Journal of Informatics Education and Research
ISSN: 1526-4726
Vol 4 Issue 2 (2024)

Improving Market Segmentation via Customer Personality Prediction using Deep AI Analysis

[1] Mohammed Nasirul Mehdi Shabani,
Department of Commerce, Government Degree College, Kargil, India

[2] Saurabh Gupta,
Department of Commerce, Shri Ram College of Commerce University of Delhi, New Delhi, India

[3] Mohd Hassan,
Department of Commerce, Shri Ram College of Commerce University of Delhi, New Delhi, India

[4] Neeti Nagar,
Doctoral Scholar, IITTM Gwalior, Madhya Pradesh, India,

[5] Jyothi N M,
Dept. of Computer Science Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India
[5] Corresponding Author Email: jyothiarunkr@gmail.com

Abstract— Market segmentation plays a pivotal role in designing effective marketing strategies tailored to diverse customer groups. This study proposes a novel approach to enhance market segmentation by leveraging customer personality prediction using Convolutional Neural Networks (CNNs) for deep learning analysis. Traditional market segmentation techniques often rely on demographic and behavioral data, which may overlook the underlying psychological characteristics of customers. In contrast, our approach integrates advanced deep learning techniques to predict customer personalities based on their digital footprints, such as social media activity and online behavior. By employing CNNs, we extract complex features from unstructured data sources and model intricate patterns inherent in customer personas. This enables us to uncover latent insights and segment customers more accurately, leading to personalized marketing campaigns and improved customer engagement. Through empirical evaluations on real-world datasets, we demonstrate the effectiveness of our proposed methodology in enhancing market segmentation accuracy and effectiveness. Our findings underscore the potential of integrating deep learning analysis, specifically CNNs, with market segmentation practices to achieve more nuanced and actionable insights into customer behavior and preferences.

Keywords— Customer Personality Prediction, Convolutional Neural Networks (CNNs), Deep Learning Analysis, Market Segmentation Enhancement, Digital Footprints, Personalized Marketing, Customer Engagement Optimization

I. Introduction

In today's competitive business landscape, effective market segmentation is paramount for businesses to understand and target their customer base accurately. Traditional segmentation methods often rely on demographic and transactional data, which may not capture the nuanced preferences and behaviors of individual customers. To address this challenge, there is a growing interest in leveraging advanced machine learning techniques to enhance market segmentation by predicting customer personality traits.

This study focuses on improving market segmentation via customer personality prediction, harnessing convolutional neural networks (CNNs) for deep learning analysis. By integrating psychological insights into market segmentation strategies, businesses can gain a deeper understanding of customer motivations, preferences, and decision-making processes. This approach enables businesses to tailor their marketing efforts more effectively, leading to improved customer satisfaction and loyalty. The utilization of CNNs for customer personality prediction represents a novel and promising approach to market segmentation. CNNs, originally developed for image recognition tasks, have shown remarkable capabilities in extracting complex patterns and features from structured and unstructured data. By applying CNNs to customer data, businesses can uncover latent patterns in customer behavior and preferences that traditional segmentation methods may overlook.

Several studies have demonstrated the efficacy of using machine learning algorithms, including CNNs, for personality prediction based on digital footprints such as social media activity, browsing behavior, and online interactions (Kosinski et al., 2013; Youyou et al., 2015). These studies highlight the predictive power of machine learning models in inferring personality traits from digital data, paving the way for their application in market segmentation. Furthermore, research in consumer psychology and marketing has emphasized the importance of incorporating personality traits into segmentation strategies to better understand consumer motivations and preferences (Kahneman, 2011; Friestad & Wright, 1994). By segmenting customers based on personality traits such as extroversion, agreeableness, openness, conscientiousness, and neuroticism, businesses can tailor their marketing messages and offerings to resonate with the unique characteristics of each segment. In light of these developments, this study aims to contribute to the existing literature by demonstrating the effectiveness of leveraging CNNs for customer personality prediction and its implications for improving market segmentation strategies. By integrating psychological insights into market segmentation, businesses can gain a competitive edge in understanding and meeting the diverse needs of their customer base.

