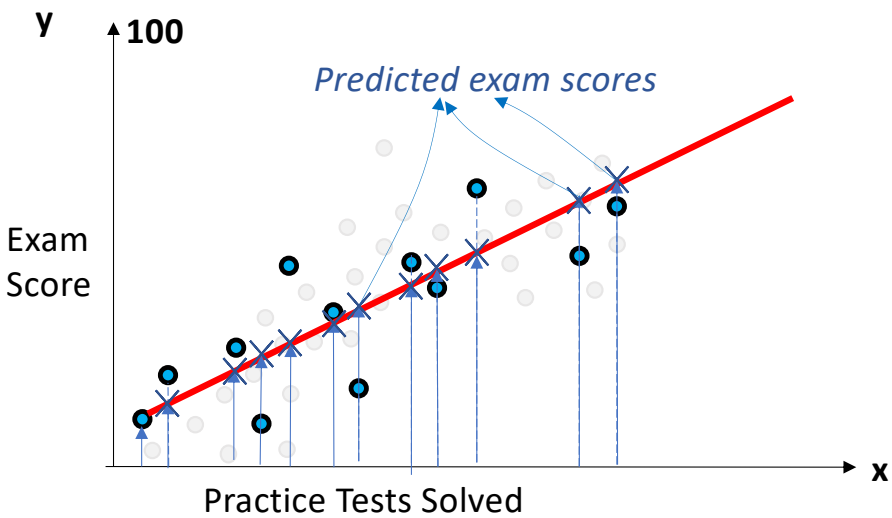
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Test:

Testing **data set is used to** Test model *error*

Trained Model: $Y = f(\mathbf{x}) = 8\mathbf{x} + 10$



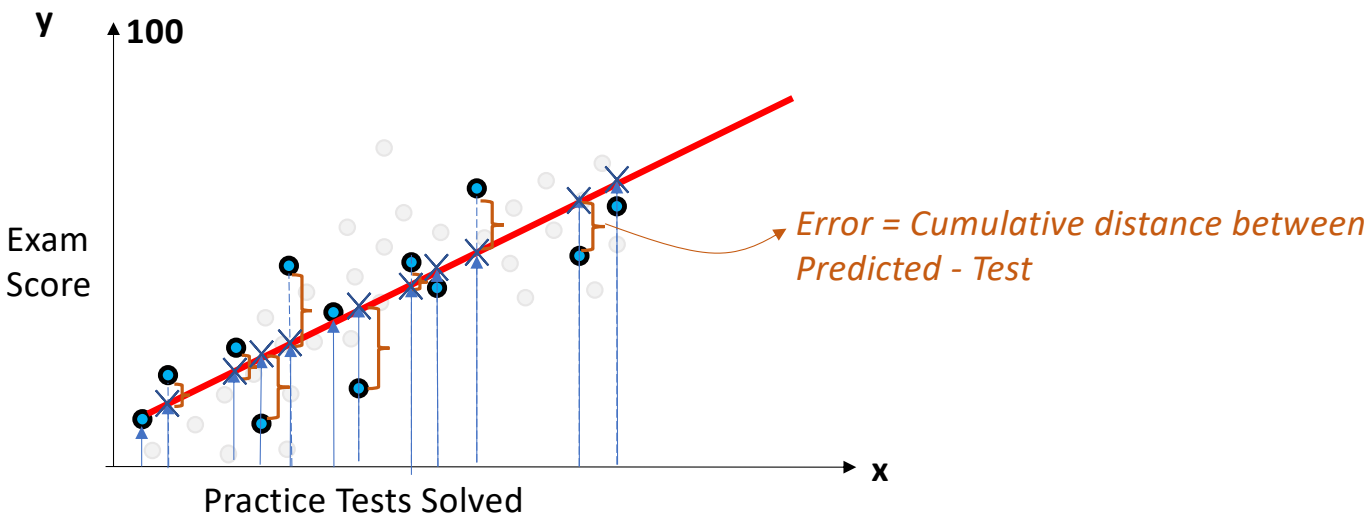
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Test:

Testing **data set is used to** Test model *error*

Trained Model: $Y = f(\mathbf{x}) = 8\mathbf{x} + 10$



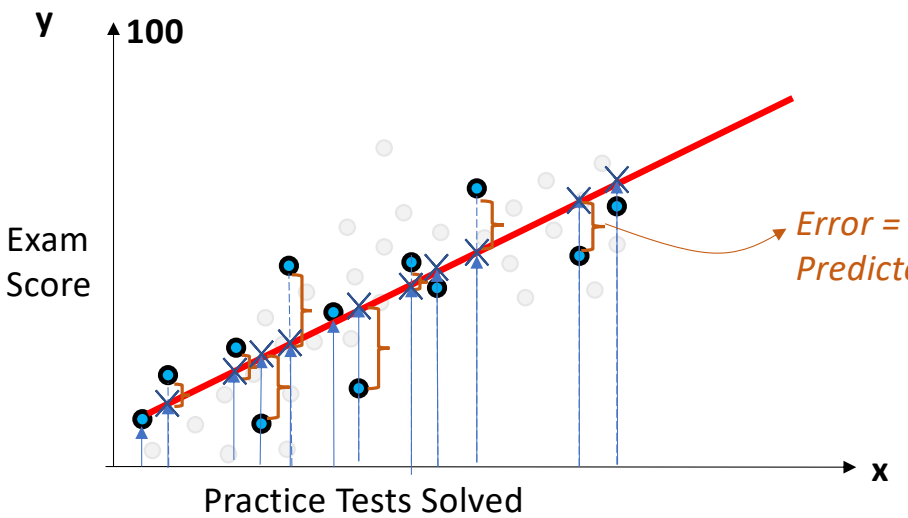
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Test:

Testing **data set is used to** Test model *error*

Trained Model: Y = f(**x**) = 8**x** + 10



Error = Cumulative distance between Predicted - Test

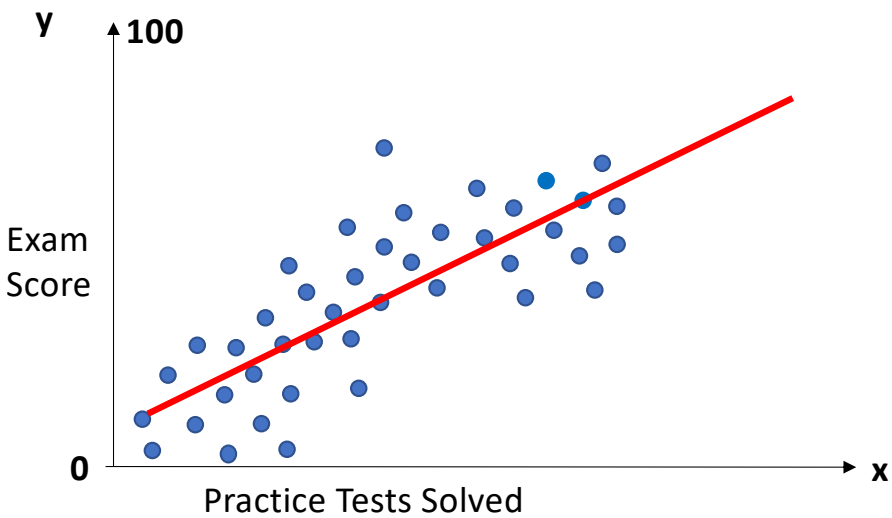
Error = 10%

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Decision: *10% error* is **acceptable** **so we** **deployed our model into production**!

Trained Model: $Y = f(\mathbf{x}) = 8\mathbf{x} + 10$

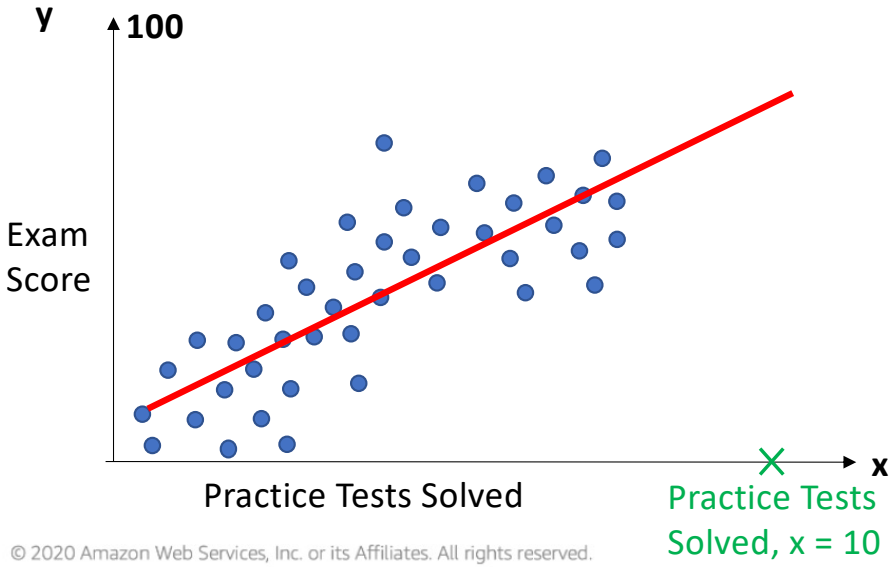


Example: Data Scientist in an Education Institute

aws training and certification

"I solved 10 practice tests, what will be my predicted exam score?"

Trained Model: $Y = f(\mathbf{x}) = 8\mathbf{x} + 10$



© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

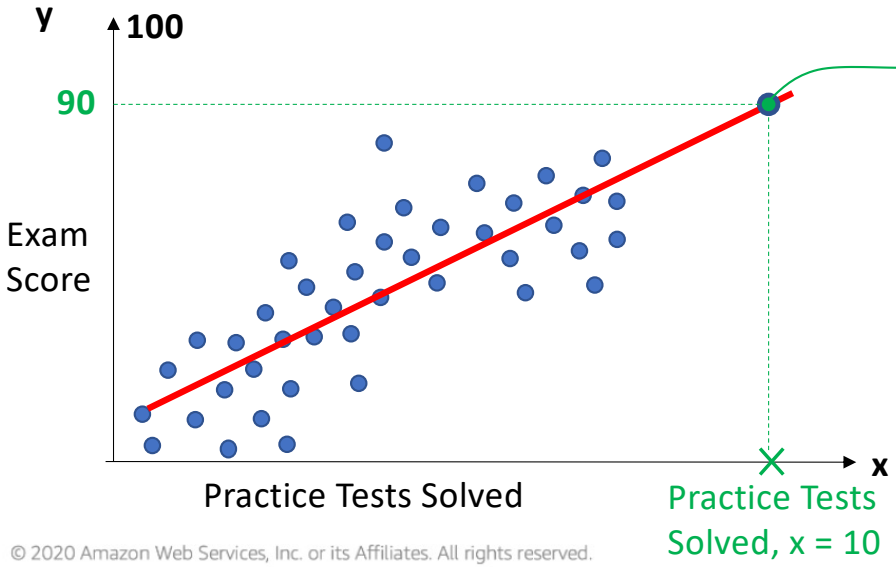
aws training and certification

"I solved 10 practice tests, what will be my predicted exam score?"

Trained Model: $Y = f(\mathbf{x}) = 8\mathbf{x} + 10$

For x = 10,

$Y = f(\mathbf{10}) = 8 * \mathbf{10} + 10 = \underline{90}$

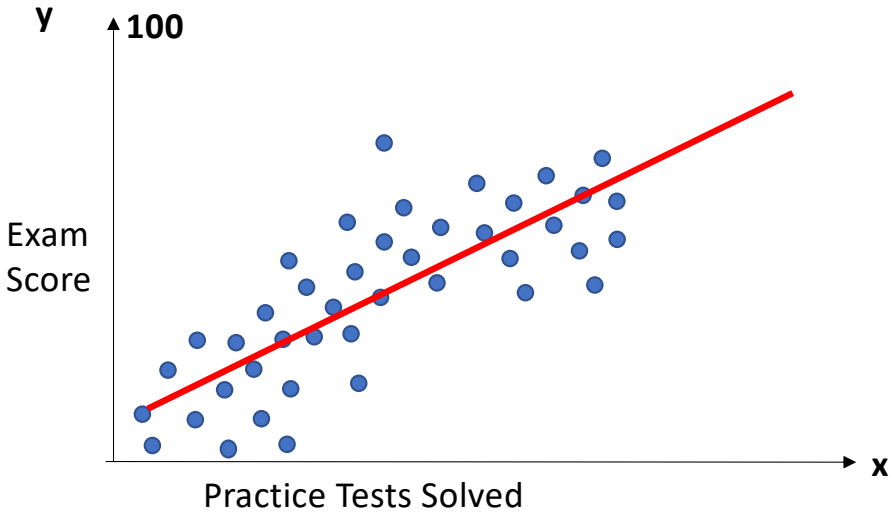


© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Two weeks later, "Hey, I solved 20 practice tests, what will be my predicted exam score?"

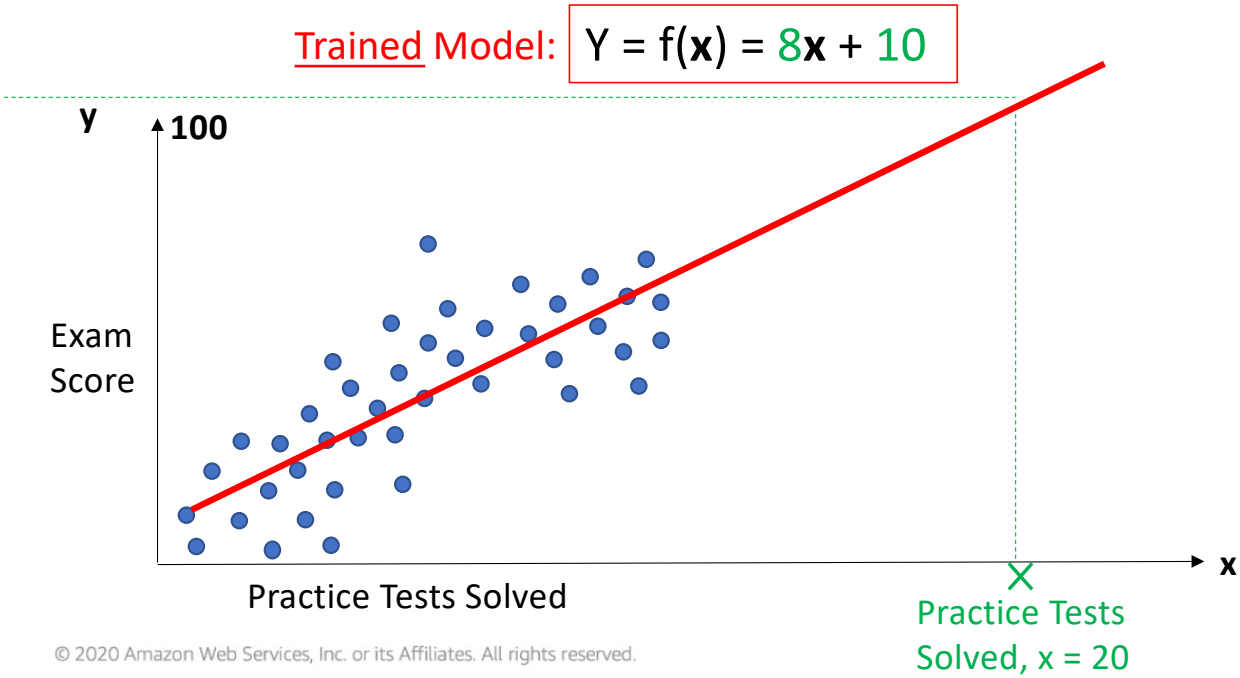
Trained Model: $Y = f(\mathbf{x}) = 8\mathbf{x} + 10$



© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Two weeks later, a different student; "Hey, I solved 20 practice tests, what will be my predicted exam score?"

