**SOCIAL\_NETWORK\_ADS.CSV USING LOGISTIC REGRESSION**

**INTRODUCTION:**

Social network advertising, also social media targeting, is a group of terms that are used to describe forms of online advertising that focus on social networking services. One of the major benefits of this type of advertising is that advertisers can take advantage of the users’ demographic information and target their ads appropriately . Advantages are advertisers can reach users who are interested in their products, allows for detailed analysis.

**PROBLEM:**

If a company launches a new brand or model of car ,the company needs to know about the customers to obtain the maximum sale on the particular product by doing so the company can increase the production and invest more on the new launch ,so there is need to analyze our customers from the given dataset social\_network\_ads using any one of the Machine learning algorithm in python

**DATASET:**

The Social\_Networks\_Ads dataset contains five columns and 400 entries

* UserID - Each person has a unique ID from which we can identify the person uniquely.
* Gender - Person can male or female. This field is very important for our hypothesis.
* Age - Age of the person. Because our product can be useful to some ages only.
* EstimatedSalary - This column contains salary of a person as salary can affect the shopping of a person.
* Purchased - Contains two numbers ‘0’ or ‘1’. ‘0’ means not purchased and ‘1’ means purchased.This variable is our dependent variable.

**ALGORITHM :**

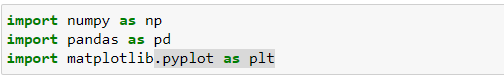
In this problem we are using logistic regression using python

**About logistic regression :**

Logistic Regression is a Machine Learning algorithm which is used for the classification problems, it is a predictive analysis algorithm and based on the concept of probability. It is the go-to method for binary classification problems (problems with two class values),that is 0=No (they don’t buy a car) and 1=YES (they will buy the car) we have in our dataset .

**SOLUTION:**

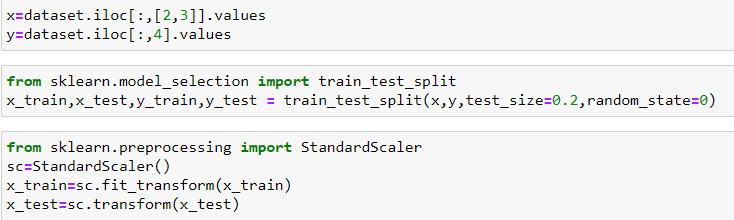
* Import Packages



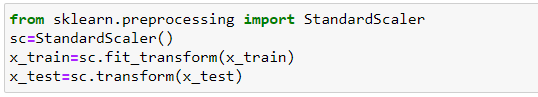
* Load The Dataset



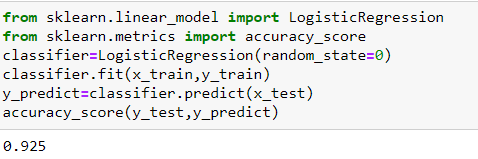
* Splitting The Dataset Into Training Dataset And Test Dataset



* Preprocessing The Data



* Using Logistic Regression classification

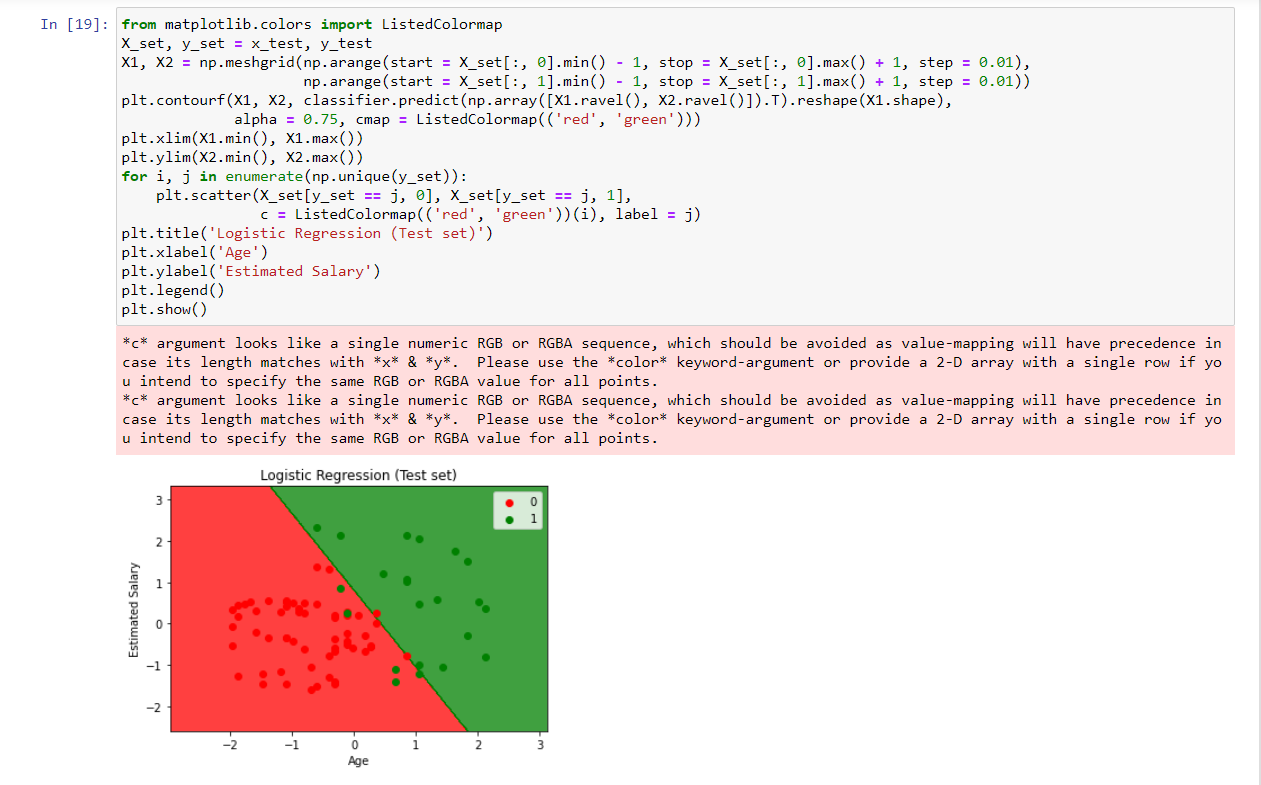


* Plotting The Result

TRAINING DATASET



TESTING DATASET



**CONCLUSION:**

The data visualization of the training set and test set in the above figure ,gives the logistic regression is linear model and the green dots represent the customers they buy the car and the red dots represent the customers they do not buy the car and we have a result that the logistic regression model give 92% of accuracy score .