Machine Learning

Abdul Haseeb Shaikh

Agenda

- Introduction to Machine Learning
- Importance of Machine Learning
- Types of Machine Learning
- Applications of Machine Learning
- Machine Learning: State-of-the-art
- Data Mining versus Machine Learning
- Role of Data Analyst and Machine Learning Engineer
- Introduction to Python
- Recap of the Lecture

What is Machine Learning

- Certain tasks are extremely difficult to program by hand:

Spam filtering

What is Machine Learning

- Certain tasks are extremely difficult to program by hand:

Spam filtering

Face recognition

What is Machine Learning

- Certain tasks are extremely difficult to program by hand:

Spam filtering

Face recognition

Machine translation

What is Machine Learning

- Certain tasks are extremely difficult to program by hand:

Spam filtering
Face recognition
Machine translation
Speech recognition

What is Machine Learning

- Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.
- A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.
- Machine learning is programming computers to optimize a performance criterion using example data or past experience.

What is Machine Learning

- Suppose your twitter program watches which tweet you do or do not mark as cyberbullying, and based on that learns how to better filter bullying tweet. What is the task T in this setting?

a. Classifying tweets as bully or no bully.
b. Watching you label tweets as bully or no bully.
c. The number (or fraction) of tweets correctly classified as bully or no bully.
d. None of the above—this is not a machine learning problem.

Generalities

- Virtually all learning problems can be formulated as (complex) mappings between inputs and outputs
  - We are trying to learn what is the best output **o** to produce for each possible input ***i***
  - Mathematically speaking, we search for a «good» function **F: I → O**, where **I** is the set of possible inputs, and **O** the set of possible outputs

Generalities - Examples

	Input ***i***	Output ***o***
Spam filtering	An email	{spam, non-spam}
Face recognition	An image	Identified faces
Machine translation	A sentence in language A	A sentence in language B
Speech recognition	A speech signal	A (text) sentence
Data mining	A financial transaction	{fraud, non-fraud}
Robot motion	Sensory data	Motor control

Learning Methods

- But how do we learn this mapping?
- The learning method depends on the kind of data that we have at our disposal
  - We can have examples of data where we have both the inputs and outputs: (***i***,***o***)
  - For some data, we only have the inputs ***i***

Learning Methods

- But how do we learn this mapping?
- The learning method depends on the kind of data that we have at our disposal
  - We can have examples of data where we have both the inputs and outputs: (***i***,***o***) → supervised learning
  - For some data, we only have the inputs ***i*** → unsupervised learning

Types of Machine Learning

Supervised

- Regression
- Classification

Unsupervised

- Clustering

Semi-Supervised

- Active Learning

Case-Based Reasoning

Reinforcement Learning

Supervised Learning

- In supervised learning, we have *training data* encoded as pairs (i,o), where the «correct» output is often manually annotated
  - E.g. spam filtering, machine translation, face recognition, etc.
- The function **F: I → O** is often dependent on a (sometimes large) set of parameters
- ... and the learning goal is to «adjust» these parameters in order to fit the data

Supervised Learning

- A spam filtering system might for instance have «weights» associated to each possible English word
  - The higher the weight, the more it contributes to the probability that the email is a spam
- The learning algorithm will then *adjust* these weights to fit the data

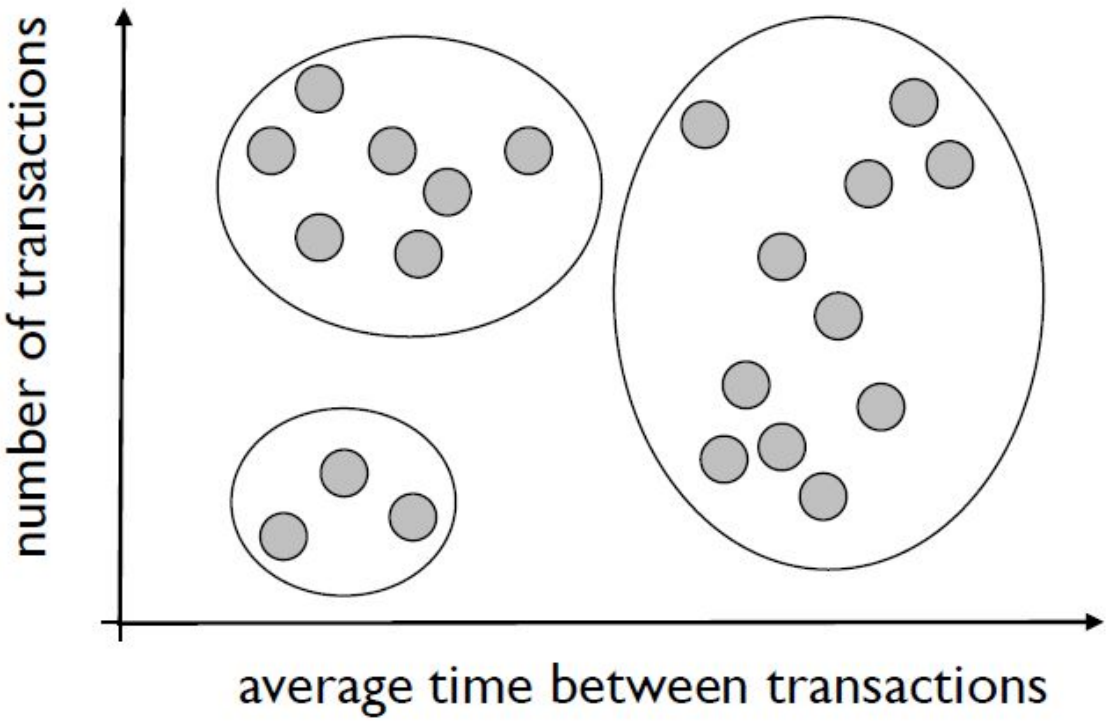
feature of the input, like presence/absence of a word

$$P(\text{email is spam}) \propto \sum_{w_i \in \text{weights}} w_i \, f_i(\mathbf{i})$$

Supervised Learning

- A good learning model is a model that **generalises** well to new data
  - In other words, it is able to *abstract* over its experience to detect underlying patterns
  - The design and test of such models is a crucial part of machine learning
- Else, the model is said to be *overfitted*
  - In other words, it is very well «fitted» to the examples it has processed, but perform very poorly with unseen data

Unsupervised Learning



- Sometimes, we don't have access to any output value **o**, we simply have a collection of input examples ***i***
- In this case, what we try to do is to learn the *underlying patterns* of our data
  - is there any *correlations* between features?
  - can we *cluster* our data set in a few groups which behave similarly, and detect *outliers*?