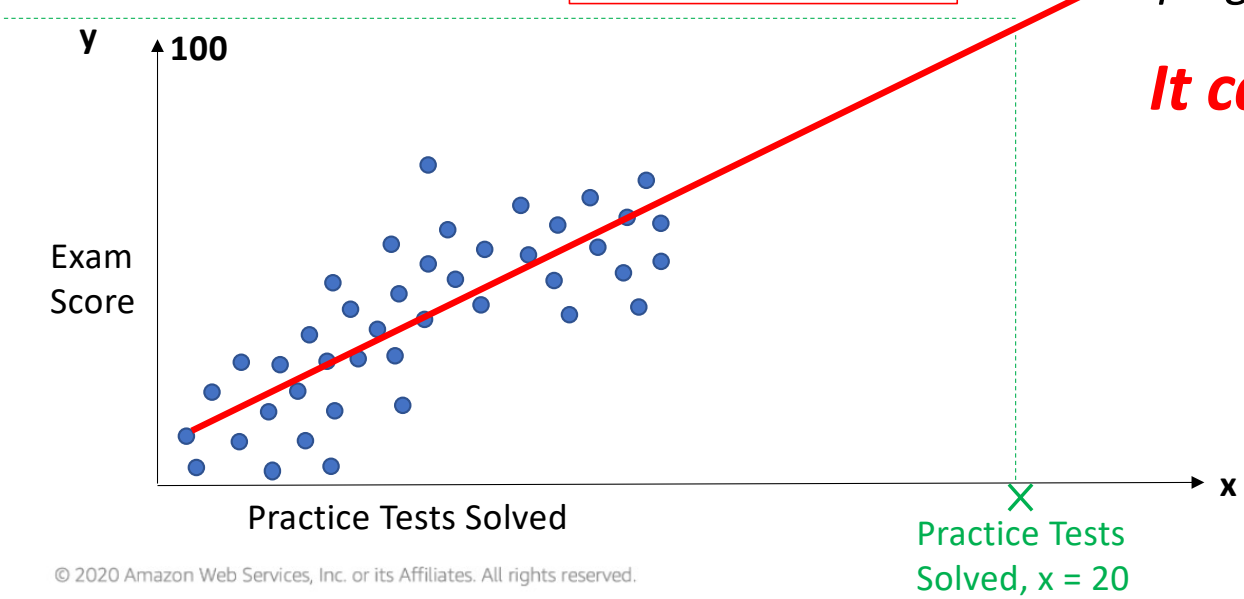


© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Two weeks later, a different student; "Hey, I solved 20 practice tests, what will be my predicted exam score?"

Trained Model: $Y = f(\mathbf{x}) = 8\mathbf{x} + 10$



For x = 20,

*Y = 8 * 20 + 10 = 170!*

It cant be greater than 100!

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Action: Our model looks too naïve***.*** *Let's try to increase our model complexity!*

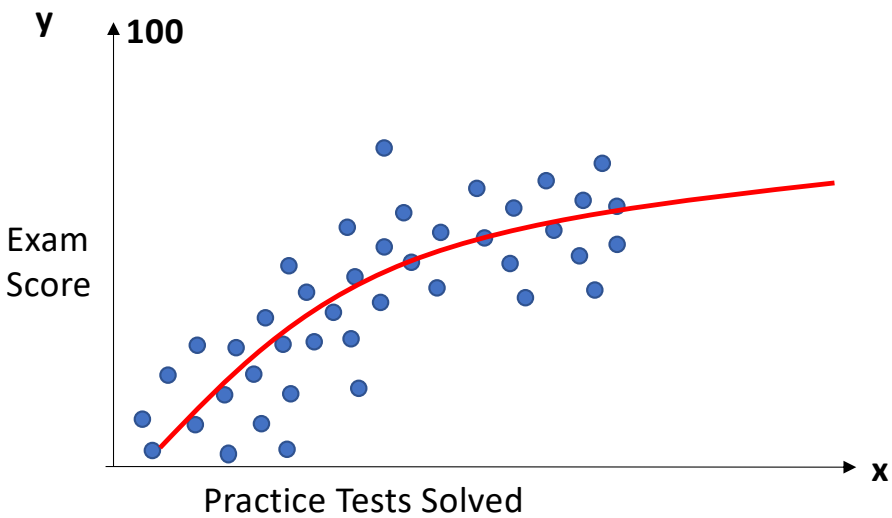
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Machine Learning Model: *Second Order Regression*

New Model: $Y = f(\mathbf{x}) = a_2\mathbf{x}^2 + a_1\mathbf{x} + a_0$

Our model has changed!

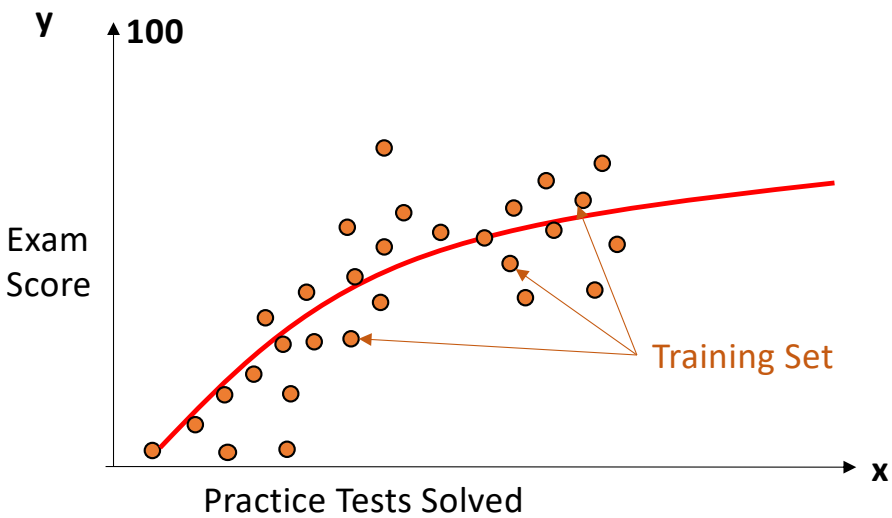


© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Machine Learning Model: *Second Order Regression*

New Model: $Y = f(\mathbf{x}) = a_2\mathbf{x}^2 + a_1\mathbf{x} + a_0$



Re-Training:

minimize(distance) using Training set

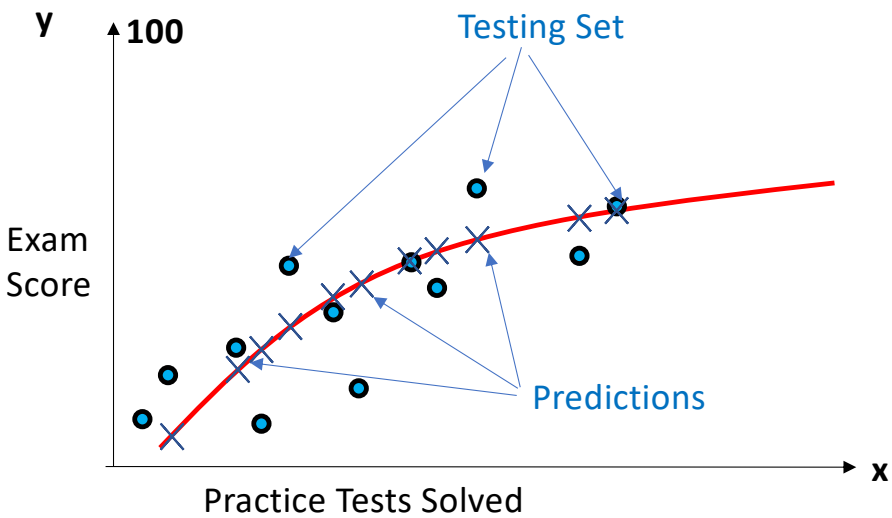
Find $a_0 = \ldots$, $a_1 = \ldots$, $a_2 = \ldots$

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Machine Learning Model: *Second Order Regression*

New Model: $Y = f(\mathbf{x}) = a_2\mathbf{x}^2 + a_1\mathbf{x} + a_0$



Re-Training:

minimize(distance) using Training set

Find $a_0 = \ldots$, $a_1 = \ldots$, $a_2 = \ldots$

Test/Validate:

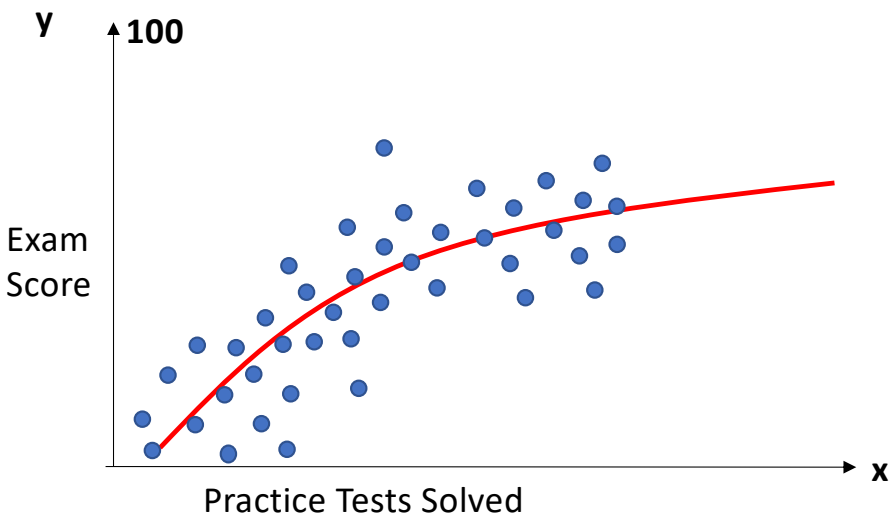
Find *Error* = …% for Testing Set

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Machine Learning Model: *Second Order Regression*

General Model: $Y = f(\mathbf{x}) = a_2\mathbf{x}^2 + a_1\mathbf{x} + a_0$



Re-Training:

minimize(distance) using Training set

Find $a_0 = \ldots$, $a_1 = \ldots$, $a_2 = \ldots$

Test/Validate:

Find *Error* = …% for Testing Set

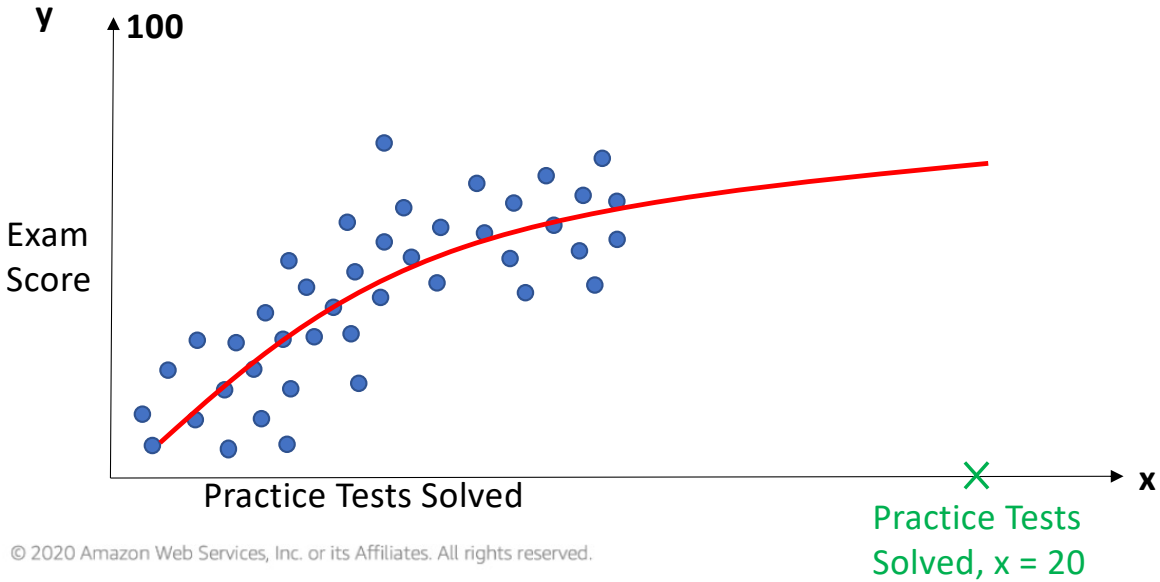
Deploy!!

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Two weeks later, "Hey, I solved 20 practice tests, what will be my predicted exam score?"

Trained Model: $Y = f(\mathbf{x}) = \ldots\mathbf{x}^2 + \ldots\mathbf{x} + \ldots$

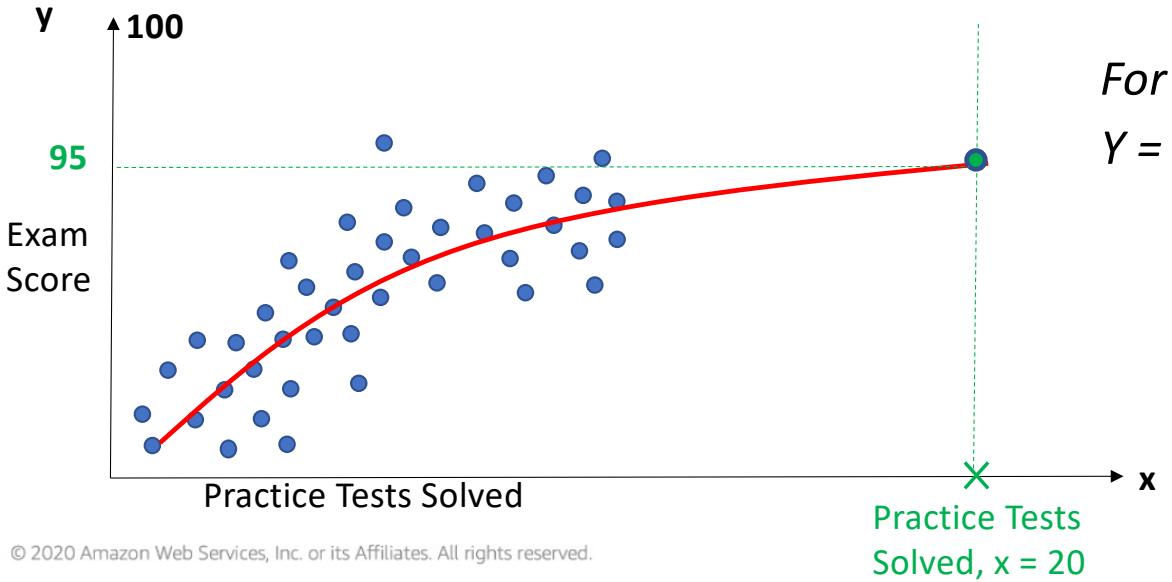


© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Two weeks later, "Hey, I solved 20 practice tests, what will be my predicted exam score?"

