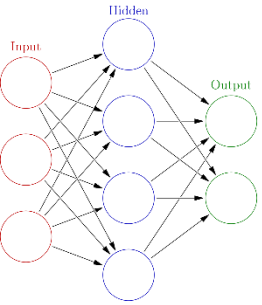
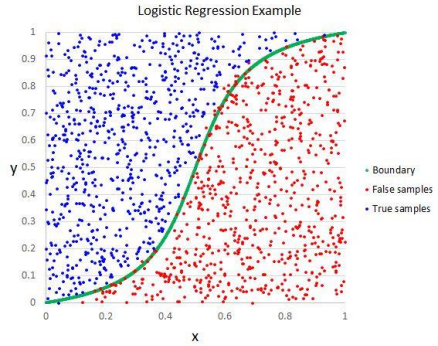
Deep Learning

Logistic Regression, one of the basic and popular algorithm to solve a classification problem

Dr. Sarwan Singh

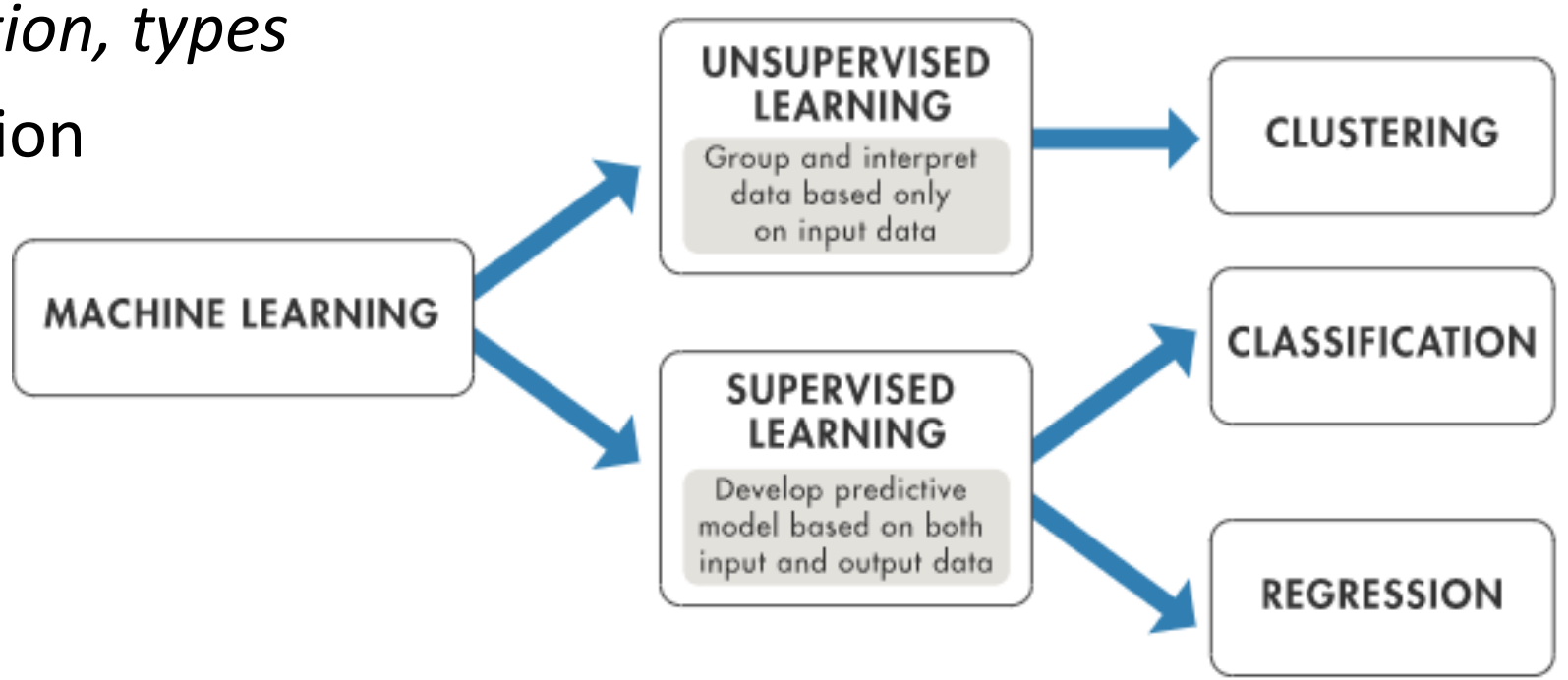
NIELIT Chandigarh

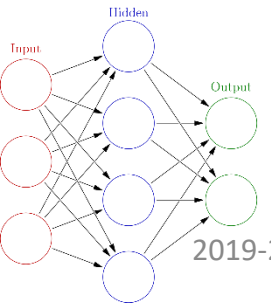




agenda

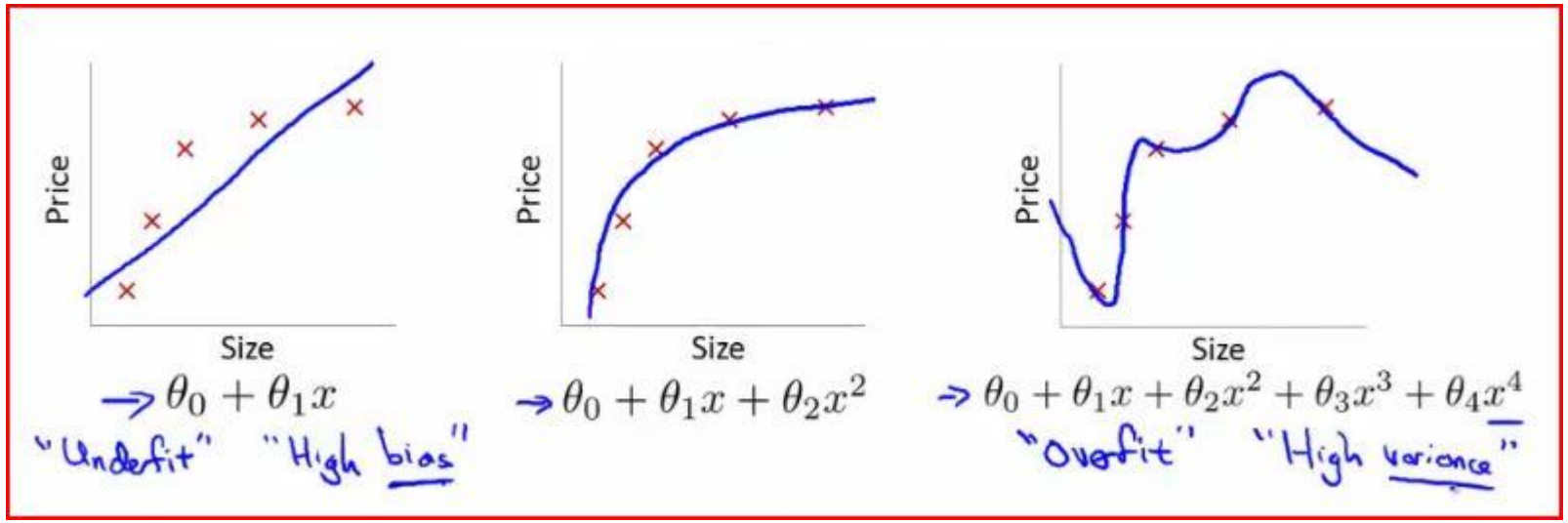
- classification – *introduction, types*
- Regression vs Classification
- Sigmoid function
- Decision boundary
- Making predictions
- Confusion Matrix

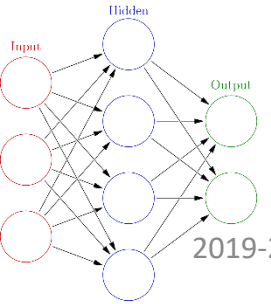




References

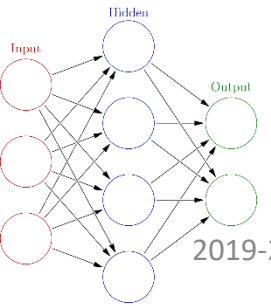
- Medium.com - Blockchain
- neuralnetworksanddeeplearning.com
- Fundamentals of Deep Learning - *Designing Next-Generation Machine Intelligence Algorithms* … Nikhil B
- Websites: Wikipedia, javapoint, ml-cheatsheet, realpython





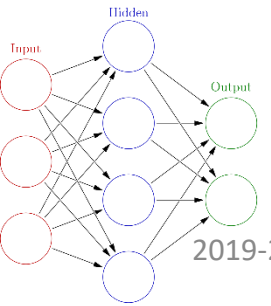
Classification

- Supervised machine learning algorithms define models that capture relationships among data.
- **Classification** is an area of supervised machine learning that tries to predict which class or category some entity belongs to, based on its features.
