

Rasel1435 / Advanced-Data-Science-Machine-Learning-Projects

Q Type to search

>

+

<> Code Issues Pull requests Actions Projects Wiki Security Insights Settings

main

Advanced-Data-Science-Machine-Learning-Projects / Social Media Ads Classification with Machine Learning

Go to file

t

Rasel1435 Social Media Ads Classification548a6bd · 2 minutes agoHistory

PreviewCodeBlame764 lines (764 loc) · 89.5 KBRawDownloadEdit

Social Media Ads Classification with Machine Learning

The classification of social media ads is all about analyzing the ads for classifying whether your target audience will buy the product or not. It's a great use case for data science in marketing. So, if you want to learn how to analyze social media ads to classify your target audience, then this article is for you. In this article, I will walk you through the task of social media ads classification with machine learning using Python.

Social Media Ads Classification

Classifying social media ads means analyzing your social media ads for finding the most profitable customers for your product who are more likely to buy the product. Sometimes the product you are offering is not suitable for all people when it comes to age and income. For example, a person between the ages of 20 and 25 may like to spend more on smartphone covers than a person between the ages of 40 and 45.

Likewise, a high-income person can afford to spend more on luxury goods than a low-income person. So this is how a business can determine whether a person will buy their product or not by classifying their social media ads. In the section below, I will walk you through social media ads classification with Machine Learning using Python.

Social Media Ads Classification using Python

The dataset I am using for the task of Social Media Ads Classification is downloaded from Kaggle. It contains data about a product's social media advertising campaign. It contains features like

1. the age of the target audience
2. the estimated salary of the target audience
3. and whether the target audience has purchased the product or not

Import Necessary Libraries

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

%matplotlib inline
```

Gathering Data

```
In [3]: data = pd.read_csv(r'./data/social.csv')
data.head()
```

```
Out[3]:
```

	Age	EstimatedSalary	Purchased
0	19	19000	0
1	35	20000	0
2	26	43000	0
3	27	57000	0
4	19	76000	0

Data Pre-Processing

```
In [4]: print(data.head())
```

```
Out[4]:
```

	Age	EstimatedSalary	Purchased
0	19	19000	0
1	35	20000	0
2	26	43000	0
3	27	57000	0
4	19	76000	0

```
In [6]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 400 entries, 0 to 399
Data columns (total 3 columns):
 #   Column          Non-Null Count  Dtype
---  ---
 0   Age             400 non-null    int64
 1   EstimatedSalary  400 non-null    int64
 2   Purchased       400 non-null    int64
dtypes: int64(3)
memory usage: 9.5 KB
```

```
In [9]: data.describe()
# print(data.describe())
```

```
Out[9]:
```

	Age	EstimatedSalary	Purchased
count	400.000000	400.000000	400.000000
mean	37.655000	69742.500000	0.357500
std	10.482877	34096.960282	0.479864
min	18.000000	15000.000000	0.000000
25%	29.750000	43000.000000	0.000000
50%	37.000000	70000.000000	0.000000
75%	46.000000	88000.000000	1.000000
max	60.000000	150000.000000	1.000000

Now let's take a look at some of the insights from the data to see if we need to make any changes to the dataset

```
In [12]: data.isnull()
# print(data.isnull())
```

```
Out[12]:
```

	Age	EstimatedSalary	Purchased
0	False	False	False
1	False	False	False
2	False	False	False
3	False	False	False
4	False	False	False
...
395	False	False	False
396	False	False	False
397	False	False	False
398	False	False	False
399	False	False	False

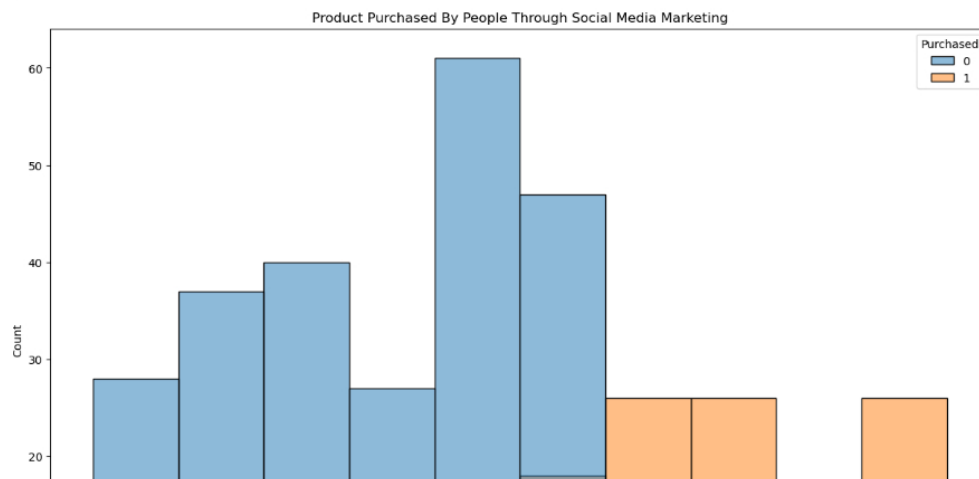
400 rows × 3 columns

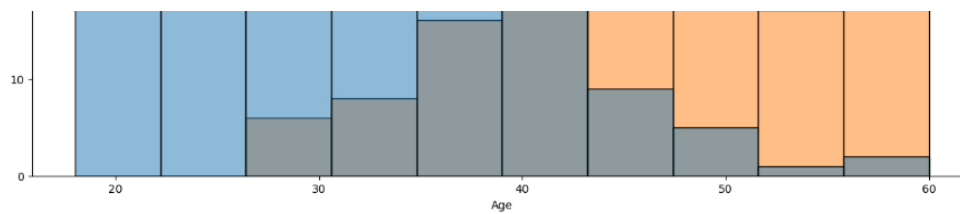
```
In [16]: data.isnull().sum()
# print(data.isnull().sum())
```