New Model: $Y = f(\mathbf{x}) = \ldots\mathbf{x}^2 + \ldots\mathbf{x} + \ldots$



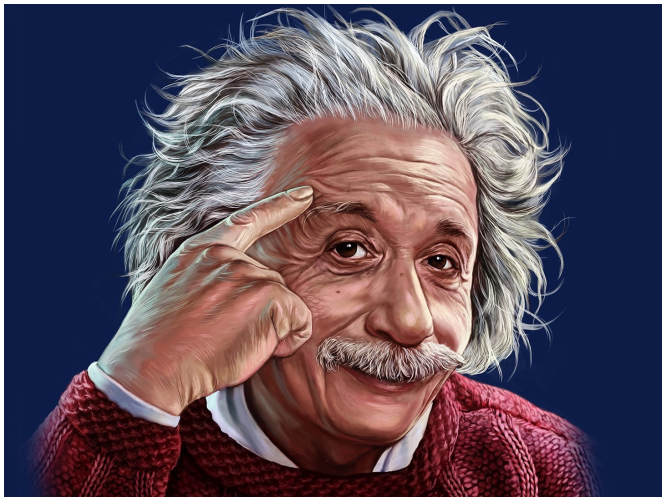
For x = 20,

$Y = \ldots * 20^2 + \ldots * 20 + \ldots =$ ***95!***

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Management: *It seems like increasing model complexity works fine. Let's make it more complex, so increase the order of your curve.*

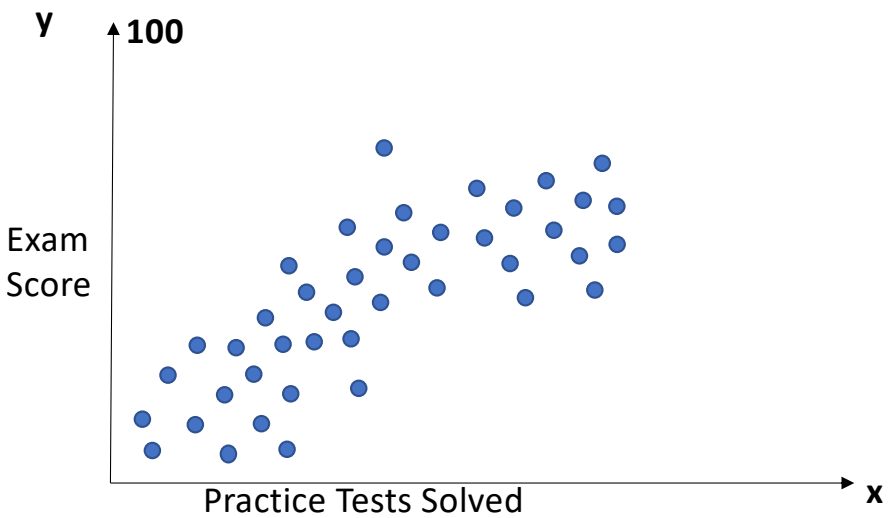


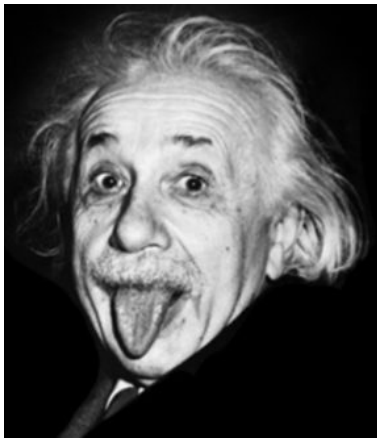
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Machine Learning Model: *10th Order Regression*

Model: $Y = f(\mathbf{x}) = a_{10}\mathbf{x}^{10} + a_9\mathbf{x}^9 + \ldots + a_1\mathbf{x} + a_0$

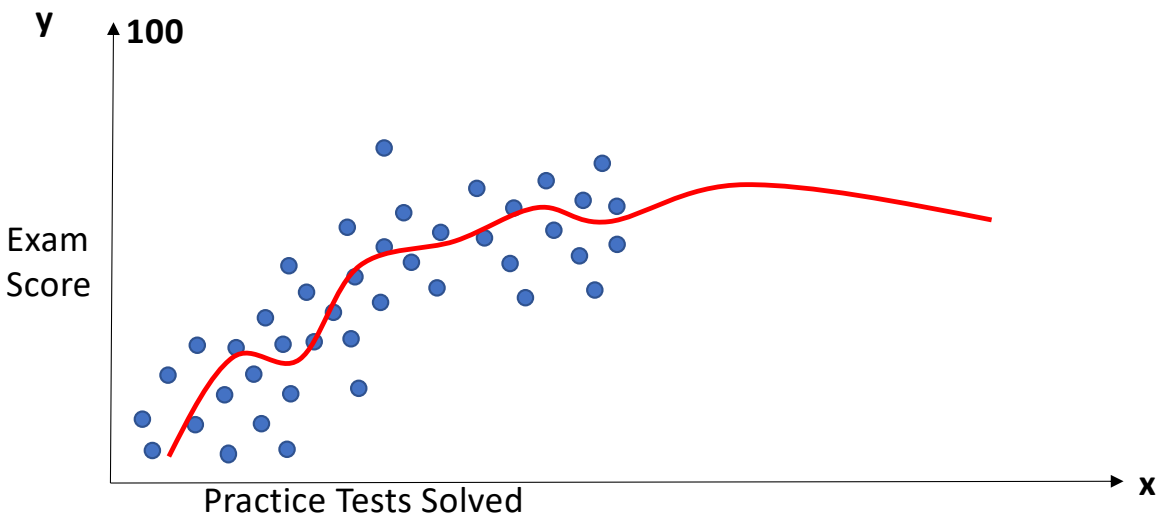




Example: Data Scientist in an Education Institute

Machine Learning Model: *10th Order Regression*

Model: $Y = f(\mathbf{x}) = a_{10}\mathbf{x}^{10} + a_9\mathbf{x}^9 + \ldots + a_1\mathbf{x} + a_0$



© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

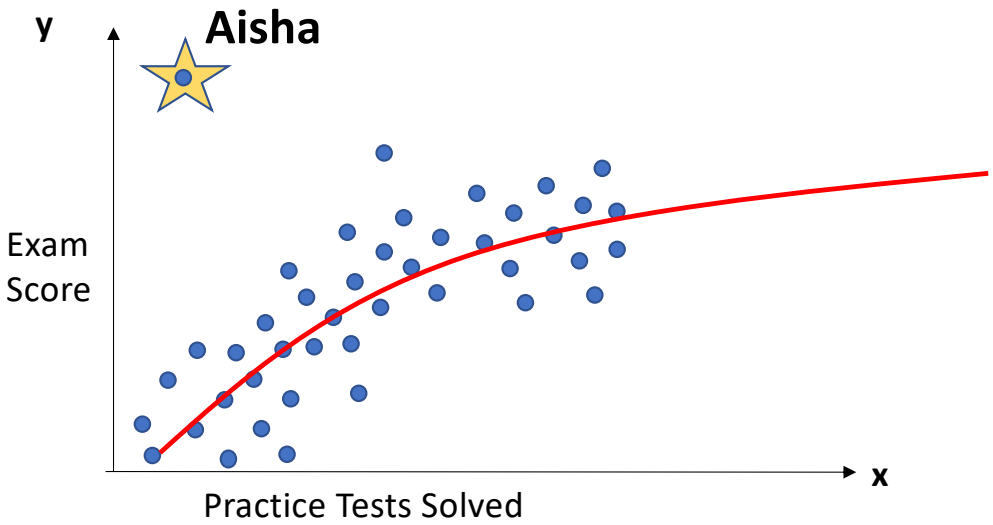
- It is called "overfitting" (No worries, you will learn this)
  → **There is no short-cut**

- **Ok we selected second degree regression, our students were happy!**

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Two weeks later, Aisha got 95 but only solved 1 Practice test! → Model catastrophically failed

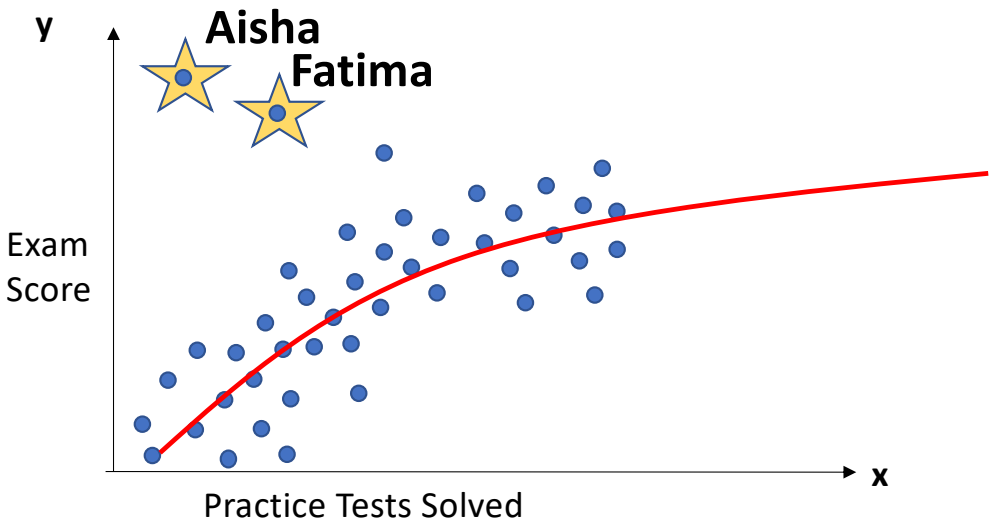


Answer: *Perhaps it is an "anomaly"! The data is "outlier", we can delete?*

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Five weeks later, Fatima got 90 but only solved 2 Practice test! → Model catastrophically failed



© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Data Deep Dive:

- Aisha 10 years of experience in that field.
- Fatima had a Master's Degree in that field.

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

Answer:

- Our model's feature scope is inadequate.
- We need to include more features!
- How can I quantify degree's?

→ **"Feature Engineering"**

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Example: Data Scientist in an Education Institute

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Three categories of machine learning algorithms

aws training and certification

Supervised

Unsupervised

Reinforcement

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Three categories of machine learning algorithms

aws training and certification

Supervised

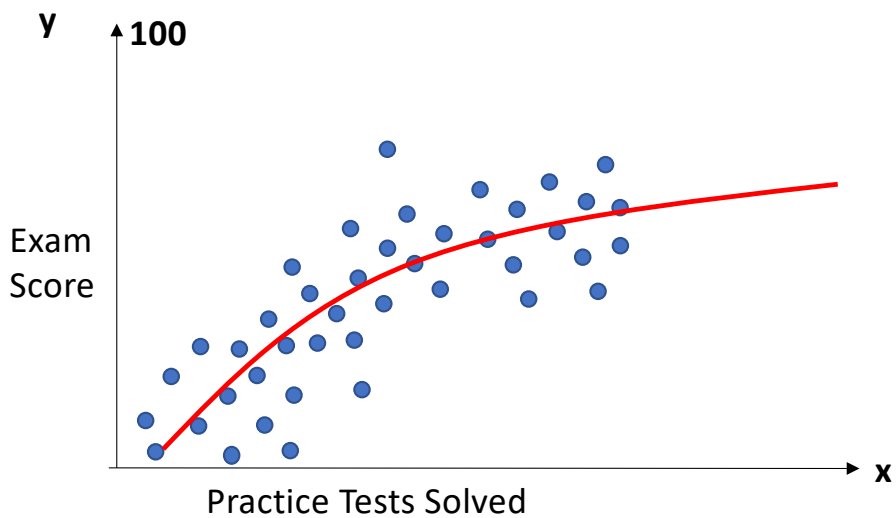
Unsupervised

Reinforcement

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

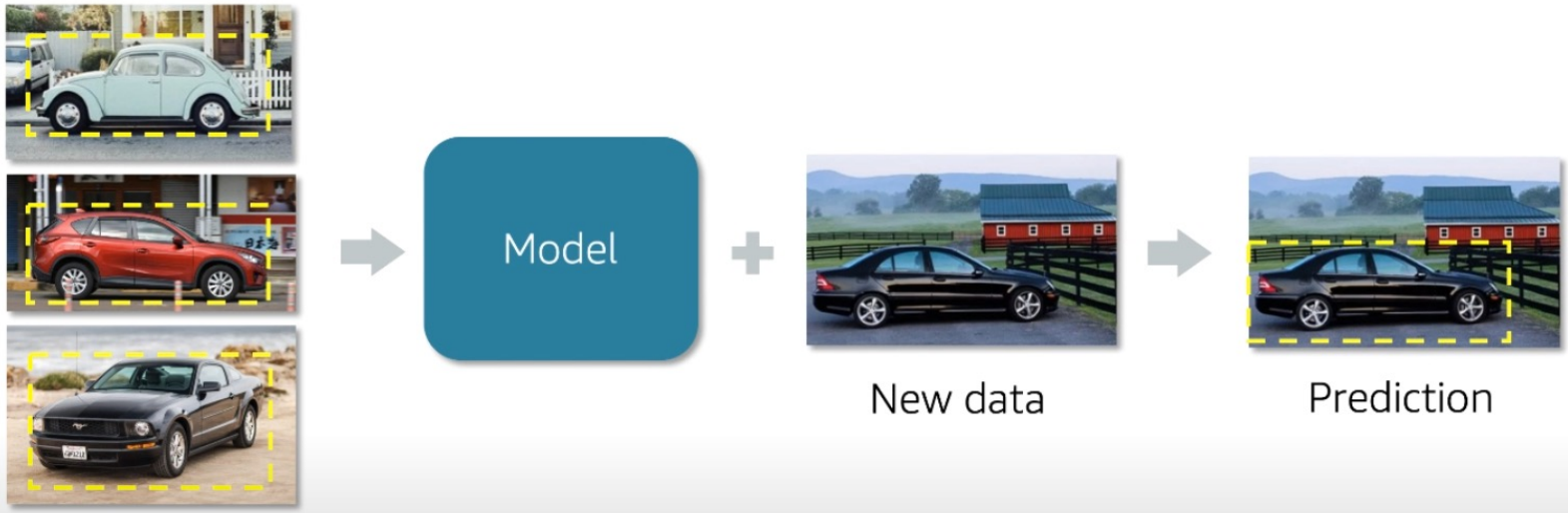
Supervised

$$Y = f(\mathbf{x}) = a_2\mathbf{x}^2 + a_1\mathbf{x} + a_0$$



Supervised

Learn by identifying patterns in data that's **already been labeled**.



Pictures of cars, labeled by people

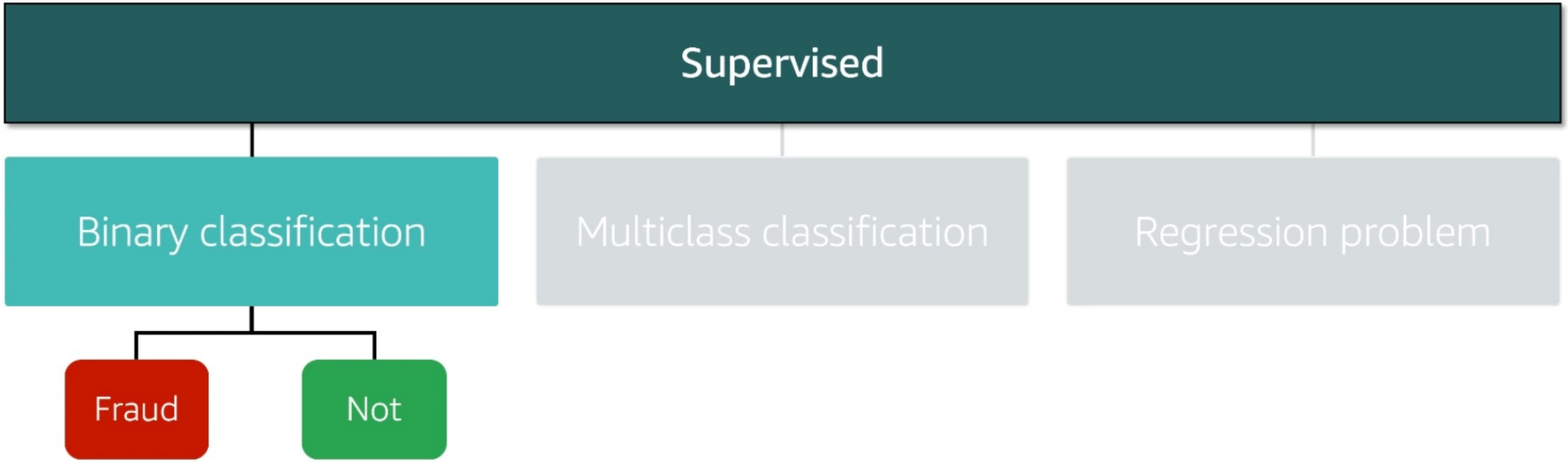
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