- For example, you might analyze the employees of some company and try to establish a dependence on the **features** or **variables**, such as the level of education, number of years in a current position, age, salary, odds for being promoted, and so on. The set of data related to a single employee is one **observation**. The features or variables can take one of two forms:
  - **Independent variables -** level of education, time in a current position, and age
  - **Dependent variables –** salary, promotion



Regression vs Classification

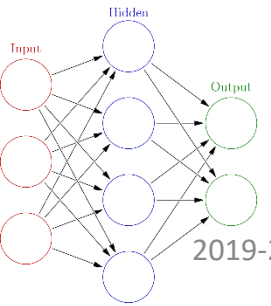
- **Regression** problems have continuous and usually unbounded outputs. An example is when you're estimating the salary as a function of experience and education level.
- On the other hand, **classification** problems have discrete and finite outputs called **classes** or **categories**. For example, predicting if an employee is going to be promoted or not (true or false) is a classification problem.



Types of classification

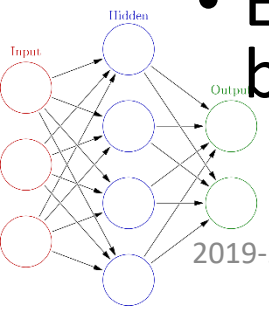
- **Binary** or **binomial classification:** exactly two classes to choose between (usually 0 and 1, true and false, or positive and negative)
- **Multiclass** or **multinomial classification:** three or more classes of the outputs to choose from

