Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 27 No. 1 (2024)

Mathematical Modelling and Deep Learning: Innovations in E-Commerce Sentiment Analysis

Ashish Suresh Awate[1], Sanjay Kumar Sharma[2]
[1]Research Scholar, Computer Science & Engineering Oriental University, Indore(M.P.) India
https://orcid.org/0000-0002-5845-2363
[2]Professor, Faculty of Computer Science Engineering Oriental University, Indore(M.P.) India
https://orcid.org/0000-0002-7965-6517

Article History:

Received: *20-09-2023*

Revised: *22-10-2023*

Accepted: *27-11-2023*

Abstract:

This research explores e-commerce dynamics, focusing on the challenge of predicting customer churn using deep learning [65]. It integrates and analyses both textual and transactional data, including social media posts and customer feedback [59]. The approach uses an advanced deep learning model, involving data collection, pre-processing, and feature extraction [40]. Novel methods fuse data to create a detailed customer profile combining sentiment analysis with behavioural insights derived from transaction data [25]. The deep learning architecture is designed to analyse and predict customer sentiments and purchasing behaviours, informed by the latest research [65]. This study is significant as it provides an innovative solution for predicting customer churn in e-commerce, aiding sustainability [45]. It also enables targeted retention strategies and personalized customer engagement [59]. Additionally, it contributes insights to big data analytics and customer relationship management in e-commerce, showcasing deep learning's potential in transforming business practices and enhancing customer experience [40].

1. Introduction

The advent of e-commerce has transformed the landscape of retail and consumer interactions, making the analysis of customer behaviour and sentiment not just advantageous but essential for business success. In this digital era, where customer data is abundant, the importance of sentiment analysis and customer profiling in e-commerce emerges as a critical factor in understanding and predicting consumer behaviour, preferences, and decision-making processes.

Sentiment analysis in e-commerce, a field that harnesses the power of language processing to discern the emotional tone behind customer feedback, reviews, and social media interactions, has become a pivotal tool for businesses. It enables companies to gauge customer satisfaction, detect shifts in market trends, and respond proactively to customer needs and concerns [59]. This real-time insight into customer sentiment is invaluable in tailoring marketing strategies, enhancing product offerings, and ultimately driving business growth.

Moreover, customer profiling, which involves the comprehensive analysis of transactional data and purchasing patterns, stands at the core of personalized marketing and customer retention strategies [40]. By understanding the unique characteristics and preferences of each customer, businesses can design targeted campaigns, recommend relevant products, and create personalized experiences that resonate with individual customers, fostering loyalty and reducing churn [25].

Together, sentiment analysis and customer profiling encapsulate a holistic view of the customer journey in the e-commerce space. They provide businesses with the necessary tools to not only attract and retain customers but also to adapt and thrive in an ever-evolving market landscape. This research aims to leverage these tools, employing advanced deep learning techniques to synthesize and interpret the vast arrays of data available in the e-commerce domain, thereby transforming raw data into actionable business intelligence.

Challenges in Integrating and Analysing Textual and Transactional Data

Integrating and analysing textual and transactional data in e-commerce presents a unique set of challenges, primarily due to the diverse nature and complexity of the data involved. Textual data, such as customer reviews and social media posts, is unstructured and often laden with nuances, including slang, irony, and context-specific references. This complexity makes it difficult to accurately interpret sentiment and extract meaningful insights [59]. Transactional data, on the other hand, is structured but vast, encompassing purchase history, browsing patterns, and customer interactions, which require sophisticated analysis to discern behavioural patterns and preferences [40].