Semi-Supervised Learning

- The crux of the *semi-supervised machine learning or Active Learning* approaches is that machine learning algorithms can obtain optimum classification accuracy with only a few labeled instances.

- These approaches are beneficial where un- labeled data can be obtained easily in huge volumes; nonetheless, the labeling of the collected data is difficult, laborious, and expensive.

Reinforcement Learning

Playing Atari with Deep Reinforcement Learning

Volodymyr Mnih **Koray Kavukcuoglu** **David Silver** **Alex Graves** **Ioannis Antonoglou**

Daan Wierstra **Martin Riedmiller**

DeepMind Technologies

{vlad,koray,david,alex.graves,ioannis,daan,martin.riedmiller} @ deepmind.com

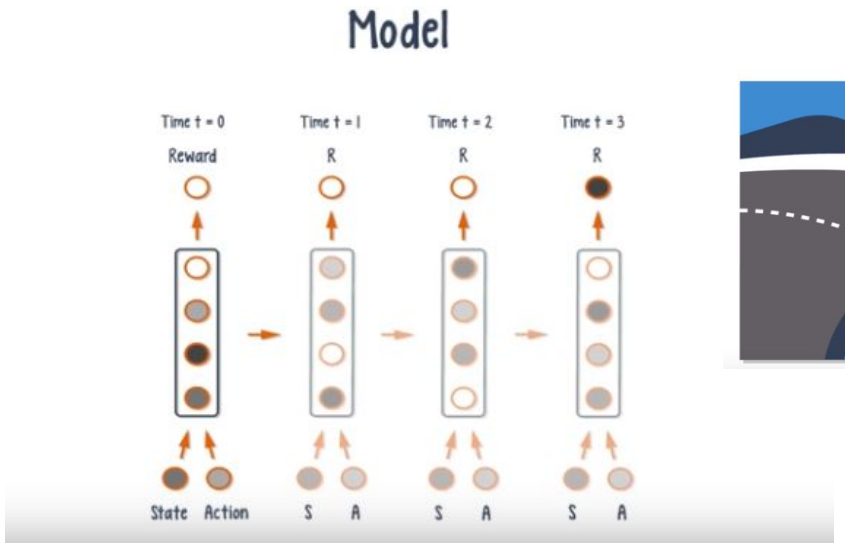
Abstract

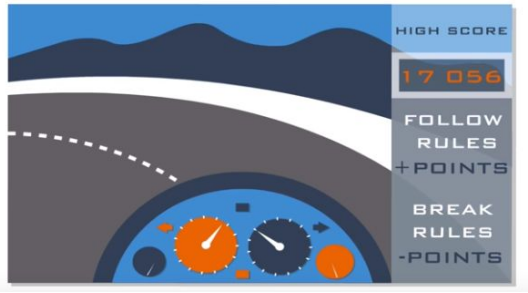
We present the first deep learning model to successfully learn control policies directly from high-dimensional sensory input using reinforcement learning. The model is a convolutional neural network, trained with a variant of Q-learning, whose input is raw pixels and whose output is a value function estimating future rewards. We apply our method to seven Atari 2600 games from the Arcade Learning Environment, with no adjustment of the architecture or learning algorithm. We find that it outperforms all previous approaches on six of the games and surpasses a human expert on three of them.

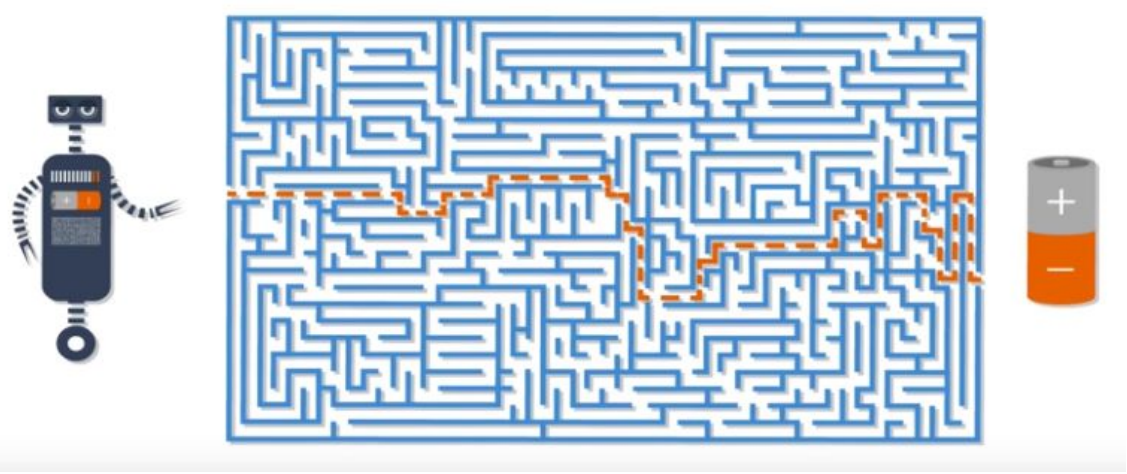
- December, 2013
- Google bought DeepMind

Reinforcement Learning

- Involvement of AI, animal psychology, and control theory
- Reinforcement learning (RL) can learn from experience and interaction with the environments.
- Agent interacts with environment and its goal is to maximize the numerical reward







Case-based Reasoning (CBR)

Case is the problem-Solution Pair

Case base is the Collection of cases

Process of solving new problems based on the solutions of previous similar problems.

