



MDT: Deepak Mishra  
GitHub: [github.com/deepakmishra](#)

## Introduction to Logistic Regression

- What is machine learning?
- Why is machine learning?
- What is machine learning?
- Why is machine learning?
- What is machine learning?
- Why is machine learning?

**NOTE:** Implementing these algorithms is much easier but understanding them is a bit challenging.

### What is a Classification problem?

Classification is the problem of identifying which set of categories or class does a particular sample belong. In machine learning, classification refers to a predictive modeling problem where a class label is predicted for a given example of input data.

- Example:
- Identify if a mail is spam or not?
  - Identify if a person has cancer or not?
  - Identify if animal is a dog or cat?

Structured classification dataset

(Example dataset for diabetes test)

	Age	Glucose	BMI	OUTCOME
1)	34	135	34.3	YES
2)	58	85	25.1	NO
3)	25	97	44.8	NO
4)	29	102	23.1	YES

In Classification the "Target" variable is always Categorical

In Binary  $\rightarrow$  2 Classes  
In Multiclass  $\rightarrow$  More than 2 classes

### Types of Classification problems

**Binary Classification**  
(Binary classification refers to predicting one of two classes)



**Multiclass Classification**  
(Multiclass classification involves predicting one of more than two classes)



Iris Dataset

	sepal.length (cm)	sepal.width (cm)	petal.length (cm)	petal.width (cm)	species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	4.4	1.5	0.4	versicolour
7	4.8	3.9	1.4	0.3	versicolour
8	5.2	4.7	1.6	0.4	versicolour
9	5.2	3.4	1.6	0.4	versicolour
10	5.3	3.7	1.5	0.3	versicolour
11	5.8	4.0	1.6	0.3	versicolour
12	6.1	4.3	1.5	0.3	versicolour
13	5.7	4.4	1.5	0.3	versicolour
14	6.3	5.8	1.6	0.3	versicolour
15	7.0	6.2	1.8	0.3	versicolour
16	6.4	4.7	1.6	0.3	versicolour
17	6.5	5.0	1.8	0.3	versicolour
18	6.2	4.9	1.5	0.3	versicolour
19	6.7	5.7	1.7	0.3	versicolour
20	6.3	5.6	1.4	0.3	versicolour
21	6.8	5.5	2.0	0.3	versicolour
22	7.1	6.3	2.0	0.3	versicolour
23	7.0	6.8	2.2	0.3	versicolour
24	6.9	6.1	2.1	0.3	versicolour
25	7.3	6.8	2.1	0.3	versicolour
26	7.4	6.4	2.1	0.3	versicolour
27	7.5	6.6	2.3	0.3	versicolour
28	7.7	7.0	2.3	0.3	versicolour
29	7.7	6.9	2.4	0.3	versicolour
30	7.9	7.4	2.6	0.3	versicolour
31	6.4	3.9	1.4	0.4	virginica
32	6.5	4.0	1.5	0.4	virginica
33	6.3	3.3	1.5	0.4	virginica
34	7.2	4.7	1.6	0.4	virginica
35	7.7	5.0	1.8	0.4	virginica
36	7.7	5.4	1.8	0.4	virginica
37	7.4	4.6	1.6	0.4	virginica
38	7.9	5.4	1.8	0.4	virginica
39	7.4	4.8	1.8	0.4	virginica
40	7.8	5.1	1.8	0.4	virginica

**NOTE:** For Unstructured data we use Deep Learning.

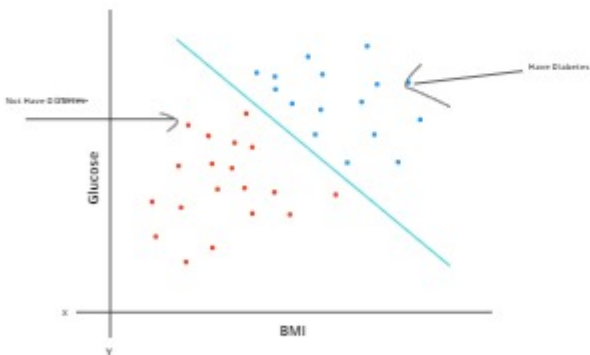
### Some common ML model for classification:

- Logistic Regression
- Naive Bayes
- Stochastic Gradient Descent
- K-Nearest Neighbours
- Decision Tree
- Random Forest
- Support Vector Machine

## Introduction to Logistic Regression

It is a machine learning model which is used to solve classification problems. By default it only supports "binary" classification and cannot be used for multiclass classification.

**NOTE:** In Linear Regression we try to find the "Line of Best Fit", but in Logistic Regression we try to find the "Line of Best Separation". This line perfectly separates the two classes.



The Logistic Regression model in simple terms, takes the "DATA" and output a probability value. It is upon us to set a **THRESHOLD**, most common threshold is 0.5

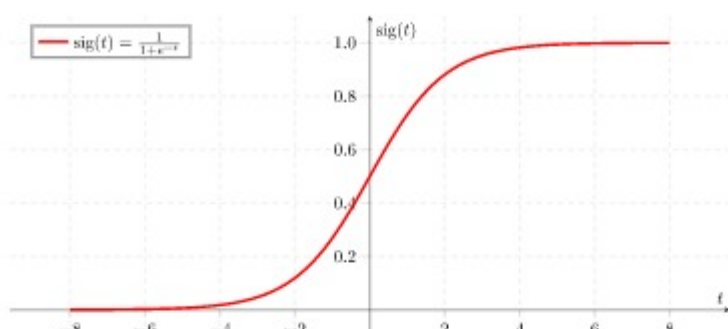


Logistic Regression at its core uses "Sigmoid Function"

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

This function convert any real value in the range of (0,1), Higher values are given higher probabilities

The Logistic Regression model uses this function to predict the probabilities



## Cost Function in Classification

The most common cost function for regression problem is "MSE". But for classification problems we have different cost functions. The most common cost function for classification is "CROSS ENTROPY"

**NOTE:** We have different cost function for Binary & Multiclass Classification

### CROSS ENTROPY

**Binary Cross Entropy**

**Categorical Cross Entropy**

### Binary Cross Entropy (BCE)

(Loss Function)

$$BCE(t, p) = \sum -(t * \log(p) + (1 - t) * \log(1 - p))$$

t = target value  
p = predicted value

$$BCE(t, p) = \begin{cases} -\log(1-p) & \text{if } t = 0 \\ -\log(p) & \text{if } t = 1 \end{cases}$$

(Example on how to calculate BCE)

	Age	Glucose	BMI	t	P	
				OUTCOME	PREDICTED	
1)	34	135	34.3	1	0.6	→ 0.22
2)	58	85	25.1	0	0.2	→ 0.09
3)	25	97	44.8	1	0.5	→ 0.3
4)	29	102	23.1	0	0.8	→ 0.6

LOSS = 1.21