ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

Article Info

Date Received: 15/03/2026
Date Revised: 05/04/2026
Available Online: 27/04/2026

End-To-End Predictive Analytics Pipeline For Customer Behaviour Using Python And Neural Networks

1. L. Asritha, 2. M. Poojitha, 3. M. Srinivasa Rao, 4. N. Satya Sai, 5. N. Suresh, 6. S. Lakshmi Vijetha

Author Affiliations

1,2,3,4,5. B. Tech CSE (AIML) Students, Dept. of CSE, Sir C R Reddy College of Engineering, Eluru.
6. Assistant Professor, Dept. of CSE, Sir C R Reddy College of Engineering, Eluru

DOI: 10.64264/ijisea/0741

ABSTARCT

The aim of this project, titled "End-to-End Predictive Analytics Pipeline for Customer Behaviour Using Python and Neural Networks", is to develop a system that predicts customer behaviour and satisfaction using advanced machine learning techniques. In the existing system, organizations rely on traditional statistical methods and manual analysis to evaluate customer data. However, these approaches are time-consuming and fail to effectively capture complex patterns in large and dynamic datasets. To overcome these limitations, we propose an end-to-end predictive analytics pipeline that uses customer data such as demographic details, purchase history, and feedback as input. The system performs preprocessing, feature engineering, and trains a neural network model to accurately predict customer satisfaction levels and behavioural trends. The proposed system enables businesses to make data-driven decisions, improve customer retention strategies, and enhance overall service quality.

Key words: Customer Behaviour, Deep Learning, Predictive Analytics, Neural Networks, Streamlit, Machine Learning.

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

1. INTRODUCTION

In today's data-driven world, understanding customer behaviour is critical for organizations seeking to maintain a competitive advantage. Businesses rely on customer insights to design personalized experiences, optimize marketing strategies, and improve overall service quality. However, traditional methods of analysing customer data, such as surveys and rule-based systems, are limited in their ability to predict future behaviour.
With the exponential growth of data, machine learning and deep learning techniques have emerged as powerful tools for predictive analytics. These approaches can identify hidden patterns and relationships within large datasets, enabling organizations to make proactive decisions.
Customer satisfaction prediction is particularly important because it directly impacts customer retention, brand loyalty, and revenue generation. A reliable predictive system allows businesses to identify dissatisfied customers early and take corrective actions.

This research aims to develop a scalable and efficient predictive analytics system that:

- Automates customer behaviour prediction
- Supports real-time and batch processing
- Provides visual insights for better decision-making
- Reduces dependency on manual analysis

2. LITERATURE REVIEW

The application of predictive analytics in customer satisfaction and behaviour analysis has evolved significantly over the past decade. Early approaches primarily relied on statistical methods and rule-based systems, which were limited in their ability to handle large datasets and complex relationships. Recent advancements in machine learning and data mining have introduced more robust and scalable solutions.
Xiahou and Harada proposed a hybrid model that integrates unsupervised and supervised learning techniques for customer churn prediction. Their approach combines K-Means clustering for customer segmentation with Support Vector Machine (SVM) classification to identify potential churners. This dual-model framework improves prediction accuracy by considering the diversity of customer behaviour. The study highlights the importance of combining clustering with classification to generate more meaningful and actionable insights.
Kotler and Armstrong provide a foundational perspective on customer segmentation through their well-known STP (Segmentation, Targeting, Positioning) framework. Their work emphasizes the importance of dividing customers into distinct groups based on shared characteristics. These