Case-based reasoning can be used for classification and regression.

If the cases are simple, one algorithm that works well is to use the ***k*-nearest neighbors** for some given number *k*.

A Simple Example: Diagnosis of Car Faults

- Given: Symptoms
  e.g. engine doesn't start
  and measured values
  e.g. battery voltage = 6.3V
- Goal: Find cause for fault
  e.g. dead battery
  and repair strategy
  e.g. charge battery

Example Cases

Case 1

Problem & Features
- Problem: Front light not working
- Car: VW Golf, 2.0L
- Year: 1999
- Battery voltage: 13.6V
- State of lights: OK
- State of light switch: OK

Solution
- Diagnosis: Front light fuse defect
- Repair: Replace front light fuse

Case 2

Problem & Features
- Problem: Front light not working
- Car: Passat
- Year: 2000
- Battery voltage: 12.6V
- State of lights: surface damaged
- State of light switch: OK

Solution
- Diagnosis: Bulb defect
- Repair: Replace front light

New Problem

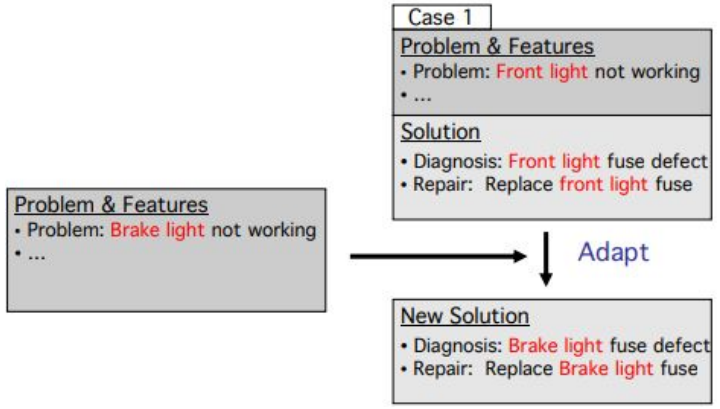
- Observations define a new problem
- Not all feature values may be known
- New problem = case without solution

Problem & Features
- Problem: Brake light not working
- Car: Passat V6
- Year: 2002
- Battery voltage: 12.9V
- State of lights: OK
- State of light switch: ?

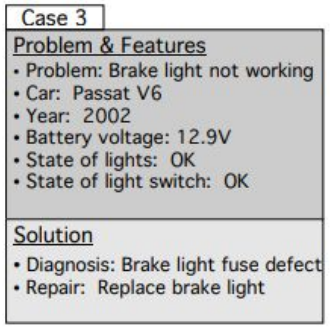
Case-based Reasoning

Case-based Reasoning

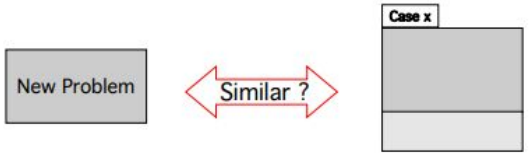
Reuse Case 1



Store New Case



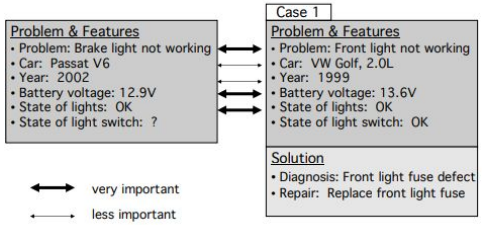
Find Similar Case



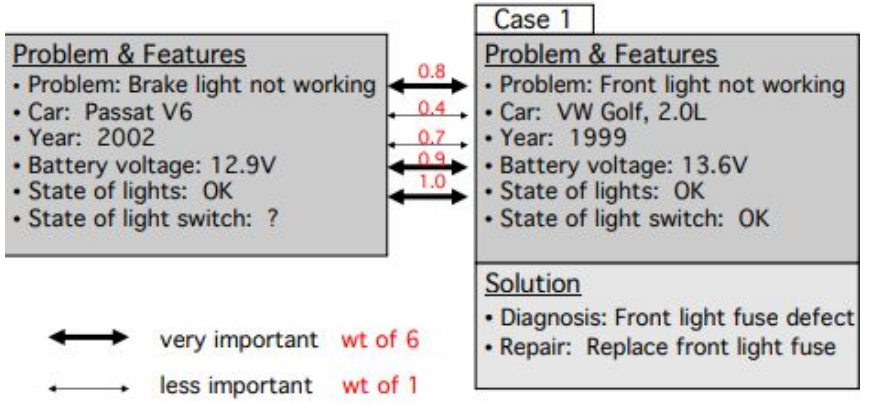
Find Similar Case

- Compare similarity of each feature
- But some features may be more important

Compare with Case 1

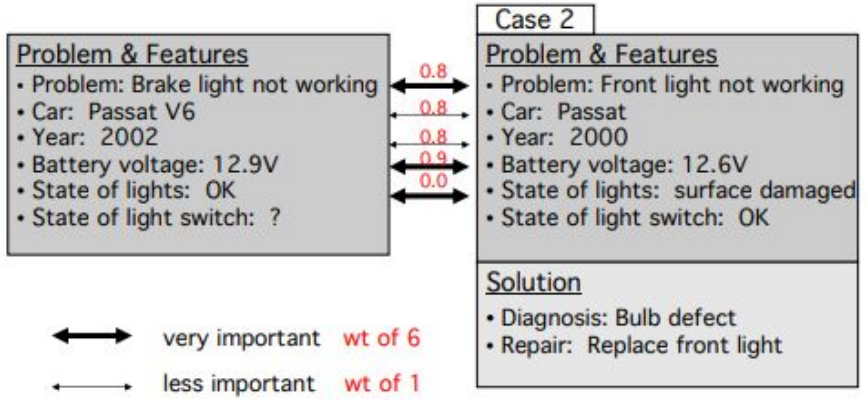


Compare with Case 1



Similarity by wted avg = 1/20 (6*0.8 + 1*0.4 + 1*0.7 + 6*0.9 + 6*1.0) = 0.87

Compare with Case 2



Similarity by wted avg = 1/20 (6*0.8 + 1*0.8 + 1*0.8 + 6*0.9 + 6*0.0) = 0.59

Growth of Machine Learning

Machine learning is preferred approach to

- Speech recognition, Natural language processing
- Computer vision
- Medical outcomes analysis
- Robot control
- Computational biology

Data Mining versus Machine Learning

Data Mining is used to look for emerging patterns that can help shape our decision-making processes.