The different types of supervised ML problems

aws training and certification

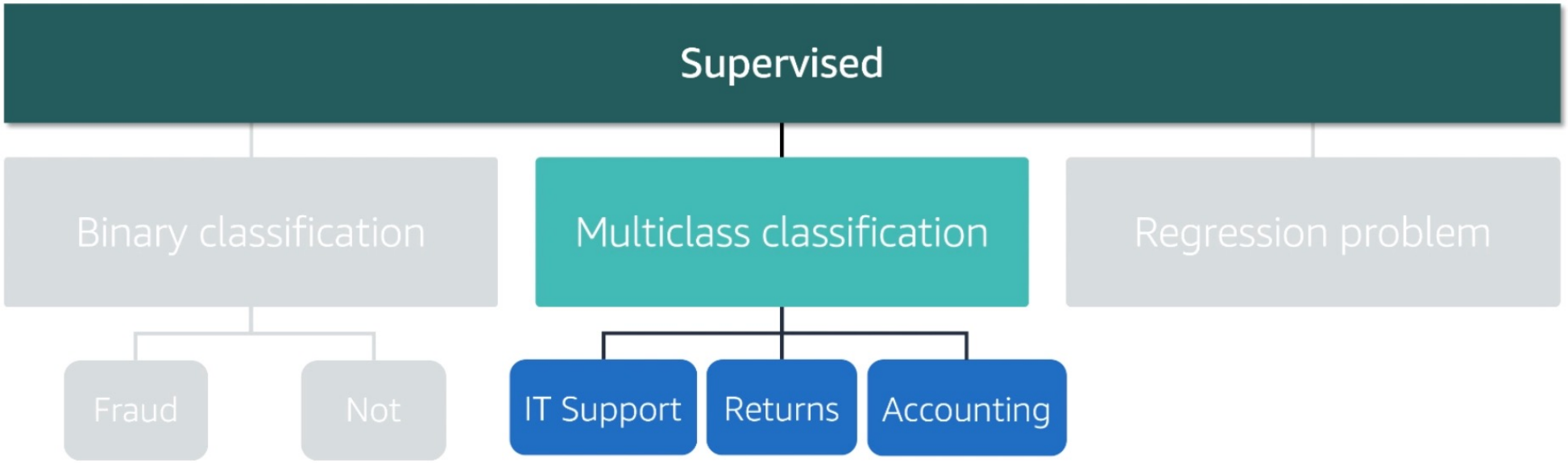
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Binary classification classifies an observation into one of two categories



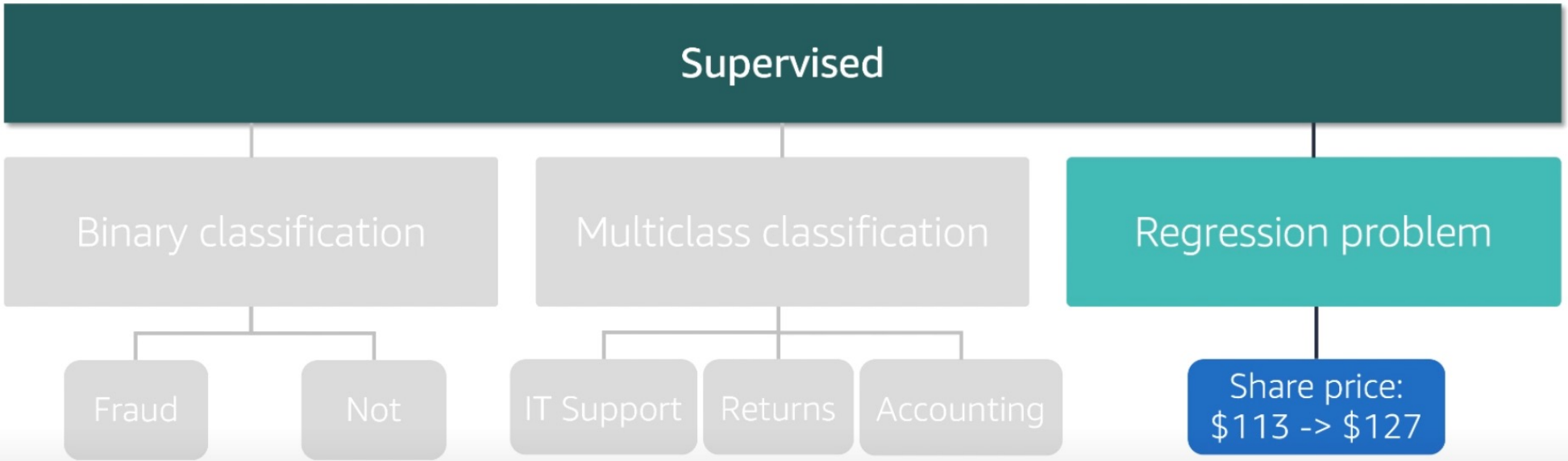
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Multiclass classification classifies into one of three or more categories



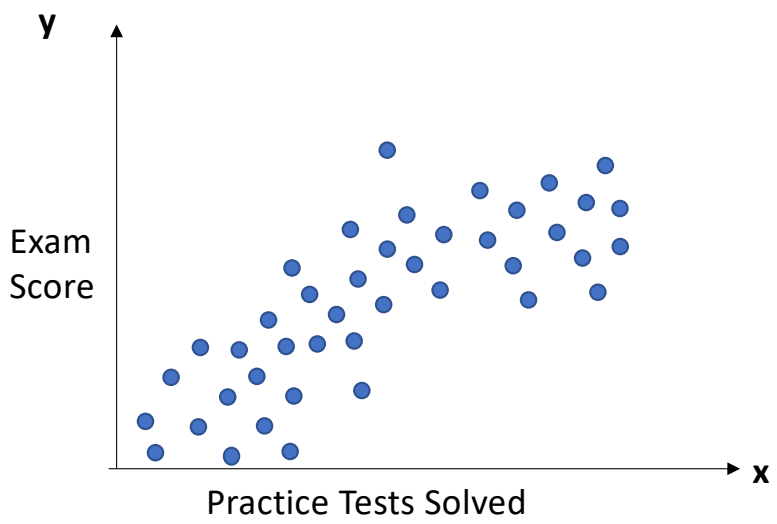
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Regression problems map an input to a continuous value, like an integer



© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

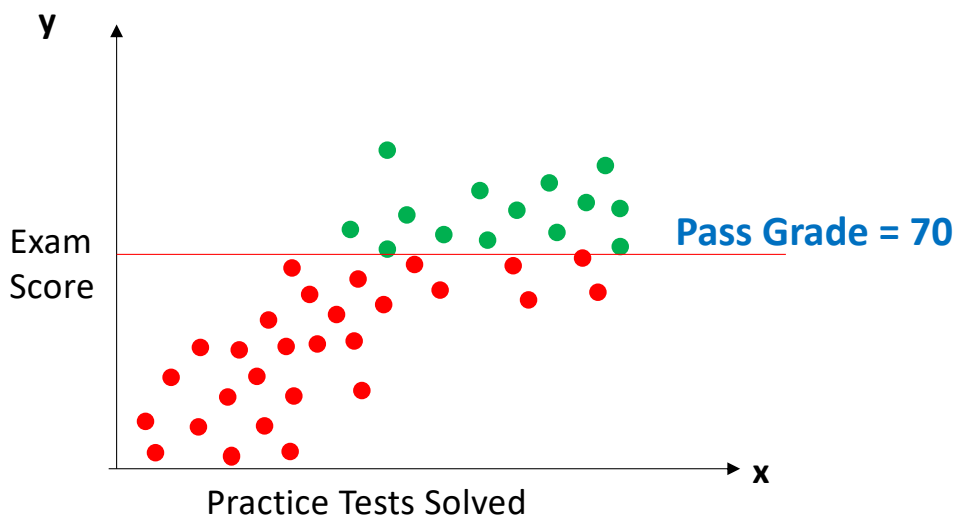
Converting our Grade Predictor to a Binary Classifier



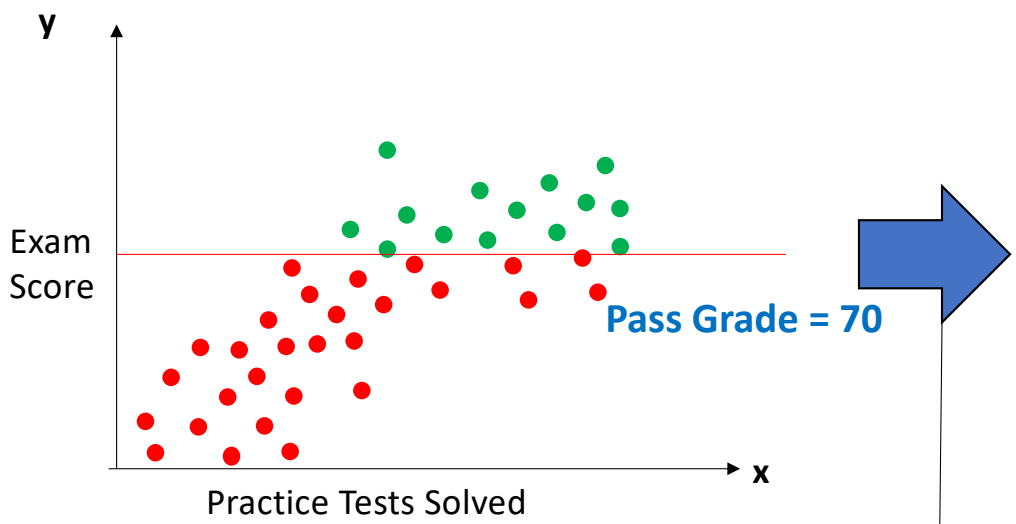
Problem: Based on number of practice tests solved, we *just need to predict whether you will pass the exam or not?*

>> Passing Grade = 70

Converting our Grade Predictor to a Binary Classifier

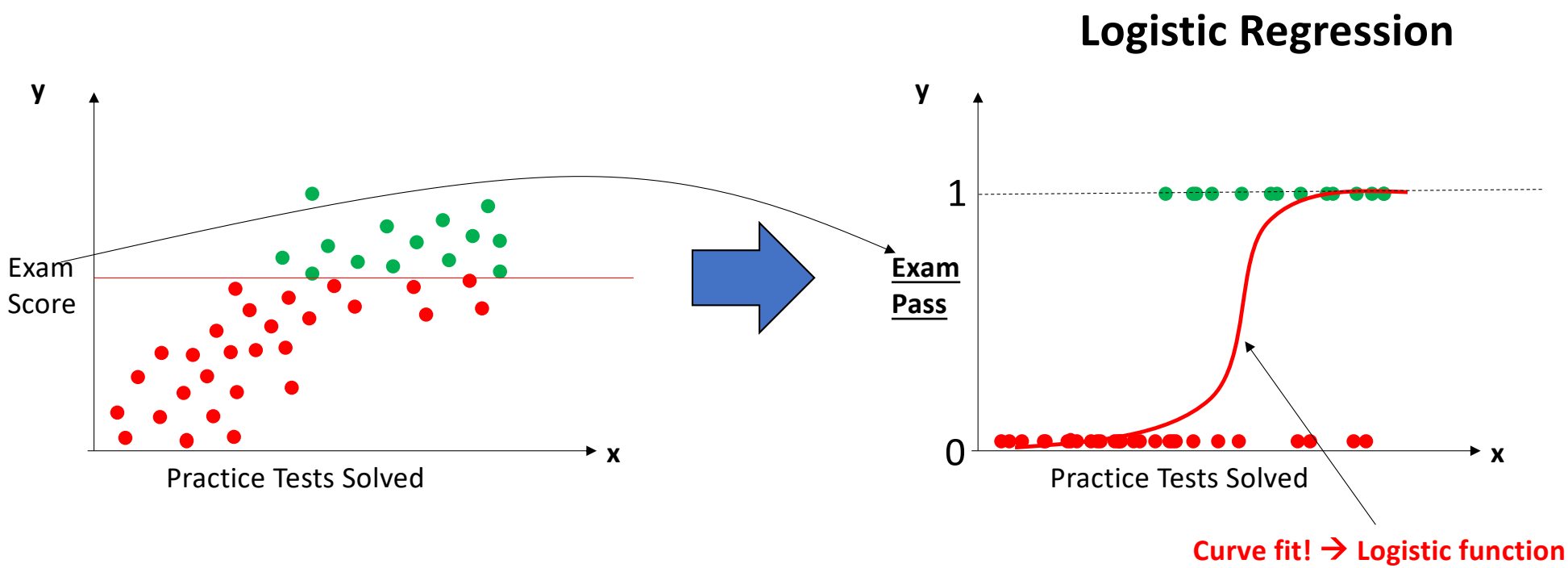


Converting our Grade Predictor to a Binary Classifier



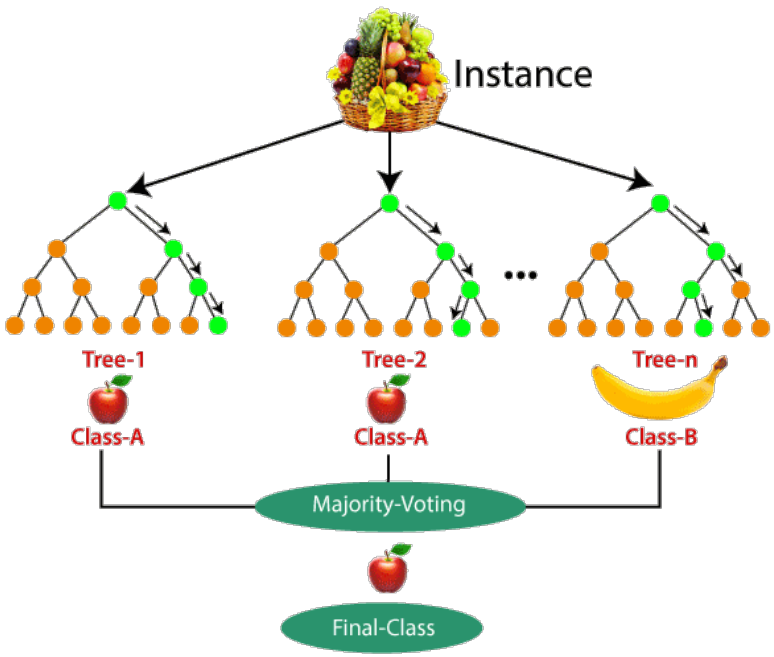
For all data;
    If Exam Score > 70;
        Exam Pass = 1
    Else
        Exam Pass = 0