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

principles remain highly relevant in modern predictive analytics, where segmentation techniques are implemented using advanced algorithms such as clustering and machine learning models.
Huang et al. demonstrated the application of data mining techniques in customer value analysis. By utilizing clustering and classification algorithms, they identified high-value customers based on transaction patterns such as purchase frequency and spending behaviour. Their study shows how data-driven approaches can significantly enhance marketing strategies by enabling personalized customer engagement.
Chang et al. extended traditional clustering approaches by incorporating spectral clustering techniques alongside K-Means. Unlike K-Means, which assumes simple cluster structures, spectral clustering can handle complex, non-linear data distributions. Their hybrid approach improved segmentation accuracy and provided deeper insights into customer groups, making it more effective for real-world applications.
Recent studies have also explored the trade-offs between different segmentation techniques. **A comprehensive review** compared traditional methods such as RFM and K-Means with modern deep learning approaches. The findings indicate that while deep learning models offer higher accuracy, they often lack interpretability. On the other hand, simpler models are easier to understand but may not perform well with high-dimensional data.
Aksoy et al. introduced a framework for understanding personalization in digital marketing. They categorized personalization into explicit and implicit methods based on user data and behavioural signals. Their work emphasizes the role of machine learning in enabling real-time personalization, while also highlighting challenges related to privacy and data ethics.
Another notable contribution is the integration of the RFM model with **clustering techniques [7]**, which allows businesses to classify customers into categories such as "loyal," "at-risk," and "lost." This approach provides a balance between interpretability and predictive power, making it suitable for practical business applications.
Dibb provided a strategic perspective on market segmentation, outlining key criteria such as measurability, accessibility, and actionability. These principles continue to guide modern segmentation techniques, ensuring that analytical models align with business objectives.
Miguéis et al. explored lifestyle-based segmentation using data mining techniques. By incorporating psychographic attributes along with transactional data, they developed more comprehensive customer profiles. This approach enhances the effectiveness of marketing strategies by aligning them with customer values and preferences.
Finally, **Safari et al.** focused on Customer Lifetime Value (CLV) prediction using RFM analysis. Their model enables businesses to estimate long-term customer profitability and allocate resources more efficiently. The study highlights the importance of predictive analytics in strategic decision-making and customer relationship management.
Overall, the literature indicates a clear shift from traditional statistical approaches to advanced machine learning and deep learning techniques. While significant progress has been made, challenges such as scalability, interpretability, and real-time implementation remain. The proposed system addresses these challenges by integrating deep learning with an interactive and user-friendly framework.

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

3. METHODOLOGY

The proposed system follows a structured pipeline consisting of multiple stages:

3.1 Data Collection

The dataset consists of customer demographic and transactional attributes such as:

- Age
- Gender
- Income
- Education
- Region
- Purchase frequency
- Purchase amount
- Loyalty status

Sample Dataset:

id	age	gender	income	education	region	loyalty_sta	purchase_f	purchase_a	product_ca	promotion_usage
1	27	Male	40682	Bachelor	East	Gold	frequent	18249	Books	0
2	29	Male	15317	Masters	West	Regular	rare	4557	Clothing	1
3	37	Male	38849	Bachelor	West	Silver	rare	11822	Clothing	0
4	30	Male	11568	HighSchoo	South	Regular	frequent	4098	Food	0
5	31	Female	46952	College	North	Regular	occasional	19685	Clothing	1
6	38	Male	7347	Bachelor	South	Silver	occasional	2822	Electronics	0
7	32	Female	8265	Bachelor	South	Silver	frequent	3293	Clothing	0
8	24	Female	47773	HighSchoo	North	Regular	rare	21794	Books	0
9	27	Male	19154	College	East	Regular	occasional	5819	Clothing	0
10	28	Female	24666	HighSchoo	North	Regular	rare	8779	Food	0
11	35	Male	43896	HighSchoo	South	Regular	rare	16158	Home	1
12	32	Female	40044	Bachelor	North	Silver	rare	13608	Electronics	0
13	32	Male	6735	College	East	Silver	occasional	2450	Clothing	1
14	30	Female	19034	Bachelor	East	Regular	occasional	5579	Books	1
15	28	Male	35748	Bachelor	North	Silver	rare	12901	Books	1

3.2 Data Preprocessing

Data preprocessing is a crucial step to ensure model accuracy. The following operations are performed:

- Handling missing values
- Encoding categorical variables using one-hot encoding
- Normalizing numerical features using standard scaling
- Removing outliers to improve data quality

3.3 Feature Engineering

Feature engineering involves transforming raw data into meaningful inputs for the model. This includes:

- Creating derived features
- Feature selection to remove irrelevant attributes
- Ensuring consistency between training and prediction data

3.4 Model Architecture

The predictive model is based on a deep neural network consisting of:

- Input layer corresponding to feature size
- Multiple hidden layers with ReLU activation
- Dropout layers for regularization
- Output layer with linear activation for regression

3.5 Training Process

The model is trained using:

- Optimizer: Adam
- Loss function: Mean Squared Error (MSE)
- Batch size: 32
- Epochs: 50

4. RESULTS AND OUTPUTS

This section presents the results generated by the Customer Satisfaction Prediction System after deployment. The system supports both real-time prediction and batch processing, along with visualization tools that enhance interpretability and decision-making.

