

Machine Learning 101









Outline

What is machine learning?

Importance of data

Learning types and algorithms

Examples and demo

About me

Mark Kalal

Software development / technology solutions

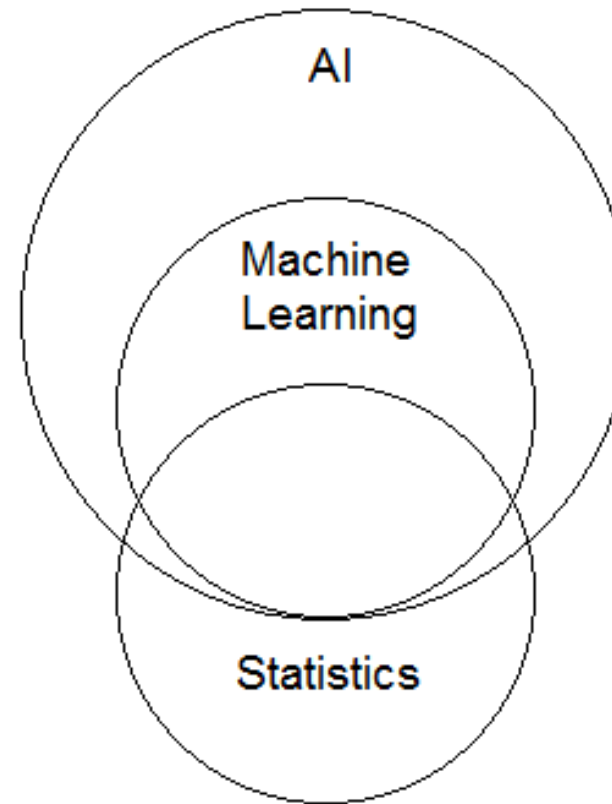
mdkalal@gmail.com

@MarkKalal

What is Machine Learning

Statistics?

Artificial Intelligence?



Learns by experience



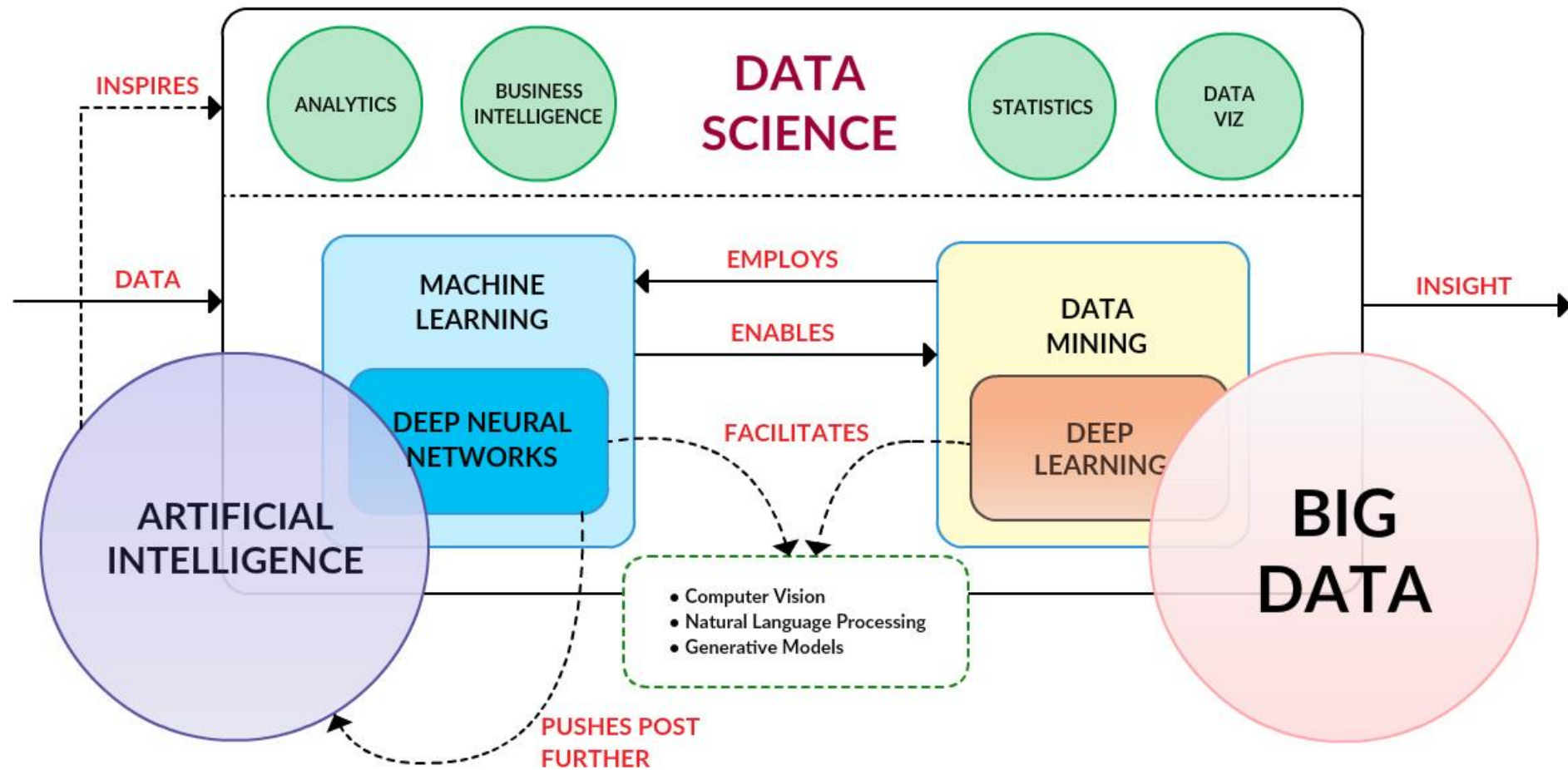
Gets specific instructions



Learns by ~~experience~~ data



Application of artificial intelligence (AI) that provides systems the ability to learn and improve from “experience” (data) without being explicitly programmed



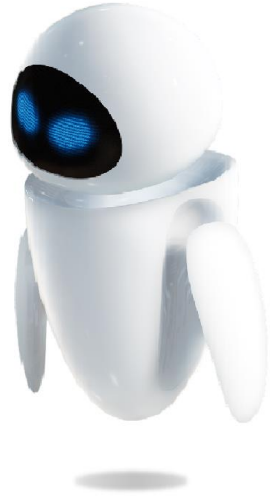
So what?

Many benefits

- Speed of analyzing complex data, revolutionizing business and data processing
- Greatly increased memory handling and computational powers (past barriers to implementation)

Some concerns

- Consequences, potential for misuse – intentional and non-intentional
- Opaque processing



My computer suddenly started
singing "Hello from the other side"!

Of course it did, after all ...



