Machine Learning

By Pratinav Seth and Siddharth Singh

What is Machine Learning?

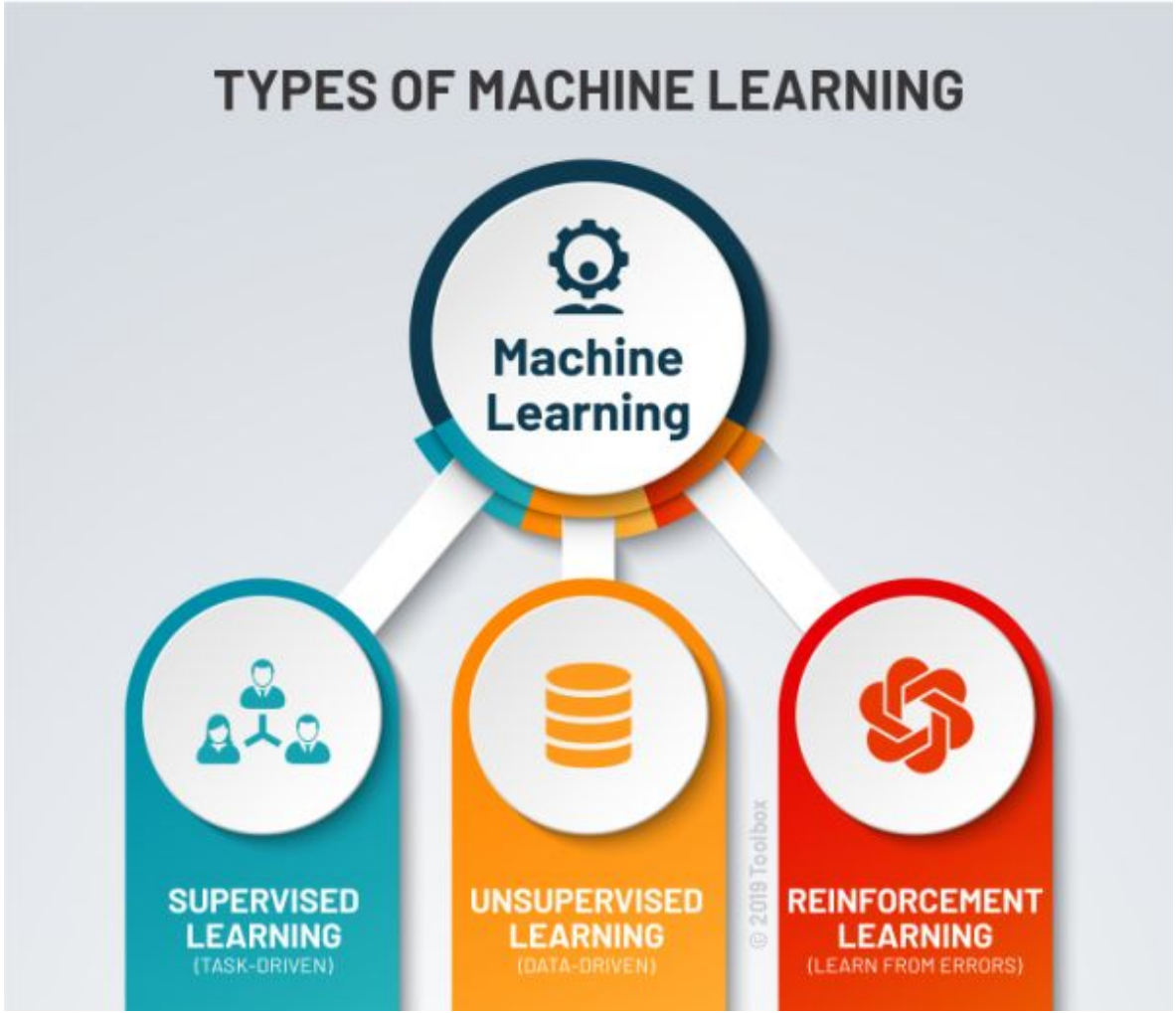
Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.