Converting our Grade Predictor to a Binary Classifier



Supervised Learning Methods

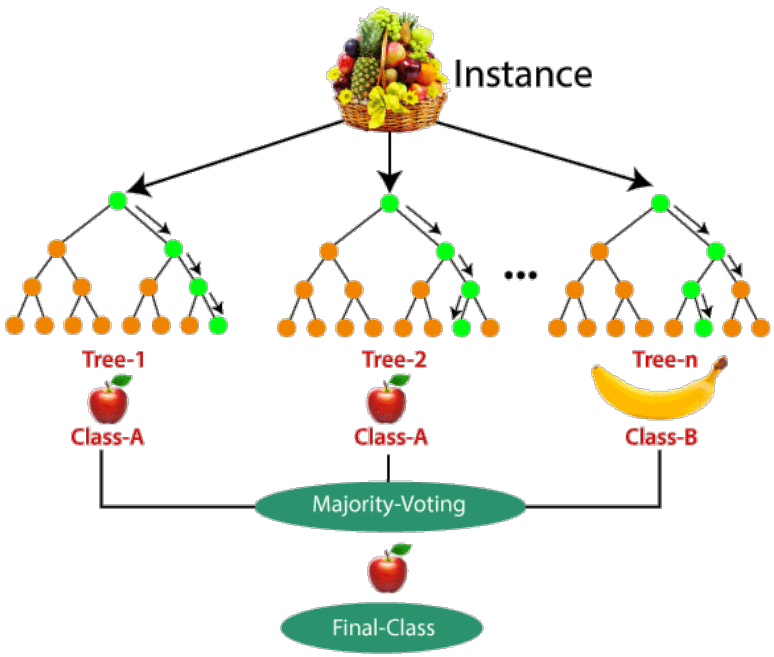
Random Forest



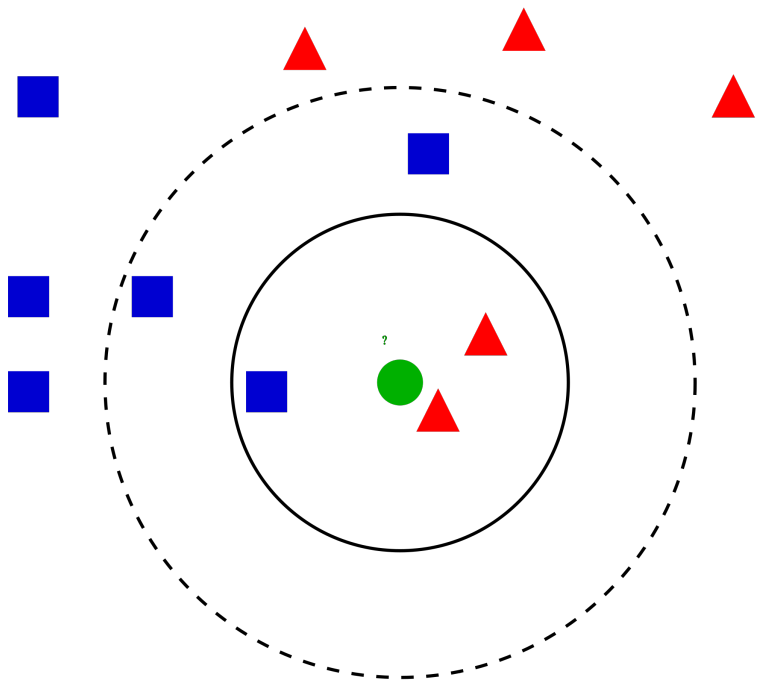
Supervised Learning Methods

aws training and certification

Random Forest

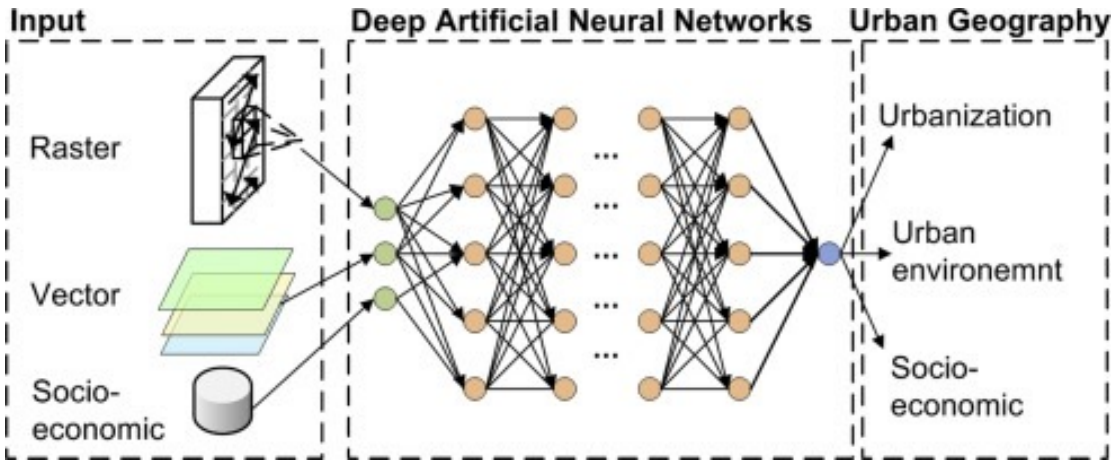


K-Nearest Neighbor



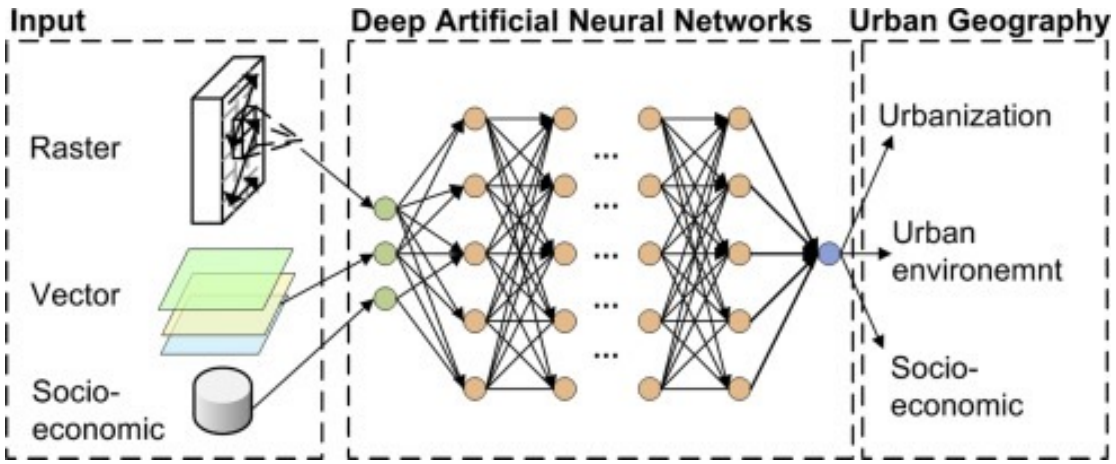
Supervised Learning Methods

Artificial Neural Networks

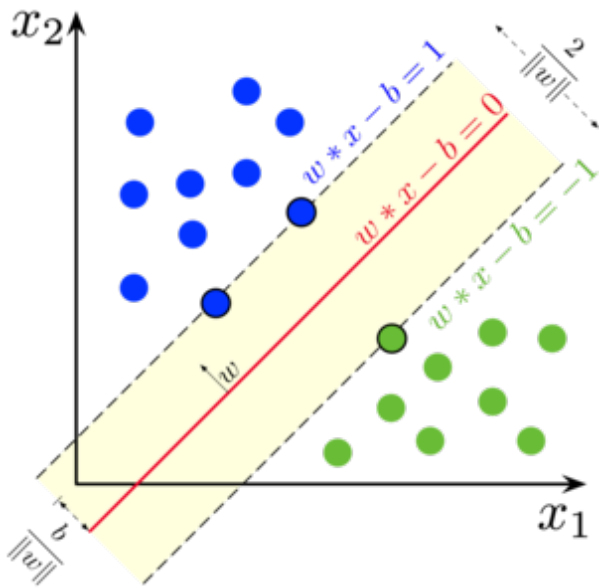


Supervised Learning Methods

Artificial Neural Networks



Support Vector Machine



Supervised Learning Methods

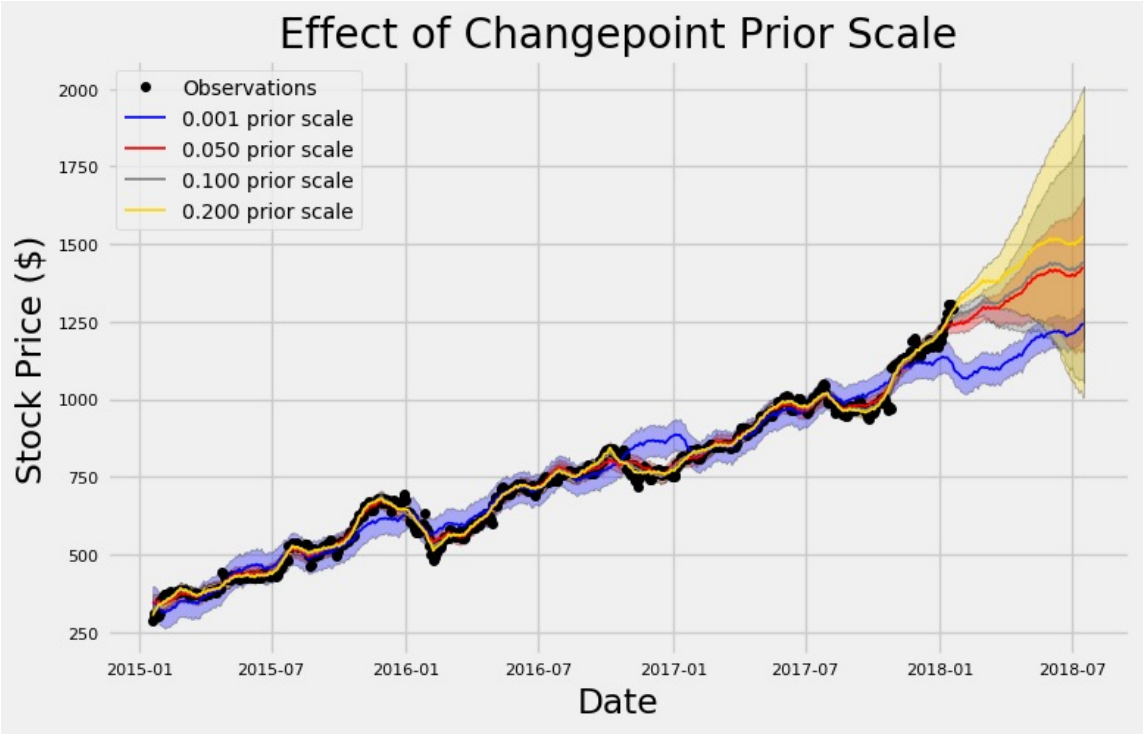
NLP (Natural Language Processing)

amazon alexa

Supervised Learning Methods

Time Series Analysis via

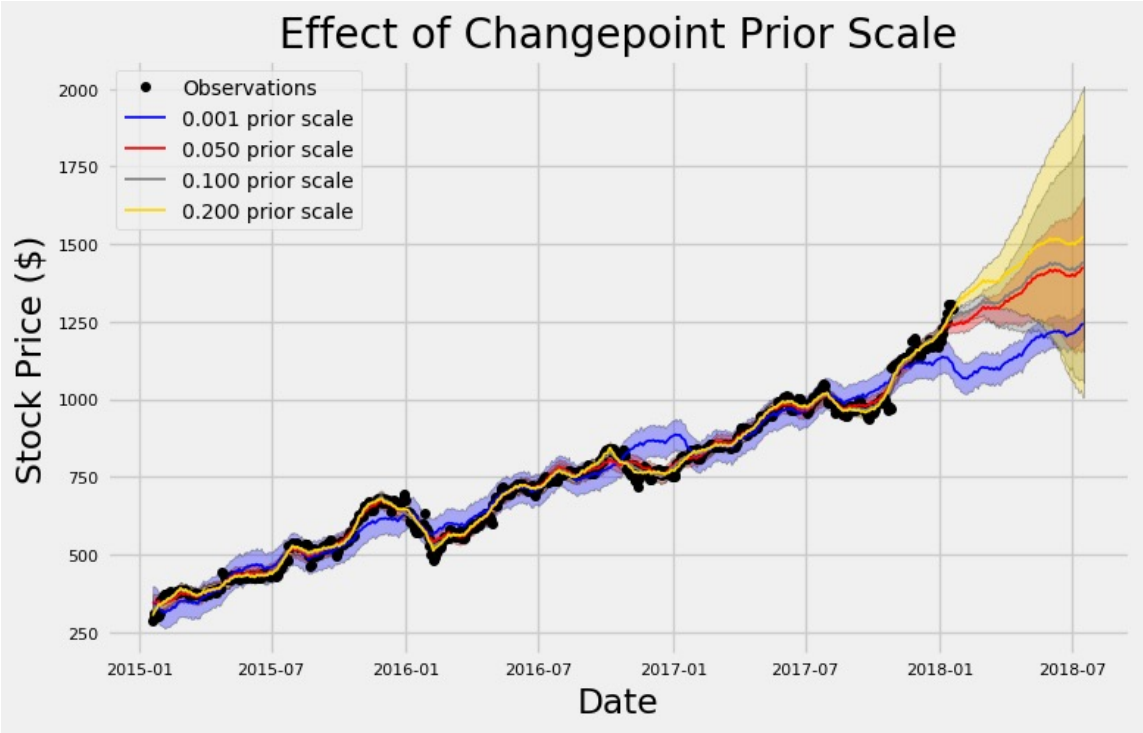
Recurrent Neural Networks (RNNs) or Statistics (ARIMA)

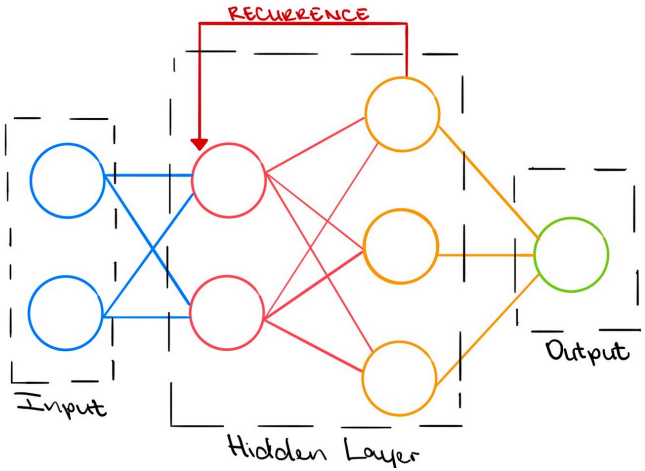


Supervised Learning Methods

Time Series Analysis - Forecasting

Recurrent Neural Networks (RNNs) or Statistics (ARIMA)

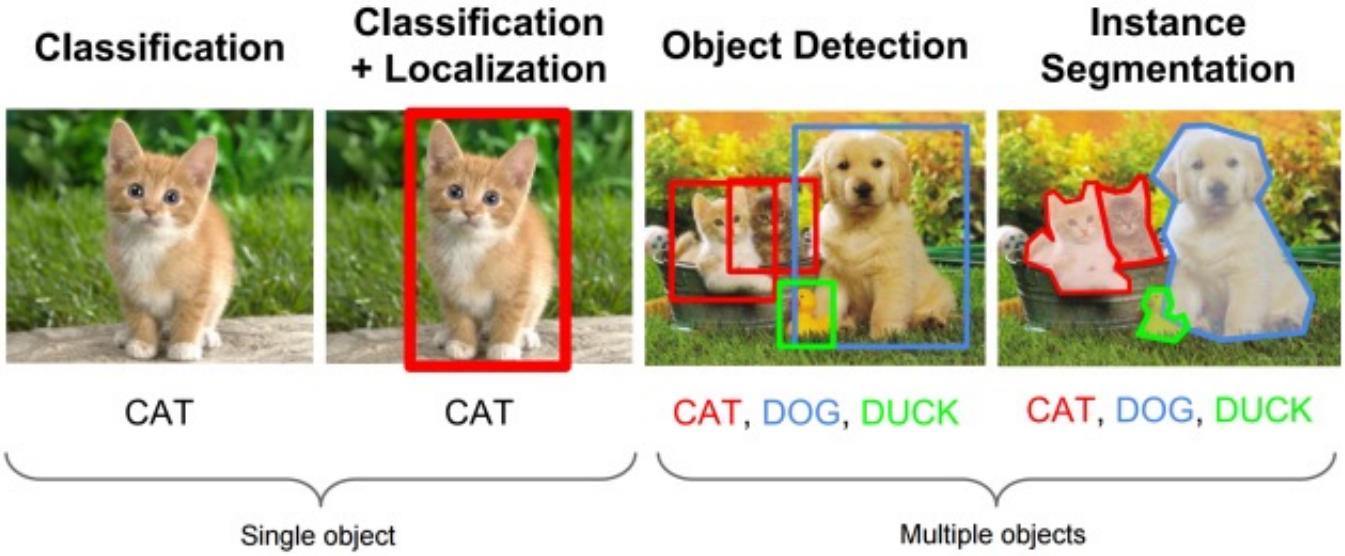




Supervised Learning Methods

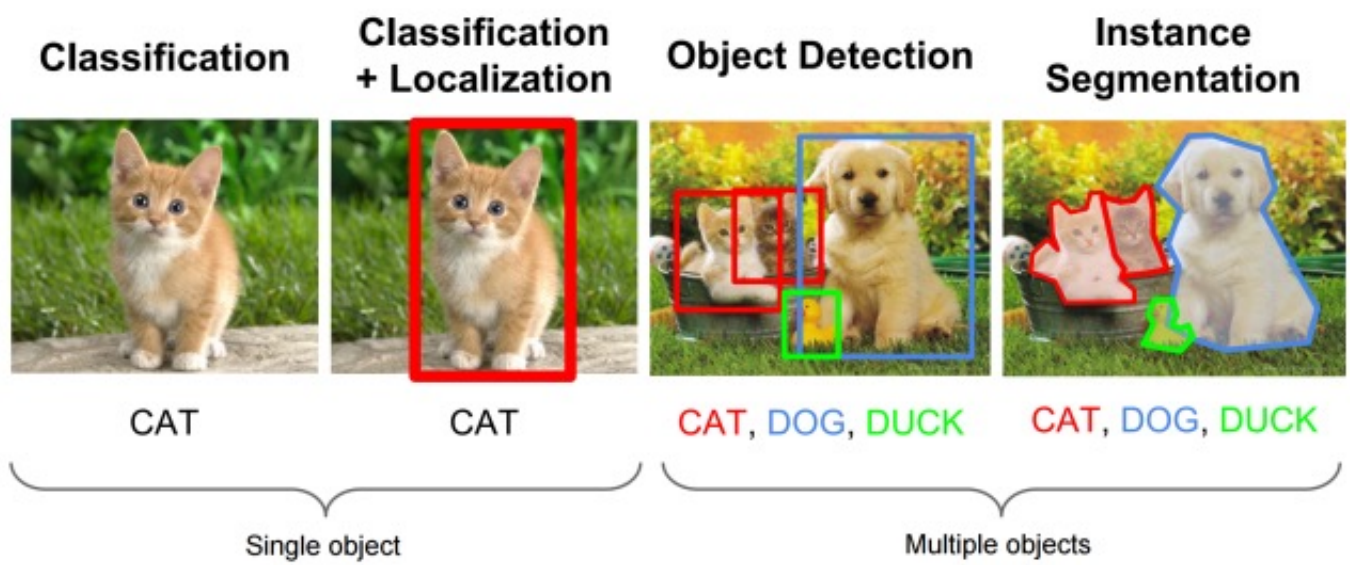
aws training and certification

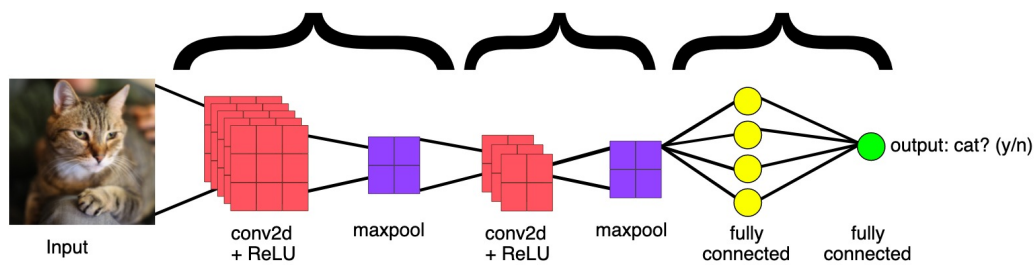
Image Recognition (*Convolutional Neural Networks - CNNs)*