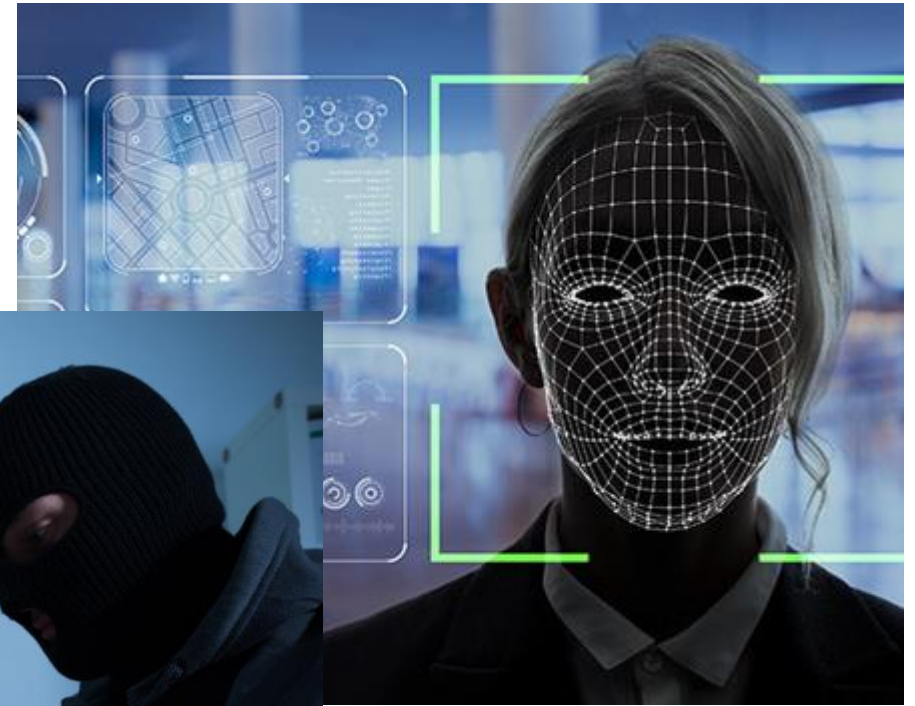
It's A Dell!

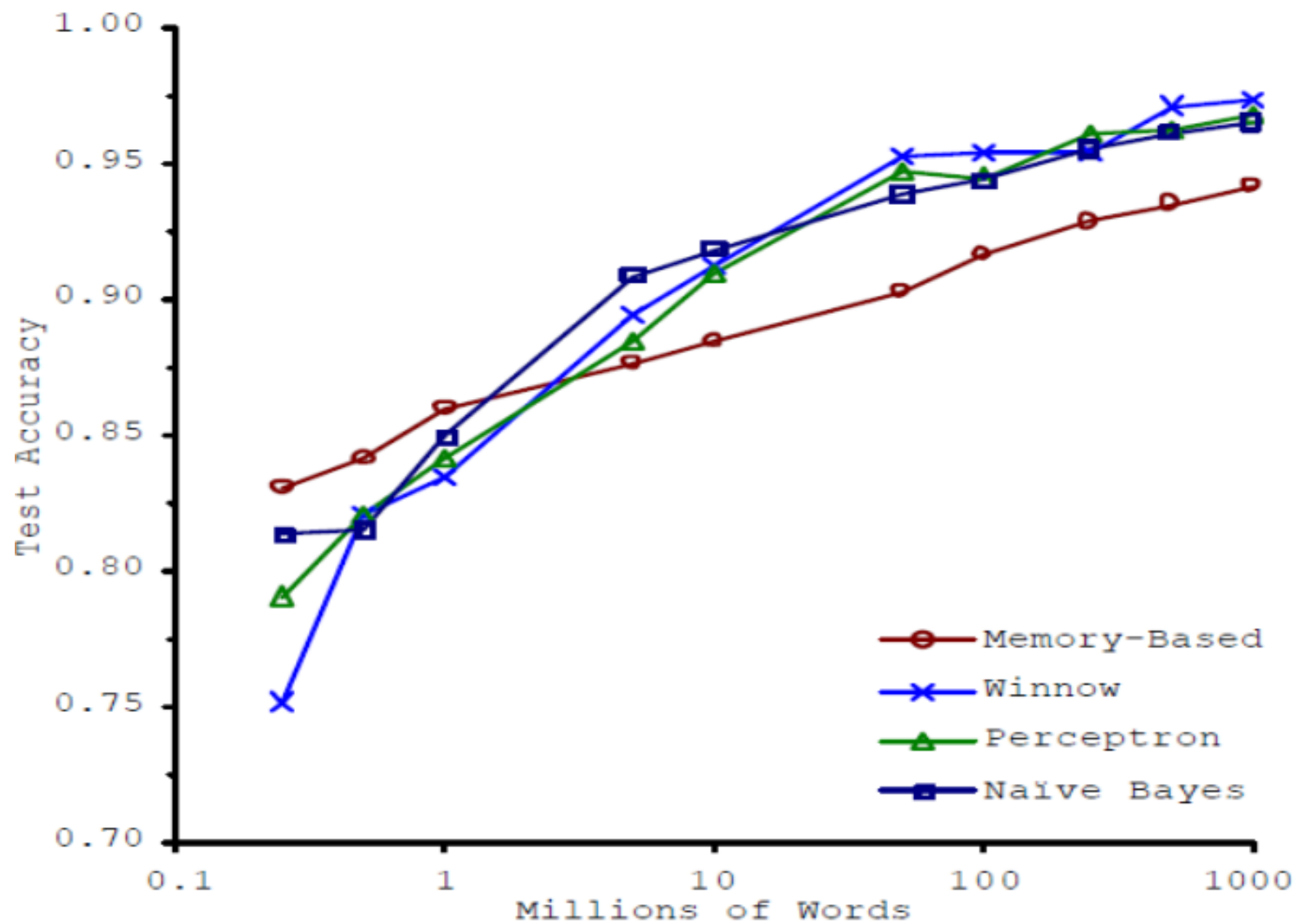
It's all about the data

“Fuel” for machine learning

Data, data, and more data

Many useful things are being done





“Big Data”