Rules
Data
Traditional programming
Answers


Answers
Data
Machine Learning
Rules

Different Types of Machine Learning

- **Supervised Learning**
- **Unsupervised Learning**
- **Reinforcement Learning**

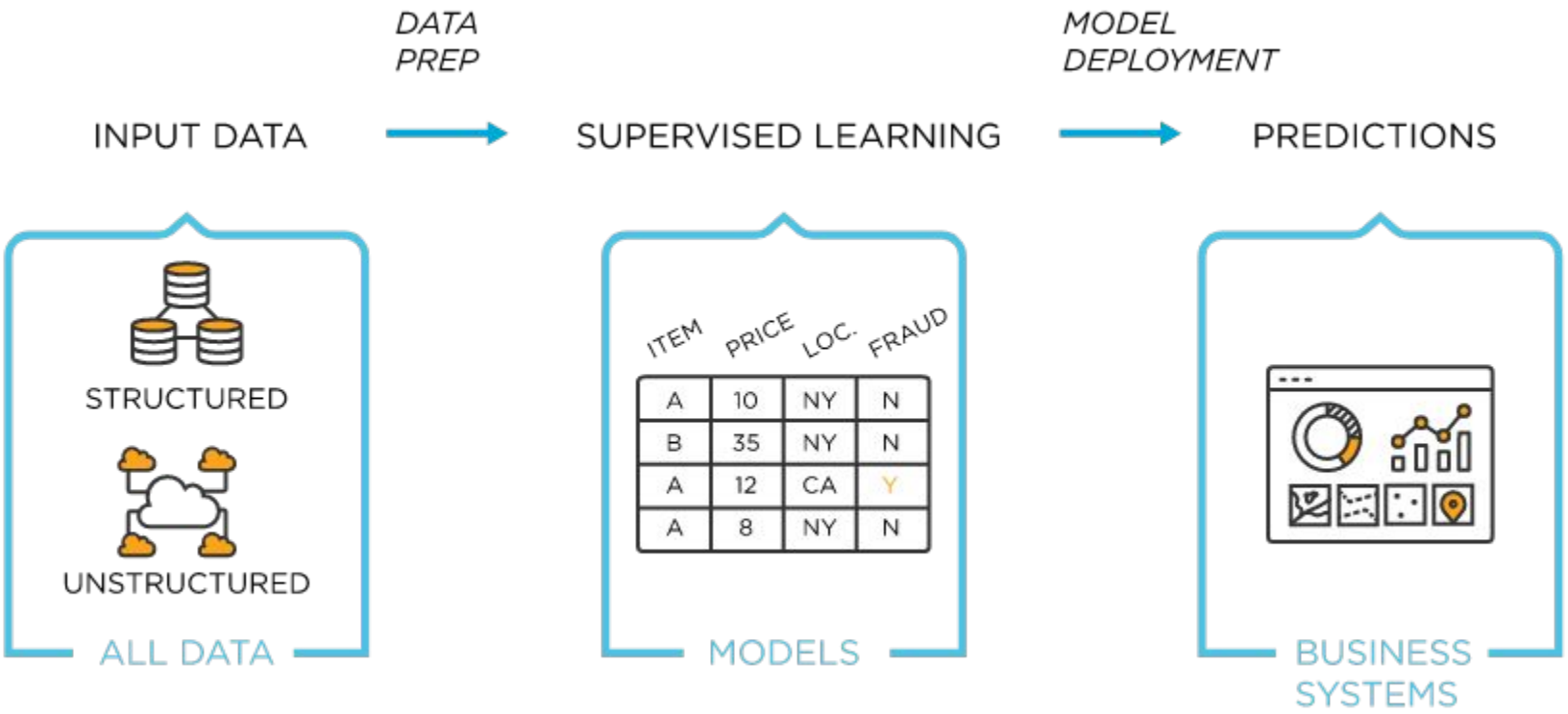

TYPES OF MACHINE LEARNING
Machine Learning
SUPERVISED LEARNING
(TASK-DRIVEN)
UNSUPERVISED LEARNING
(DATA-DRIVEN)
REINFORCEMENT LEARNING
(LEARN FROM ERRORS)
© 2019 Toolbox

Supervised Learning

- Supervised learning, as the name indicates, has the presence of a supervisor as a teacher. Basically supervised learning is when we teach or train the machine using data that is well labeled. Which means some data is already tagged with the correct answer.

