

**NAME: M.KANCHANA**  
**DEPT: CSE**  
**SEC: “C”**  
**YEAR: IV**  
**ROLLNO: 19BCS126**  
**SUBJECT: COMPUTATIONAL INTELLIGENCE**  
**SUBJECT CODE: 19CS701**

### **ASSIGNMENT-1**

---

#### **Logistic Regression:**

##### **Introduction:**

- Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique.
- It is used for predicting the categorical dependent variable using a given set of independent variables.
- Logistic regression predicts the output of a categorical dependent variable.
- It can be either Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1.
- Logistic Regression can be used to classify the observations using different types of data and can easily determine the most effective variables used for the classification.

- Logistic Regression is a significant machine learning algorithm because it has the ability to provide probabilities and classify new data using continuous and discrete datasets.

### **Significance of Logistic Regression:**

- Logistic regression is used for solving the classification problems.
- Logistic regression is used to obtain odds ratio in the presence of more than one explanatory variable.
- Logistic regression helps us estimate a probability of falling into a certain level of the categorical response given a set of predictors.

### **Logistic Function (Sigmoid Function):**

- In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1).
- The sigmoid function is a mathematical function used to map the predicted values to probabilities.
- It maps any real value into another value within a range of 0 and 1.
- The value of the logistic regression must be between 0 and 1, which cannot go beyond this limit, so it forms a curve like the "S" form. The S-form curve is called the sigmoid function or the logistic function.
- In logistic regression, we use the concept of the threshold value, which defines the probability of either 0 or 1. Such as values above the threshold value tends to 1, and a value below the threshold values tends to 0.

### **Assumptions for Logistic Regression:**

- The dependent variable must be categorical in nature.
- The independent variable should not have multi-collinearity.

## Logistic Regression Equation:

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

## LINEAR REGRESSION VS LOGISTIC REGRESSION.

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 find the best fit line, by which we can 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.