Volume

Velocity

Variety



Where does it come from

Some you already have

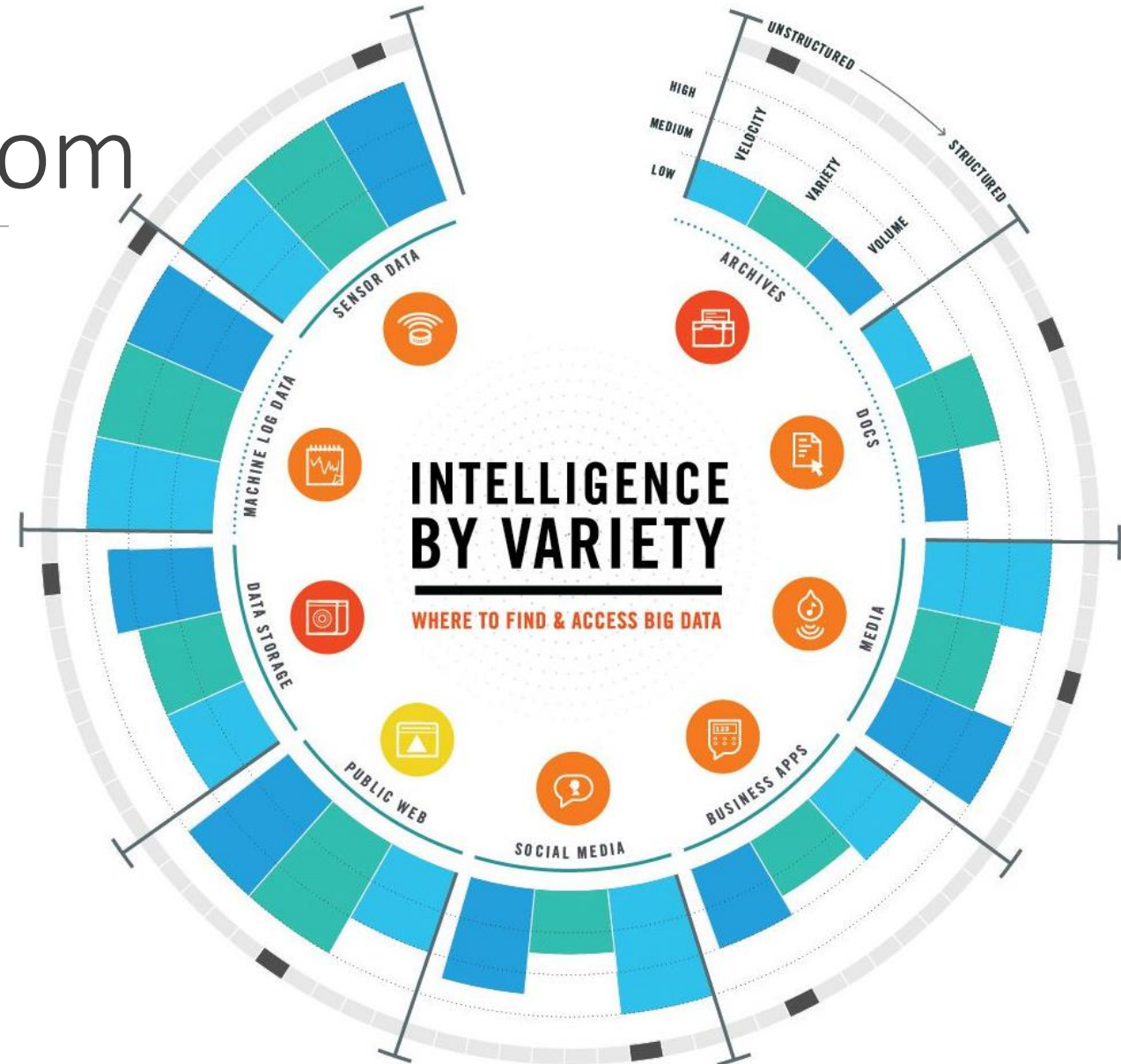
- Database
- Logs

Some you can get

- Web, public sources
(data.gov, kaggle.com/datasets)
- IOT sensors

Some you can ask for

- Social media
- Anything a user can provide



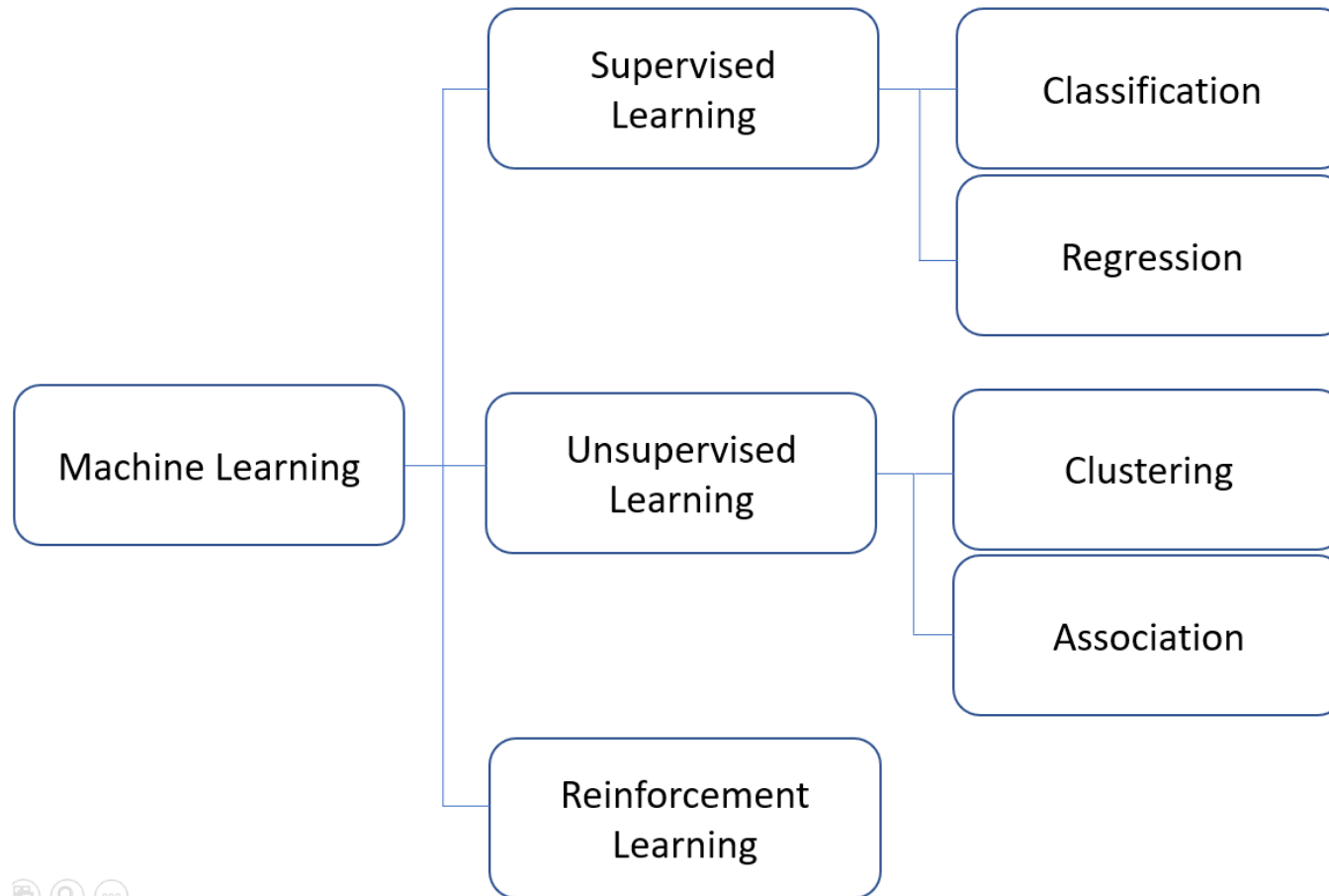


The journalist asked a programmer: "What makes bad code"? His reply?

