Machine Learning

Muhammad Affan Alim

Learning

- The ability to improve the behavior based on experience

Example

- Human learning
- Machine learning

Machine Learning

By Tom Mitchell (1998)

- 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- Example (1/3)

- Suppose your email watches with emails you do or don't mark as spam, and based on that learns how to better filter spam.
- What is the task T in the setting

Machine Learning (ML) (2/3)

- Classifying emails as spam or not spam
- Watching your label emails as spam or not spam
- The number (or fraction) of emails correctly classified as spam/not spam
- This is not machine learning problem

Machine Learning – Example (3/3)

- Classifying emails as spam or not spam Task
- Watching your label emails as spam or not spam Experience
- The number (or fraction) of emails correctly classified as spam/not spam Performance
- This is not machine learning problem

Machine Learning- Example2 (1/2) (checkers)

- A computer program that learns to play checkers might improve its performance as measured by its ability to win at the class of tasks involving playing checkers games, through experience obtained by playing games against itself.

Machine Learning- Example2 (2/2) (checkers)

- A checkers learning problem:
  - Task T : playing checkers
  - Performance measure P: percent of games won against opponents
  - Training experience E: playing practice games against itself

Machine Learning- Example3 (1/1) (handwriting)

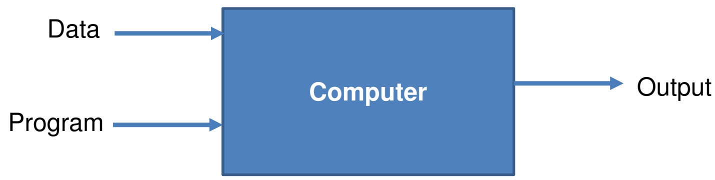
- A handwriting recognition learning problem:
  - Task T: recognizing and classifying handwritten words within images
  - Performance measure P: percent of words correctly classified
  - Training experience E: a database of handwritten words with given classifications

Machine Learning- Example4 (1/1) (robot)

- A robot driving learning problem:
  - Task T: driving on public four-lane highways using vision sensors
  - Performance measure P: average distance traveled before an error
  - Training experience E: a sequence of images and steering commands recorded while observing a human driver

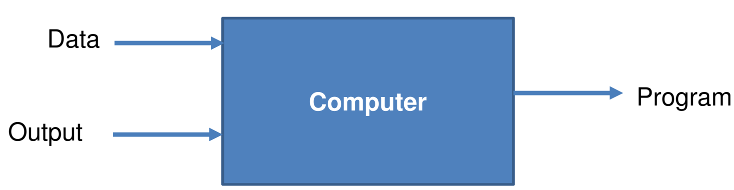
Traditional Programming vs ML (1/2)

- Traditional Programming



Traditional Programming vs ML (2/2)

- Machine Learning



Domain and application

- Computer vision

  (i) Say what objects appear in an image
  (ii) Convert handwritten digits to characters 0…9
  (iii) Detect where object appear in an image

Domain and application

- Robot control:

  - Design autonomous mobile robots that learn to navigate from their own experience

Domain and application

- NLP
  - Detect where entities are mentioned in NL
  - Detect what facts are expressed in NL
  - Detect if product/movies review is positive, negative or neutral (sentiment analysis)

  Speech recognition

  Machine translation

Difference b/w classification and regression ML

- Learning algorithm is called
  - classification problem when target function Y is discrete
  - regression problem when target function Y is continuous

- *Fundamentally, classification is about predicting a label and regression is about predicting a quantity*.

Classification Predictive Modeling

- Classification predictive modeling is the task of approximating a mapping function (f) from input variables (X) to discrete output variables (y).
- The output variables are often called labels or categories.
- The mapping function predicts the class or category for a given observation.

Classification Predictive Modeling - example

- an email of text can be classified as belonging to one of two classes: “spam“ and “not spam“.
- A classification can have real-valued or discrete input variables.
- A problem with two classes is often called a two-class or binary classification problem.
- A problem with more than two classes is often called a multi-class classification problem.

Regression Predictive model

- Regression predictive modeling is the task of approximating a mapping function (f) from input variables (X) to a continuous output variable (y).
- A continuous output variable is a real-value, such as an integer or floating point value.
- These are often quantities, such as amounts and sizes.

Regression Predictive model - example

- a house may be predicted to sell for a specific dollar value, perhaps in the range of $100,000 to $200,000.
- A regression problem requires the prediction of a quantity.
- A regression can have real valued or discrete input variables.
- A regression problem where input variables are ordered by time is called a time series forecasting problem.

Question

You're running a company, and you want to develop learning algorithms to address each of two problems.