- After that, the machine is provided with a new set of examples(data) so that the supervised learning algorithm analyses the training data(set of training examples) and produces a correct outcome from labeled data.

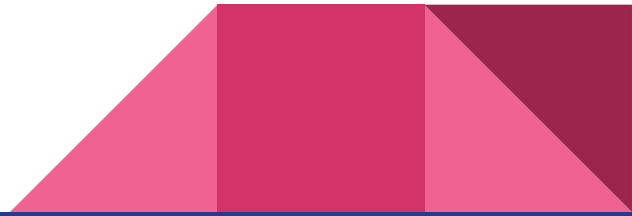
Representation of a Supervised Learning Program



Supervised Learning can be further divided into two categories:

- **Classification:**
  - A classification problem is when the output variable is a category, such as “Red” or “blue” or “disease” and “no disease”.

- **Regression:**
  - A regression problem is when the output variable is a real value, such as “dollars” or “weight”.



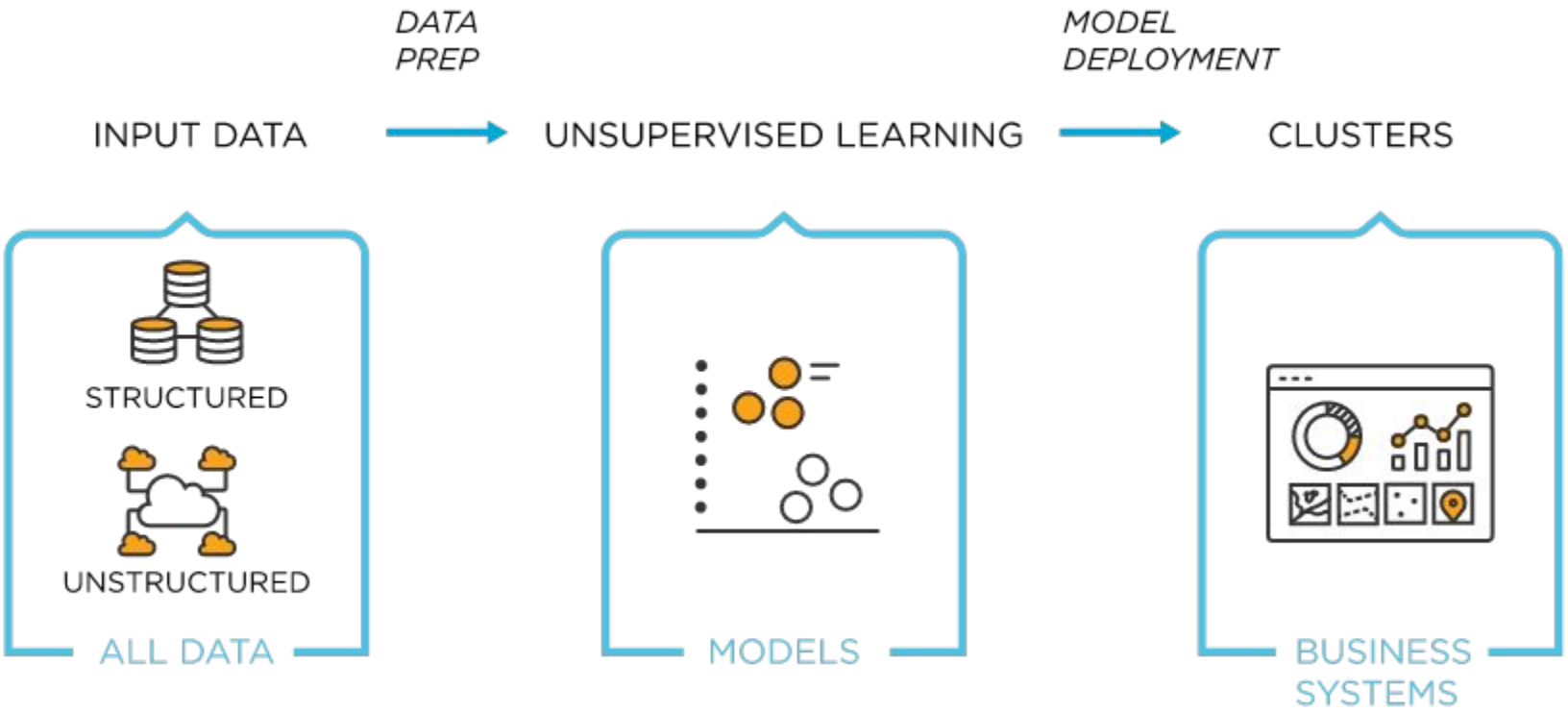
Unsupervised Learning

- Unsupervised learning is the training of a machine using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance.

- Here the task of the machine is to group unsorted information according to similarities, patterns, and differences without any prior training of data.

- Therefore the machine is restricted to find the hidden structure in unlabeled data by itself.

Representation of an Unsupervised Learning Pipeline



NOTE:

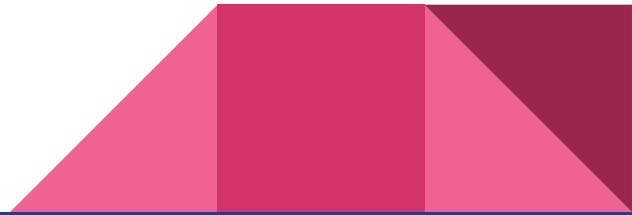
- **Structured data** is data that has been predefined and formatted to a set structure before being placed in data storage, which is often referred to as schema-on-write.

- **Unstructured data** is data stored in its native format and not processed until it is used, which is known as schema-on-read. It comes in a myriad of file formats, including email, social media posts, presentations, chats, IoT sensor data, and satellite imagery.

Unsupervised Learning can be further divided into categories:

- **Clustering:**
  - A clustering problem is where you want to discover the inherent groupings in the data, such as grouping customers by purchasing behavior.

- **Association:**
  - An association rule learning problem is where you want to discover rules that describe large portions of your data, such as people that buy X also tend to buy Y.