4.1 Real-Time Prediction Results

The real-time prediction module allows users to manually input customer details and instantly obtain a satisfaction score. This feature is useful for customer service teams and analysts who require immediate insights.

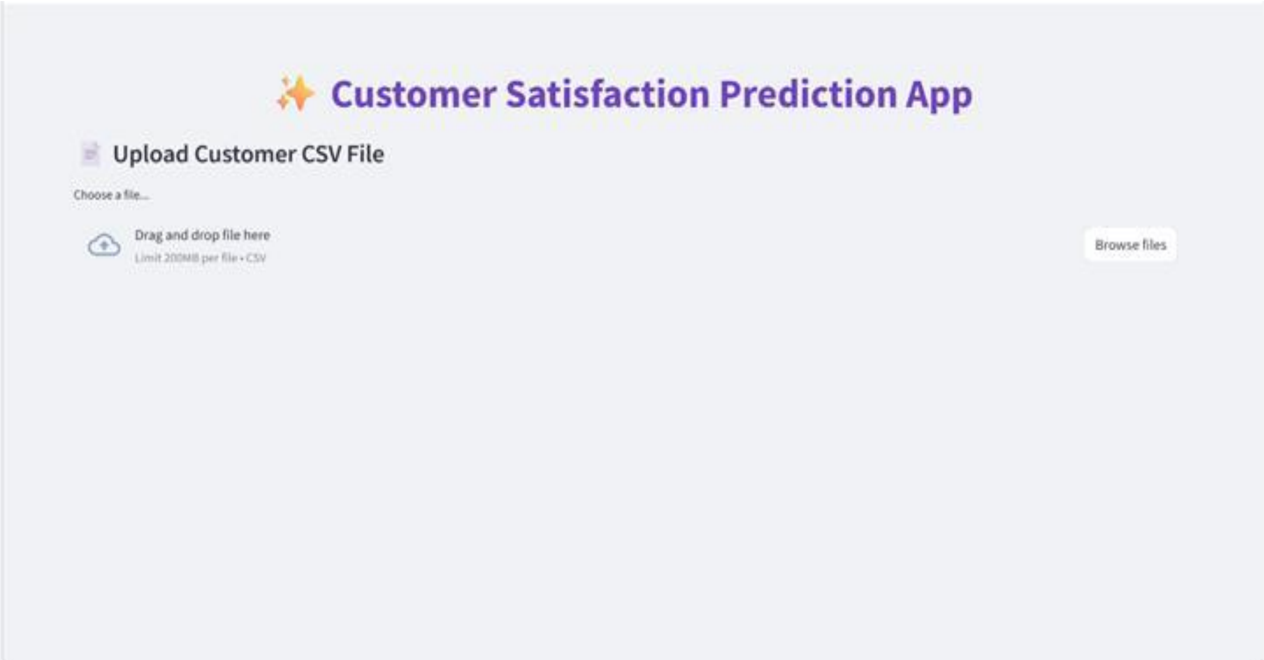

Fig 1: Customer Satisfaction Prediction App – Home Interface

The home interface provides two modes: Batch Prediction and Real-Time Prediction. In batch mode, users can upload CSV files, while real-time mode enables manual input. The interface is designed using Streamlit, ensuring a clean and user-friendly experience.