Problem 1: You have a large inventory of identical items. You want to predict how many of these items will sell over the next 3 months.
Problem 2: You'd like software to examine individual customer accounts, and for each account decide if it has been hacked/compromised.

Should you treat these as classification or as regression problems?

- ◯ Treat both as classification problems.
- ◯ Treat problem 1 as a classification problem, problem 2 as a regression problem.
- ◯ Treat problem 1 as a regression problem, problem 2 as a classification problem.
- ◯ Treat both as regression problems.

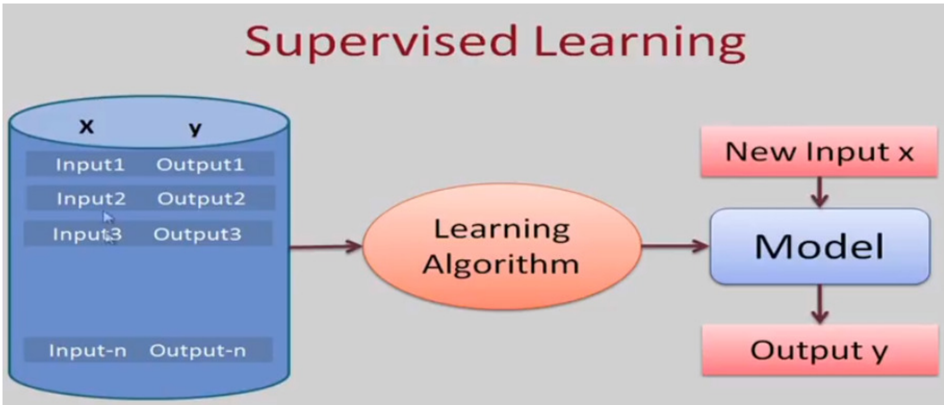
Types of learning algorithms

- There are several types of learning algorithms

  (i) Supervise learning
  (ii) Unsupervise learning
  (iii) Semi-supervise learning
  (iv) reinforcement learning

Supervised learning

- X, Y (pre-classified training examples)
- Given an observation X, what is the best label for Y

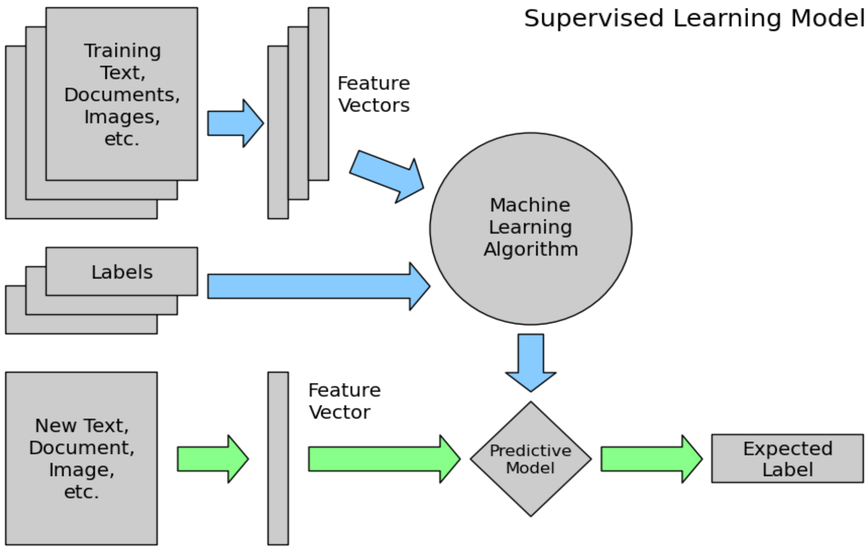


Supervised learning

- In supervise learning we use data that comprise of input and corresponding output. For every data instance we can have the input X and the corresponding output y.
- the machine learning system will build a model so that given new observation x_{new} will try to find out what is the corresponding y.
- this is called supervised learning because for labeled data