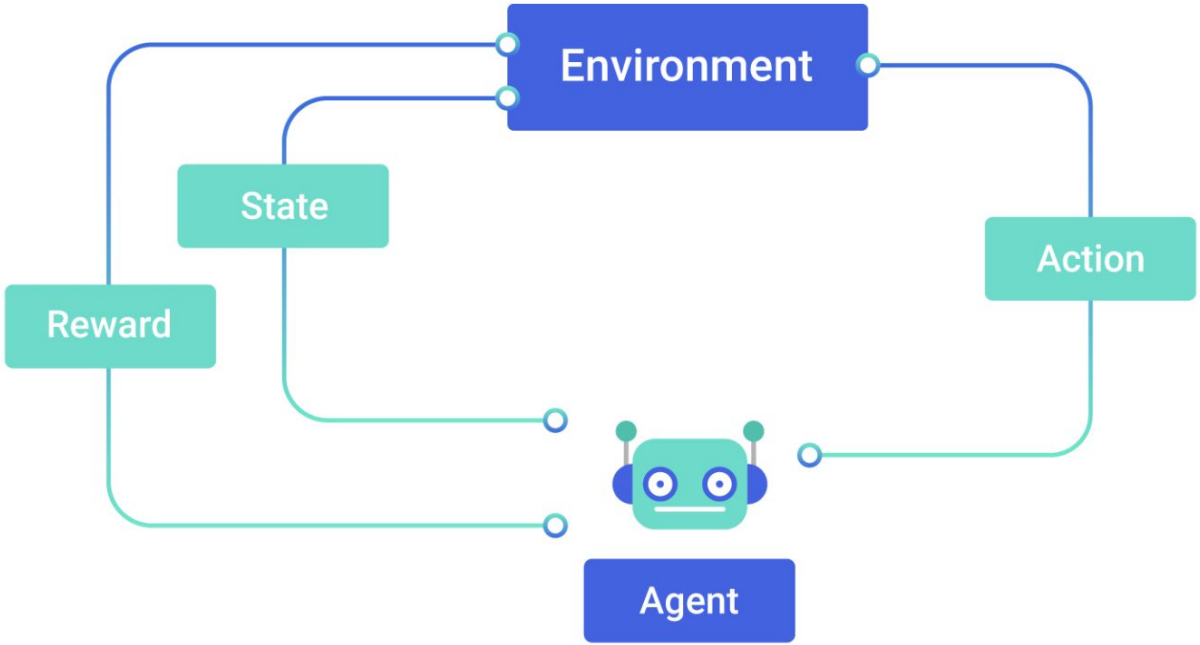
Reinforcement Learning

- It is a learning method that interacts with its environment by producing actions and discovers errors or rewards.

- Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning.

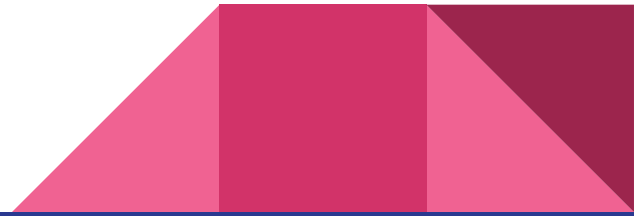
- This method allows machines and software agents to automatically determine the ideal behavior within a specific context in order to maximize its performance.

Representation of Reinforcement Learning

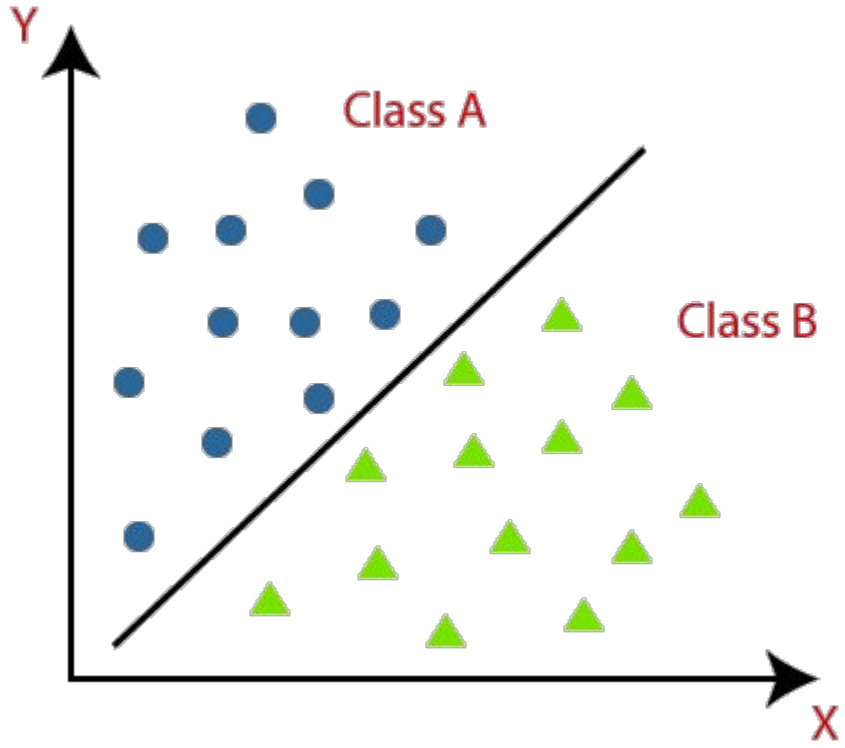


What is Classification?