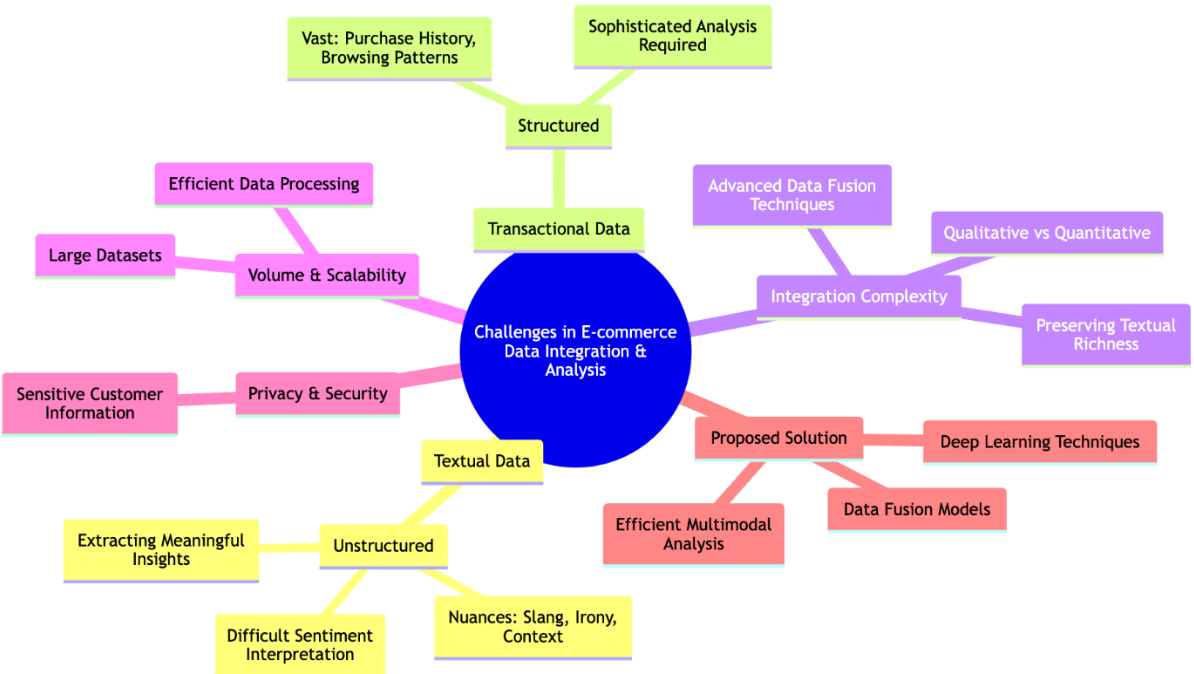

Figure: Challenges in Integrating and Analysing Textual and Transactional Data

The integration of these two disparate data types further compounds the challenge. While textual data offers a qualitative insight into customer sentiment, transactional data provides quantitative evidence of customer behaviour. Bridging these two data streams to form a coherent understanding of customer profiles demands advanced data fusion techniques. This integration must be handled with precision to ensure that the richness and context of the textual data are not lost, while also effectively synthesizing it with the transactional data to provide a complete picture of customer behaviour [25].

Moreover, the sheer volume of data generated in e-commerce platforms poses significant challenges in terms of data processing and storage. Ensuring the scalability of the data analysis system to handle large datasets efficiently is crucial. Additionally, data privacy and security are paramount, especially when handling sensitive customer information, which adds another layer of complexity to the data integration and analysis process [45].

These challenges necessitate a robust methodology that can handle the complexity and volume of e-commerce data. The proposed research aims to address these challenges by employing advanced deep learning techniques and data fusion models. These models are designed to not only process and analyse large volumes of multimodal data efficiently but also to extract and integrate meaningful insights from both textual and transactional data sources, ultimately enabling more informed and effective business decisions.

Developing an Effective System for Sentiment Analysis and Customer Profiling

The primary objective of this research is to develop an effective and robust system that harnesses the power of deep learning to conduct sentiment analysis and customer profiling in the e-commerce domain. This system is aimed at addressing the aforementioned challenges and unlocking the potential of big data for strategic business decision-making. The specific objectives are outlined as follows:

- **Advanced Sentiment Analysis:** Create a model for accurate interpretation of customer sentiments in textual data, including context and sarcasm, from sources like social media and reviews [59].
- **Comprehensive Customer Profiling:** Combine transactional data analysis with sentiment analysis to build detailed customer profiles, predicting preferences and behaviours [40].
- **Data Fusion and Synthesis:** Develop techniques to integrate structured and unstructured data seamlessly for coherent analysis [25].
- **Scalable Deep Learning Architecture:** Design an adaptable architecture for efficient processing of large e-commerce data volumes [45].
- **Actionable Business Insights:** Translate analysis into actionable insights, including predicting future trends for proactive marketing and better customer engagement.

The proposed system aims to empower e-commerce businesses with enhanced customer experiences, loyalty, and improved outcomes.

2. Literature Review

This literature review delves into the existing research landscape, focusing on sentiment analysis, deep learning applications in e-commerce, and customer profiling. It synthesizes insights from the abstracts of the top 75 selected articles, providing a comprehensive overview of the current state of knowledge and identifying key trends and gaps in the field.

