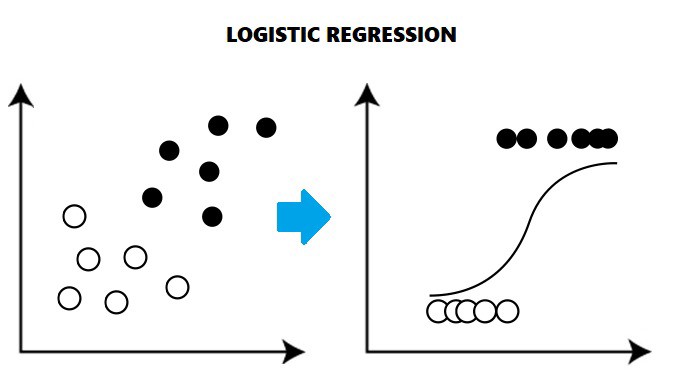
Assignment 13

1. What are the different types of logistic Regression?

Ans:

**Logistic Regression** : Logistic regression is **a statistical analysis method to predict a binary outcome, such as yes or no, based on prior observations of a data set**. A logistic regression model predicts a dependent data variable by analyzing the relationship between one or more existing independent variable.



Types of Logistic Regression:

There are three main types of logistic regression: binary, multinomial and ordinal. They differ in execution and theory. Binary regression deals with two possible values, essentially: yes or no. Multinomial logistic regression deals with three or more values. And ordinal logistic regression deals with three or more classes in a predetermined order.

### Binary logistic regression

Binary logistic regression was mentioned earlier in the case of classifying an object as an animal or not an animal—it’s an either/or solution. There are just two possible outcome answers. This concept is typically represented as a 0 or a 1 in coding. Examples include:

* Whether or not to lend to a bank customer (outcomes are yes or no).
* Assessing cancer risk (outcomes are high or low).
* Will a team win tomorrow’s game (outcomes are yes or no).

### Multinomial logistic regression

Multinomial logistic regression is a model where there are multiple classes that an item can be classified as. There is a set of three or more predefined classes set up prior to running the model. Examples include:

* Classifying texts into what language they come from.
* Predicting whether a student will go to college, trade school or into the workforce.
* Does your cat prefer wet food, dry food or human food?

### Ordinal logistic regression

Ordinal logistic regression is also a model where there are multiple classes that an item can be classified as; however, in this case an ordering of classes is required. Classes do not need to be proportionate. The distance between each class can vary. Examples include:

* Ranking restaurants on a scale of 0 to 5 stars.
* Predicting the podium results of an Olympic event.
* Assessing a choice of candidates, specifically in places that institute ranked-choice voting

1. What is the difference between the outputs of the Logistic Model and the Logistic function?

Ans:

**Output of Logistic Model:**

The output of a logistic regression model is **the probability of our input belonging to the class labeled with 1**. And the complement of our model's output is the probability of our input belonging to the class labeled with 0

**Output of Logistic Function:**

The output of logistical regression is reported in terms of odds ratios, which is **the numerical odds (bounded by 0 and infinity) of the binary, dependent variable being true, given a one-unit increase in the independent variable**.

1. How do we handle categorical variables in Logistic Regression?

Logistic regression models are a great tool for analysing binary and categorical data, allowing you to **perform a contextual analysis to understand the relationships between the variables, test for differences, estimate effects, make predictions, and plan for future scenarios**

1. What are the assumptions made in Logistic Regression?

Ans:

Assumptions:

1. In a logistic regression , there is No or minimum muliticollinearity amongst independent variable(X). This means that the independent variables should not be too highly correlated with each other.
2. In logistic regression , relationship between ln=∑ᶱixi  must be linear .
3. Logistic regression includes independence of errors.
4. The error terms (residuals) do not need to be normally distributed.
5. Homoscedasticity is not required.
6. The dependent variable in logistic regression is not measured on an interval or ratio scale.
7. logistic regression assumes linearity of independent variables and log odds.  although this analysis does not require the dependent and independent variables to be related linearly, it requires that the independent variables are linearly related to the log odds.
8. Logistic regression typically requires a large sample size.  A general guideline is that you need at minimum of 10 cases with the least frequent outcome for each independent variable in your model. For example, if you have 5 independent variables and the expected probability of your least frequent outcome is .10, then you would need a minimum sample size of 500 (10\*5 / .10).
9. Why Can’t we use MSE as a cost function for Logistic Regression?