E-commerce Customer Segmentation

Abstract:

A key challenge for e-commerce businesses is to analyze the trend in the market to increase their sales. The trend can be easily observed if the companies can group the customers; based on their activity on the e-commerce site. This grouping can be done by applying different criteria like previous orders, mostly searched brands and so on.

Problem Statement:

Given the e-commerce data, use k-means clustering algorithm to cluster customers with similar interest.

Dataset Information:

The data was collected from a well known e-commerce website over a period of time based on the customer's search profile.

Variable Description:

Column	Description
Cust_ID	Unique numbering for customers
Gender	Gender of the customer
Orders	Number of orders placed by each customer in the past

Remaining 35 features (brands) contains the number of times customers have searched them

Scope:

Problem Statement – K means

- Analyzing the existing customer data and getting valuable insights about the purchase pattern
- Data pre-processing including missing value treatment
- Segmenting customer based on the optimum number of clusters ('k') with the help of silhouette score

Learning Outcome:

The students will get a better understanding of how the variables are linked to each other and will be able to apply cluster analysis to business problem such as customer segmentation.

Predicting Term Deposit Subscription by a client

Abstract:

Marketing campaigns are characterized by focusing on the customer needs and their overall satisfaction. Nevertheless, there are different variables that determine whether a marketing campaign will be successful or not. There are certain variables that we need to take into consideration when making a marketing campaign.

A Term deposit is a deposit that a bank or a financial institution offers with a fixed rate (often better than just opening a deposit account) in which your money will be returned back at a specific maturity time.

Problem Statement:

Predict if a customer subscribes to a term deposits or not, when contacted by a marketing agent, by understanding the different features and performing predictive analytics

Dataset Information:

The data is related with direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to assess if the product (bank term deposit) would be ('yes') or not ('no') subscribed.

The dataset consists of several predictor variables and one target variable, Outcome. Predictor variables includes the age, job, marital status, and so on

Variable Description:

Column	Description
age	Age of the client

Problem Statement – SVM

job	Type of job (categorical: 'admin.','blue-collar', 'entrepreneur', 'housemaid', 'management', 'retired','self-employed', 'services', 'student', 'technician', 'unemployed', 'unknown')
marital	Marital status (categorical: 'divorced','married','single','unknown'; note: 'divorced' means divorced or widowed)
education	(categorical: 'basic.4y', 'basic.6y' ,'basic.9y', 'high.school', 'illiterate', 'professional.course', 'university.degree','unknown')
default	has credit in default? (categorical: 'no','yes','unknown')
housing	has a housing loan? (categorical: 'no','yes','unknown')
loan	has a personal loan? (categorical: 'no','yes','unknown')
contact	contact communication type (categorical: 'cellular','telephone')
month	last contact month of year (categorical: 'jan', 'feb', 'mar', ..., 'nov', 'dec')
day_of_week	last contact day of the week (categorical: 'mon','tue','wed','thu','fri')
duration	last contact duration, in seconds
campaign	number of contacts performed during this campaign and for this client (numeric, includes last contact)
pdays	number of days that passed by after the client was last contacted from a previous campaign (numeric; 999 means client was not previously contacted)
previous	number of contacts performed before this campaign and for this client
poutcome	outcome of the previous marketing campaign (categorical: 'failure','nonexistent','success')
emp.var.rate	employment variation rate - quarterly indicator (numeric)
cons.price.idx	consumer price index - monthly indicator
cons.conf.idx	consumer confidence index - monthly indicator

euribor3m	euribor 3 month rate - daily indicator
nr.employed	number of employees - quarterly indicator
y	has the client subscribed a term deposit? (binary: 'yes','no')

Scope:

- Sentiment analysis in a variety of forms
- Data Pre-processing
- Training data using SVM
- Hyperparameter Tuning

Learning Outcome:

The students will get a better understanding of how the variables are linked to each other and should be able to predict whether a client of a bank will subscribe to a term deposit or not such that banks can plan the different campaigns for different customers according to the patterns.