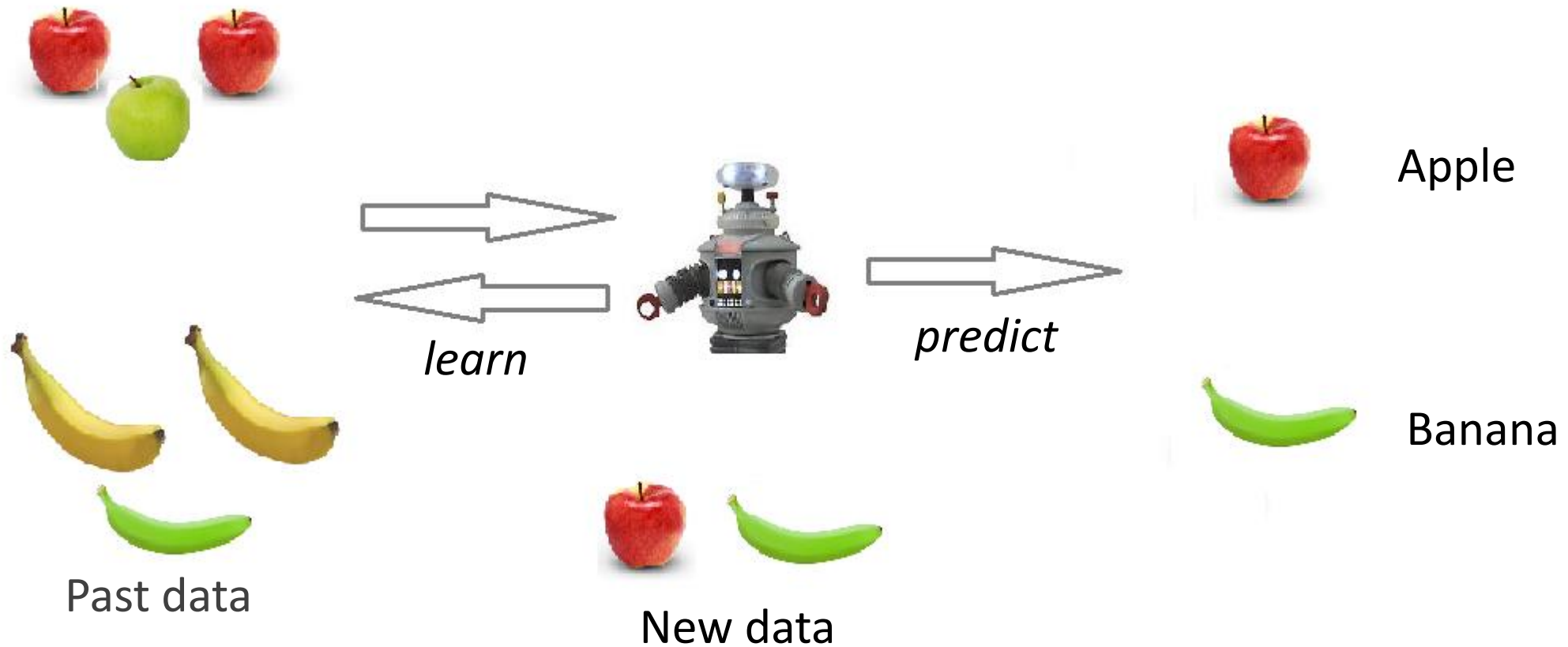
No Comment!



Learning types



Supervised Learning – make predictions



Supervised Learning

Classification

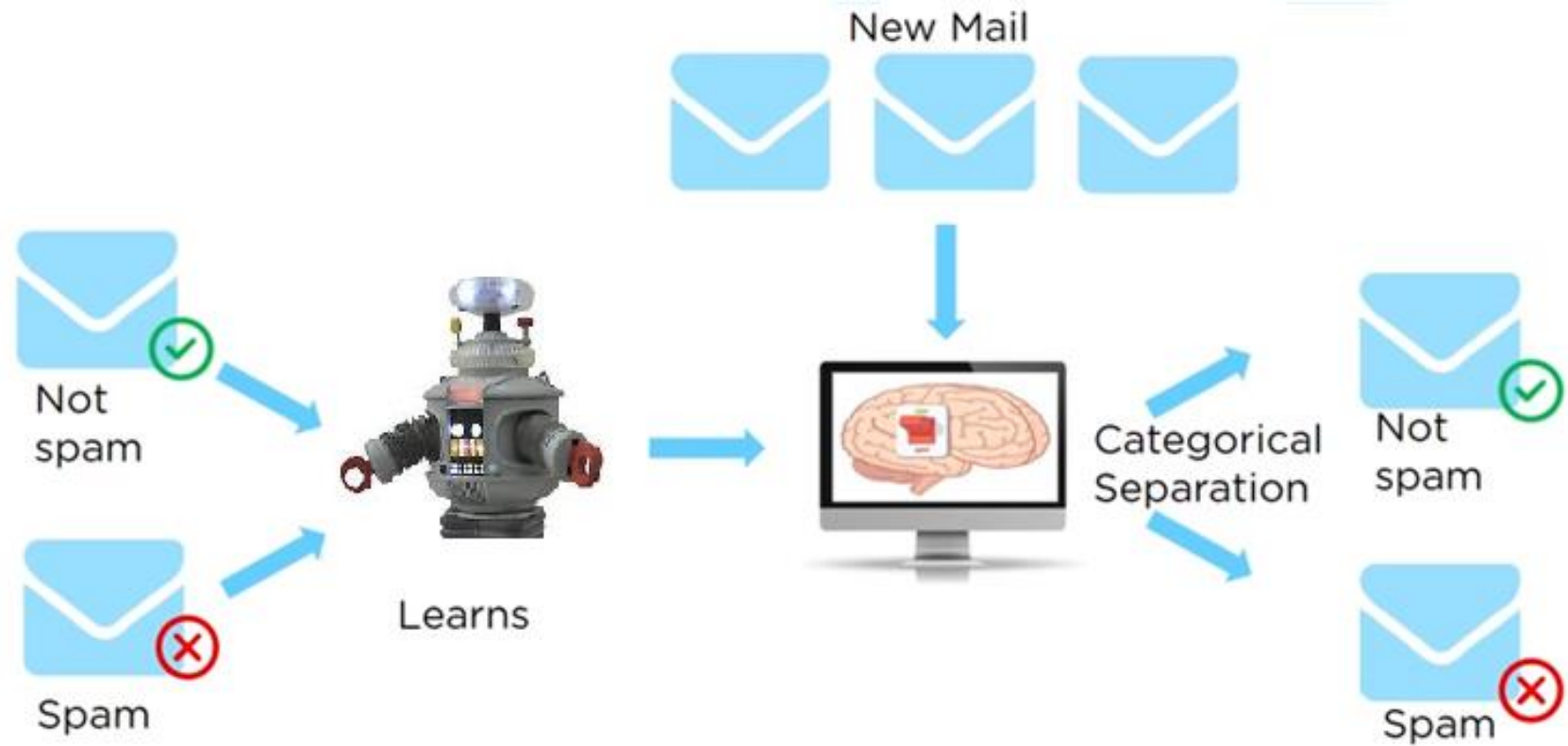
- Question or output is categorical, i.e. True/False