Supervised Learning Methods

Image Recognition (*Convolutional Neural Networks - CNNs)*





Three categories of machine learning algorithms

aws training and certification

Supervised

Unsupervised

Reinforcement

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

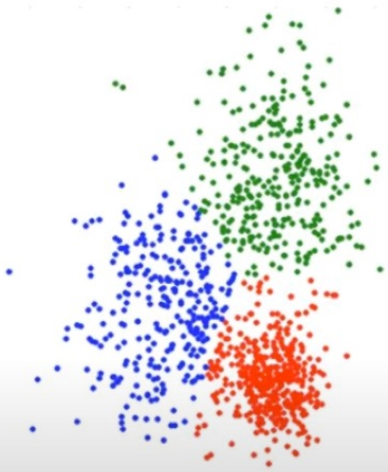
Unsupervised

aws training and certification

Clustering

The machine has to uncover and create the labels itself.

Original data

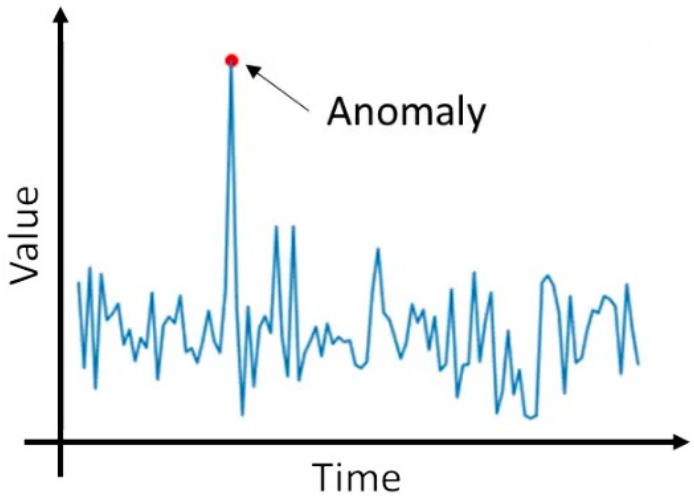


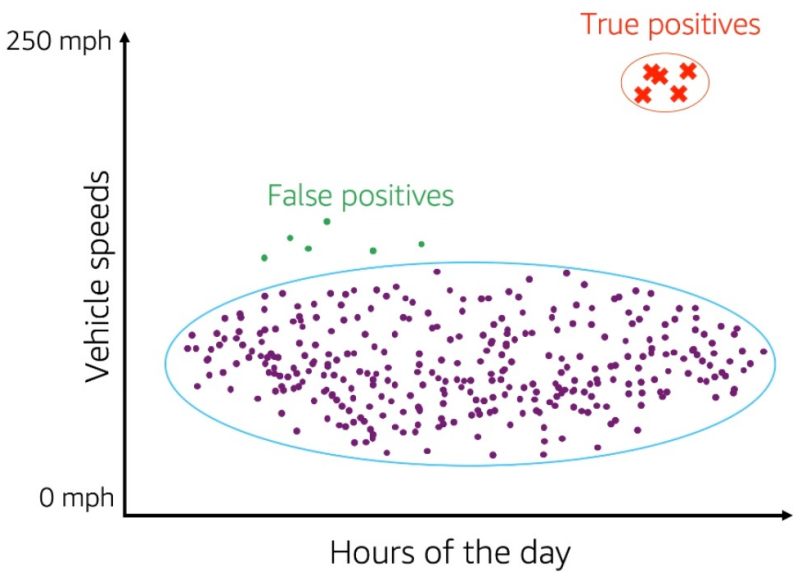
Clustered data

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Anomaly Detection

Using Random Cut Forest





Three categories of machine learning algorithms

aws training and certification

Supervised

Unsupervised

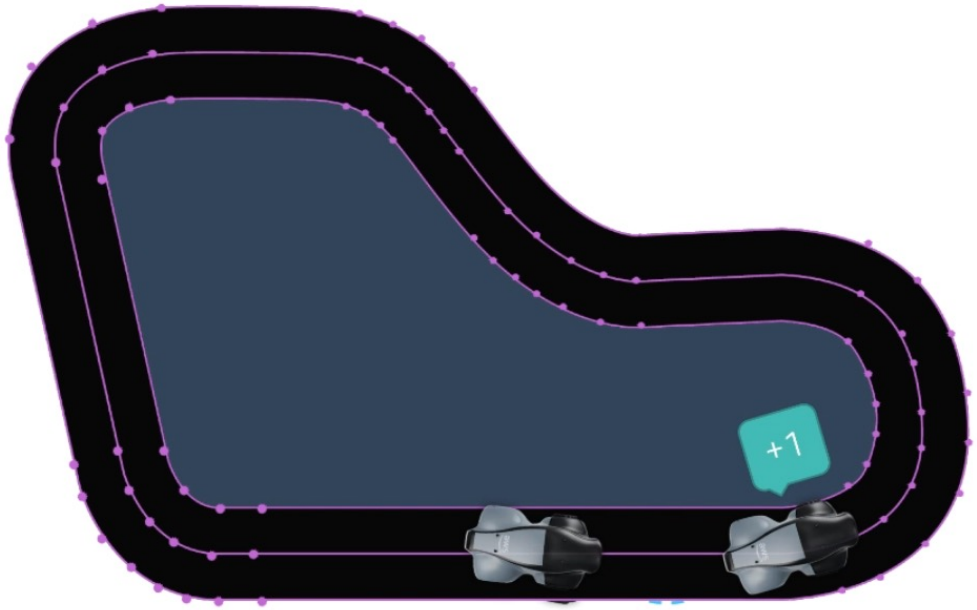
Reinforcement

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Reinforcement Learning

Learning through **trial and error**

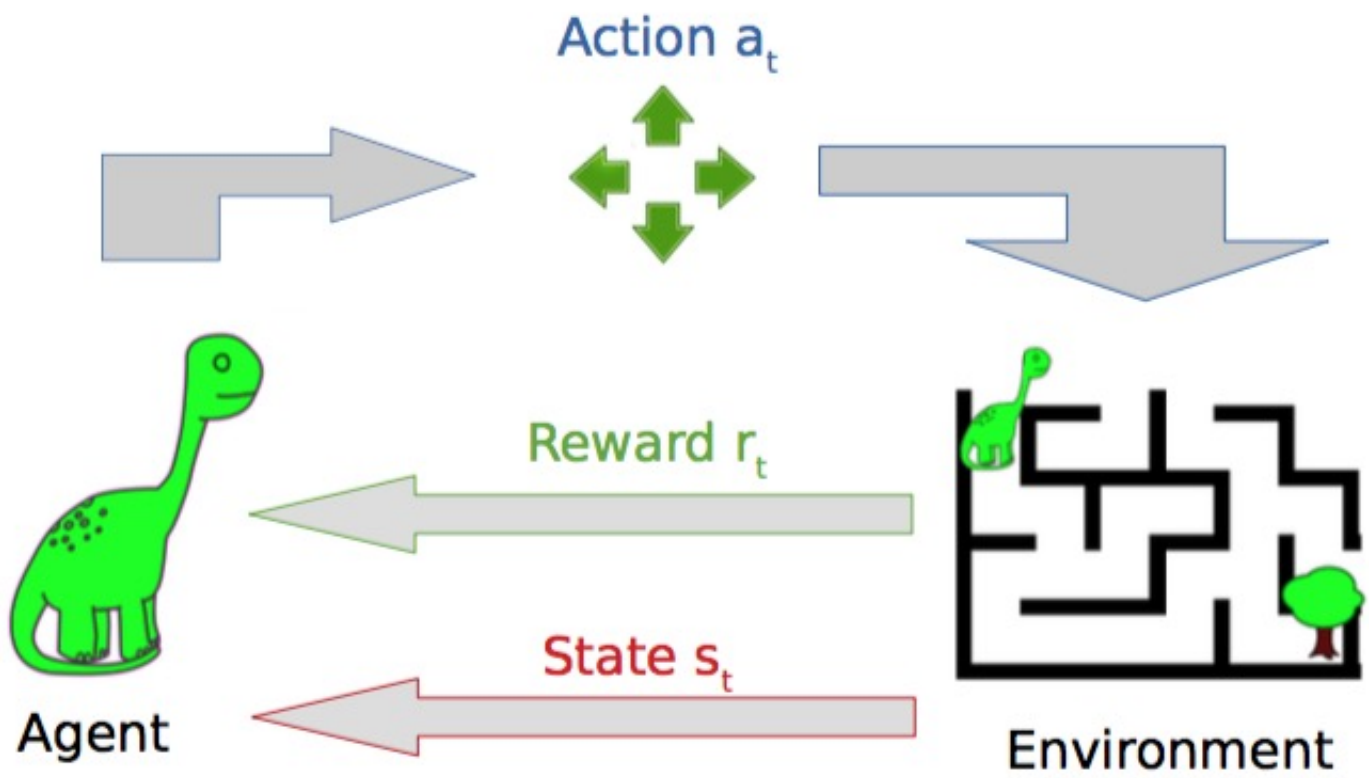
Best when the desired outcome is known but the exact path to achieving it isn't.



