

ASSIGNMENT 6

NAME: FALAK KHAN
BATCH: Machine Learning and AI Batch A3

1.Explain the Logistic Function in detail.

>>Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.

In simple words, the dependent variable is binary in nature having data coded as either 1 (stands for success/yes) or 0 (stands for failure/no).

Mathematically, a logistic regression model predicts $P(Y=1)$ as a function of X . It is one of the simplest ML algorithms that can be used for various classification problems such as spam detection, Diabetes prediction, cancer detection etc.

- The [logistic function](#) also called the sigmoid function was developed by statisticians to describe properties of population growth in ecology, rising quickly and maxing out at the carrying capacity of the environment.
- It's an S-shaped curve that can take any real-valued number and map it into a value between 0 and 1, but never exactly at those limits.
- Formula: $1 / (1 + e^{(-\text{value})})$, where e is the base of natural algorithms (Euler's number or the EXP() function in your spreadsheet) and value is the actual numerical value that you want to transform.

2. Is decision boundary linear in Logistic Regression?

>> Logistic regression is known and used as a linear classifier. It is used to come up with a hyper plane in feature space to separate observations that belong to a class from all the other observations that do not belong to that class. The decision boundary is thus linear.

3. Why can't we use Linear Regression instead of Logistic Regression for Binary Classification?

>> There are two things that explain why Linear Regression is not suitable for classification. The first one is that Linear Regression deals with continuous values whereas classification problems mandate discrete values. The second problem is regarding the shift in threshold value when new data points are added. Let us take a simple Linear Regression example and fit a line to it. Hence to avoid this we use an algorithm called the Logistic Regression which is a binary classification algorithm to stepover these practical problems that hold back Linear Regression for classification. Logistic Regression deals with discrete values unlike Linear Regression and also maintains the value of the threshold even when new data points are added.

4. What is Maximum Likelihood Estimator?

>> Maximum likelihood estimation is a method that determines values for the parameters of a model. The parameter values are found such that they maximize the likelihood that the process describes by the model produces the data were actually observed.

5. Which metric will you use for judging a Logistic Regression model?

>> Log Loss is the most important classification metric based on probabilities. It's hard to interpret raw log-loss values, but log-loss is still a good metric for comparing models. For any given problem, a lower log loss value means better predictions.

Mathematical interpretation: Log Loss is the negative average of the log of corrected predicted probabilities for each instance.