Regression

- Question or output is a real or continuous value

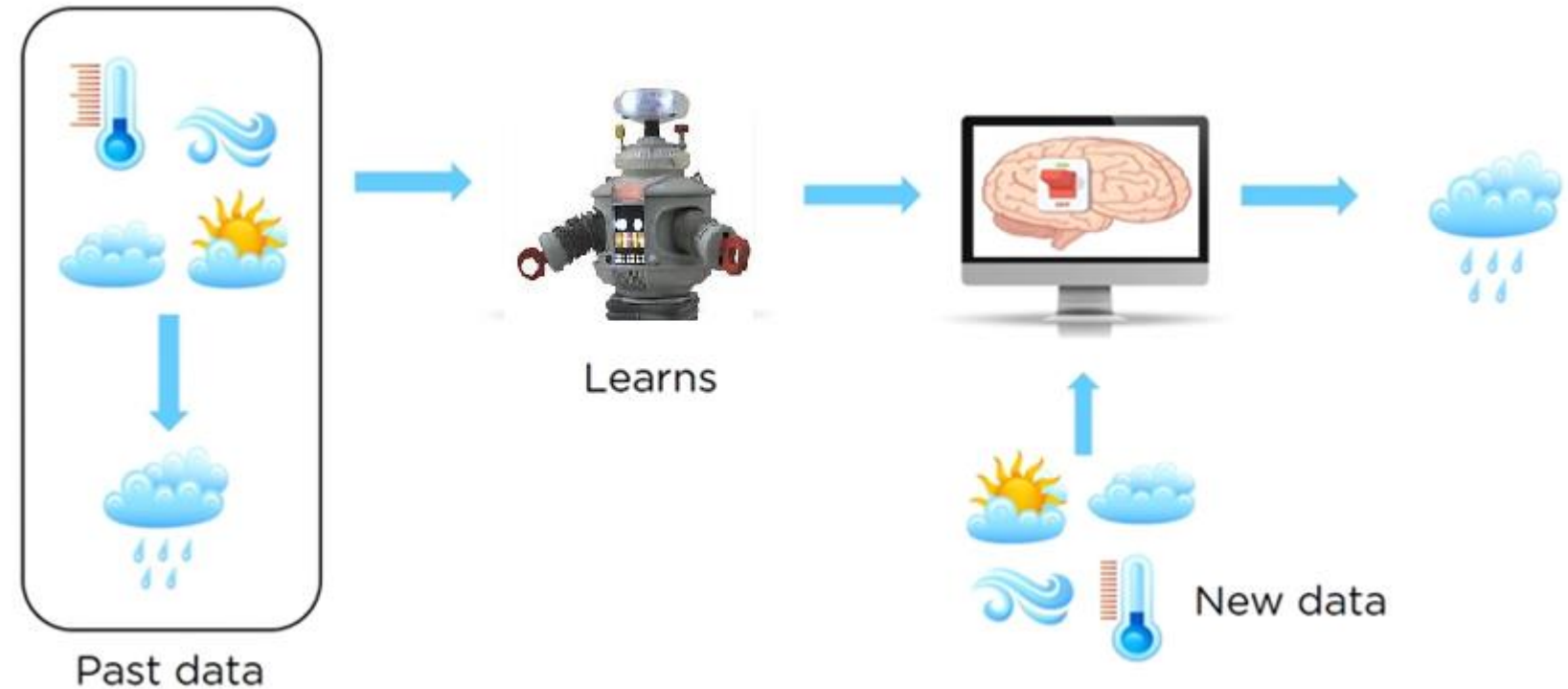
Supervised Learning

Classification

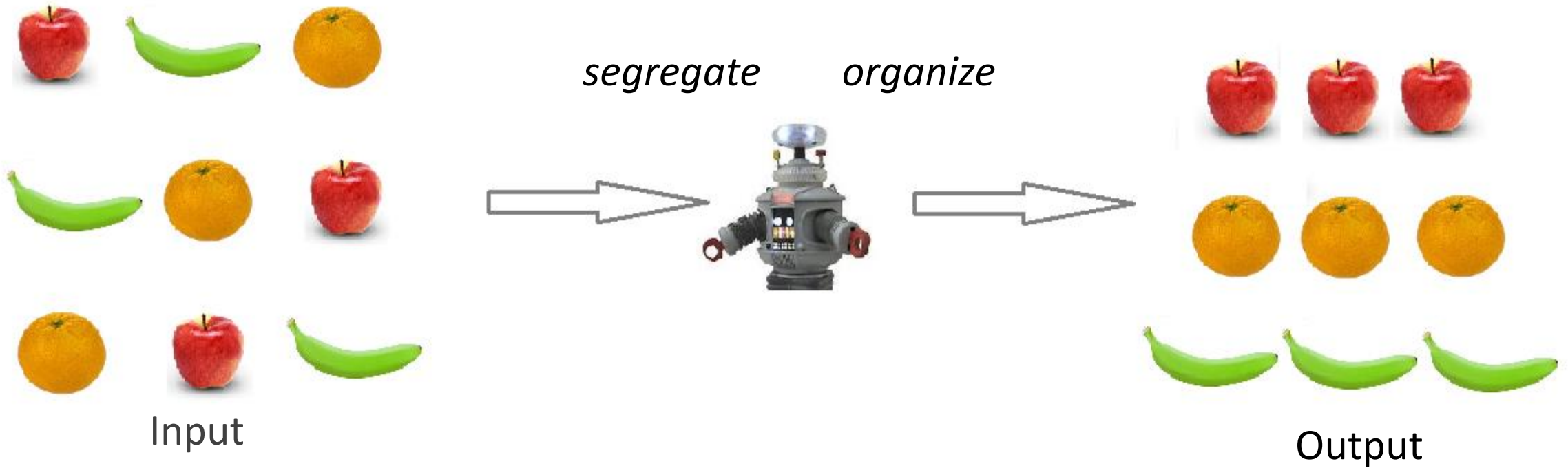


Supervised Learning

Regression



Unsupervised Learning – finding hidden patterns



Unsupervised Learning

Clustering

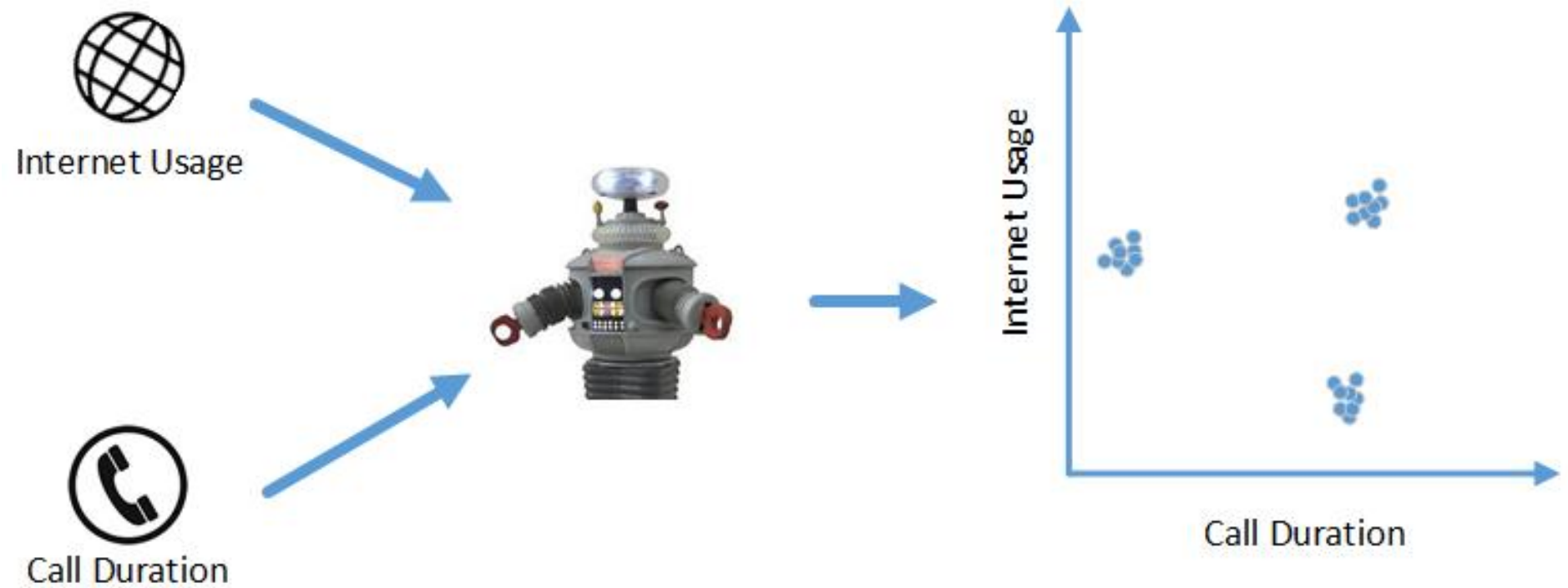
- Groups things based on similarities between them, and differences between others