- Instances are described in term of features

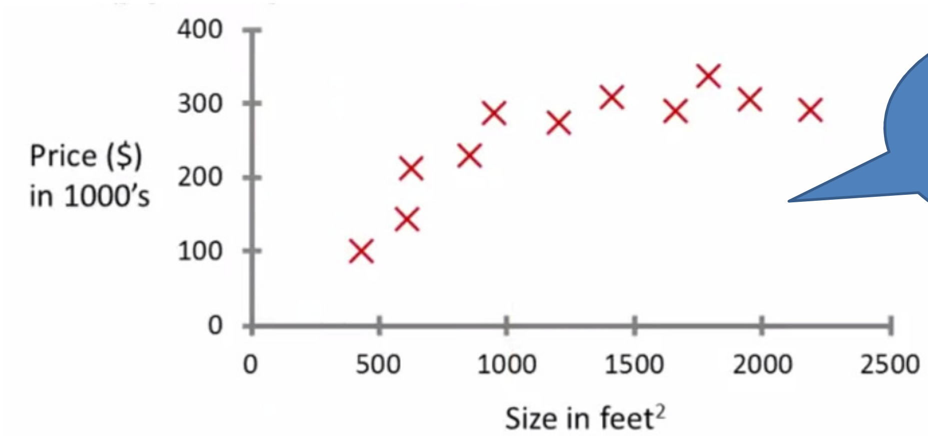
Supervised learning (classification example)

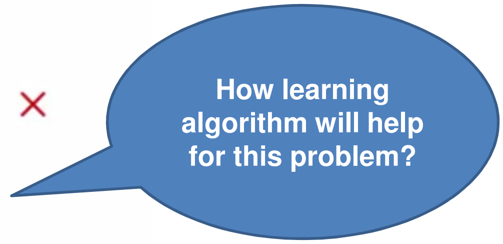


Supervised learning (classification example)

- *Supervised classification*: The algorithm learns to assign labels to types of webpages based on the labels that were inputted by a human during the training process

Supervised learning – House prediction

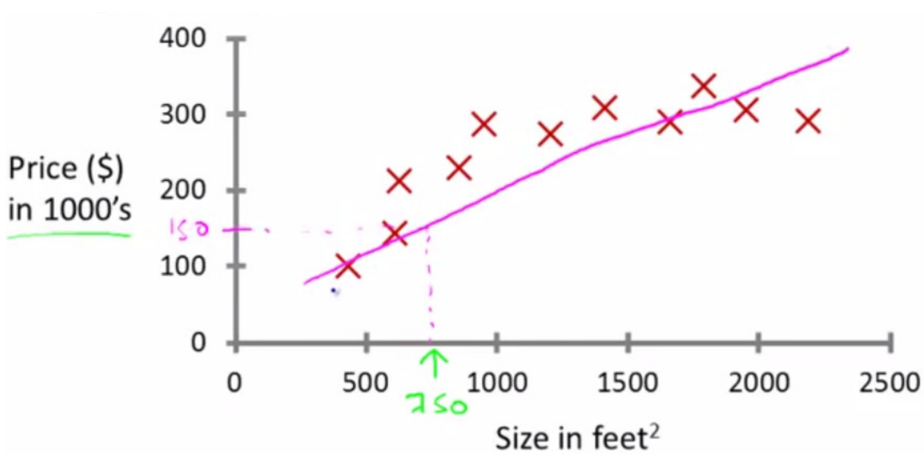




- Given data like this, how can we learn to predict the prices of other houses in Karachi, as a function of the size of their living areas?

Supervised learning – house prediction

-



Supervised learning – house prediction

- Suppose $x^{(i)}$ to denote the "*input*" variables, also called input features, and $y^{(i)}$ to denote the "*output*" or target variable that we are trying to predict

- A pair $(x^{(i)}, y^{(i)})$ is called a training example, and the dataset that we will be using to learn

- a list of m training examples $\{x^{(i)}, y^{(i)}; i = 1,...m\}$ is called a training set

Supervised learning – house prediction

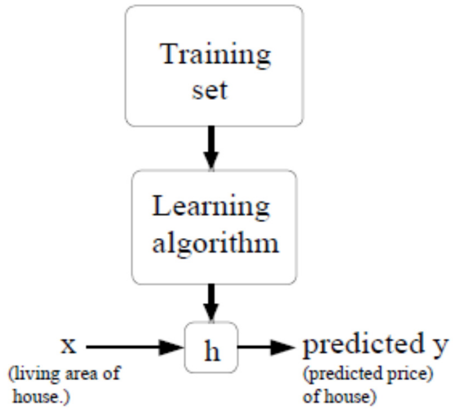
- We will also use X denote the space of input values, and Y the space the output values.