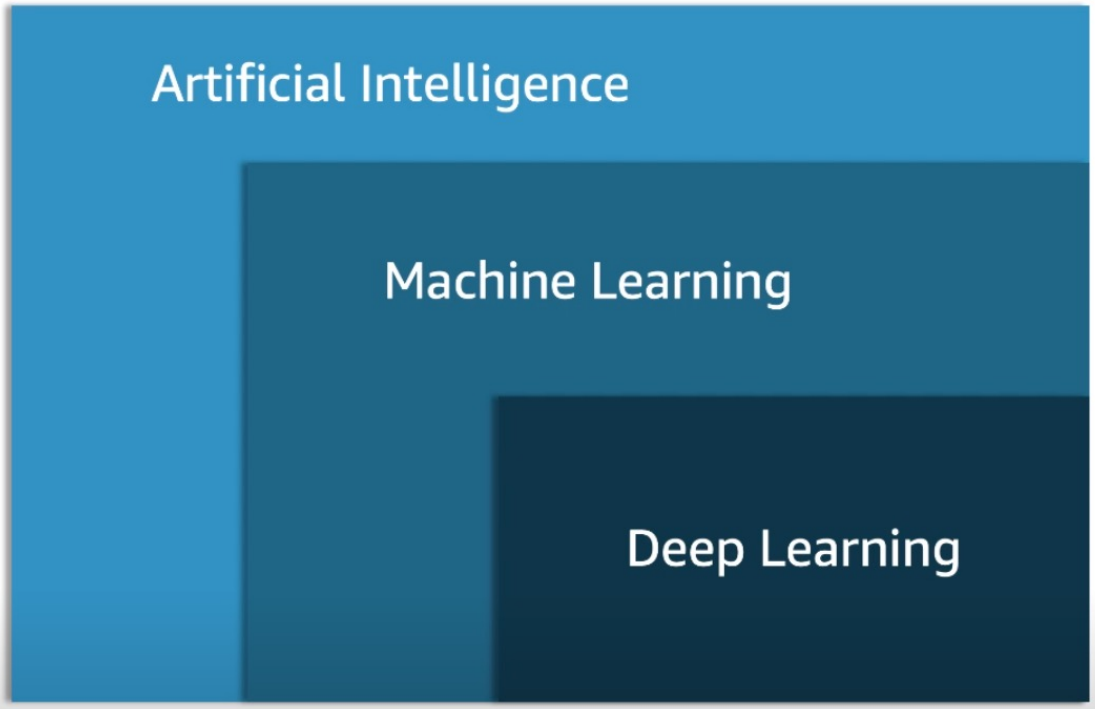
AWS DeepRacer

© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Reinforcement Learning

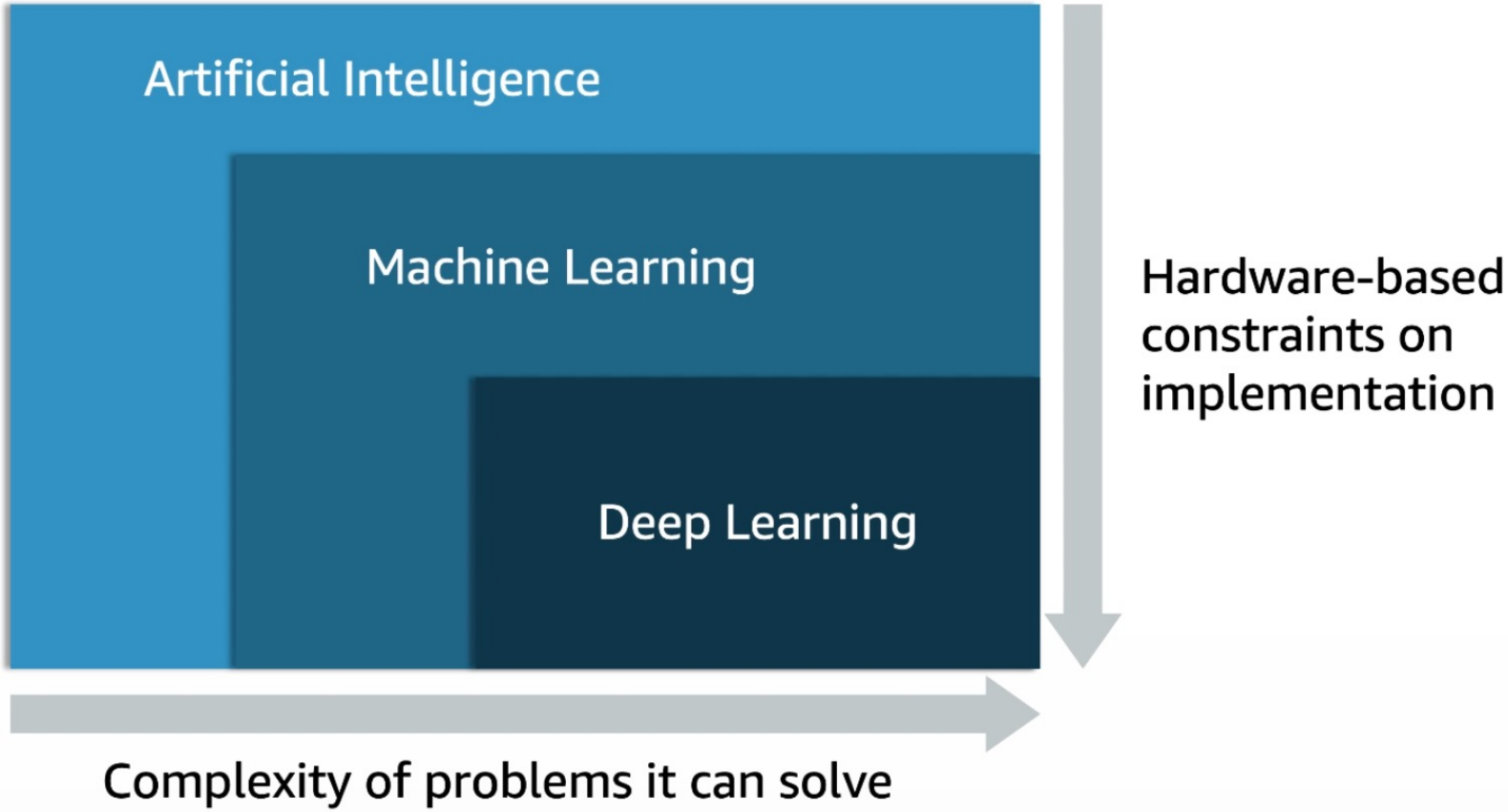


A sub-category of machine learning



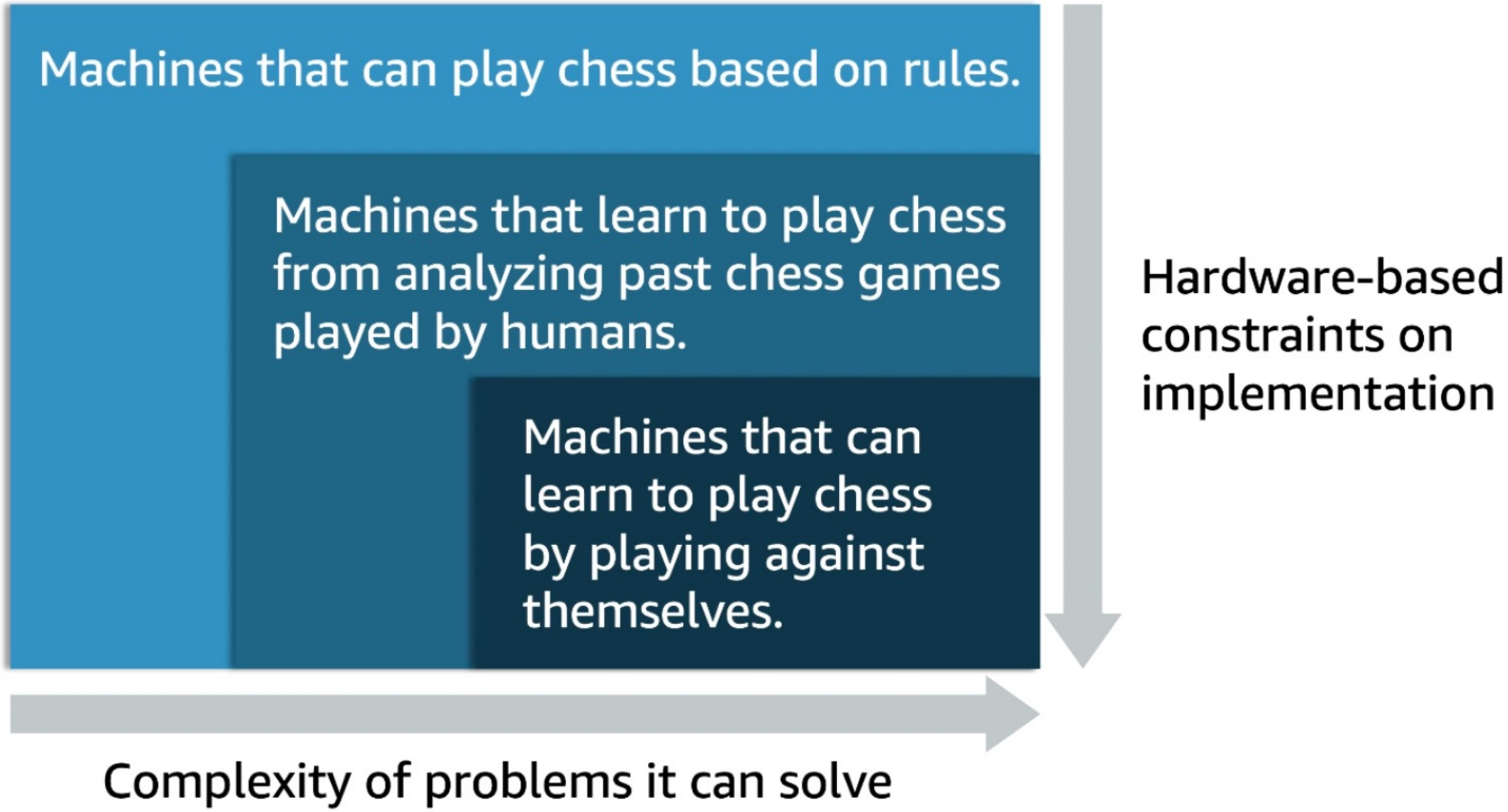
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

A sub-category of machine learning



© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

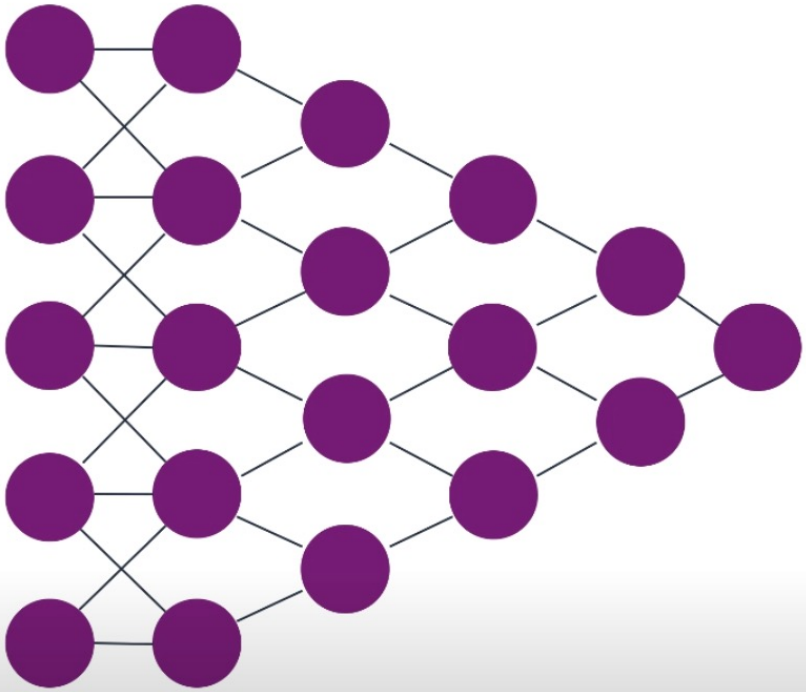
Examples



© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

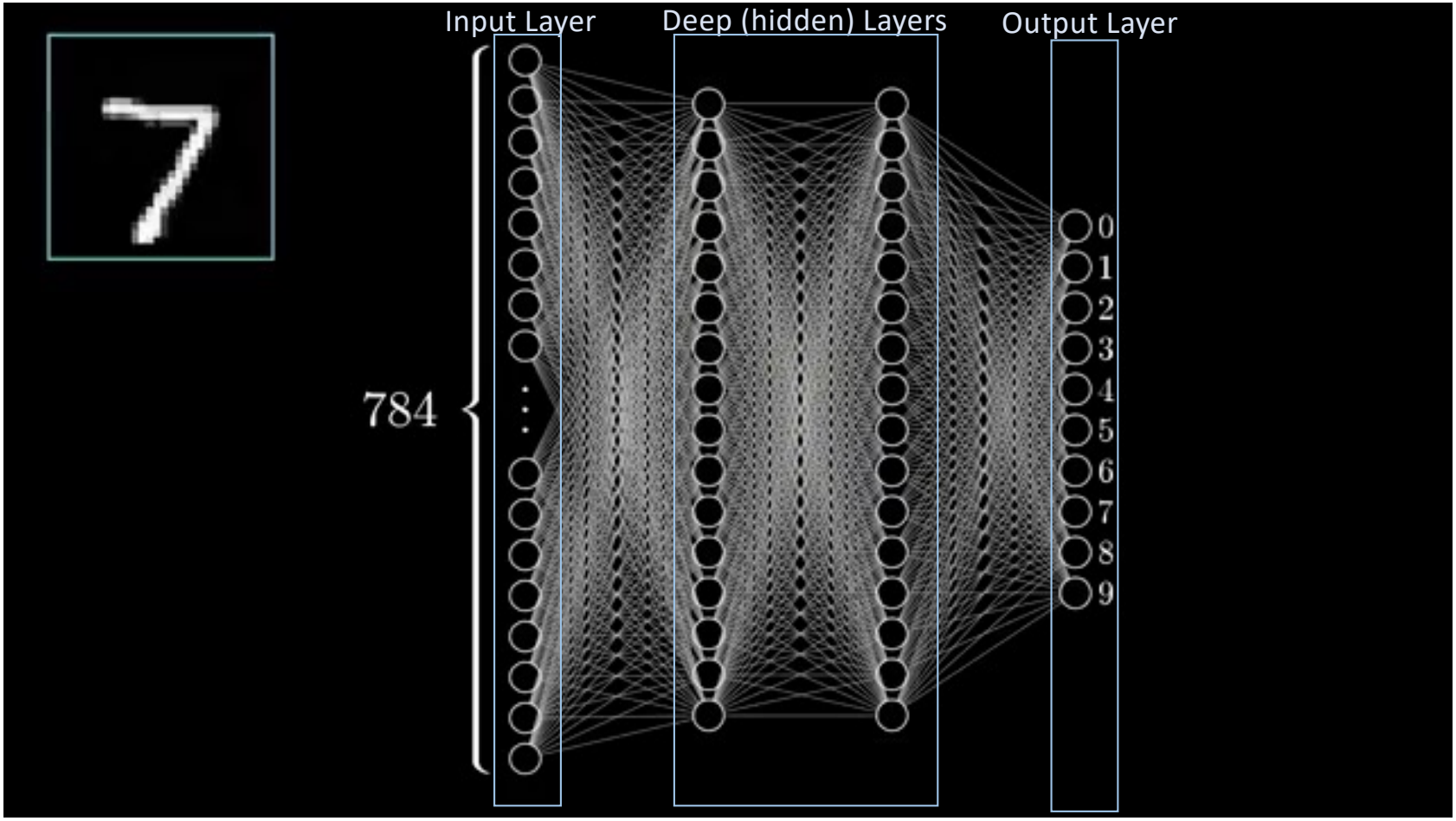
So how does DL learn?

- Tasks are learned using Artificial Neural Networks
- Models can be trained on raw data
- Feature extraction is automatically performed; process is iterative, using multiple layers