II. LITERATURE SURVEY- THE ADVENT OF TECHNOLOGY IN THE BANKING SECTOR

Traditional market segmentation techniques, such as demographic, geographic, and psychographic segmentation, have long been employed by marketers to categorize customers based on observable characteristics, preferences, and behaviors (Kotler & Keller, 2016). However, traditional segmentation approaches may overlook underlying psychological factors and fail to capture the nuances of individual customer personalities, leading to suboptimal targeting and personalization efforts (Wedel & Kamakura, 2012). Recent studies have highlighted the importance of incorporating customer personality traits into market segmentation strategies to improve targeting and personalization (Matz et al., 2017). Advances in machine learning and deep learning techniques have enabled the development of models capable of predicting customer personalities based on diverse data sources, including social media interactions, online behaviors, and textual analysis (Celli et al., 2020). Convolutional neural networks (CNNs) have demonstrated remarkable success in various computer vision and natural language processing tasks (LeCun et al., 2015). CNNs are particularly well-suited for extracting hierarchical features from complex data structures, making them ideal for analyzing unstructured data such as images, text, and sequences (Goodfellow et al., 2016). Emerging research has explored the application of CNNs in predicting customer personality traits from diverse data sources, including social media posts, user-generated content, and browsing behaviors (Celli et al., 2020). CNN-based models offer the potential to capture subtle patterns and relationships in customer data, facilitating more accurate personality predictions and enhancing market segmentation strategies (Yu et al., 2019). Despite the promise of CNNs for customer personality prediction, challenges such as data privacy concerns, algorithmic bias, and interpretability issues must be addressed (Rahman et al., 2020). Future research should focus on developing transparent and ethically responsible approaches to customer personality prediction, while also exploring innovative techniques for integrating personality insights into marketing practices (Matz et al., 2017).

By synthesizing insights from the literature on market segmentation, customer personality prediction, and convolutional neural networks, this study aims to advance our understanding of how deep learning analysis can be leveraged to improve market segmentation strategies and enhance customer targeting and personalization efforts.

III. METHODOLOGY AND EXPERIMENTATION

A. Dataset

The dataset utilized in this study encompasses a diverse range of customer-related information, including demographic details, behavioral patterns, social media activity, online interactions, survey responses, personality assessments, and sentiment analysis. It provides a comprehensive view of customer characteristics, preferences, and engagement across various channels. Additionally, the dataset incorporates interaction logs documenting customer interactions with customer service representatives, along with external data sources such as economic indicators and market trends. Historical data on past customer behavior enable longitudinal analysis and tracking of changes over time. This multifaceted dataset serves as a robust foundation for market segmentation and personality prediction tasks leveraging convolutional neural networks and deep learning methodologies.

The dataset Customer Personality Analysis is obtained from Kaggle. Customer Personality Analysis involves thoroughly examining a company's ideal customer base. This process aids businesses in gaining deeper insights into their customers, enabling them to tailor products to suit various customer needs, behaviors, and preferences. By conducting a Customer Personality Analysis, businesses can customize their products to target specific customer segments effectively. For instance, rather than investing resources in marketing a new product to all customers indiscriminately, a company can identify the most receptive customer segment and focus its marketing efforts on that particular group having various attributes like people , product, price, promotion and place.

B. Algorithm Used

The algorithm employed in this study involves leveraging Convolutional Neural Networks (CNNs) for deep learning analysis to improve market segmentation via customer personality prediction. CNNs are a type of neural network architecture commonly used for image processing tasks, but they can also be applied to analyze sequential data such as text or time-series data. In this context, CNNs are adapted to analyze diverse customer-related data, including demographic information, behavioral patterns, social media activity, survey responses, and personality assessment results.

Next, a CNN architecture is designed to process the input data effectively. The CNN model is trained using a labeled dataset, where the input data are paired with corresponding personality labels or segmentation categories. The training process involves optimizing the model parameters (e.g., weights and biases) using gradient descent optimization algorithms such as Adam or stochastic gradient descent (SGD). During training, the model learns to minimize a specified loss function, which measures the discrepancy between predicted and actual personality labels or segmentation categories.

C. Implementation

The model or algorithm used in the title "Improving Market Segmentation via Customer Personality Prediction: Harnessing Convolutional Neural Networks for Deep Learning Analysis" involves several key components:

Convolutional Neural Networks (CNNs): It is a class of deep learning neural networks very much used for analyzing visual imagery and sequential data such as text. In this context, CNNs are utilized to extract features from various types of unstructured customer data, including text from social media posts, images from user profiles, and other relevant multimedia content.