- In this example $X=Y=R$

- To describe the supervised learning problem slightly more formally, out goal is, given a training set, to learn a function $h: X \rightarrow Y$ so that $h(X)$ is a good predictor for corresponding value of y

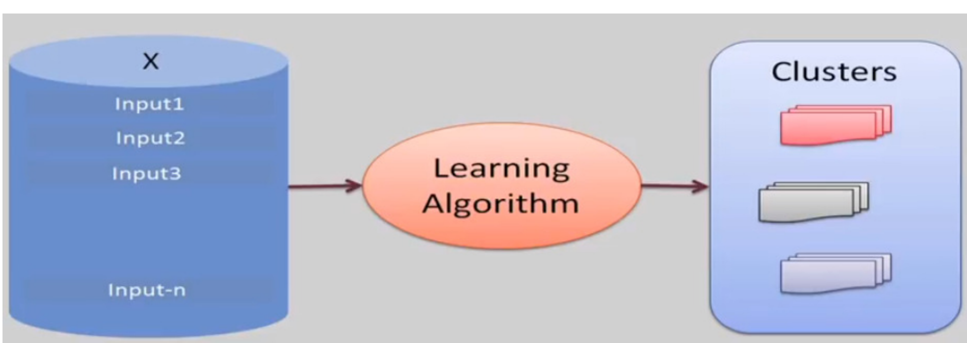
Supervised learning – house prediction

•



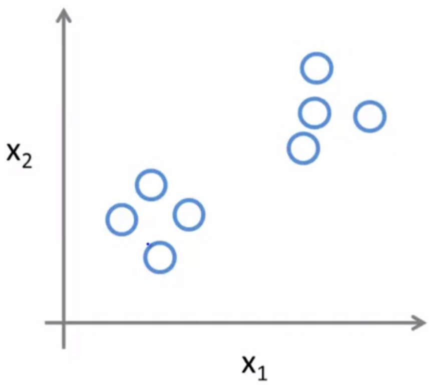
Unsupervised learning

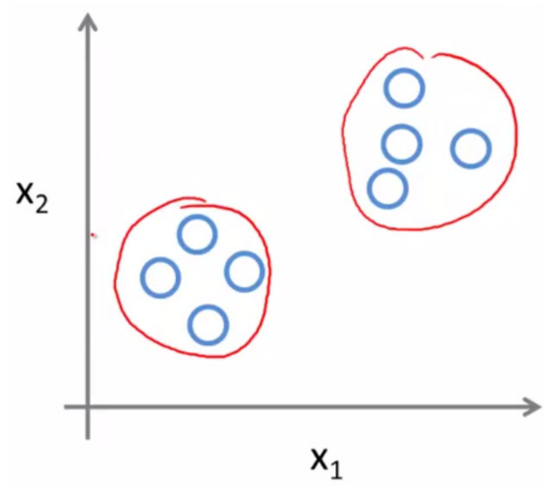
- In unsupervised learning you are only given X, there is no label to the data
- It is required to find some patterns in given data