Applications - spam emails, image recognition/classification



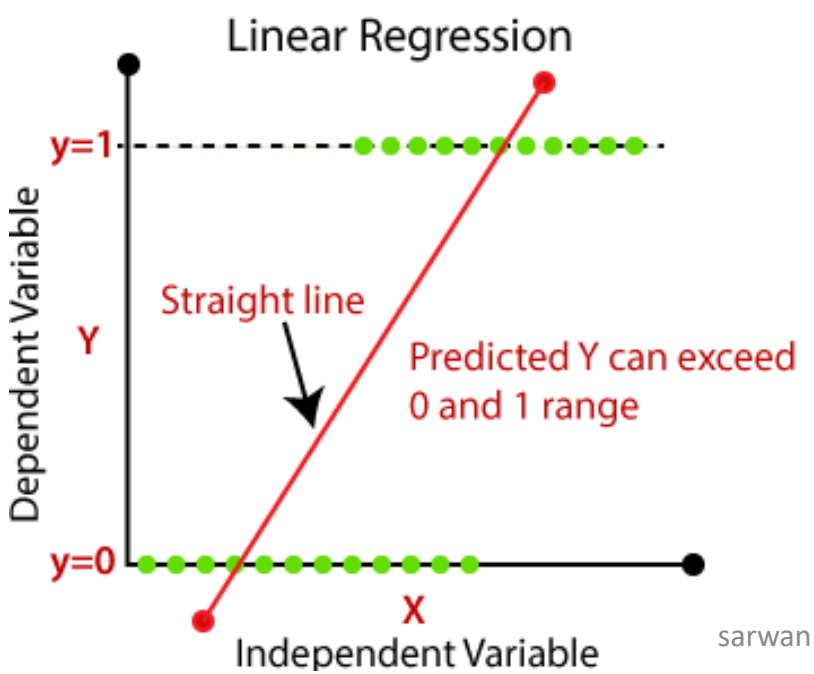
Logistic Regression….*overview*

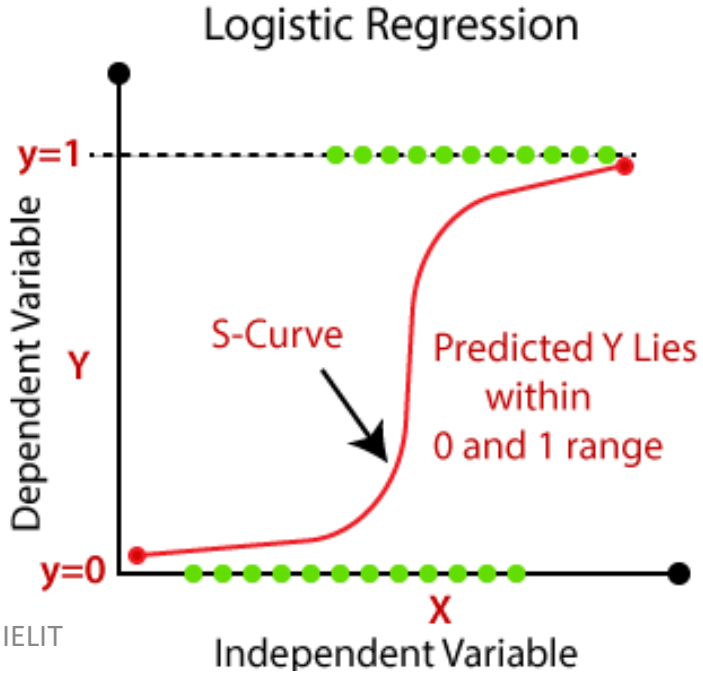
- Logistic regression is a fundamental classification technique. It belongs to the group of linear classifiers and is somewhat similar to polynomial and linear regression.
- Logistic regression is fast and relatively uncomplicated, and it's convenient for you to interpret the result
- *Def* … In statistics, *the logistic model (or logit model) is used to model the probability of a certain class or event existing such as pass/fail, win/lose, alive/dead or healthy/sick*.
- This can be extended to model several classes of events such as determining whether an image contains a cat, dog, lion, etc.
- Each object being detected in the image would be assigned a probability between 0 and 1, with a sum of one.

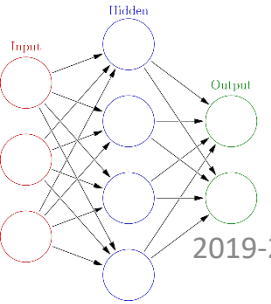


Linear vs Logistic Regression

- Linear Regression and Logistic Regression are the two famous Machine Learning Algorithms which come under supervised learning technique.
- main difference between them is how they are being used
- Linear Regression is used for solving Regression problems whereas Logistic Regression is used for solving the Classification problems.

