Customer retention analysis

This is a dataset that tells about how the e-shopping sites try to retain its customers .For this number of factors are responsible :-

- The content on the website must be easy to understand.

- Information about the product must be stated clearly to compare between products.

- Navigation must be easy in website .

- User friendly interface.

- Lot of options in payment methods.

- Trust that online store should create on customers.

- Try to solve the queries of their customers.

- Guarantee the privacy of customer.

- Displaying of quality information on products.this will improve customer satisfaction.

- Wide variety of products.

- The webpage should be appealing.

- Reliability of website.

- Speedy order delivery .

- Security of customer s financial information.

- And many more-------------------.

Success of online store depends on customers point of view also .There are number of other factors for increasing the popularity of website.Some of these factors are:-

- Availability of payment options.

- According to customers if shopping gives them monetary benefits.

- If they enjoy their shopping.

- Return and Replace policy also motivates customers to shop online.

- If they trust the site.

- Monetary benefits .

- Gives them sense of adventure.

- If they get satisfaction in the quality for what they pay.

- Ease to use app.

- If they get delivery on time.

- And many more

There are many more features for retaining customers.

For the analysis we have to work on jupyter notebook

Before starting we will import necessary libraries using import

Load the dataset

Check the nulls using.sum().null()

See the description using .describe method().

Check the datatypes to convert it into numeric data because computer only understands numeric data.

Plotting of data for data visualization using countplot.Also we will count the number of options in a particular column.

For that some necessary libraries need to be imported such as matplotlib.pyplot and seaborn.

Plotting map for each column using .count plot and observe what all factors affect to hold its customers.

Also used the.value_counts () method to count the number of options in particular column.

Data visualization is necessary as it makes easier to identify the trends and patterns in data.also it is easier to find which features are responsible for customer retention.

After data visualization we will encode the data the dataset has number of string values .We will use encoding techniques .

See the correlation of columns with target column using df.corr() method.

And plotting heatmap and if the correlation value is greater than 0.5 than it is considered as greatest correlation with label.

Deleting the column which have less relation with column.

Separating features and label

Scale the data

Building Model starts