Association

- Discovers relations or probability of occurrences within data

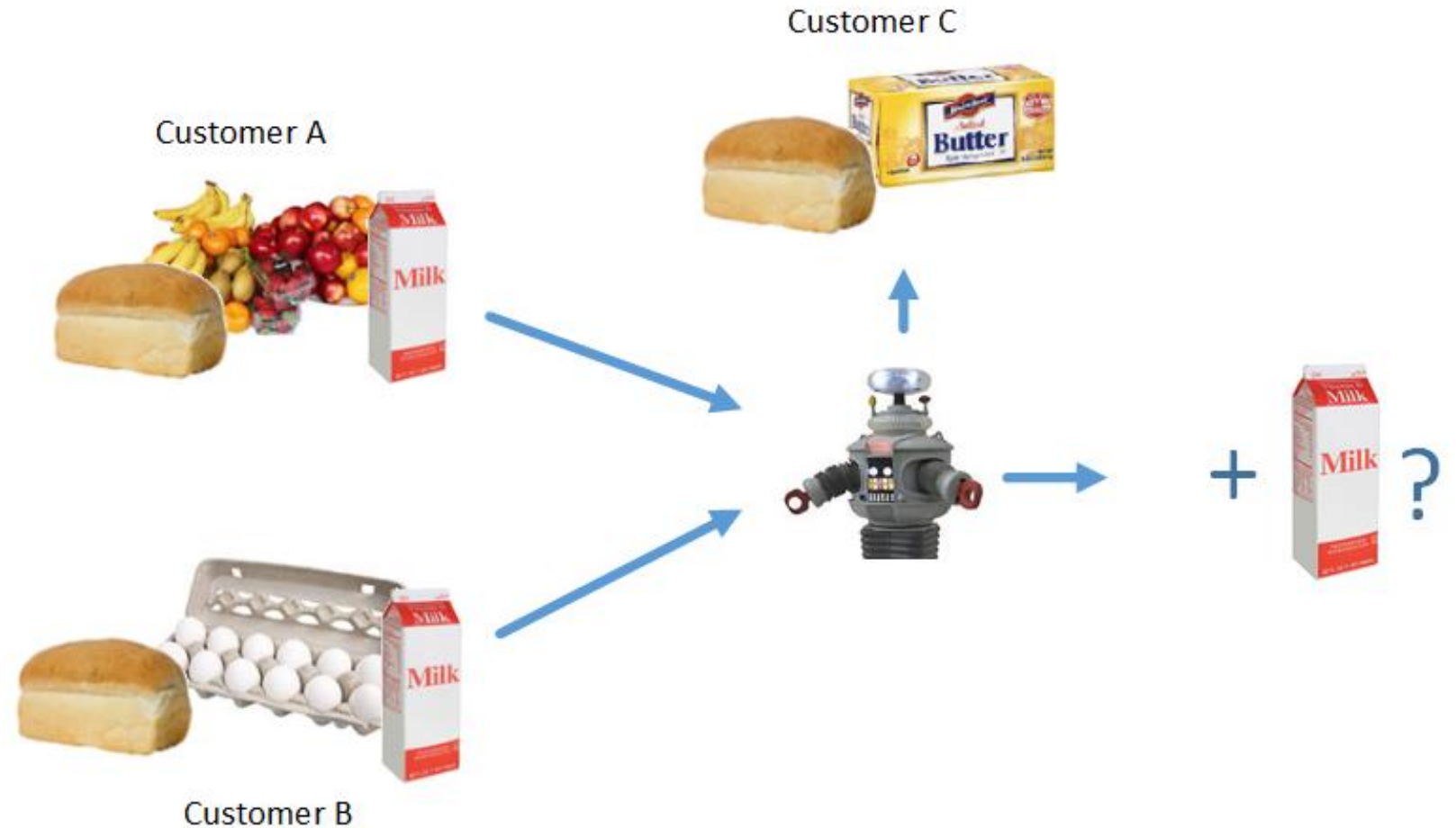
Unsupervised Learning

Clustering

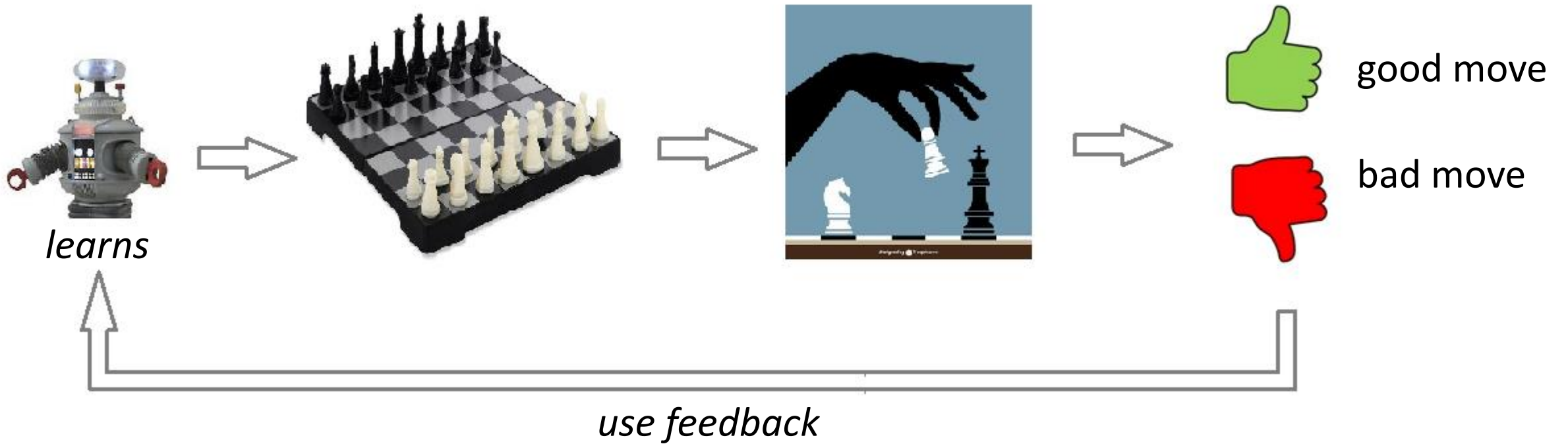


Unsupervised Learning

Association



Reinforcement Learning - Decisions based on rewards for past actions



Algorithms



Algorithms

Decision Tree

- Represents data that is divided/"branched" by conditions (questions and answers)