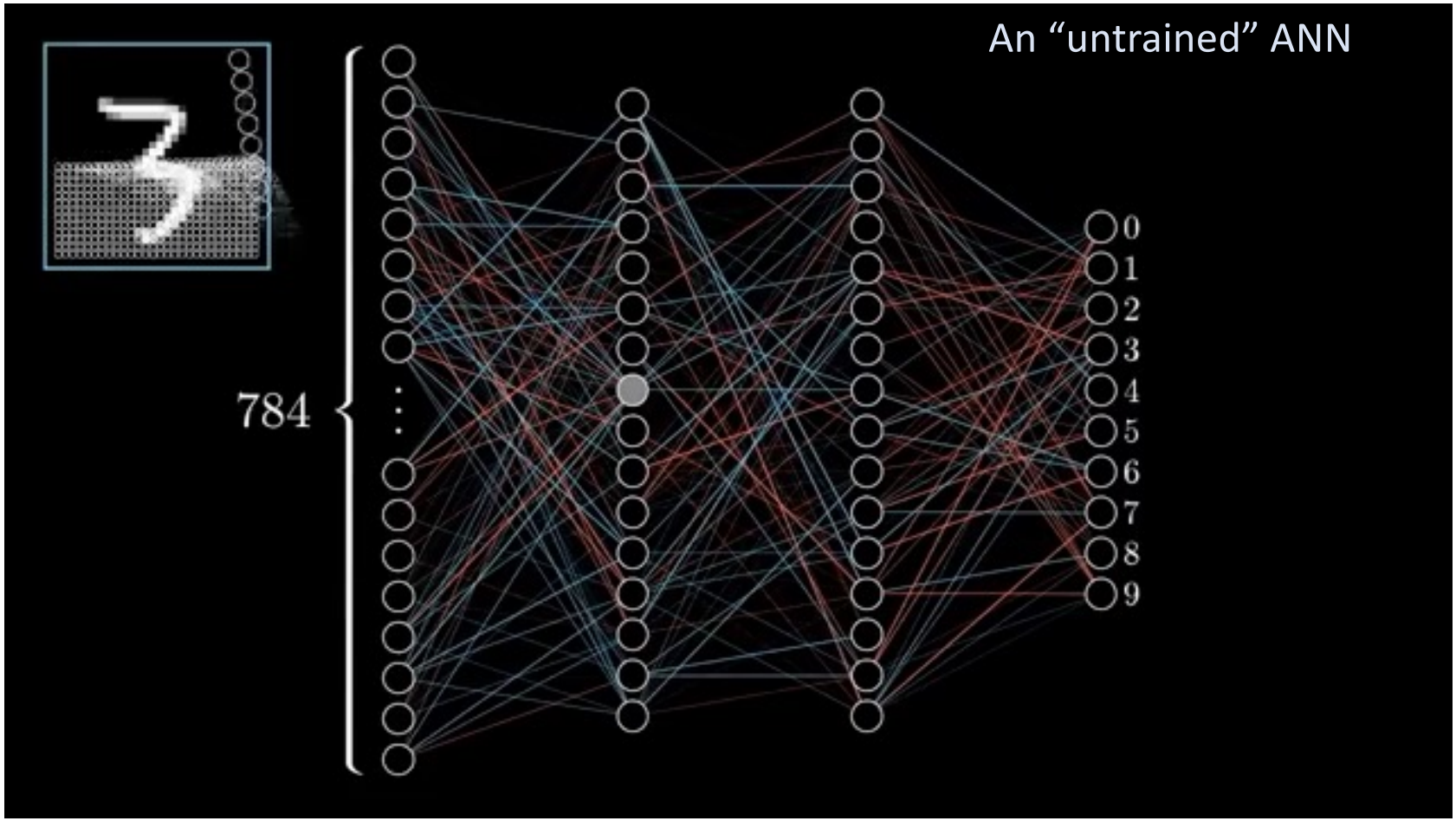


© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

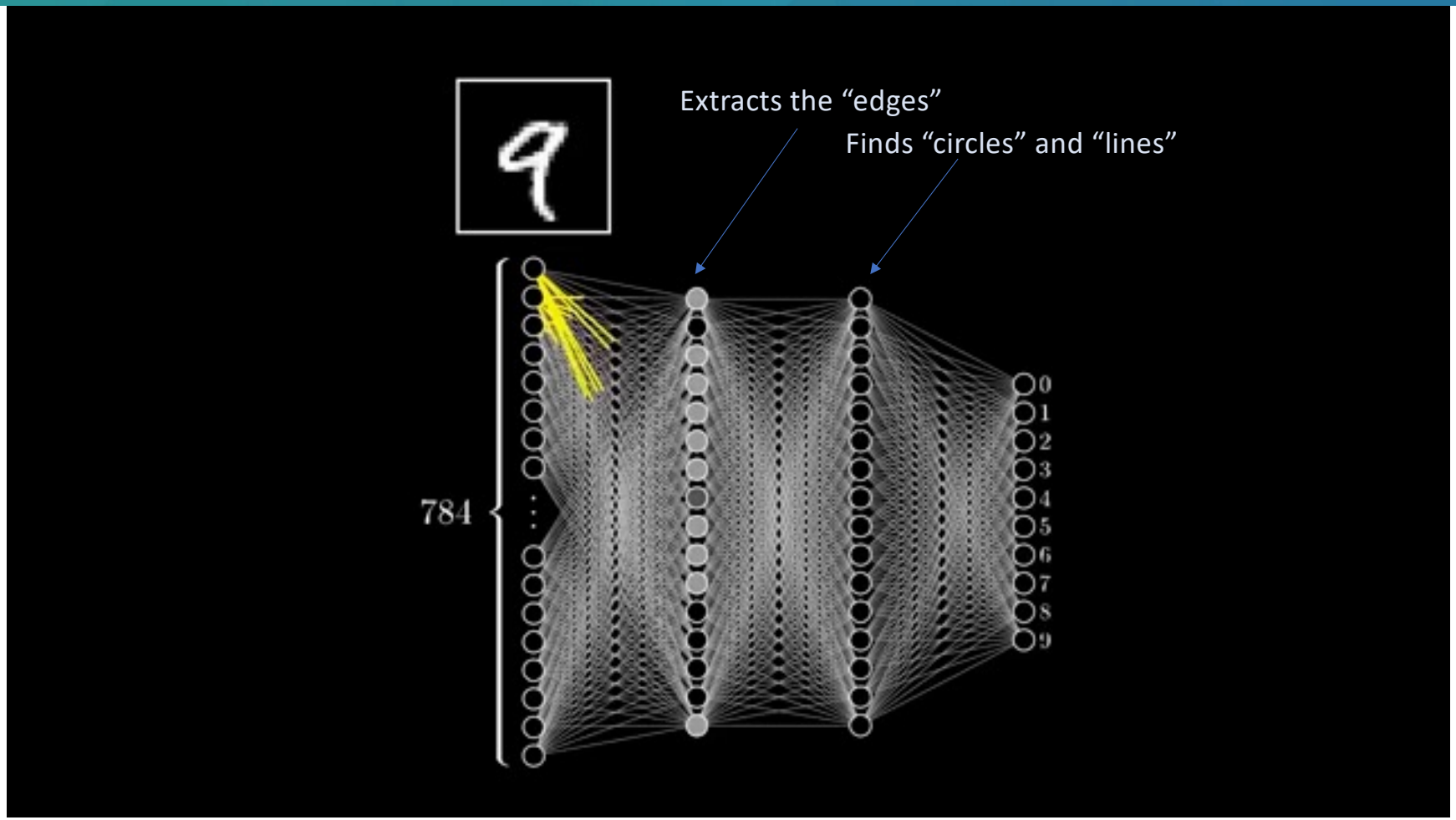
Deep Learning



Deep Learning

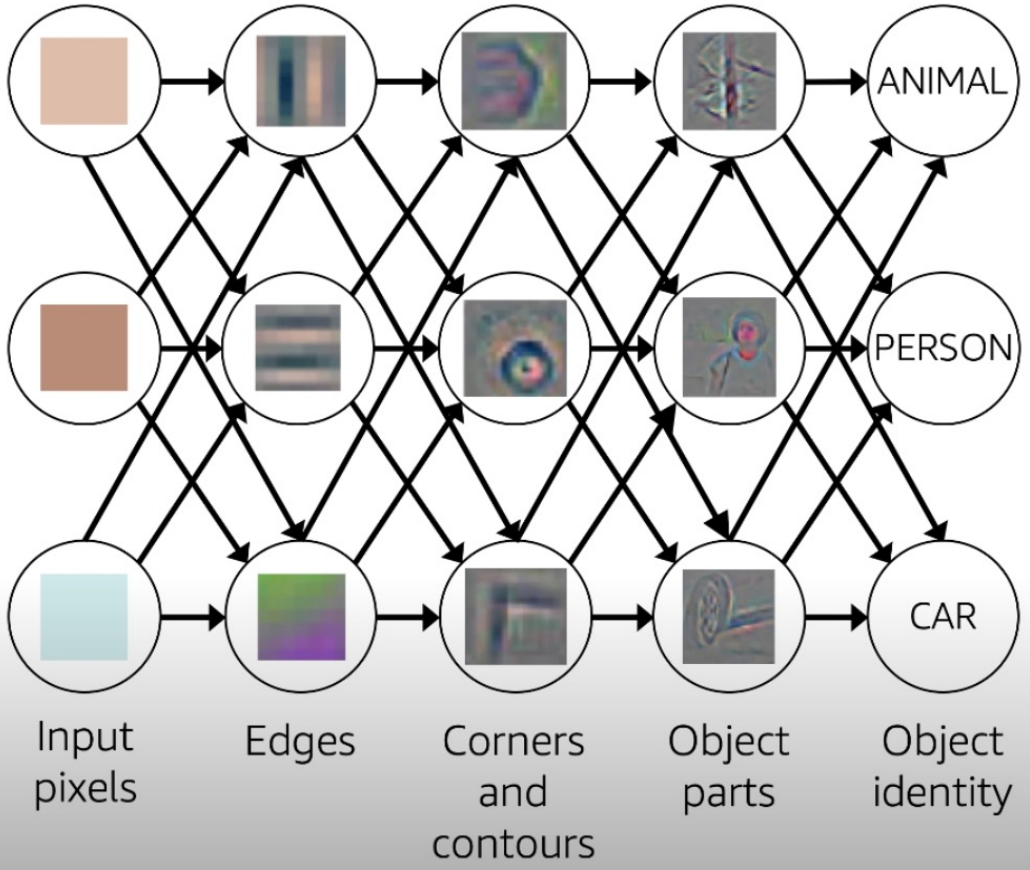


Deep Learning



Deep Learning

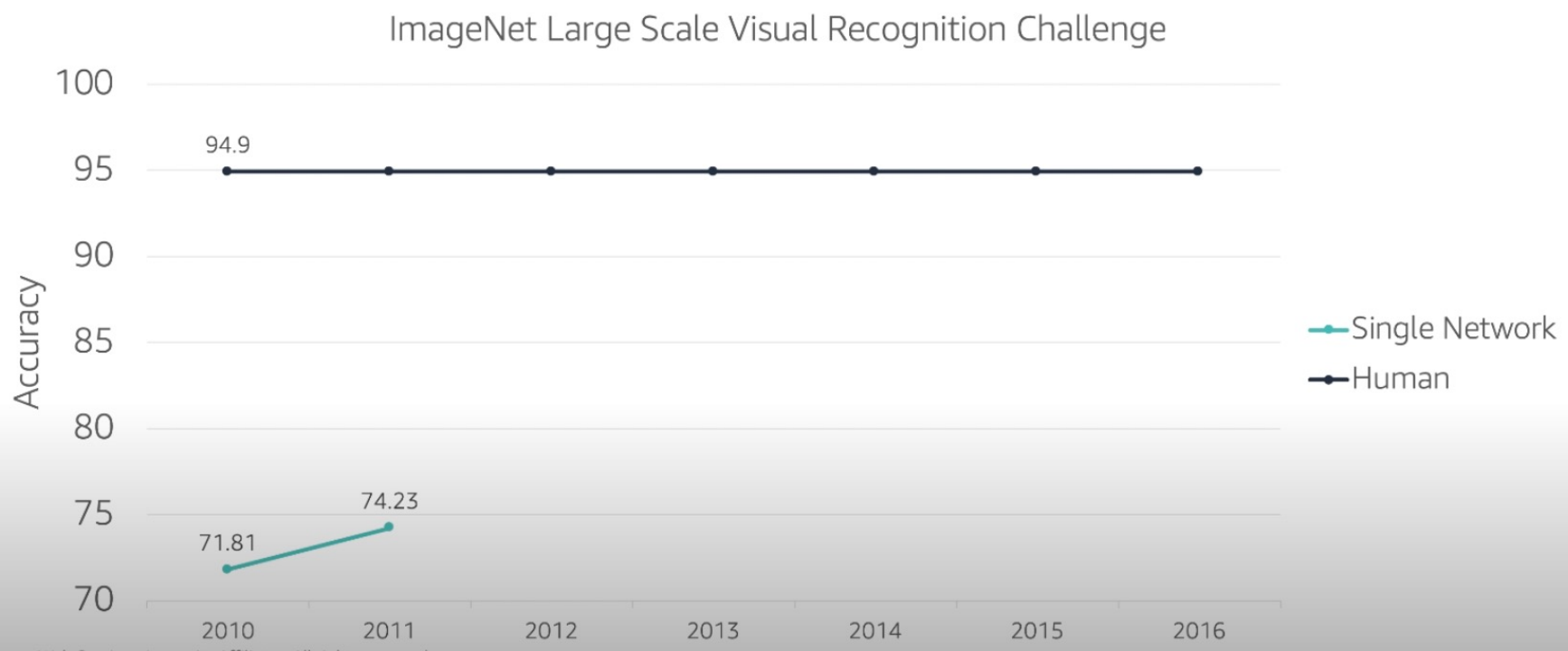
Example: Image classification using ANN



© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

An example of the impact of deep learning

2010 & 2011: Even the best programs submitted were still well below human accuracy.



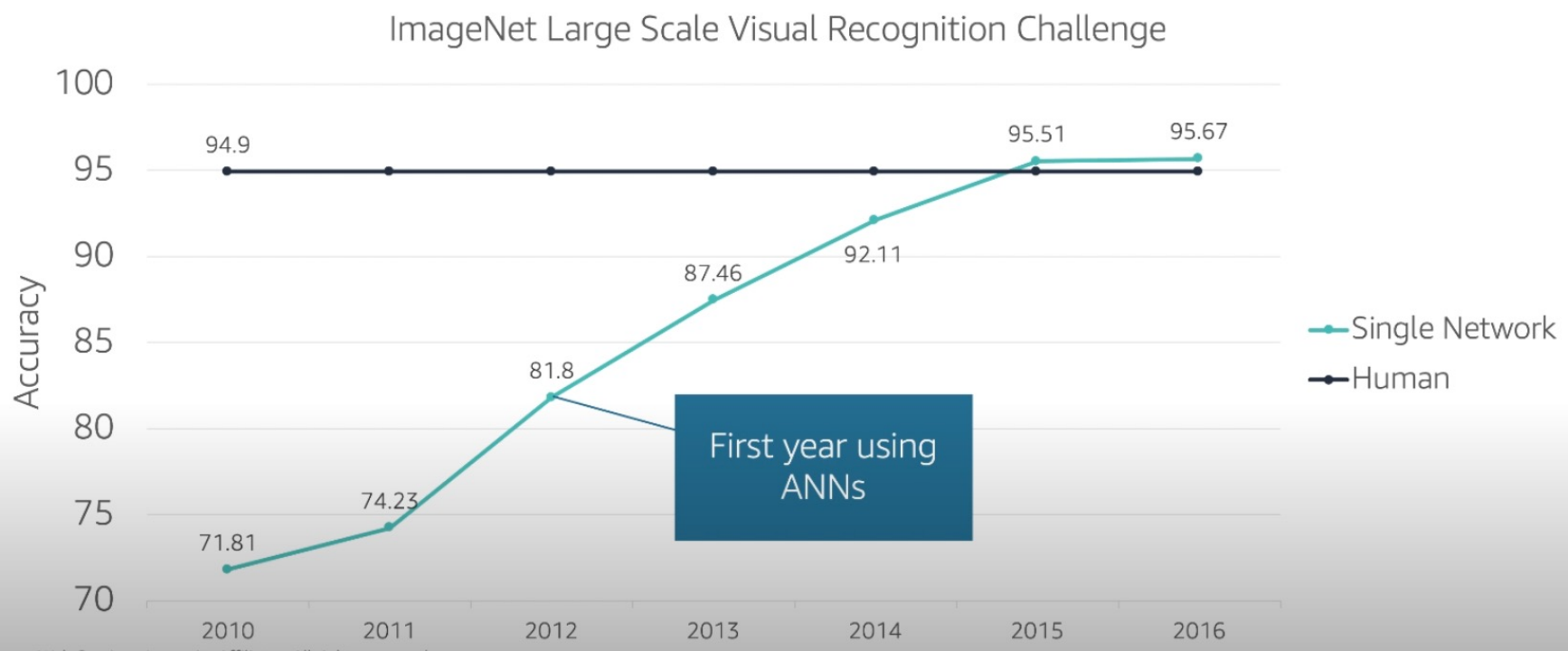
© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

An example of the impact of deep learning

aws training and certification

2010 & 2011: Even the best programs submitted were still well below human accuracy.

2012+: Accuracy jumps every year until surpassing human accuracy in 2015.



© 2020 Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Relax.. Nobody expects you to know everything

- You are **not** expected to know the underlying mathematics for AI/ML models.

- Models have already been built and which model works in what condition and when they should be used. You should only know this!

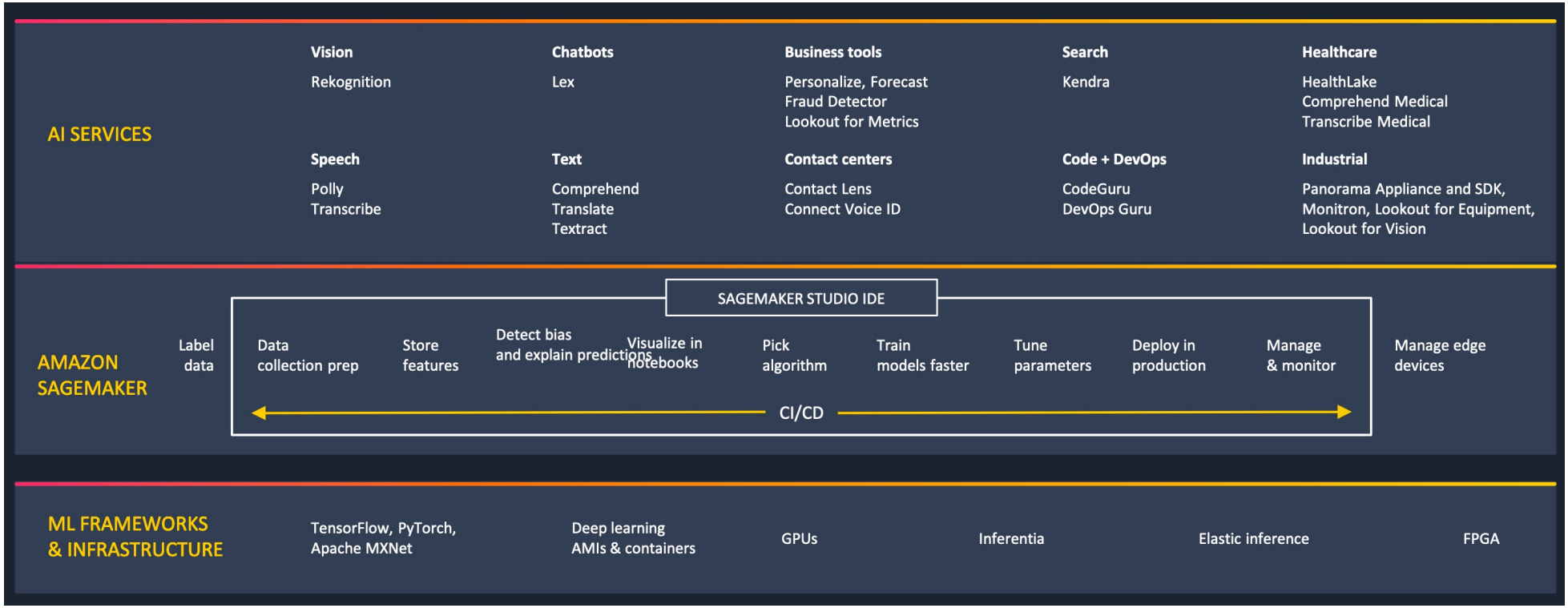
- You should know how to perform ("Feature Engineering"), statistical basics (distributions, types of errors, confusion matrix, etc.)

→ You should know which Python libraries are helpful.

- You can find plenty of material online, incredible collaboration in data science. And we are here to help! ☺

ML Stack on AWS

aws training and certification



Let’s Try!