Classification is a task that requires the use of machine learning algorithms that learn how to assign a class label to examples from the problem domain. An easy to understand example is classifying emails as “spam” or “not spam.”



Example of Classification

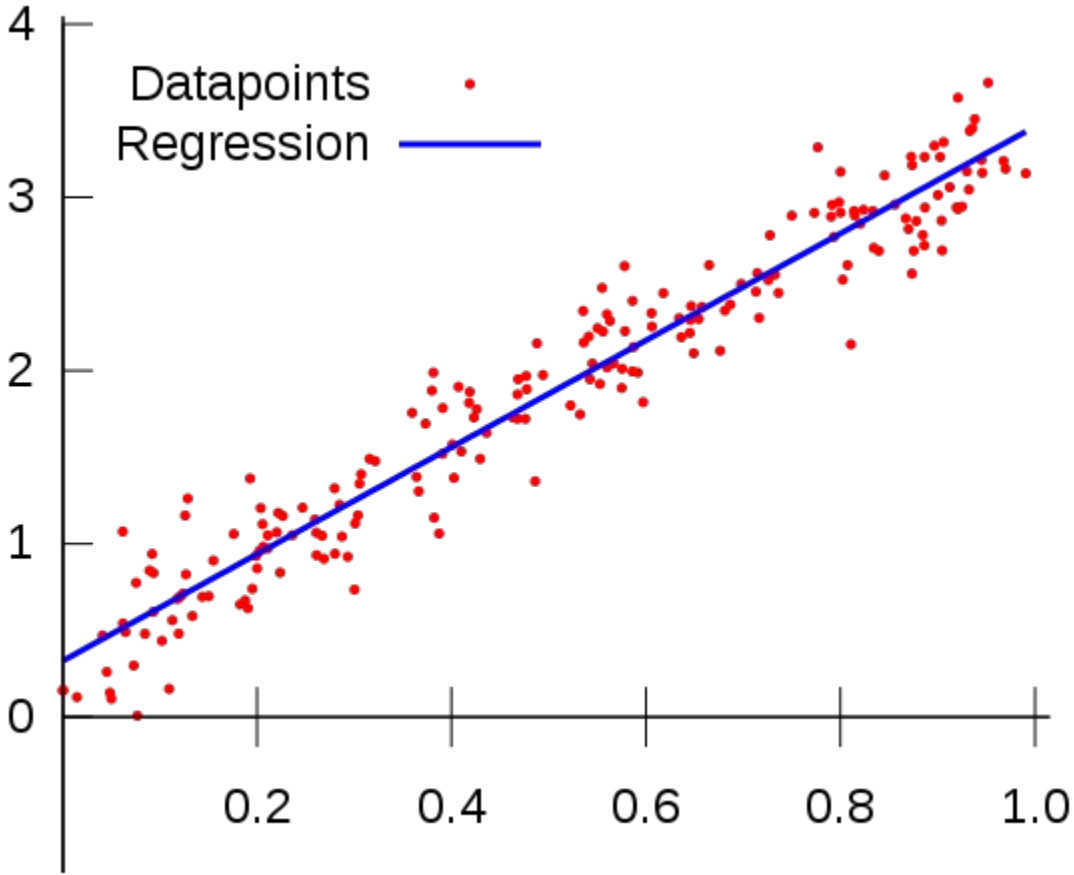


What is Regression?

- Regression analysis is a statistical method to model the relationship between a dependent (target) and independent (predictor) variables with one or more independent variables.