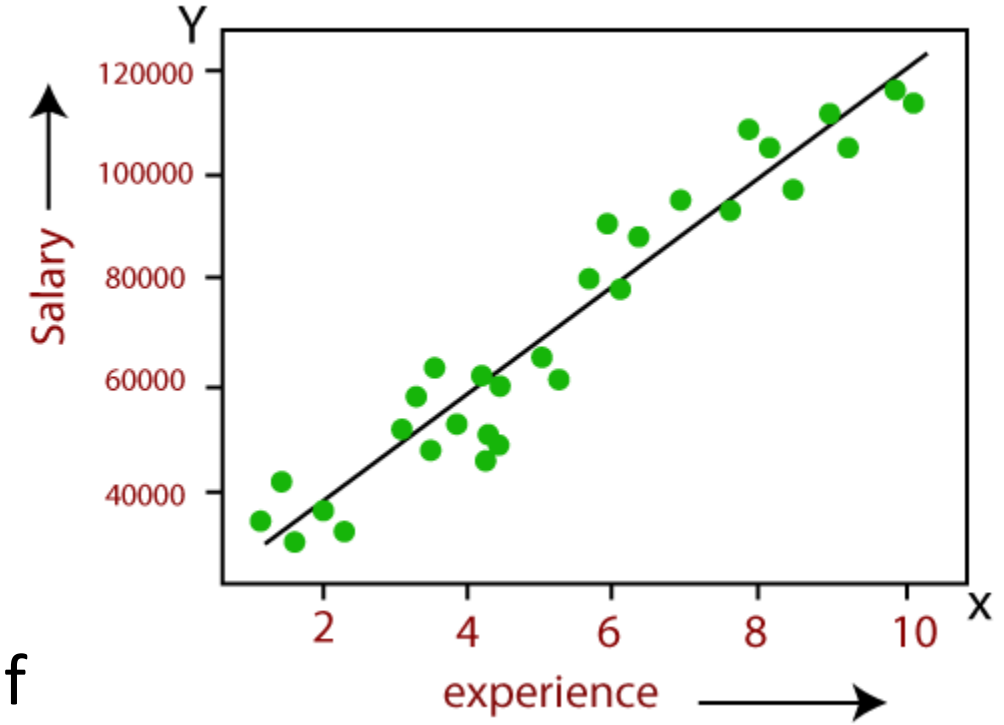


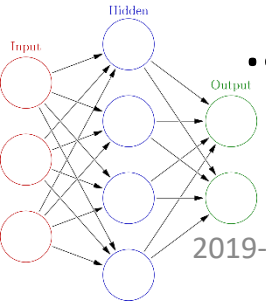




Linear Regression … *summary*

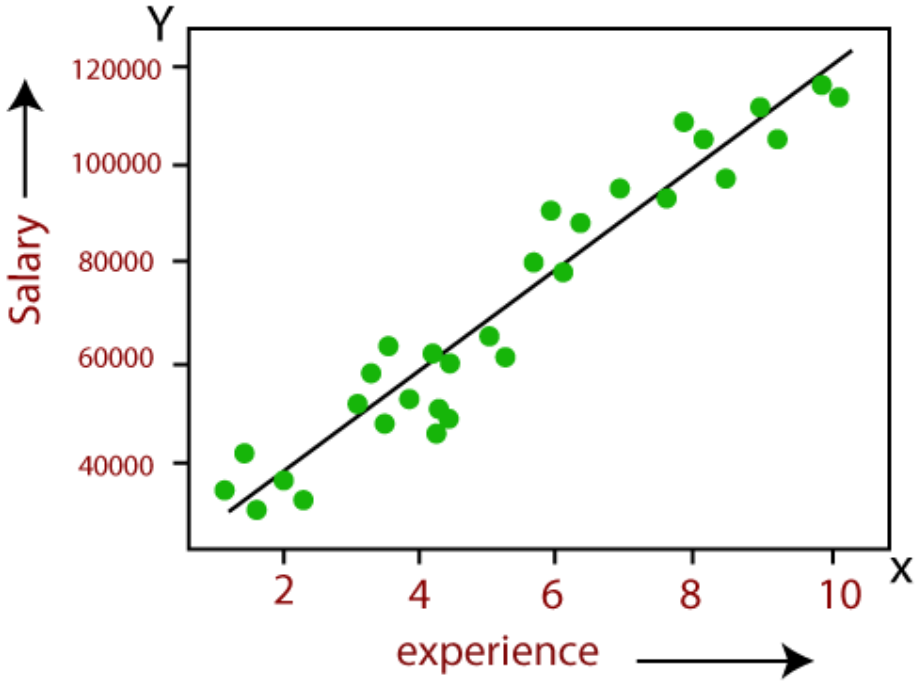
- Most simple Machine learning algorithm under Supervised Learning
- It is used for predicting the continuous dependent variable with the help of independent variables.
- goal is to find the best fit line that can accurately predict the output for the continuous dependent variable.
- If single independent variable is used for prediction … Simple Linear Regression and if there are more than two independent variables …Multiple Linear Regression.



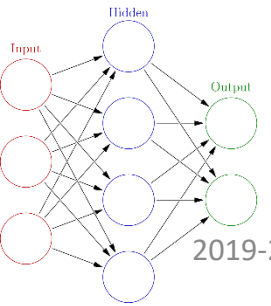


Linear Regression ... *summary*

- By finding the best fit line, algorithm establish the relationship between dependent variable and independent variable. And the relationship should be of linear nature.
- The output for Linear regression should only be the continuous values such as price, age, salary, etc
- dependent variable is on Y-axis (salary) and independent variable is on x-axis(experience)



