Sumon Singh
07.05.22

Aim: Write a program to implement Logistic Regression Algorithm

1) Dataset Description:

A) Name: Social Network Ads

B) Description: This is a catagorical dataset to determine whether a user purchased a particular product, based on the user's gender, age and salary. The dataset is available on Kaggle website.

c) Size: The dataset consists of 400 rows and 5 columns.

D) Attributes: The dataset has 4 attributes

1) User ID: Unique numbers of the user
2) Gender: Gender of the user
3) Age: Age of the user
4) EstimatedSalary: Salary of user

E) Label: The target is 'Purchased' which is 0 or 1, if 0 then the user didn't bought the product else bought it.

Sumon Singh
~~2+~~
07.09.22

2) ML Model Description:

A) Name of model :- Logistic Regression

B) Type of model :- It is a supervised learning model which is used for classification. ~~mod~~

C) Algorithm :-

1) Using linear regression calculate output based on features.

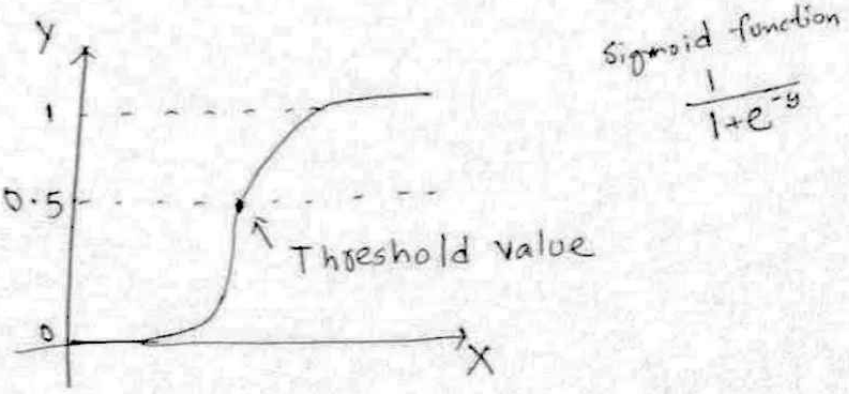
Example: Let us assume, we have a dataset consisting of n features $X_1, X_2 \cdots X_n$, of real values
∴ output will be -

$$y = b_0 + b_1X_1 + b_2X_2 \cdots + b_nX_n$$

2) Using Sigmoid function find the ~~prob~~ probablity of the output of previous step
Sigmoid function : $\frac{1}{1+e^{-y}}$ (y is previous output)

3) If the probablity output is more than 0.5 then the actual classification will be of type 1 else type 0.

D) Draw :-



Sumon Singh
15.05.22

3) Result Analysis : The outcomes are 0 and 1

A) Accuracy : Our model identifies 0.8875 data correctly which means our model is pretty good.

B) Precission :

For 0 the model's precission is 0.89

For 1 the model's precission is 0.88

C) Recall :

For 0 the recall is 0.94 and for 1 the recall is 0.79

D) f1-score :

For 0 the f1-score is 0.92 and for 1 f1-score is 0.83

* The model identifies 94% '0' labeled data correctly among all '0' labeled data, similarly it identifies 79% '1' labeled data correctly among all '1' labled data

So, our model works fine on the dataset