



Advisor Name:  
Dr. Brian Letort  
  
Grand Canyon University

# Predicting customer behavior using machine learning :Customer Insight Segmentation App

Abadit Weldeclassie

the intuitive app empowers users to gain insights without technical expertise

### Context and background

- In today’s competitive market, businesses struggle to understand and predict customer behavior, which can lead to inefficient marketing strategies and reduced customer retention.
- Accurately predicting customer behavior is crucial for businesses to create personalized marketing strategies, optimize resource allocation, and ultimately enhance customer satisfaction and loyalty.
- The primary objective of our project is to develop a predictive model that segments customers based on their behavior, enabling businesses to implement targeted marketing strategies and improve overall customer engagement.

A screenshot of the app's input form. It contains several fields: 'How old are you?' with a slider from 1 to 89; 'Graduated' with a dropdown menu set to 'Yes'; 'Profession' with a dropdown menu set to 'Healthcare'; 'Work\_Experience' with a slider from 0.00 to 7.50; 'Spending\_Score' with a dropdown menu set to 'Low'; 'Family\_Size' with a slider from 1 to 9; and 'Var\_1' with a dropdown menu set to 'Cat\_1'.

### Customer Insight Segmentation App

This app predicts customer segments using a Random Forest Classifier, a powerful machine learning algorithm

- Data Exploration
- Data Visualization
- Input features
- Data Preparation
- Model Prediction
- Model Evaluation

### Methodology

**Algorithms used:** Random Forest, Logistic regression, support vector machine, K-NN and decision tree

**Data:** The model was trained on dataset of (10,695,11) of demographic and behavioral features.

**Data preparation:** data was cleaned, missing values were imputed, and features were encoded for model training."

**Visualization and Analysis:** Our app includes histograms for univariate analysis, bar graphs for value counts, and Plotly graphs for bivariate analysis."

### Model Performances

- Model performance is evaluated using accuracy scores and confusion matrices to ensure reliable predictions
- Random Forest was selected due to its robustness in handling diverse and large datasets. It reduces overfitting and improves prediction performance through ensemble learning.

### Impact and Benefits

- By leveraging these predictions, businesses can create targeted marketing campaigns and improve customer satisfaction

### References

- <https://www.kaggle.com/datasets/kaushiksuresh147/customer-segmentation/data>
- Sharma, Udit (May 29, 2024). OSEM Framework for Data Science. LinkedIn . <https://www.linkedin.com/pulse/osem-framework-data-science-pronounced-awesome-udit-sharma-26blc/>
- Kumar, Dhairya. (December 25, 2018). Introduction to data preprocessing in machine learning. Medium. . <https://towardsdatascience.com/introduction-to-data-preprocessing-in-machine-learning-a9fa83a5dc9d>
- Practical data science. A quick guide to customer segmentation for data scientists. Practical data science. . <https://practicaldatascience.co.uk/data-science/a-quick-guide-to-customer-segmentation>