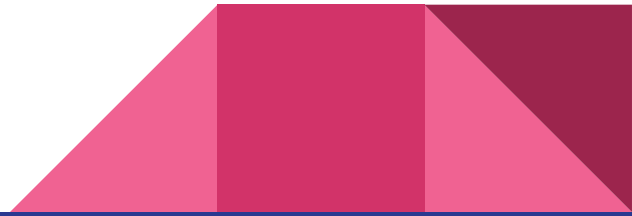
- More specifically, Regression analysis helps us to understand how the value of the dependent variable is changing corresponding to an independent variable when other independent variables are held fixed. It predicts continuous/real values such as temperature, age, salary, price, etc.

Example of Linear Regression



What is Regression?

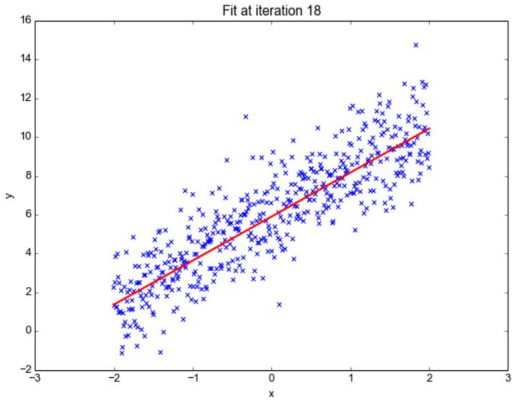
- Regression refers to method for modelling relationship between one or more independent variable and one dependent variable.
- In fields pertaining to machine learning its basically a area where one wants to predict a continuous numerical value based on features of a given data.
- Its considered a supervised ml problem where we have both the features and labels in the dataset through which we train model.



Linear Regression

- This is a type of regression where we predict a dependent variable y which is based on independent variables(also called features) x.
- In linear regression case y has a linear relationship with x
- Here we can model hypotheses as-

$$y = w.x+b$$



- So its similar to equation of line y=mx+c where m-slope,c-intercept

Linear Regression

- Our model hypotheses is- $y = w.x+b$
- Where we define w as weights and b as bias
- So what we do here is to predict/model a curve similar to features /data given.
- Thereby for data X having 9 different features like

	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	population	households	median_income	ocean_proximity
0	-122.23	37.88	41.0	880.0	129.0	322.0	126.0	8.3252	3
1	-122.22	37.86	21.0	7099.0	1106.0	2401.0	1138.0	8.3014	3
2	-122.24	37.85	52.0	1467.0	190.0	496.0	177.0	7.2574	3
3	-122.25	37.85	52.0	1274.0	235.0	558.0	219.0	5.6431	3

- We consider w to be a vector having dimensions (1,9) such that X has dimensions (9,4)
- Then $W^T.X$ has shape (1,4) which is equal to y which should be of shape (1,4).

Linear Regression

- Thus we can define it as $y = W^T.X + b$
- So now we see how it works -
- We initialize w and b randomly between 0 and 1
- We calculate y using the equation denoted as pred.
- In order to see how wrong we were we calculate the cost/loss function

$$\frac{1}{n}\sum_{i=1}^{n}(pred_i - y_i)^2$$

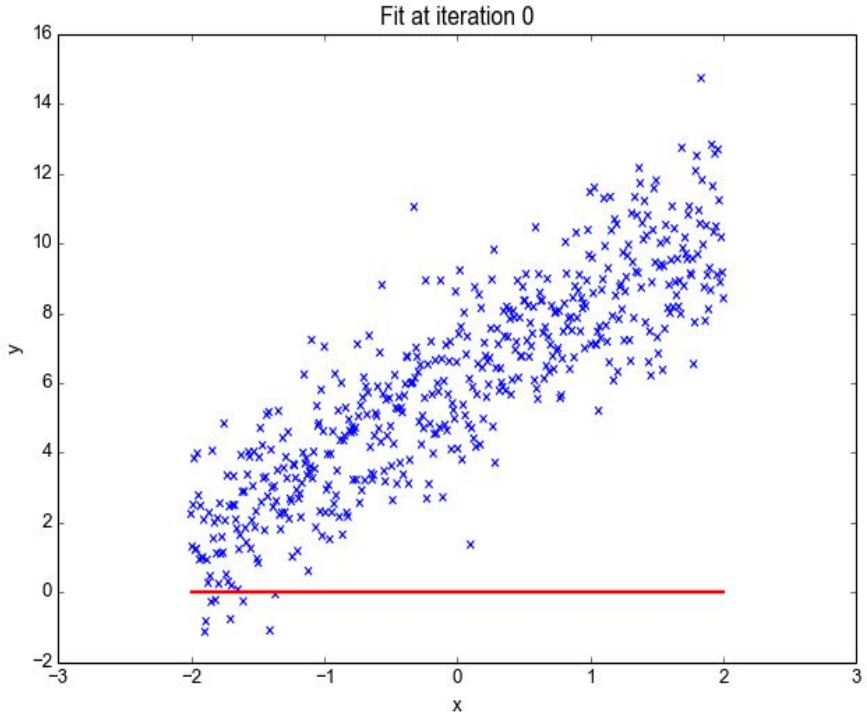
- Here pred refers to prediction made by us/model
- y refers to actual value or true value
- Now to minimize this error we update weights using gradient descent