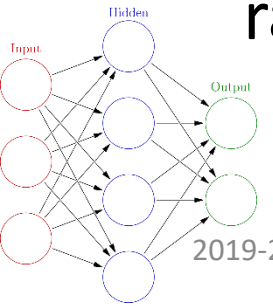
$$y = a_0 + a_1 x + \varepsilon$$



Where a_0 and a_1 are the coefficients and ε is the error term

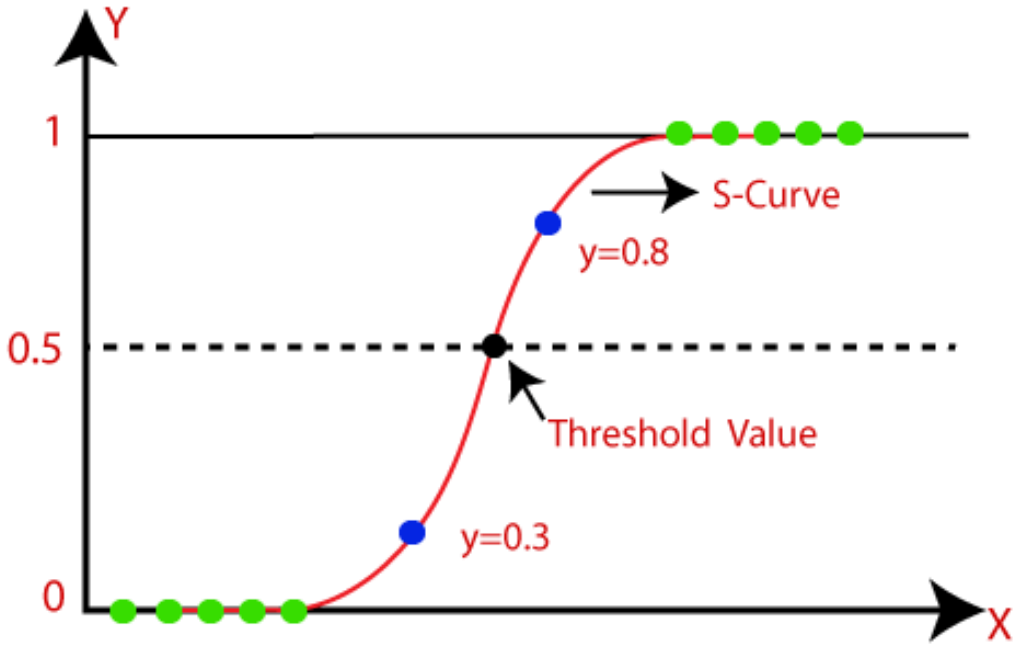
Logistic Regression

- Another popular Machine learning algorithm under Supervised Learning techniques.
- mainly used for Classification problems.
- Logistic regression is used to predict the categorical dependent variable with the help of independent variables
- The output of Logistic Regression problem can be only between the 0 and 1.
- Logistic regression can be used where the probabilities between two classes is required. Such as whether it will rain today or not, either 0 or 1, true or false etc.

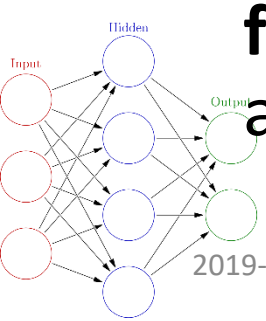


Logistic regression

- Logistic regression is based on the concept of Maximum Likelihood estimation. According to this estimation, the observed data should be most probable.
- In logistic regression, we pass the weighted sum of inputs through an activation function that can map values in between 0 and 1. Such activation function is known as **sigmoid function** and the curve obtained is called as sigmoid curve or S-curve

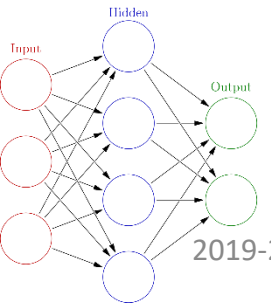


$$\log\left[\frac{y}{1-y}\right] = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$$

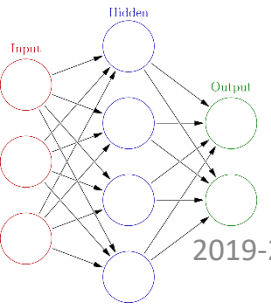


Linear vs Logistic Regression

- **Linear Regression** could help us predict the student's test score on a scale of 0 - 100. Linear regression predictions are continuous (numbers in a range).
- **Logistic Regression** could help use predict whether the student passed or failed. Logistic regression predictions are discrete (only specific values or categories are allowed). We can also view probability scores underlying the model's classifications.

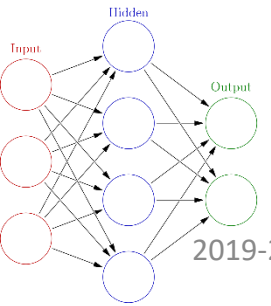


Linear Regression	Logistic Regression
Linear regression is used to predict the continuous dependent variable using a given set of independent variables.	Logistic Regression is used to predict the categorical dependent variable using a given set of independent variables.
Linear Regression is used for solving Regression problem.	Logistic regression is used for solving Classification problems.
In Linear regression, we predict the value of continuous variables.	In logistic Regression, we predict the values of categorical variables.
In linear regression, we find the best fit line, by which we can easily predict the output.	In Logistic Regression, we find the S-curve by which we can classify the samples.
Least square estimation method is used for estimation of accuracy.	Maximum likelihood estimation method is used for estimation of accuracy.
The output for Linear Regression must be a continuous value, such as price, age, etc.	The output of Logistic Regression must be a Categorical value such as 0 or 1, Yes or No, etc.
In Linear regression, it is required that relationship between dependent variable and independent variable must be linear.	In Logistic regression, it is not required to have the linear relationship between the dependent and independent variable.
In linear regression, there may be collinearity between the independent variables.	In logistic regression, there should not be collinearity between the independent variable.