```
Out[16]: Age             0
EstimatedSalary         0
Purchased               0
dtype: int64
```

As we can see our dataset is okay! Now let's explore some of the important patterns in the dataset. The first thing I want to explore is the ages of the people who responded to the social media ads and bought the product

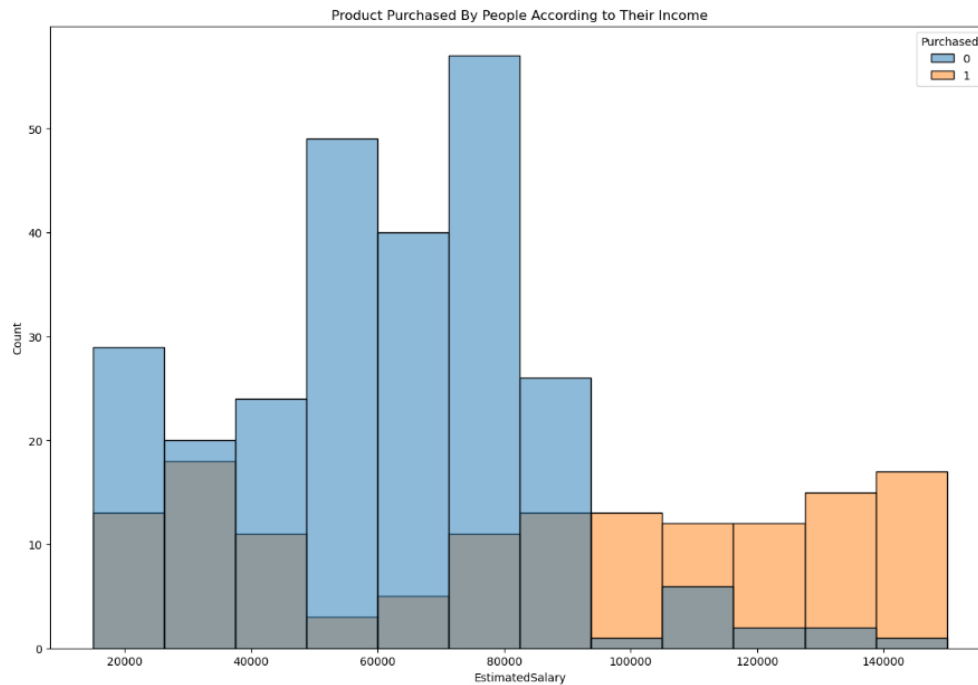
```
In [20]: plt.figure(figsize=(15, 10))
plt.title("Product Purchased By People Through Social Media Marketing")
sns.histplot(x="Age", hue="Purchased", data=data)
plt.show()
```





The visualization above shows that people over 45 among the target audience are more interested in purchasing the product. Now let's take a look at the income group of people who responded to social media ads and purchased the product

```
In [21]: plt.figure(figsize=(15, 10))
plt.title("Product Purchased By People According to Their Income")
sns.histplot(x="EstimatedSalary", hue="Purchased", data=data)
plt.show()
```



The visualization above shows that people with a monthly income of over 90,000 among the target audience are more interested in purchasing the product.

Feature Selection

Now let's train a model to classify social media ads. First I'll set the "Purchased" column in the dataset as the target variable and the other two columns as the features we need to train a model

```
In [25]: x = np.array(data[["Age", "EstimatedSalary"]])
y = np.array(data[["Purchased"]])
```

Splitting The Data

```
In [26]: from sklearn.model_selection import train_test_split
xtrain, xtest, ytrain, ytest = train_test_split(x, y, test_size=0.10, random_state=42)
```

```
In [33]: xtrain.shape, xtest.shape
```

```
Out[33]: ((360, 2), (40, 2))
```

```
In [34]: ytrain.shape, ytest.shape
```

```
Out[34]: ((360, 1), (40, 1))
```

Choosing The Model

```
In [28]: from sklearn.tree import DecisionTreeClassifier
model = DecisionTreeClassifier()
#fit the model
model.fit(xtrain,ytrain)
predictions = model.predict(xtest)
```

At last, let's have a look at the classification report of the model

```
In [36]: from sklearn.metrics import classification_report
print(classification_report(ytest, predictions))
```

	precision	recall	f1-score	support
0	0.92	0.85	0.88	27
1	0.73	0.85	0.79	13
accuracy			0.85	40
macro avg	0.83	0.85	0.84	40
weighted avg	0.86	0.85	0.85	40

Summary

So this is how you can analyze and classify social media ads about the marketing campaign of a product. Classifying social media ads means analyzing your social media ads for finding the most profitable customers for your product who are more likely to buy the product. I hope you liked this article on classifying Social Media Ads with Machine Learning using Python. Feel free to ask your valuable questions in the comments section below.

Sheikh Rasel Ahmed

Data Science || Machine Learning || Deep Learning || Artificial Intelligence Enthusiast

LinkedIn - <https://www.linkedin.com/in/shekhnirob1>

GitHub - <https://github.com/Rasel1435>

YouTube - <https://www.youtube.com/@codewithsheikhrasel>

Facebook - <https://www.facebook.com/rasel1435>

Instagram - https://www.instagram.com/shekh_nirob