**DLITHE INTERNSHIP RESEARCH PROJECT**

PROJECT ON

**SOCIAL NETWORKING ADS**

FROM

ANAGHA N

H DEEKSHA SHANBHAG

JESSICA DSOUZA

**TABLE OF CONTENTS:**

|  |  |  |
| --- | --- | --- |
| 1. No | Contents | Pg.no |
| 1. | Introduction | 2 |
| 2. | Technology used | 3 |
| 3. | Code | 4-6 |
| 4. | Screenshots | 7 |
| 5. | Conclusion | 8 |

**INTRODUCTION:**

**Social network advertising**, also **social media targeting**, is a group of terms that are used to describe forms of [online advertising](https://en.wikipedia.org/wiki/Online_advertising) that focus on [social networking services](https://en.wikipedia.org/wiki/Social_networking_service). 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.

[Social media](https://en.wikipedia.org/wiki/Social_media) targeting combines current targeting options (such as [geotargeting](https://en.wikipedia.org/wiki/Geotargeting), [behavioral targeting](https://en.wikipedia.org/wiki/Behavioral_targeting), socio-psychographic targeting, etc.), to make detailed [target group](https://en.wikipedia.org/wiki/Target_audience) identification possible. With social media targeting, advertisements are distributed to users based on information gathered from target group profiles.

Social network advertising is not necessarily the same as social media targeting. Social media targeting is a method of optimizing social media advertising by using profile data to deliver advertisements directly to individual users. Social network advertising refers to the process of matching [social network](https://en.wikipedia.org/wiki/Social_network) users to target groups that have been specified by the advertiser.

This project will be a walkthrough of a simple Logistic Regression model in an attempt to strategize a basic ad-targeting campaign for a social media network/website.

**TECHNOLOGY USED:**

PYTHON

Python is an interpreted, high-level, general-purpose programming language. Created by Guido van Rossum and first released in 1991, Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.

Python is dynamically typed and garbage-collected. It supports multiple programming paradigms, including procedural, object-oriented, and functional programming. Python is often described as a "batteries included" language due to its comprehensive standard library.

JUPYTER NOTEBOOK

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text.

Uses include:

data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

**CODE:**

Dataset contains some information about users in the social network, including their User ID, Gender, Age, and Estimated Salary. The last column of the dataset is a vector of booleans describing whether or not each individual ended up clicking on the advertisement (0 = False, 1 = True).

First we need to import the relevant libraries, the dataset, and establish which variables are either dependent or independent.

import numpy as np

import pandas as pd

dataset=pd.read\_csv(r'C:\Users\HP\Downloads\Social\_Network\_Ads (1).csv')

We must extract the relevant vectors from our dataset: the independent variables (X) and the dependent variable (Y). The following code segment describes the selection of the entire third and fourth columns for X, as well as the entire fifth column for Y.

X=dataset.iloc[:,[2,3]].values

Y=dataset.iloc[:,4].values

from sklearn.model\_selection import train\_test\_split

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X,Y,train\_size=0.8,random\_state=0)

k=int(round(np.sqrt(len(X\_train))))

if k%2==0:

k=k+1

For the result, a common tool within machine learning models is to apply Feature Scaling: "...a method used to standardize the range of independent variables or features of data. In data processing, it is also known as data normalization and is generally performed during the data preprocessing step."A helpful library called StandardScaler will quickly preprocess the data in this manner.

from sklearn.preprocessing import StandardScaler

sc=StandardScaler() #create object for that class

X\_train=sc.fit\_transform(X\_train)

X\_test=sc.transform(X\_test)

from sklearn.neighbors import KNeighborsClassifier

classifier= KNeighborsClassifier(n\_neighbors=k,metric='minkowski', p=2)

classifier.fit(X\_train, Y\_train)

y\_pred=classifier.predict(X\_test)

from sklearn.metrics import confusion\_matrix

cm=confusion\_matrix(Y\_test,y\_pred)#related to model's prediction n expected

from sklearn.model\_selection import cross\_val\_score

accuracies=cross\_val\_score(estimator=classifier, X=X\_train, y=Y\_train, cv=10)

accuracies.mean()

accuracies.std()

Now we are ready to build our Logistic Regression Model. We create an object of the LogisticRegression() class and refer to it as our 'classifier'. The random state variable allows us to all get the same outcome but can be changed to alter the results slightly. We then fit the classifier to the training set with the named .fit() method so that it can understand the correlations between X and Y. then, we will test the classifier's predictive power on the test set. The Logistic Regression's .predict() method will give us a vector of predictions for our dataset, X\_test.