For example, the data explores best-selling items, what was returned the most, and customer feedback to help sell more clothes and enhance product recommendations.

This use of data analytics can lead to an improved customer experience overall.

Data Mining versus Machine Learning

Machine learning, on the other hand, can actually learn from the existing data and provide the foundation necessary for a machine to teach itself.

For example: Ham or Spam email

Machine learning can look at patterns and learn from them to adapt behavior for future incidents, while data mining is typically used as an information source for machine learning to *pull* from.

How Machine Learning works?

Machine Learning

- Closely related to data mining and statistics

Data Mining:

Concerned with the analytics of data.

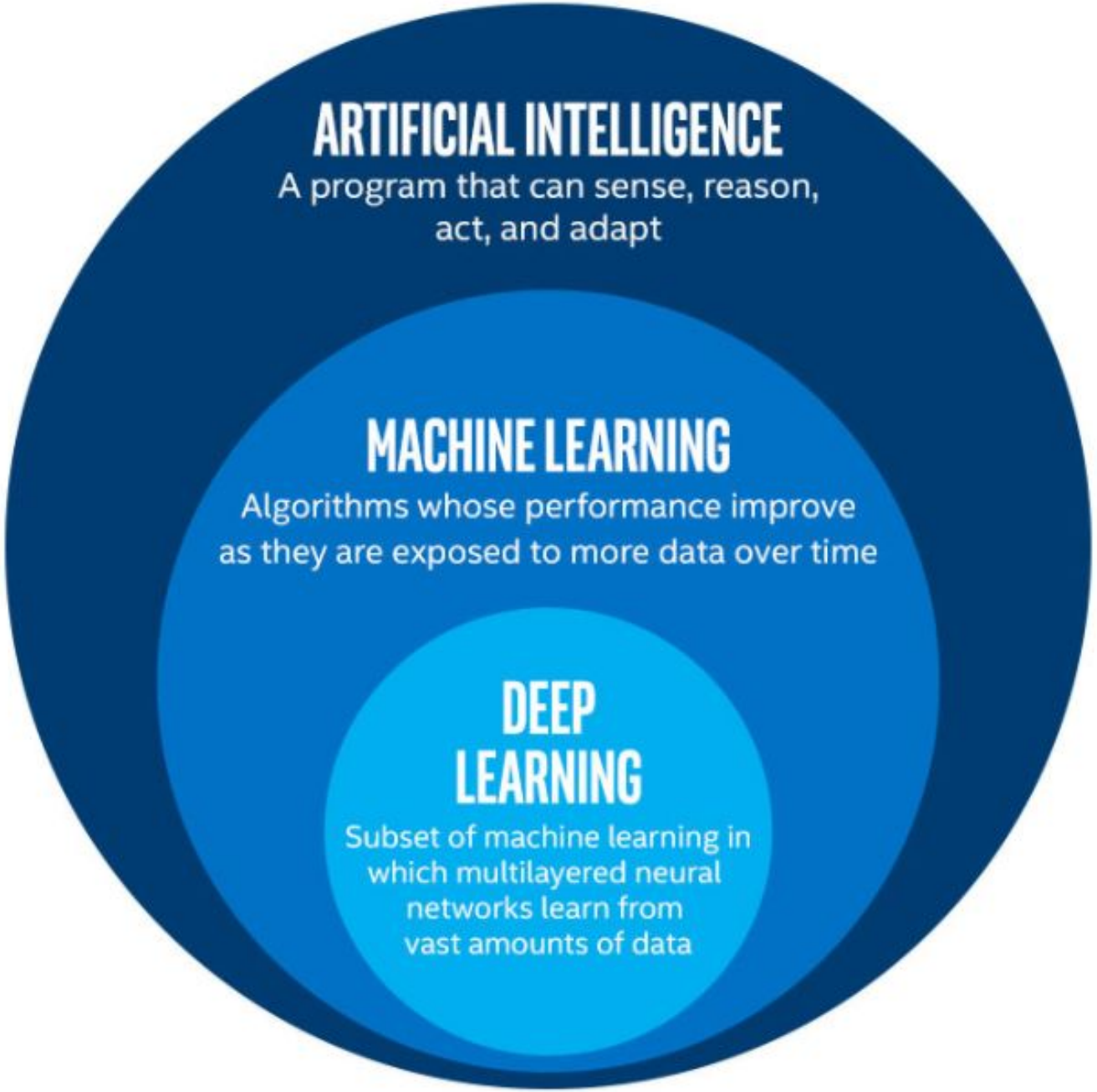
Statistics:

Concerned with prediction-making/probability.

Machine Learning

- **Machine learning** is an application of artificial intelligence that automates analytical model building by using algorithms that iteratively learn from data without being explicitly programmed.
- A system to ask questions and get answers.