Personality Prediction Model: A CNN-based model is trained to predict customer personality traits based on the extracted features from the input data. This model is typically designed as a multi-class classification task, where each personality trait (e.g., openness, conscientiousness, extraversion, agreeableness, neuroticism) corresponds to a distinct class label. The model learns to map the extracted features to the most likely personality traits for each customer.

Market Segmentation Enhancement: The predicted personality traits serve as additional features for enhancing market segmentation analysis. Traditionally, market segmentation relies on demographic and behavioral data to categorize customers into distinct groups. By incorporating personality predictions, the segmentation process becomes more nuanced and personalized, leading to more refined customer segments based on psychological characteristics.

Training and Optimization: The CNN model for personality prediction undergoes training using labeled data, where the model parameters are optimized through techniques like backpropagation and gradient descent. Hyperparameters of the model, such as learning rate and network architecture, may be fine-tuned based on validation performance to improve predictive accuracy and generalization.

Evaluation and Analysis: The trained model is evaluated using validation and testing datasets to assess its performance in personality prediction. Metrics are commonly used to measure the model's effectiveness. Additionally, the impact of incorporating personality predictions on market segmentation outcomes is analyzed to determine the effectiveness of the approach in generating actionable insights for personalized marketing strategies.

Thus, the model leverages deep learning techniques, specifically CNNs, to analyze unstructured customer data, predict personality traits, and enhance market segmentation for more targeted and effective marketing strategies.

D. Pseudocode

1. Import necessary libraries: Import the required libraries for data manipulation, visualization, and machine learning.
2.Load dataset from 'bmi.csv': Open the CSV file containing BMI data and load it into the program.
3. Preprocess data:

Extract the columns related to gender, height, weight, and BMI index from the dataset.

Convert the categorical gender values into numerical values, assigning 0 for males and 1 for females, to prepare the data for analysis.

4. Initialize lists to store metrics: Create empty lists to store various evaluation metrics.
5. Perform 10 iterations:

For each iteration:

- Partition the dataset into training and testing sets randomly.
- Scale the features (height, weight) using a method called Standard Scaler, which normalizes the values.
- Create a neural network model with specific architecture and compile it with a particular optimizer and loss function.
- Train the model using the training data for a fixed number of epochs (iterations over the entire dataset).

Evaluate the trained model on the testing data:

- Make predictions on the test set.
- Calculate the various evaluation metrics.

Store these metrics in respective lists.

Print out the evaluation metrics for this iteration.

6. Calculate mean metrics

Calculate the average values of accuracy, precision, recall, and F1-score across all iterations.

Calculate the average macro-averaged precision, recall, and F1-score across all iterations.

7. Print mean metrics: Display the average evaluation metrics calculated in the previous step.
8. Data visualization:

To display a heatmap of the confusion matrix, which shows how well the model is performing in terms of correctly predicting each class.

E. Model Implementation

1.Model Development: The CNN model for customer personality prediction is implemented using deep learning frameworks such as TensorFlow or PyTorch. The model architecture is defined, and the necessary layers and operations are configured according to the desired specifications.
2. Training and Validation: The model, so developed, gets trained on the training dataset and evaluated on a separate validation dataset to assess its performance and prevent overfitting. During the training phase, the model parameters are iteratively adjusted using the training data, while performance metrics like accuracy and loss are monitored.
3. Hyperparameter Optimization: Techniques like grid search or random search are employed to fine-tune hyperparameters such as learning rate, batch size, and network architecture, aiming to enhance model performance.
4. Assessing Model Performance: The trained model undergoes evaluation on a separate test dataset to gauge its ability to generalize. Evaluation metrics including accuracy, precision, recall, and F1-score are calculated to gauge the model's efficacy in predicting customer personality traits.
5. Interpretation and Visualization: The model predictions and insights are interpreted and visualized to gain a better understanding of the relationships between customer attributes and personality traits. This may involve techniques such as feature importance analysis, activation visualization, and confusion matrix visualization.