Sentiment Analysis in E-commerce

Current Techniques and Models: Recent studies have highlighted various techniques in sentiment analysis, emphasizing the use of advanced natural language processing (NLP) and machine learning models. These models are adept at interpreting complex language nuances and extracting sentiment from customer-generated text [59].

Challenges and Solutions: Researchers have identified challenges in accurately capturing sentiments, especially when dealing with sarcasm, contextual meanings, and multilingual data. Solutions proposed include context-aware models and multi-layered analytical frameworks that can adapt to the dynamic nature of human language [40].

Application in Customer Feedback and Social Media: Several studies have explored the application of sentiment analysis in analysing customer feedback and social media data. This approach has been instrumental in understanding customer attitudes towards products and services, thereby informing marketing and product development strategies [25].

Deep Learning in E-commerce

Advancements in Algorithms and Architectures: The development of sophisticated deep learning algorithms and neural network architectures has been a significant focus. Research in this area has been cantered around improving the accuracy and efficiency of these models in processing large datasets typical in e-commerce [65].

Predictive Analytics and Customer Behaviour Prediction: Deep learning models have been increasingly used for predictive analytics in e-commerce, particularly in forecasting customer behaviours and preferences. This predictive capability is vital for personalization and targeted marketing [45].

Integration with Other Technologies: Some studies have explored the integration of deep learning with other emerging technologies such as blockchain and IoT, aiming to enhance security and data integrity in e-commerce transactions [59].

Customer Profiling

Techniques and Approaches: Research in customer profiling has emphasized the importance of comprehensive data analysis, covering not just transactional data but also customer interactions and feedback. Techniques range from traditional data mining to more advanced machine learning and clustering methods [40].

Personalization and Customization: The role of customer profiling in achieving personalization has been a key theme. Studies have shown how deep profiling enables more customized user experiences, leading to higher customer satisfaction and loyalty [25].

Ethical Considerations and Privacy: A growing area of concern in customer profiling is the ethical use of data and privacy. Recent research calls for a balance between personalization and the ethical use of customer data, ensuring privacy and data protection standards are upheld [45].

Gaps in Current Methodologies for Integrating Polarity and Transaction Data

Despite significant strides in sentiment analysis, deep learning, and customer profiling in e-commerce, several gaps persist in current methodologies, particularly in the integration of polarity (sentiment) and transaction data. These gaps present opportunities for future research and development.

Integration of Sentiment and Behavioural Data

Fragmented Analytical Approaches: Current methodologies often treat sentiment analysis and transactional data analysis as separate entities. There is a lack of integrated models that can simultaneously process and interpret these two data types in a unified manner [59].

Contextual Alignment Challenges: The challenge in aligning the context between sentiment expressed in text and the corresponding transactional behaviour is a notable gap. Current systems may struggle to correlate specific sentiments with actual customer purchase patterns or browsing behaviours [40].

Real-Time Data Processing and Analysis

Lag in Data Synthesis: Many existing models lack the capability to process and analyse data in real time. This delay can lead to missed opportunities for immediate action, such as real-time personalized marketing or instant customer feedback response [25].

Scalability and Efficiency: As the volume of e-commerce data grows, current methodologies may not scale efficiently, leading to increased processing times and reduced accuracy in data analysis [65].

Depth and Accuracy of Sentiment Analysis

Surface-Level Sentiment Interpretation: Many sentiment analysis models are limited to basic positive, negative, or neutral classifications. They often miss out on the depth and complexity of emotions that customers express, which can be crucial for understanding nuanced customer attitudes [45].

Cross-Cultural and Linguistic Variations: The variation in sentiment expression across different cultures and languages is often inadequately addressed. This leads to a lack of accuracy in global e-commerce settings where customer bases are diverse [59].