Types of logistic regression

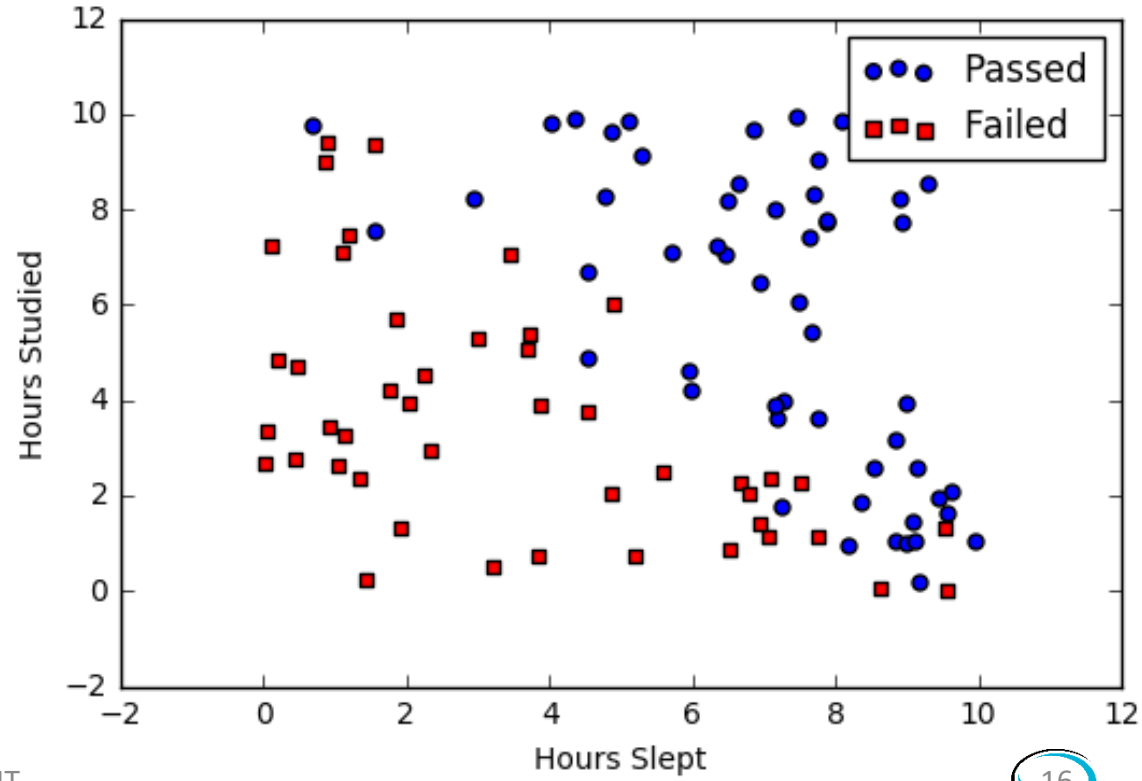
- Binary (Pass/Fail)
- Multi (Cats, Dogs, Sheep)
- Ordinal (Low, Medium, High)

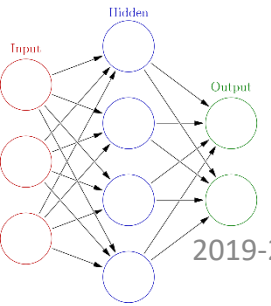


Binary Logistic Regression

- Given data on student exam results and our goal is to predict whether a student will pass or fail based on number of hours slept and hours spent studying.
- two features
  - hours slept,
  - hours studied
- two classes:
  - passed (1)
  - failed (0).

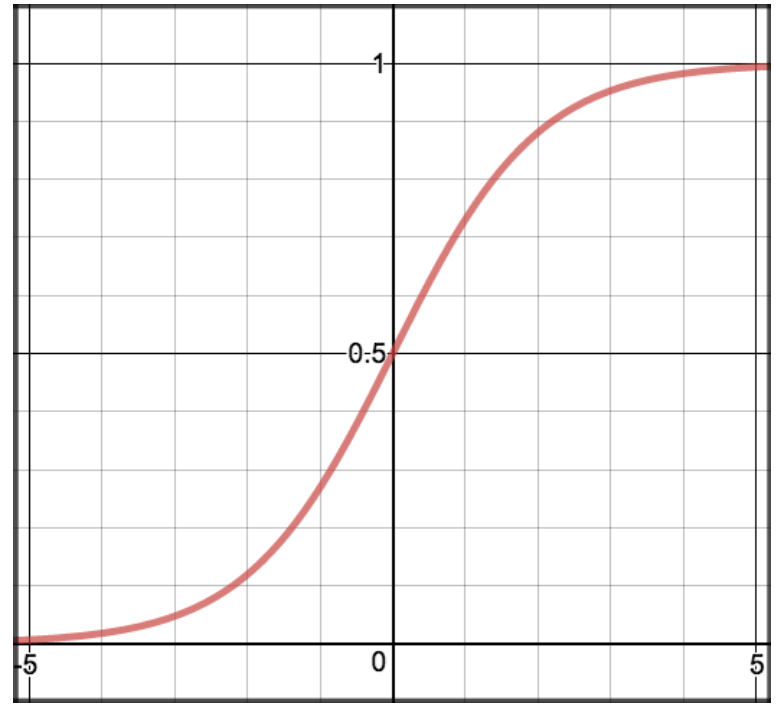
Studied	Slept	Passed
4.85	9.63	1
8.62	3.23	0
5.43	8.23	1
9.21	6.34	0



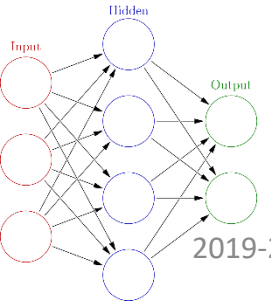


Sigmoid function

- In order to map predicted values to probabilities, the sigmoid function is used
- The function maps any real value into another value between 0 and 1.
- In machine learning, we use sigmoid to map predictions to probabilities.