Linear Regression

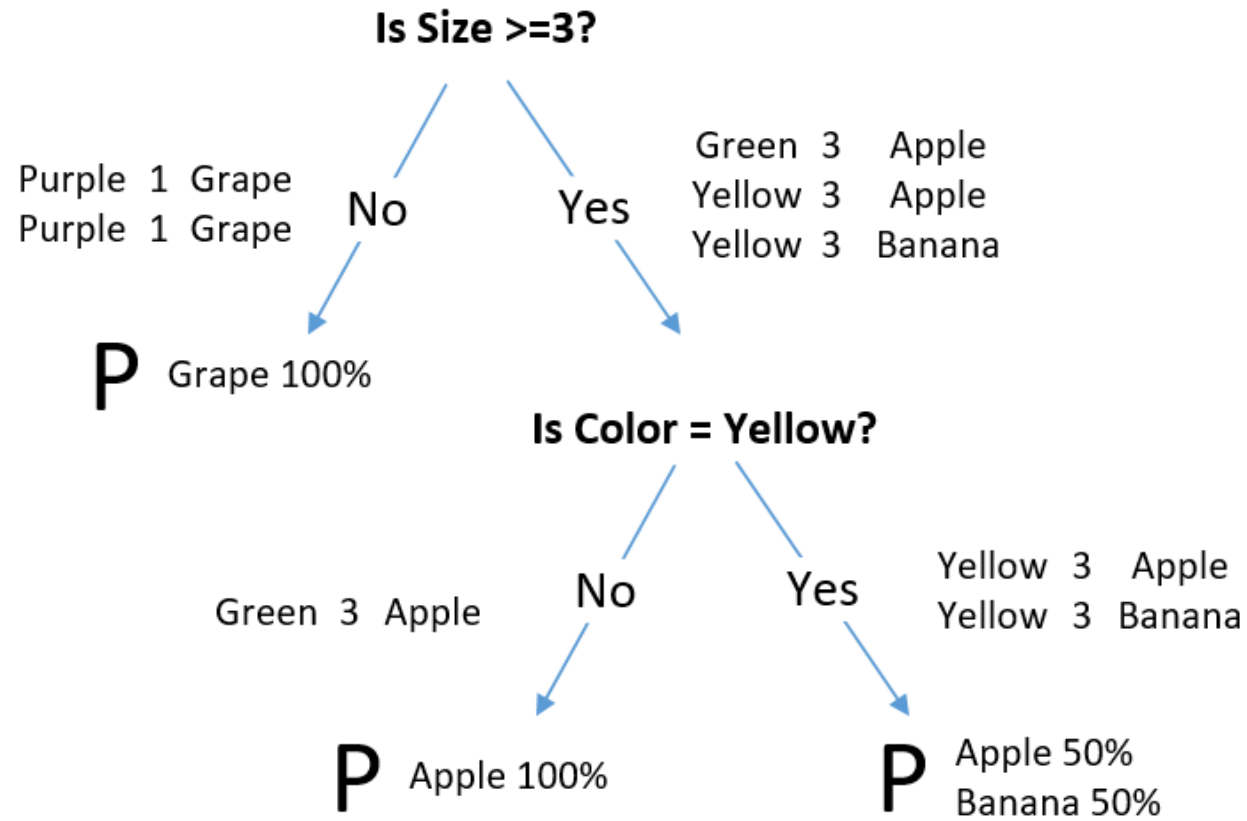
- Represents and expresses the relationship between data with a line (X-Y grid)

Decision Tree

What fruit is this? Grape? Apple? Banana?



Color	Size	Label
Green	3	Apple
Yellow	3	Apple
Purple	1	Grape
Purple	1	Grape
Yellow	3	Banana



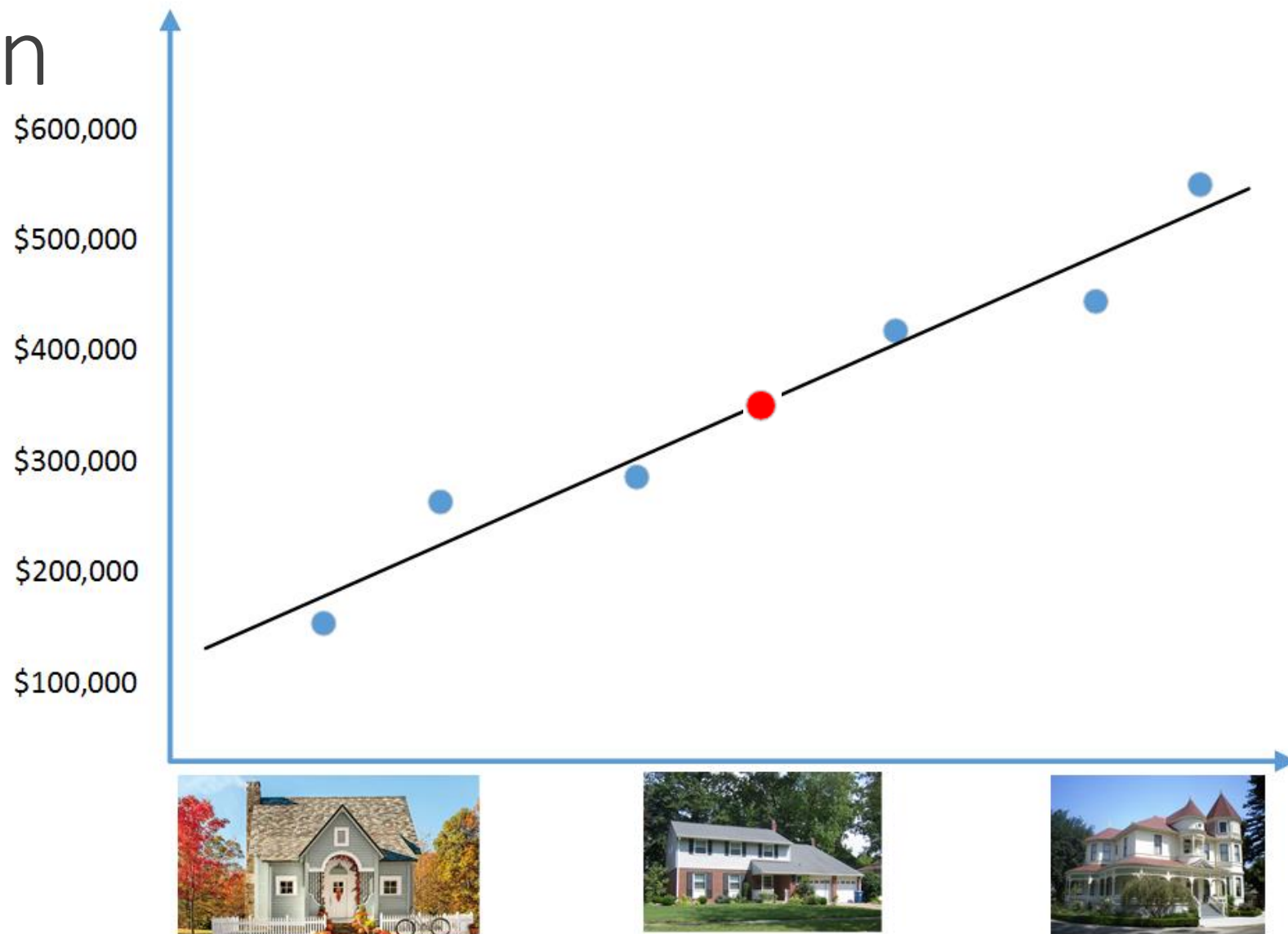
Linear Regression

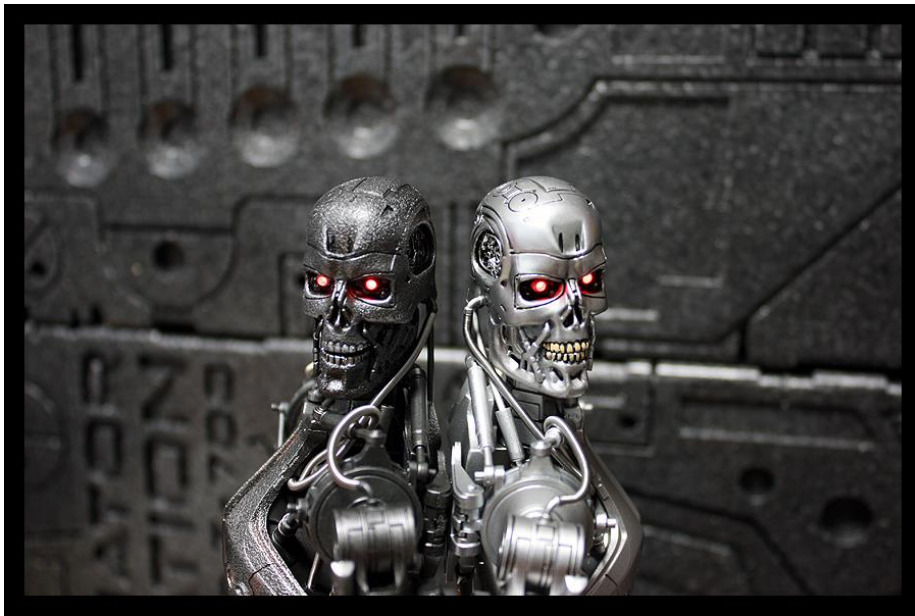
What does this house cost?