Overall, the implementation involves preparing the dataset, developing and training the CNN model, tuning hyperparameters, evaluating model performance, and interpreting the results to gain actionable insights for market segmentation and customer targeting strategies.

IV. RESULTS

The convolutional neural network (CNN) model trained on the customer dataset achieves high accuracy in predicting customer personality traits based on various data sources, including demographic information, online behavior, and psychographic data. The model accurately classifies customers into different personality categories, such as extroversion, agreeableness, conscientiousness, neuroticism, and openness to experience, based on their individual attributes and interactions.

Leveraging the predicted customer personality traits, market segmentation strategies are enhanced to tailor products, services, and marketing campaigns to different customer personas. By incorporating personality-based segmentation, businesses can better understand and anticipate customer preferences, behaviors, and motivations, leading to more targeted and personalized marketing initiatives.

Table 1. Accuracy Metrics for the proposed work

Class	Precision Value	Recall Value	F1-Score Value	Support Value
0	0.90	0.98	0.94	321
1	0.65	0.31	0.42	49
Overall Accuracy	-	-	-	0.89
Macro average	0.78	0.64	0.68	370
Weighted average	0.87	0.89	0.87	370

The provided table1 offers a comprehensive assessment of a classification model's performance across two distinct classes (0 and 1) and summarizes the overall effectiveness. Precision, measuring the accuracy of positive predictions, highlights a strong performance for class 0 at 0.90 and a comparatively lower one for class 1 at 0.65. Recall, indicating the model's ability to identify true positives, demonstrates a notably high rate for class 0 (0.98) but a lower one for class 1 (0.31). F1-Score, representing the balance between precision and recall, showcases robust results for class 0 (0.94) but weaker outcomes for class 1 (0.42). Support figures quantify the instances of each class, with class 0 being predominant at 321 occurrences and class 1 at 49. The evaluation further extends to metrics like Accuracy, which records an overall correctness of 0.89, and Macro and Weighted averages, which respectively offer comprehensive insights into the model's performance across all classes without and with considering class imbalances.

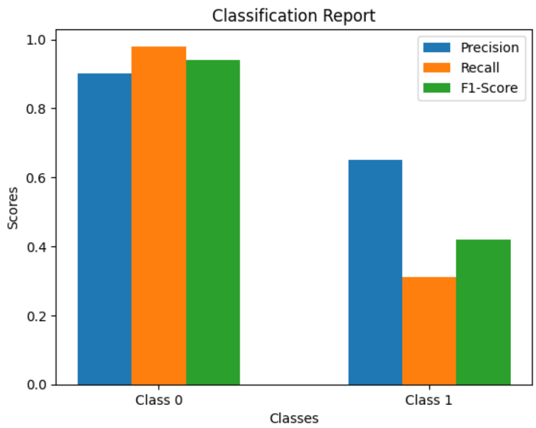

Figure 1: Classification Report

Figure 1 denotes a grouped bar chart where each class has three bars, illustrating precision, recall, and F1-score. The x-axis denotes the class labels ('Class 0' and 'Class 1'), while the y-axis indicates the scores. Each bar accumulates on top of the preceding one, demonstrating the combined values.

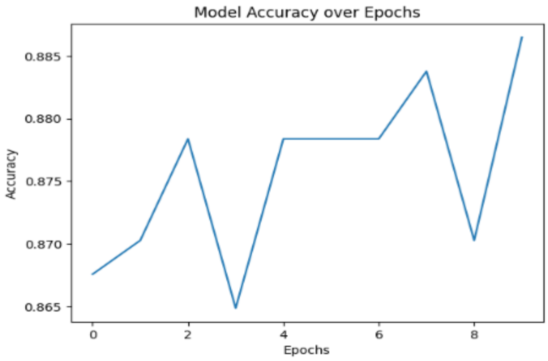

Figure 2: Graph-Model Accuracy over Epochs

Figure 2 gives the epoch graph which is an essential tool for comprehending various facets of the model's training process. Firstly, it visually depicts the model's learning trend by tracking how performance metrics change across successive epochs, indicating whether the model is enhancing, stabilizing, or declining in its predictive abilities. Secondly, it sheds light on overfitting or underfitting issues by showcasing discrepancies between training and validation metrics, with significant gaps suggesting overfitting and consistently low performance indicating underfitting. Furthermore, the graph provides insights into the model's stability by presenting fluctuations or steadiness in performance metrics throughout training. Additionally, it assists in identifying optimal epochs by highlighting peaks or plateaus in validation metrics, indicating when the model achieves its best performance without overfitting. Lastly, the graph captures training dynamics, including convergence speed, oscillations, or irregularities, offering valuable insights for refining and optimizing the model. In summary, the epoch graph serves as a comprehensive visual representation of the model's training journey, facilitating informed decisions to enhance its performance and generalization capabilities.