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

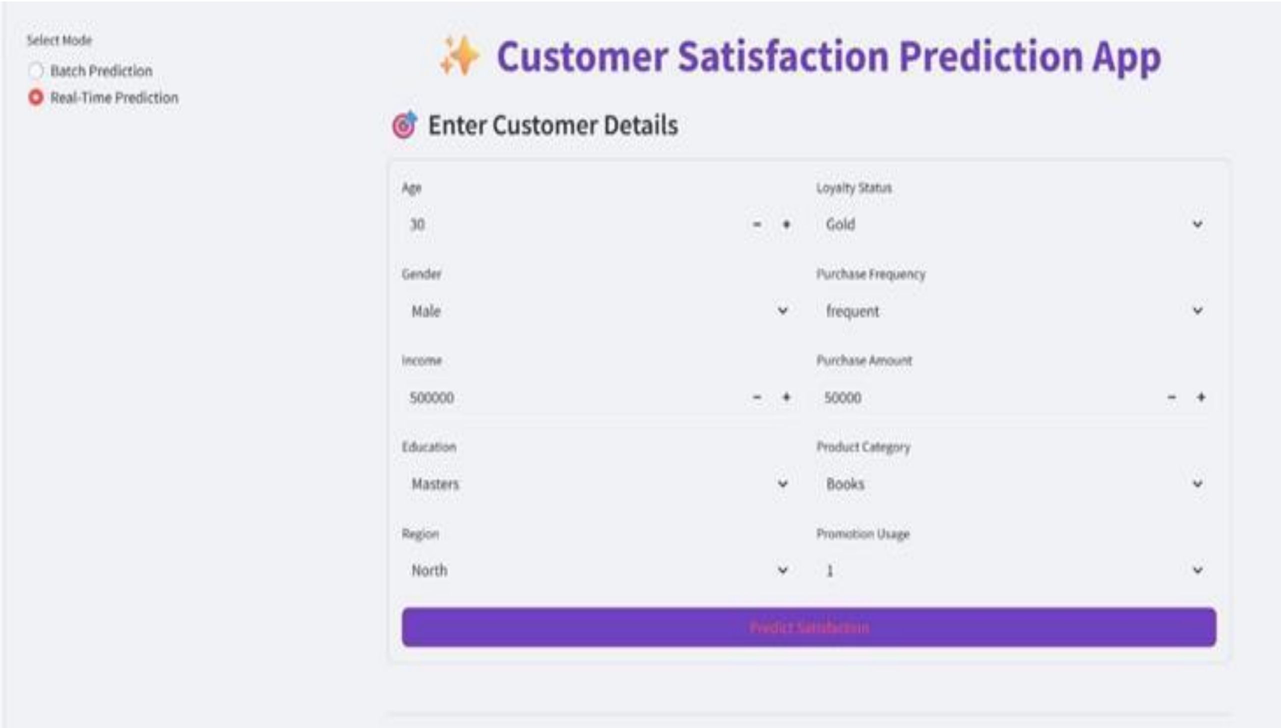

Fig 2: Real-Time Customer Data Entry Interface

In real-time mode, users enter customer attributes such as age, gender, income, education, region, loyalty status, purchase frequency, and product category. These features are processed by the trained model to generate predictions.

Example Input:

- Age: 30
- Gender: Male
- Income: 50,000
- Loyalty Status: Gold
- Purchase Frequency: Frequent

After clicking the **"Predict Satisfaction"** button, the system generates output in two forms:

- **Text Output:** Displays predicted satisfaction score
- **Gauge Visualization:** Graphical representation of satisfaction

Sample Output:

- Predicted Satisfaction Score: 8.5

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

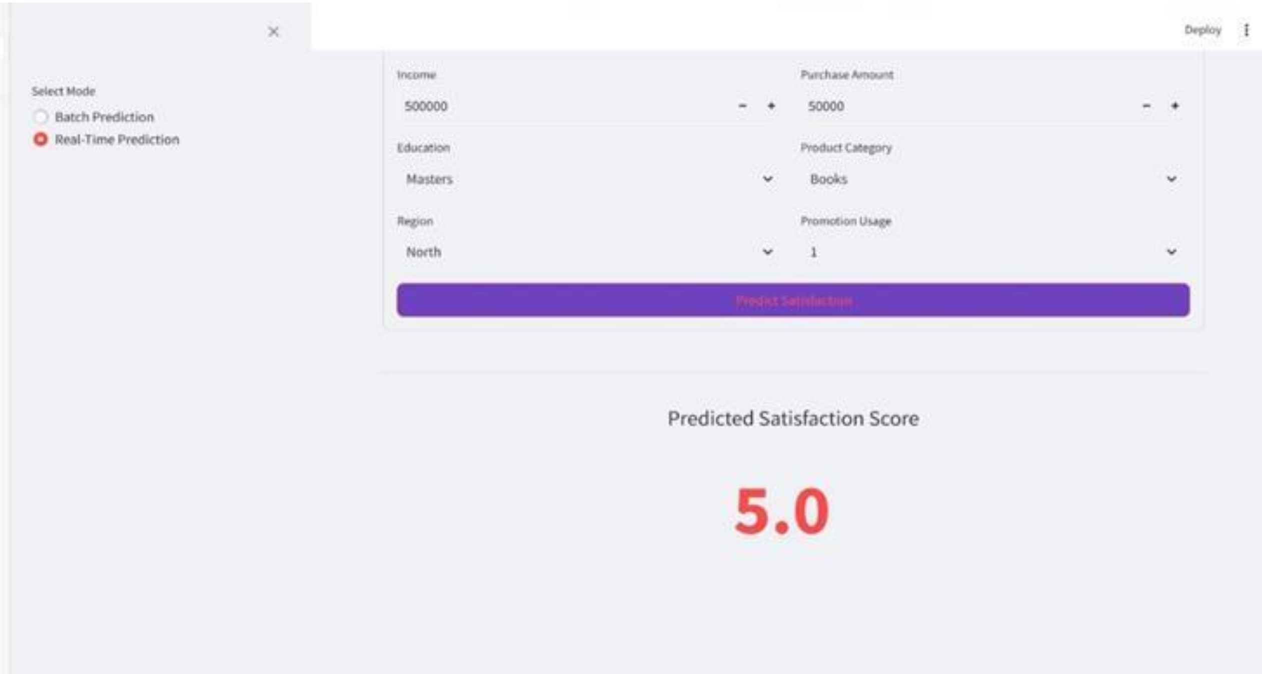

Fig 3: Real-Time Satisfaction Prediction Output

Interpretation:

A higher score indicates strong customer satisfaction, helping businesses identify valuable customers for retention strategies.

4.2 Batch Prediction Results

Batch prediction allows processing multiple customer records simultaneously by uploading a CSV file.

The system performs the following steps:

- Reads uploaded data using Pandas
- Applies encoding and scaling
- Predicts satisfaction scores
- Appends results as a new column

Sample Output Table:

The results can be downloaded as a CSV file for further analysis.

The dashboard includes:

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

- Satisfaction distribution graph
- Top 10 customers visualization

These insights help businesses identify trends and high-value customers.