Machine Learning



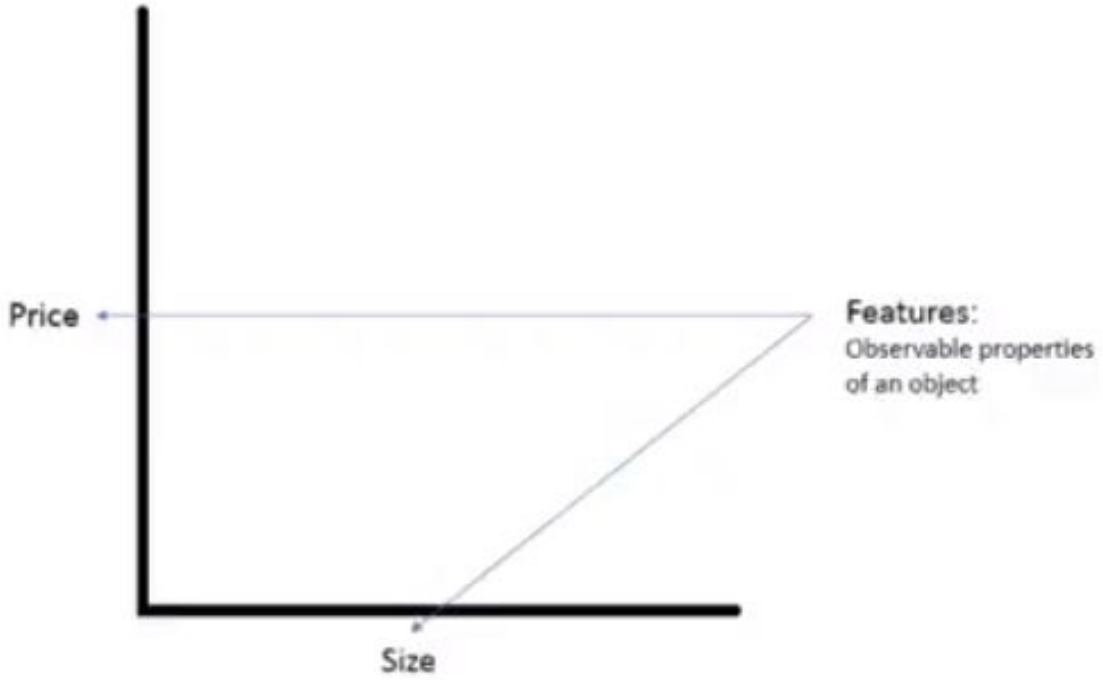
Machine Learning

- The study and construction of algorithms that can learn from and make predictions on data

Example

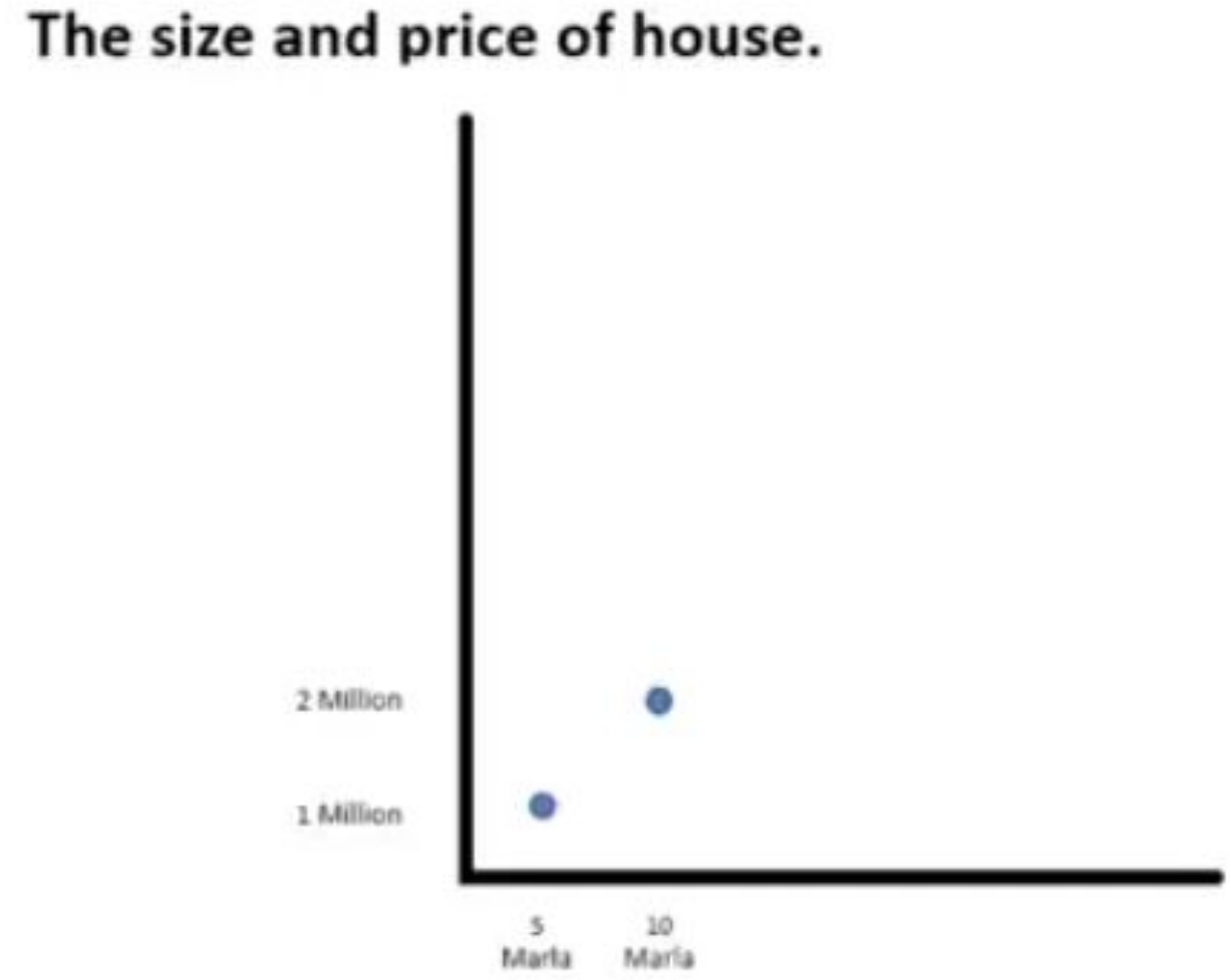
Example

The size and price of house.