Unsupervised learning

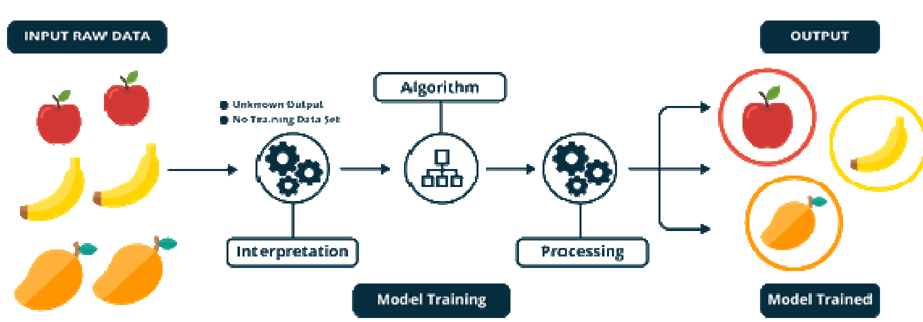
- V





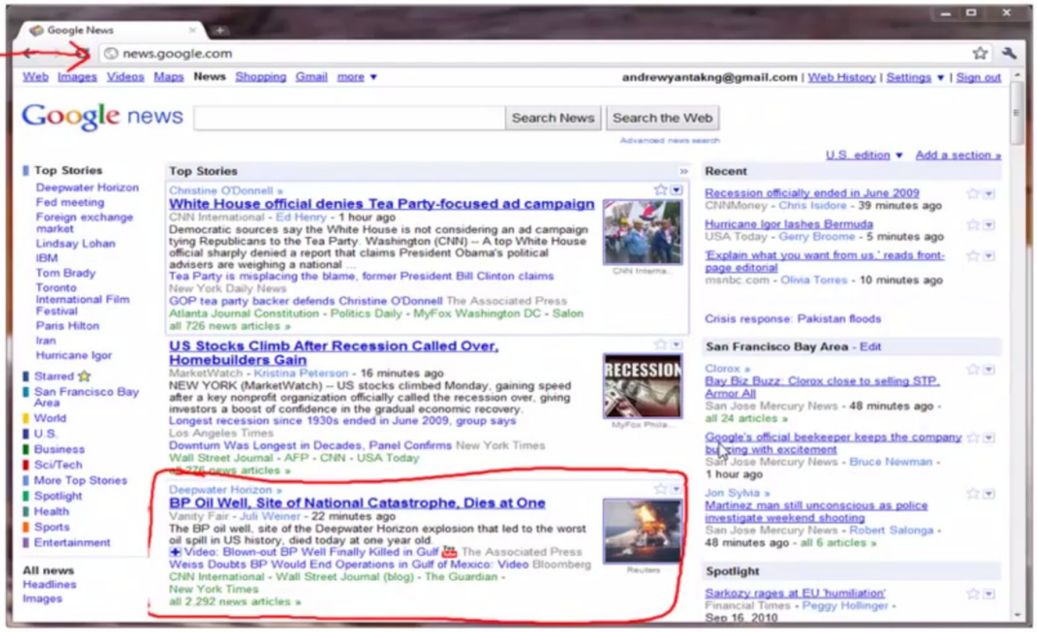
Unsupervised learning

- If somebody gives you a basket full of different fruits and asks you to separate them, you will probably do it based on their colors, shape and size, right?



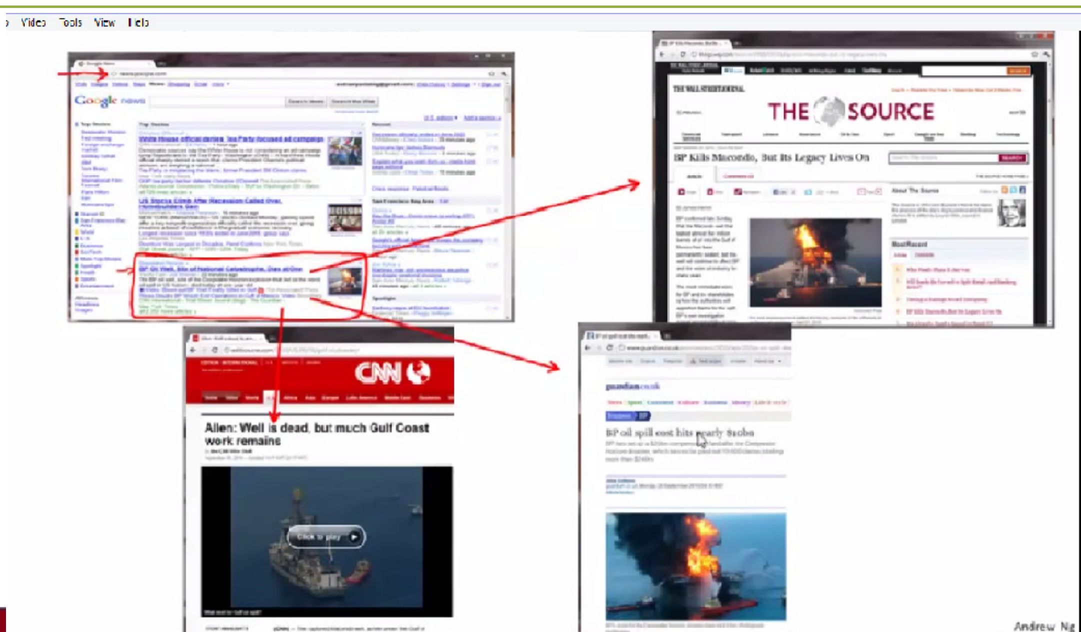
Unsupervised learning - example

•



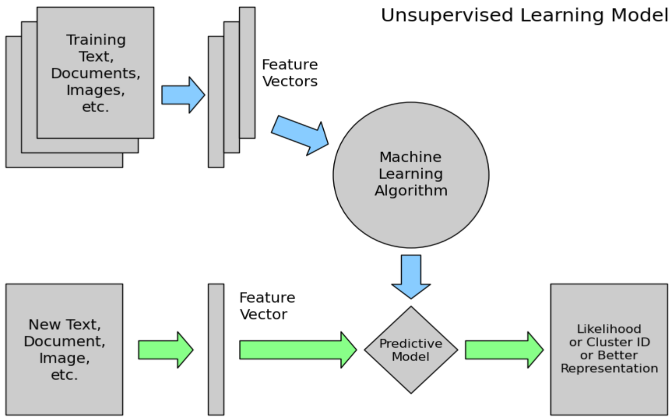
Unsupervised learning - example

•



Unsupervised learning - example

- *Unsupervised clustering*: The algorithm looks at inherent similarities between webpages to place them into groups



Unsupervised learning

- v