Linear Regression

- Gradient Descent works by updating weights based on error by a value called learning rate.
- Like if prediction >original value we reduce weights by some value and if prediction < original value we increase weights by a particular value

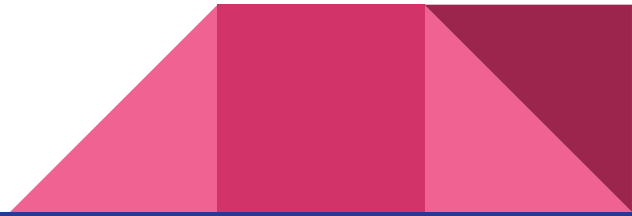
$$b = b - \alpha \cdot \frac{2}{n} \sum_{i=1}^{n} (pred_i - y_i) \qquad W = W - \alpha \cdot \frac{2}{n} \sum_{i=1}^{n} (pred_i - y_i) \cdot x_i$$

Linear Regression



Logistic Regression

- It is used for modelling relationship between features and a categorical variable
- It's used for the classification-based tasks where want to classify data based on the given data features
- Like consider we want to decide whether we should give loan to any person
- So what like these loan predictors will do they take data features from you calculate your credit score and give you a probability of you getting loan or simple yes or no
- This is simply how logistic regression works



Logistic Regression

- From logistic regression , we have –

$$y = W^T.X + b$$

- Now we want to convert it to probability (ie between 0 and 1)
- We do that by using logistic function-

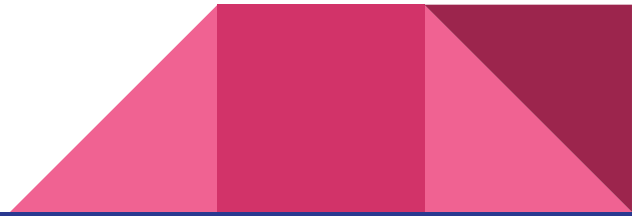
$$g(x) = \frac{1}{1 + e^{-x}}$$

- So using logistic function we convert this prediction y into probability

$$h = g(z) = \frac{1}{1 + e^{-z}} \qquad z = W^T.x + b$$

Logistic Regression

- But in classification we want a yes/no basically binary 1/0.
- This is where comes decision boundary , its basically a threshold for which prediction is considered to be true.
- If y>0.5 z=1 and If y<=0.5 z=0
- Generally 0.5 is used as threshold but can be changed .
- One of the reasons why we use logistic function is because it turns prediction into a s-shaped curve on which we can easily have a decision boundary for classification.



Logistic Regression

Logistic Regression

- To check how our model is performing we use log loss

$$-(y \log(p) + (1 - y) \log(1 - p))$$

- Here p= prediction of model and y is the actual prediction
- It works in a similar way as linear regression using gradient descent
- We use the log-loss for measuring how our model is doing and then update it weights and bias based on performances of model.

Logistic Regression