$$S(z) = \frac{1}{1 + e^{-z}}$$

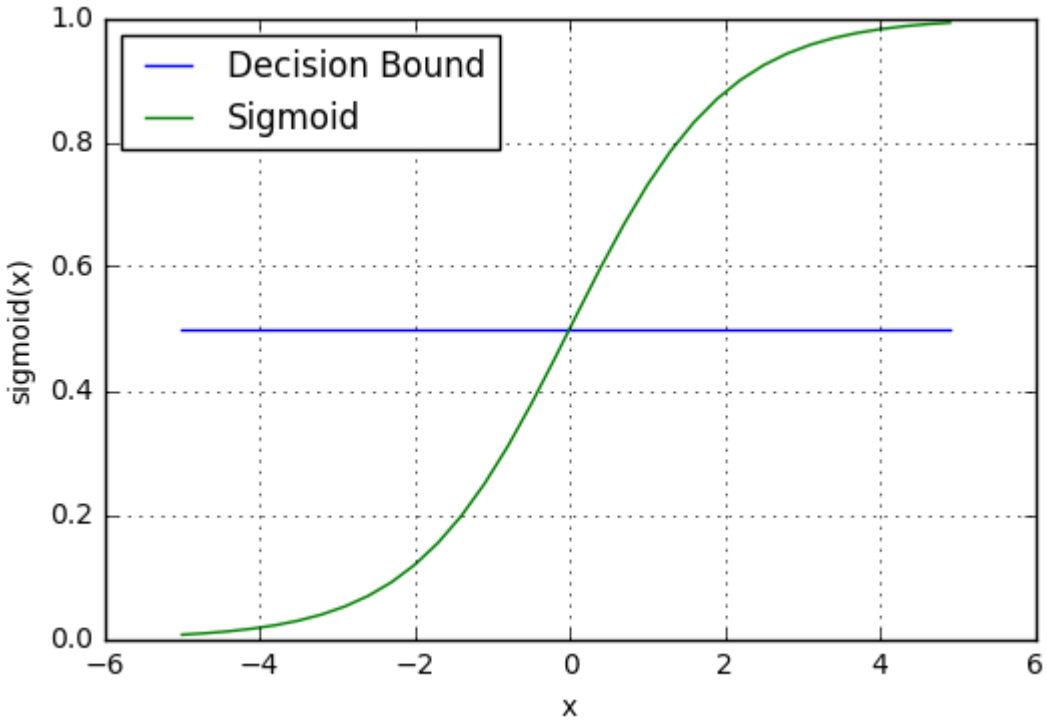


- s(z) = output between 0 and 1 (probability estimate)
- z = input to the function (your algorithm's prediction e.g. mx + b)
- e = base of natural log

```
def sigmoid(z):
    return 1.0 / (1 + np.exp(-z))
```

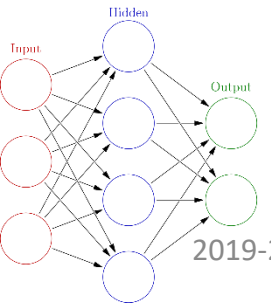
Decision Boundary

- Our current prediction function returns a probability score between 0 and 1.
- In order to map this to a discrete class (true/false, cat/dog), we select a threshold value or tipping point above which we will classify values into class 1 and below which we classify values into class 2.



- p≥0.5,class=1
- p<0.5,class=0

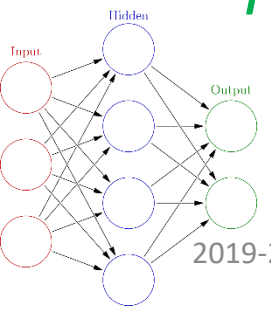
For example, if our threshold was .5 and
our prediction function returned .7, its classified as positive.
If prediction was .2 we would classify the observation as negative



Making Prediction

```
4.8 , 9.6 , 1
8.6 , 0.05, 0
3.8 , 0.7 ,  0
7.1 , 3.8 , 1
6.4 , 8.2 , 1
1.9 , 1.3 , 0
```

- Using knowledge of sigmoid functions and decision boundaries, we write a prediction function.
- A prediction function in logistic regression returns the probability of our observation being positive, True, or "Yes".
- We call this class 1 and its notation is *P(class=1)*.
- As the probability gets closer to 1, our model is more confident that the observation is in class 1
- Maths -> z = W0 + W1 * Studied + W2 * Slept
- *P(class=1) =* $\frac{1}{1+e^{-z}}$



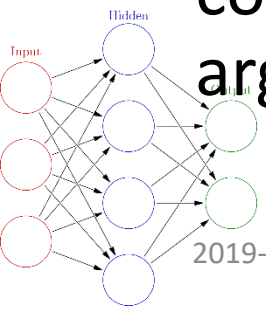

```
def predict(features, weights):
  # Returns 1D array of probabilities
  # that the class label == 1
  z = np.dot(features, weights)
  return sigmoid(z)
```

Confusion Matrix

To get more information on the accuracy of the model with a confusion matrix. In the case of binary classification, the confusion matrix shows the numbers of the following:

- True negatives in the upper-left position
- False negatives in the lower-left position
- False positives in the upper-right position
- True positives in the lower-right position