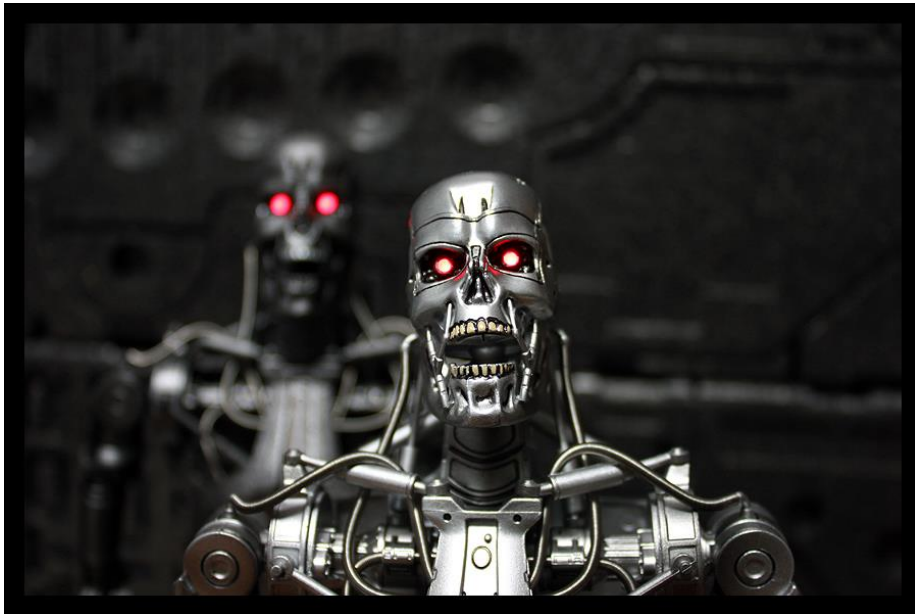
Linear Regression

What does this house cost?





I walked down a street where the houses were numbered 64k, 128k, 256k, 512k, and 1mb



It was a trip down *memory lane*!

Machine Learning Lifecycle/Pipeline

Start with a question or problem

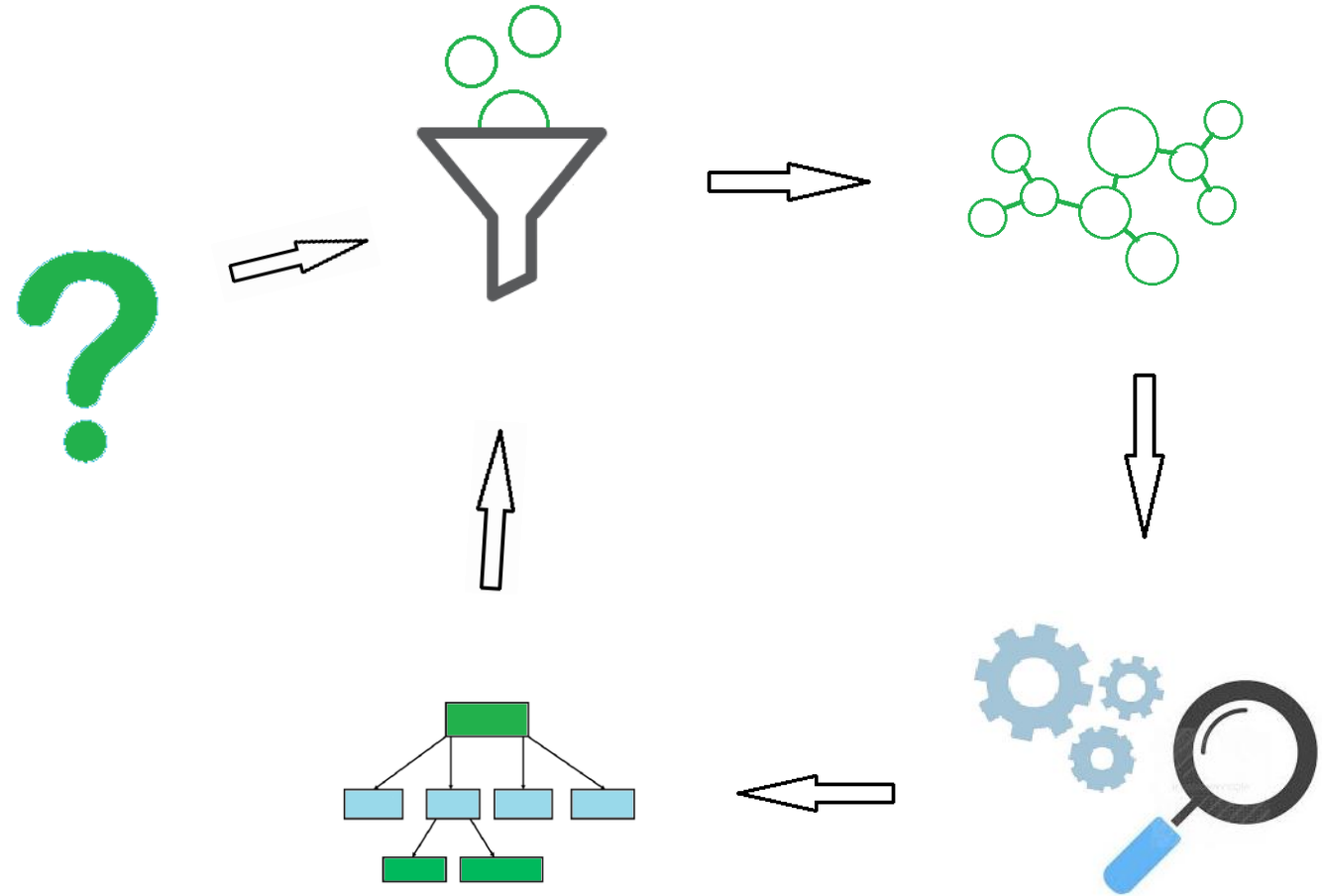
Collect data

Model

Train (and test)


Classify/Recommend/Predict

Re-train



Tools



 Microsoft
Cognitive Services



Machine Learning



PROPHET

ML.NET

 PyTorch

ML.Net Demo

- Installation
- Data files and formats
- C# API
- Model Builder
- Sample apps and code

Summary

- ✓ What is machine learning
- ✓ Importance of data
- ✓ Learning types and algorithms
- ✓ Examples and demo

For more information

ML.Net - <https://dotnet.microsoft.com/apps/machinelearning-ai/ml-dotnet>

Kaggle Datasets - <https://www.kaggle.com/datasets>

US open data - <https://www.data.gov/>

World Bank open data - <https://data.worldbank.org/>

Josh Gordon ML Recipes - <https://www.youtube.com/watch?v=cKxRvEZd3Mw>

Introduction to AI - <https://www.coursera.org/learn/ai-for-everyone>

This slide deck – <https://github.com/mdkalal/ml101>

Thank you!

Feedback welcome

Mark Kalal - mdkalal@gmail.com

@MarkKalal