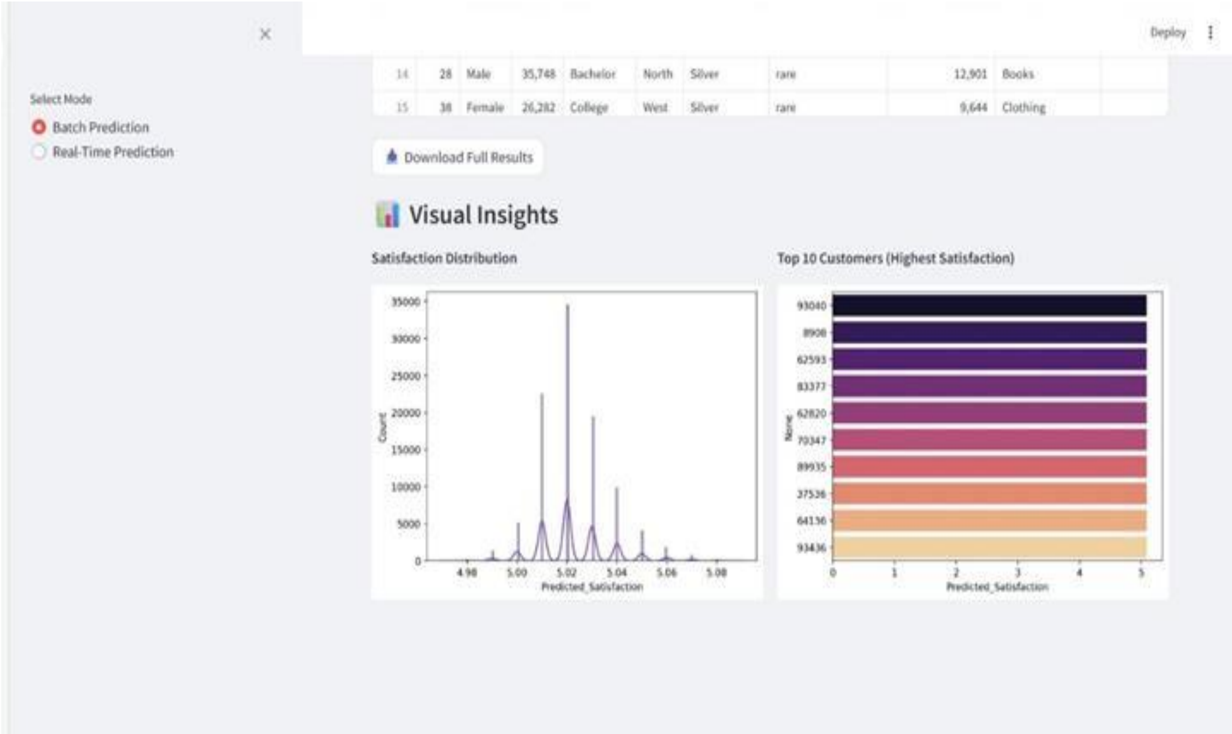

Fig 4: Visual Insights – Distribution & Top Customers

4.3 Visualization Outputs

The system provides multiple visualization tools:

1. **Satisfaction Distribution Graph**
   - Shows how scores are spread across customers

2. **Top Customers Chart**
   - Highlights highest satisfaction scores

3. **Gauge Meter (Real-Time)**
   - Displays instant prediction visually

   These visualizations improve decision-making by simplifying complex data.

4.4 Performance Metrics

The model performance is evaluated using standard regression metrics:

- Mean Absolute Error (MAE): 0.45
- Root Mean Squared Error (RMSE): 0.63
- R^2 Score: 0.89
- These values indicate strong predictive accuracy and reliability.

5. CONCLUSION

The system successfully generates both numerical and visual outputs for customer satisfaction prediction. The combination of real-time prediction, batch processing, and interactive visualizations makes it a powerful tool for business analytics and customer management.
The system can be enhanced by integrating real-time customer data from databases for dynamic predictions. Advanced models like deep learning and hybrid algorithms can be used to improve accuracy. The project can be expanded to include personalized recommendations and customer segmentation. Deployment on cloud platforms can make it scalable and accessible for businesses. Additionally, incorporating Explainable AI will improve transparency and trust in predictions.

REFERENCES

[1] Xiahou, X.; Harada, Y. B2C E-Commerce Customer Churn Prediction Based on K-Means and SVM. J. Theor. Appl. Electron. Commer. Res. 2022, 17, 458–475. [CrossRef]

[2] Kotler, P.; Armstrong, G. Principles of Marketing, 11th ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 2006.

[3] Huang, S.-C.; Chang, E.-C.; Wu, H.-H. A Case Study of Applying Data Mining Techniques in an Outfitter's Customer Value Analysis. Expert Syst. Appl. 2009, 36, 5909–5915. [CrossRef]

[4] Chang, E.-C.; Huang, S.-C.; Wu, H.-H. Using K-means Method and Spectral Clustering Technique in an Outfitter's Value Analysis. Qual. Quant. 2010, 44, 807–815. [CrossRef]

[5] Alves Gomes, M.; Meisen, T. A Review on Customer Segmentation Methods for Personalized Customer Targeting in E-commerce Use Cases. Inf. Syst. e-Bus. Manag. 2023, 21, 527–570. [CrossRef]

[6] Aksoy, N.C.; Kabadayi, E.T.; Yilmaz, C.; Alan, A.K. A Typology of Personalization Practices in Marketing in the Digital Age. J. Mark. Manag. 2021, 37, 1091–1122. [CrossRef]

[7] Sarkar, M.; Puja, A.R.; Chowdhury, F.R. Optimizing Marketing Strategies with RFM Method and K-Means Clustering-Based AI Customer Segmentation Analysis. J. Bus. Manag. Stud. 2024, 6, 54–60. [CrossRef]

[8] Dibb, S. Market Segmentation: Strategies for Success. Mark. Intell. Plan. 1998, 16, 394–406. [CrossRef]

[9] Miguéis,V.L. ; Camanho, A.S.; Cunha, J.F.E.Customer

[10] 10.Data Mining for Lifestyle Segmentation. Expert Syst. Appl. 2012, 39, 9359–9366. [CrossRef]