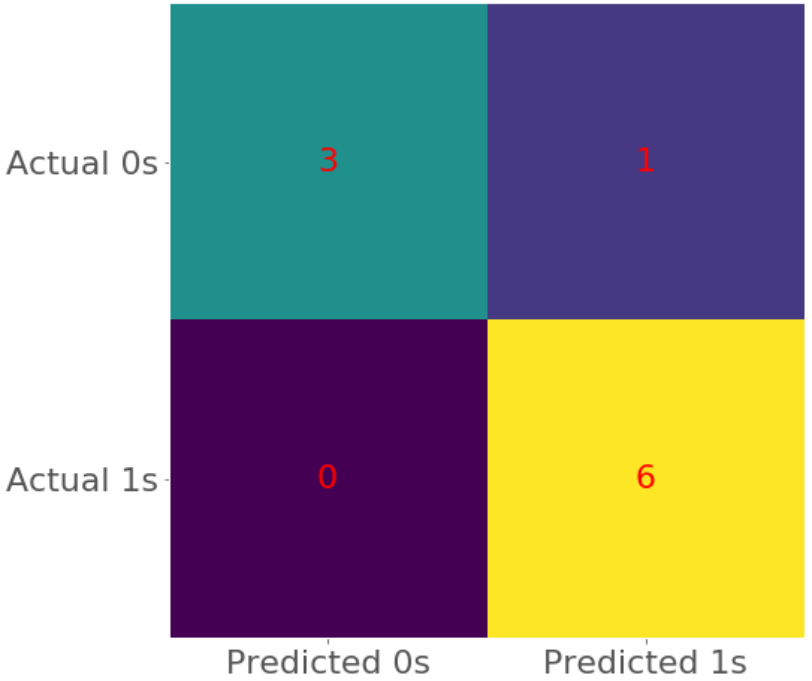
confusion_matrix() and provide the actual and predicted outputs as the arguments

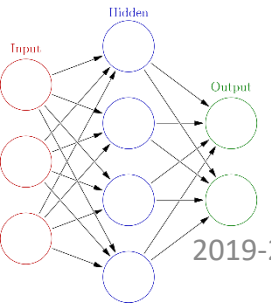


Confusion Matrix

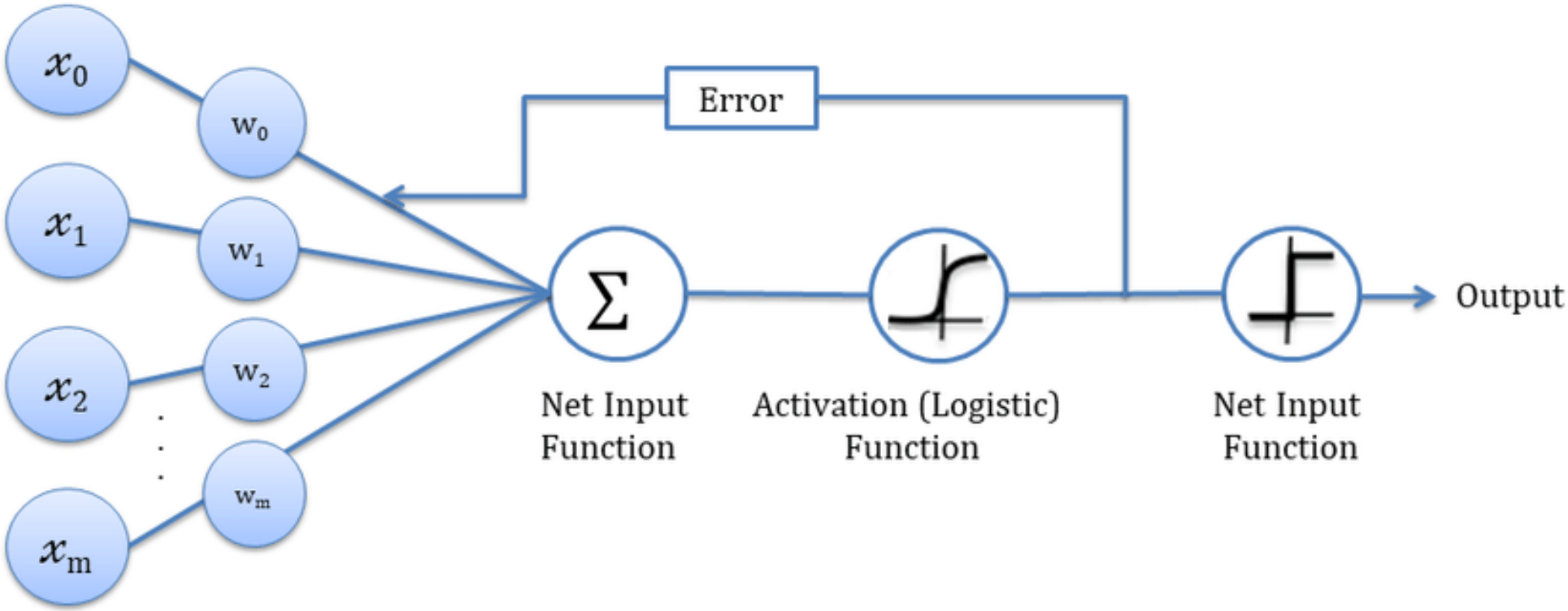
```
confusion_matrix ( y, model.predict(x))

array([[3, 1],
       [0, 6]])
```

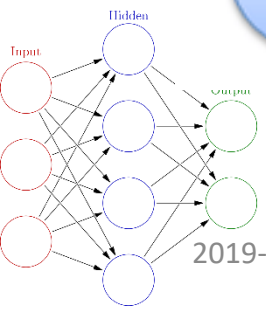
- **Three true negative predictions:** The first three observations are zeros predicted correctly.
- **No false negative predictions:** These are the ones wrongly predicted as zeros.
- **One false positive prediction:** The fourth observation is a zero that was wrongly predicted as one.
- **Six true positive predictions:** The last six observations are ones predicted correctly.





Source: realpython.com/logistic-regression-python


x_0
x_1
x_2
x_m
w_0
w_1
w_2
w_m
Error
Σ
Net Input Function
Activation (Logistic) Function
Net Input Function
Output


Input
Hidden
Output