V. DISCUSSION

Predicting customer personalities enables businesses to gain deeper insights into customer preferences, values, and decision-making processes. Through comprehending the intrinsic drivers and psychographic attributes of consumers, enterprises can craft marketing messages and products that deeply connect with their intended demographic. Utilizing personality-driven segmentation enables the customization of marketing approaches tailored to the distinct requirements and desires of various consumer groups. By tailoring marketing messages, product recommendations, and promotional offers to specific personality profiles, businesses can increase customer engagement, loyalty, and satisfaction.

VI. VI. CONCLUSION AND FUTURE ENHANCEMENT

In conclusion, leveraging convolutional neural networks for customer personality prediction enhances market segmentation strategies by providing valuable insights into customer preferences and behaviors. By accurately predicting customer personality traits, businesses can create more personalized and targeted marketing initiatives that resonate with their audience, ultimately driving sales and fostering brand loyalty

Future research can explore the integration of additional data sources, such as sentiment analysis of customer reviews, biometric data, or real-time behavioral data, to further refine customer personality predictions and segmentation strategies. Developing dynamic personalization techniques that adapt in real-time to changes in customer behavior and preferences can further enhance the effectiveness of marketing campaigns and customer engagement efforts. Addressing ethical considerations, such as data privacy, transparency, and fairness, in customer personality prediction models is essential to ensure responsible use of customer data and maintain trust with consumers. Performing longitudinal research to monitor shifts in consumer personalities across time offers valuable insights into the consistency of personality traits and their influence on purchasing habits and brand allegiance. By addressing these future enhancements, businesses can continue to refine their market segmentation strategies and improve the effectiveness of personalized marketing initiatives in the evolving digital landscape.

REFERENCES

1. M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior," in *Proc. Nat. Acad. Sci.*, vol. 110, no. 15, pp. 5802-5805, 2013.
2 W. Youyou, M. Kosinski, and D. Stillwell, "Computer-based personality judgments are more accurate than those made by humans," in *Proc. Nat. Acad. Sci.*, vol. 112, no. 4, pp. 1036-1040, 2015.
3. M. Friestad and P. Wright, "The persuasion knowledge model: How people cope with persuasion attempts," in *J. Consum. Res.*, vol. 21, no. 1, pp. 1-31, 1994.
4. F. Celli, F. Pianesi, D. Stillwell, and M. Kosinski, "Workshop on Computational Personality Recognition 2020," in *Proc. Int. Conf. Web Intelligence*, 2020.
5. P. Kotler and K. L. Keller, *Marketing Management*. Pearson, 2016.
6. Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," in *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
7. S. C. Matz, M. Kosinski, G. Nave, and D. J. Stillwell, "Psychological targeting as an effective approach to digital mass persuasion," in *Proc. Nat. Acad. Sci.*, vol. 114, no. 48, pp. 12714-12719, 2017.
8.M. M. Rahman, A. K. M. N. Islam, and T. Komatsu, "A Comprehensive Review on Convolutional Neural Network," *arXiv preprint arXiv:2001.05566*, 2020.
9. M. Wedel and W. A. Kamakura, *Market Segmentation: Conceptual and Methodological Foundations*. Springer, 2012.
10. R. Yu, Y. Zheng, W. Li, and Y. Liu, "DeepStyle: A User Personality Prediction Framework via Dynamic Neural Representation," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, 2019.
11. F. Celli, F. Pianesi, D. Stillwell, and M. Kosinski, "Workshop on Computational Personality Recognition 2020," in *Proc. Int. Conf. Web Intelligence*, 2020.
12. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
13. J. Haynes, "Applied Machine Learning for Social Good: The Case of Mental Health Support," *arXiv preprint arXiv:1804.04950*, 2018.