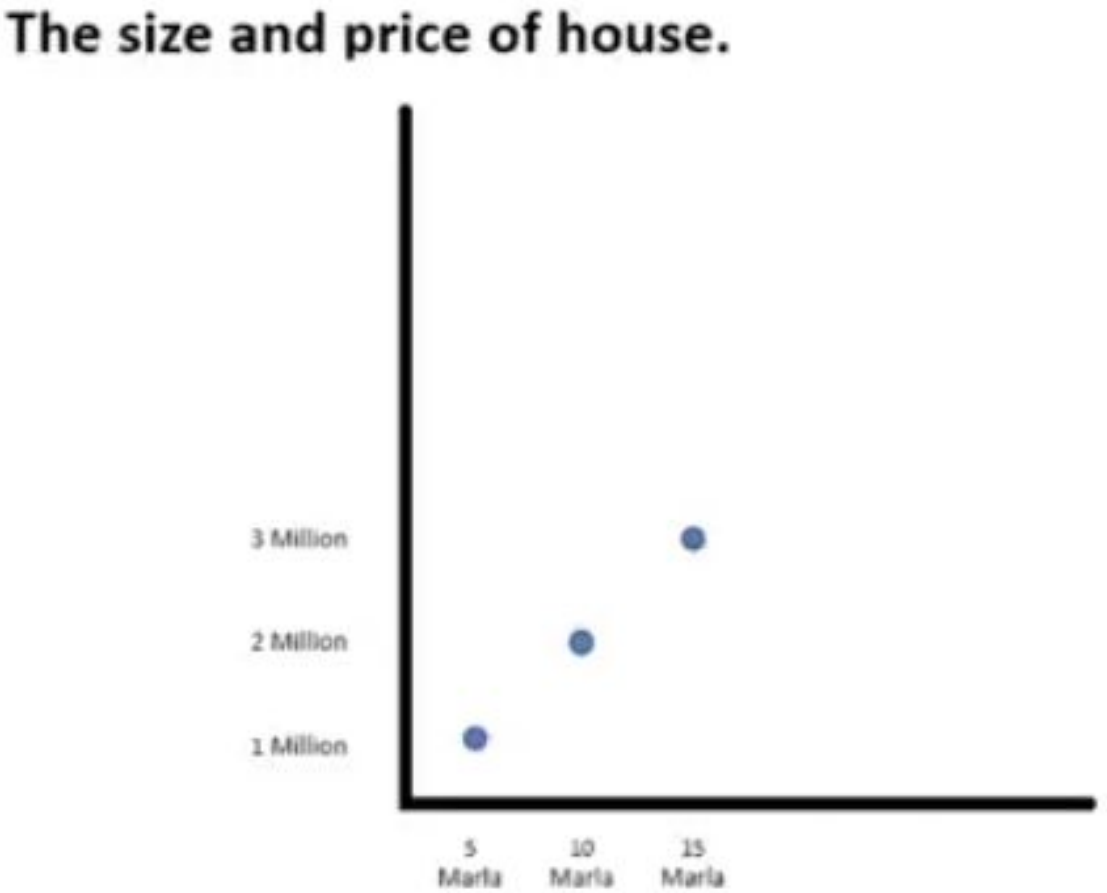


Example

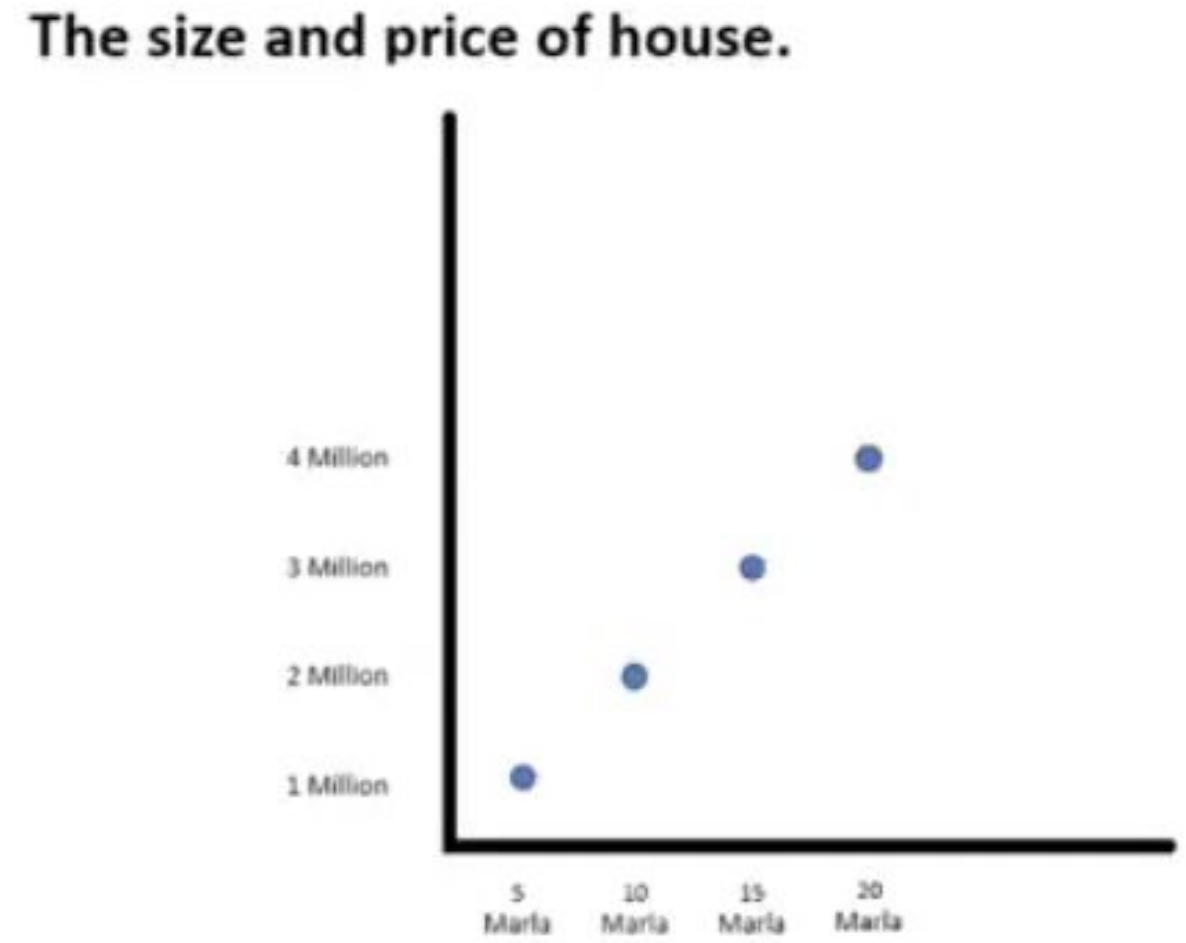
Example


The size and price of house.
2 Million
1 Million
5
Marla
10
Marla

Example


The size and price of house.
3 Million
2 Million
1 Million
5 Marla
10 Marla
15 Marla