#LINEAR REGRESSION

from sklearn.linear\_model import LinearRegression

linreg=LinearRegression()

#Training

linreg.fit(X\_train,Y\_train)

#Testing

#ACCURACY

linregacc=linreg.score(X\_test,Y\_test)

#PREDICT

linregpred=linreg.predict(X\_test)

#PLOT THE PREDICTION

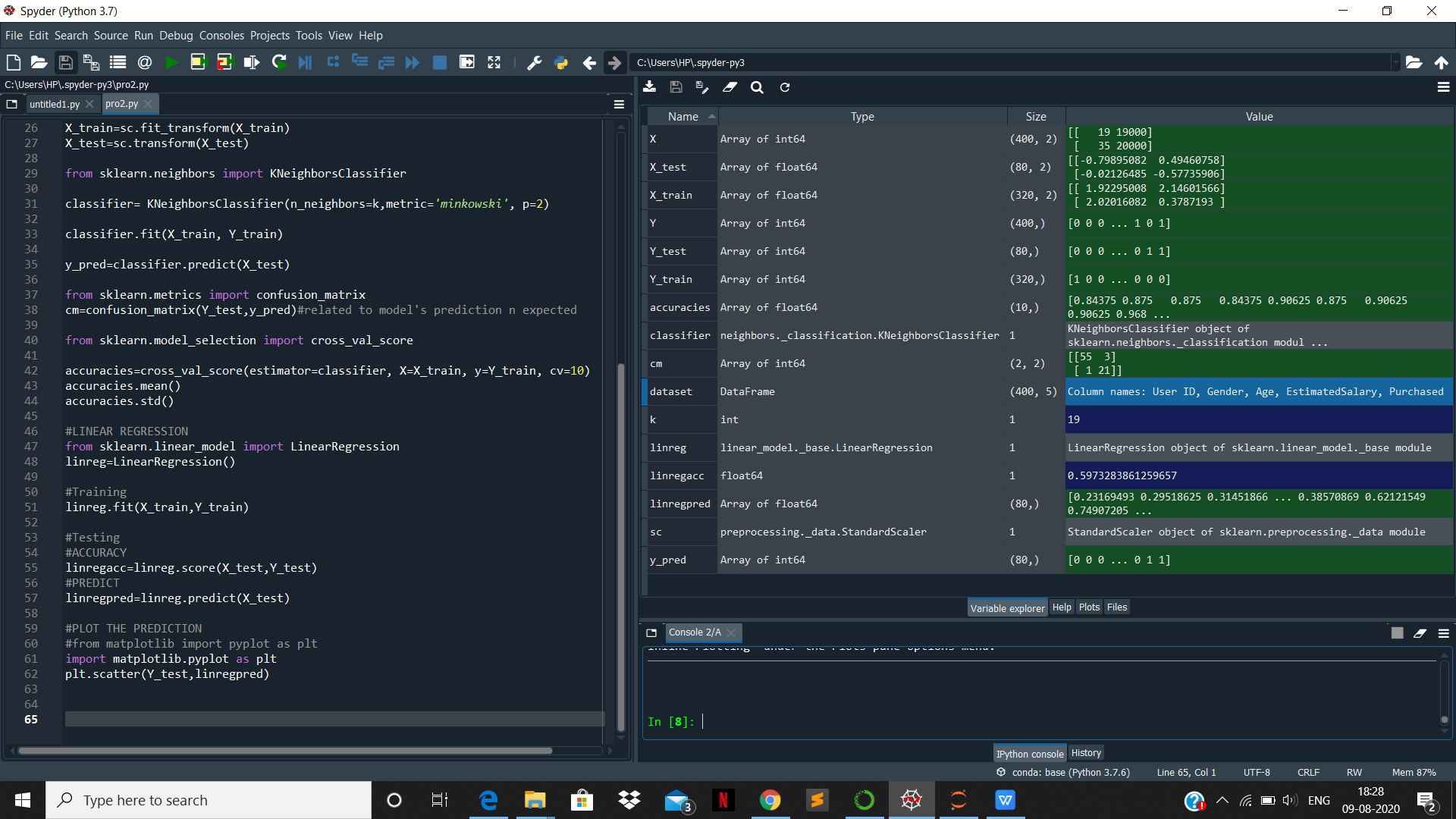
#from matplotlib import pyplot as plt

import matplotlib.pyplot as plt

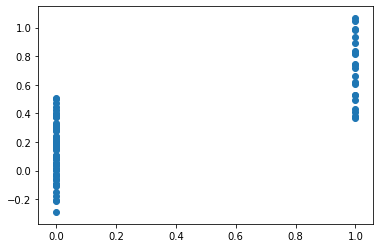
plt.scatter(Y\_test,linregpred)

**SCREENSHOTS:**

A)



B)



**CONCLUSION :**

Social network advertising is effective in terms of publicity and cost.These evidences could prove that social network advertising has high potential to take overmass network advertising in the future. Social network advertising not just lower down the cost ofspending and also raise awareness of the social network users. Besides, social network advertising can reach larger amount of viewers because there are millions and billions of people is using social network in this modern century. With all these, social network advertising shows it’s effectiveness and could be the first choice of the marketers use to advertise.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*