Example


The size and price of house.
4 Million
3 Million
2 Million
1 Million
5 Marla
10 Marla
15 Marla
20 Marla

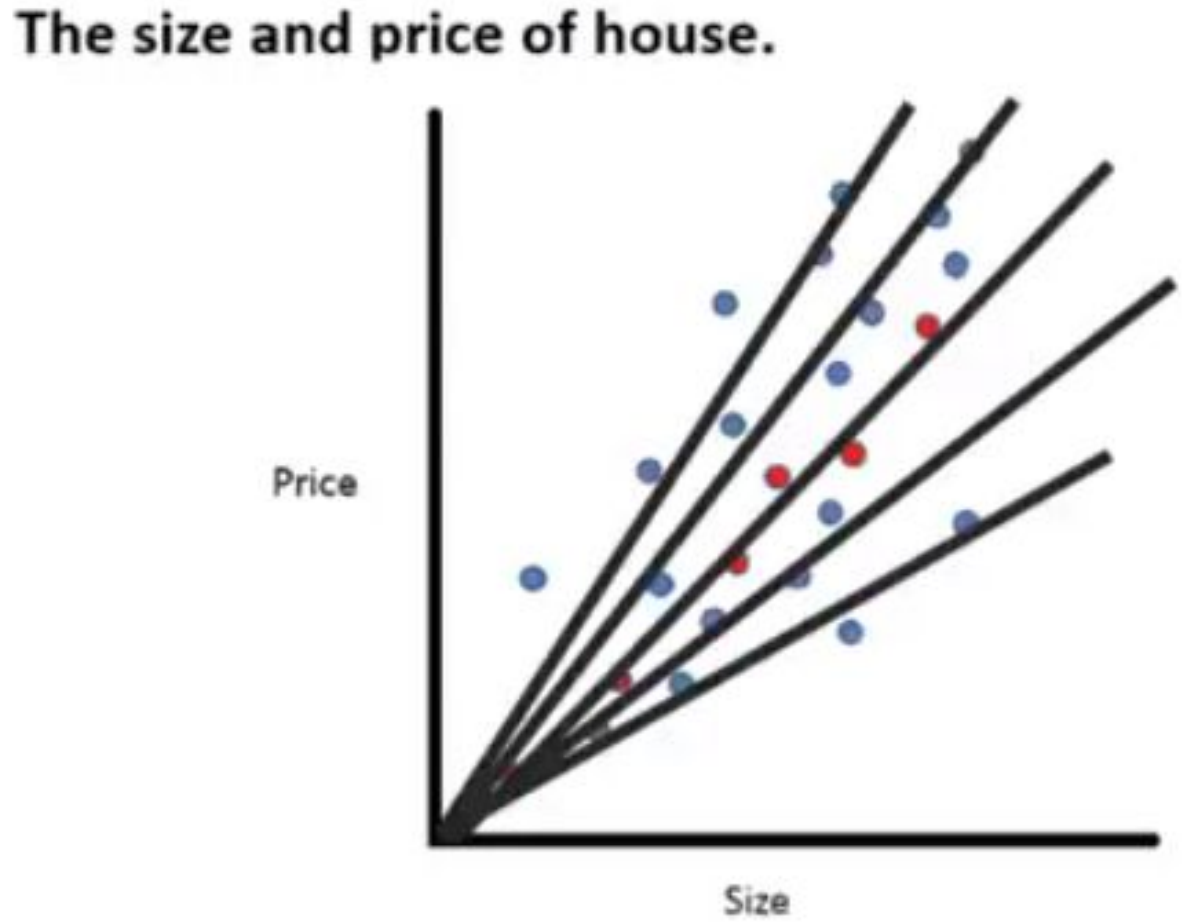
Example

Example

$$y = m * x + b$$

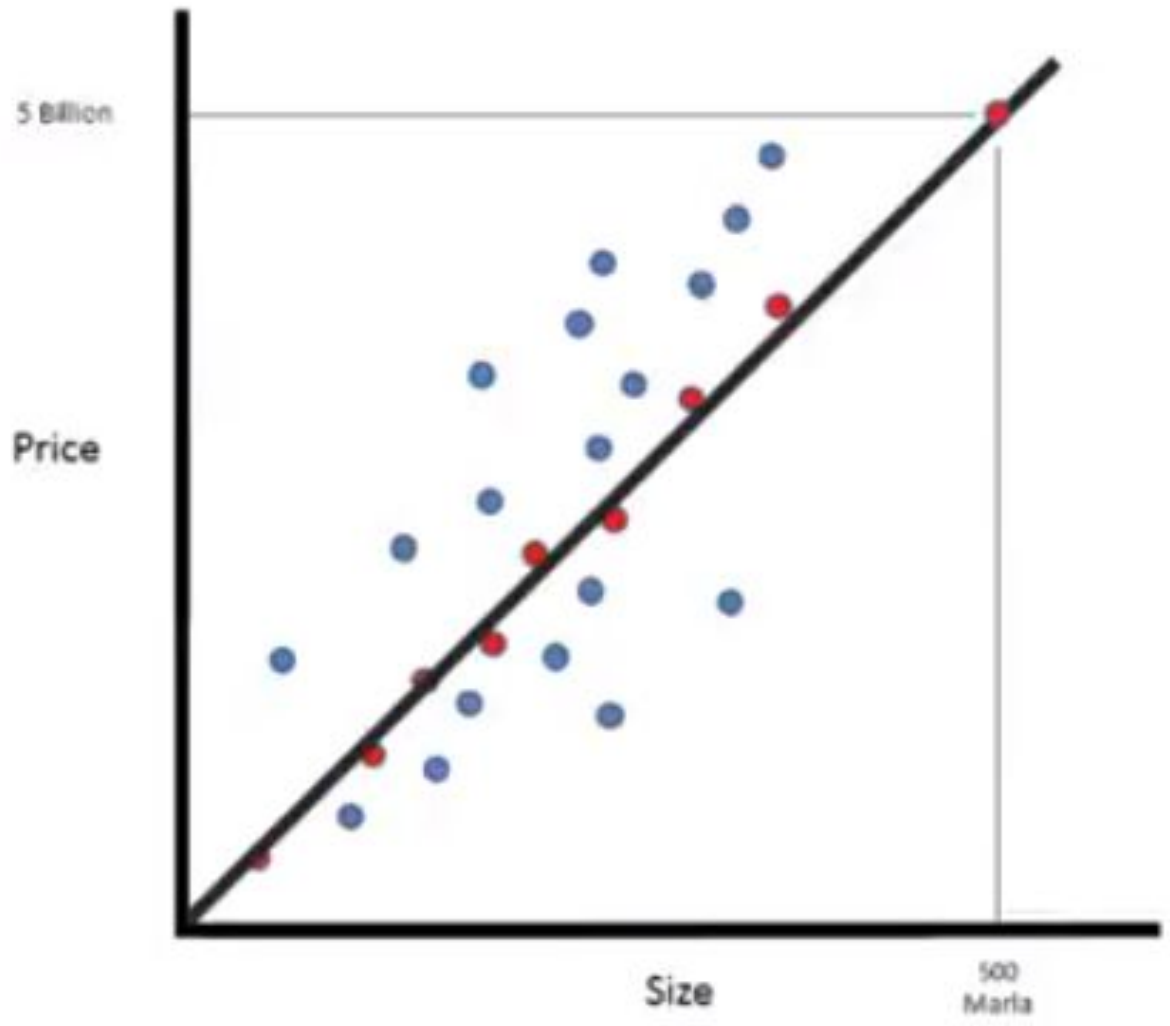
x = input
y = output
m = slope
x = y intercept

Example

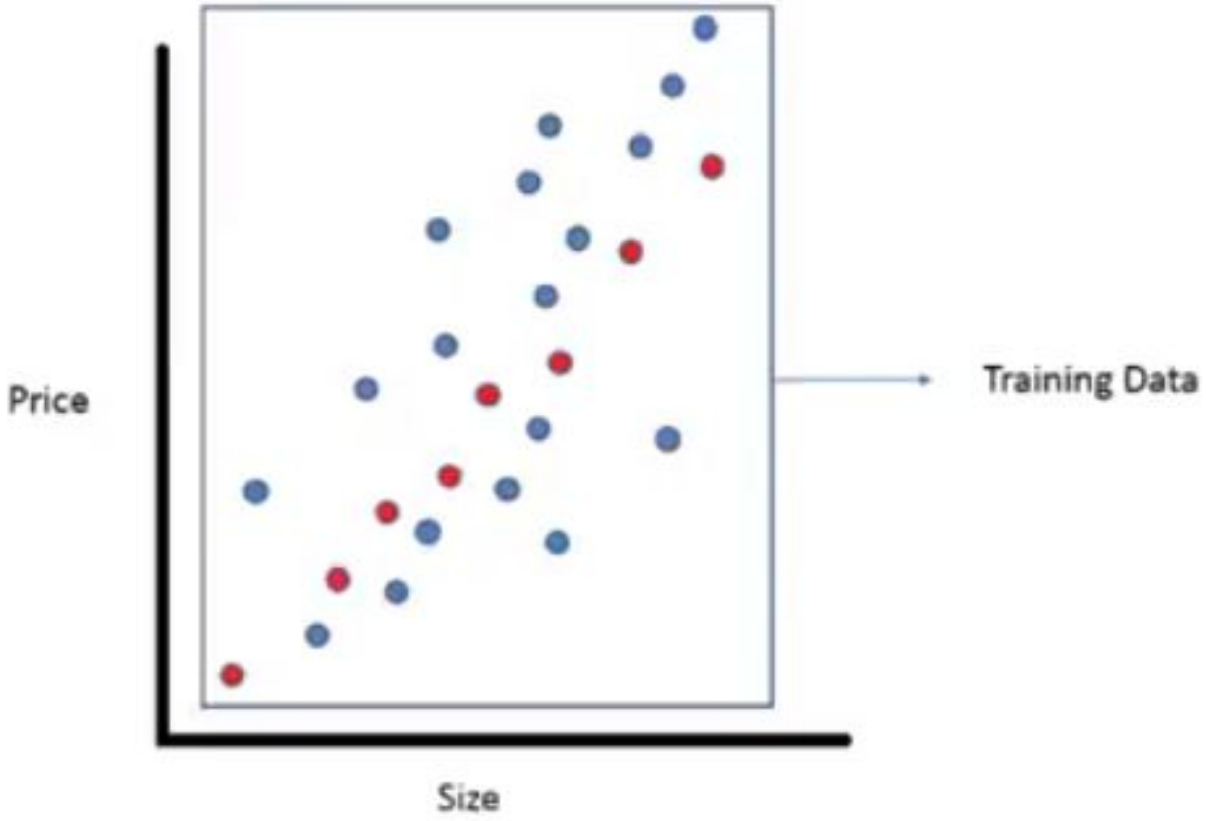


Prediction on best fit line

Example



Training data



Process of training and prediction or ML

Supervised versus Unsupervised

Supervised Learning

- Classification: Classifying labeled data
- Regression: Predicting trends using previous labeled data

Unsupervised Learning

- Clustering: Finding patterns and groupings from unlabeled data