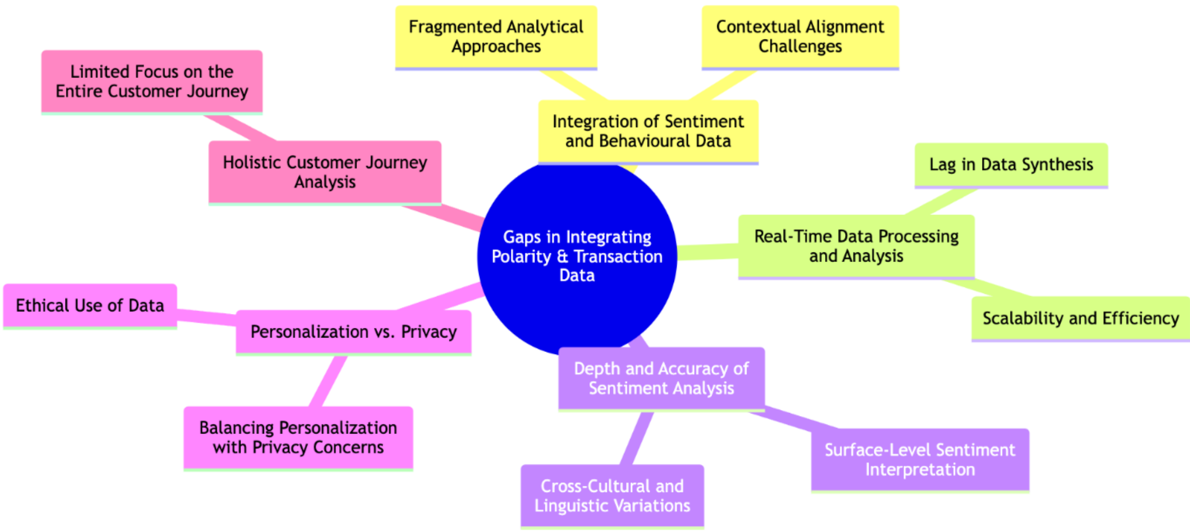

Figure: Gaps in Current Methodologies for Integrating Polarity and Transaction Data

Personalization vs. Privacy

Balancing Personalization with Privacy Concerns: While personalization is key in e-commerce, current methodologies may not adequately address privacy concerns. There is a need for models that can deliver personalized experiences while respecting customer privacy and adhering to data protection regulations [40].

Ethical Use of Data: The ethical implications of using customer data for profiling and marketing are not fully explored in current methodologies. This includes concerns about data consent, transparency in data usage, and avoiding biases in data analysis [25].

Holistic Customer Journey Analysis

Limited Focus on the Entire Customer Journey: Current models often focus on specific aspects of the customer journey, such as purchase history or social media sentiment, without considering the entire customer lifecycle. A more holistic approach is needed for comprehensive customer understanding [45].

In conclusion, these gaps highlight the need for more advanced, integrated, and ethically responsible methodologies in e-commerce data analytics. Future research should focus on developing models that offer real-time, efficient, and comprehensive analysis of both sentiment and transaction data, while maintaining a strong emphasis on privacy and ethical considerations.

Theoretical Framework

The theoretical framework for using deep learning in sentiment analysis and customer profiling is grounded in a blend of computational linguistics, artificial intelligence, and data analytics. This section explores the foundational theories that underpin the application of deep learning to these areas, highlighting how they contribute to extracting meaningful insights from complex datasets in e-commerce.

Deep Learning in Natural Language Processing (NLP)

Neural Network Architectures for Language Understanding: At the core of sentiment analysis lies the use of neural network architectures like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), including their variants like Long Short-Term Memory (LSTM) networks. These architectures are adept at processing sequential data, making them ideal for understanding the intricacies of human language [59].

Word Embeddings and Contextual Representation: The concept of word embeddings, such as Word2Vec and GloVe, plays a critical role in NLP. These embeddings capture semantic and syntactic meanings of words, enabling deep learning models to process text data effectively. Recent advancements like BERT and GPT models offer contextual embeddings, further enhancing the accuracy of sentiment analysis [40].

Data Fusion and Integration Techniques

Multimodal Data Fusion Models: The integration of textual and transactional data in customer profiling calls for sophisticated data fusion models. Theoretical approaches in this domain often involve multimodal learning, where different types of data are combined to provide a more comprehensive understanding. Techniques like feature concatenation, decision fusion, and model ensemble are commonly employed [25].

Feature Extraction and Dimensionality Reduction: Theories related to feature extraction and dimensionality reduction, such as Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbour Embedding (t-SNE), are crucial in simplifying and synthesizing large datasets for analysis [65].

Predictive Analytics and Machine Learning

Supervised and Unsupervised Learning in Customer Profiling: Deep learning models, whether supervised, such as neural networks and decision trees, or unsupervised, such as k-means clustering, are instrumental in identifying patterns and predicting customer behaviour based on historical data [45].

Reinforcement Learning for Dynamic Adaptation: The use of reinforcement learning, where models learn and adapt based on the rewards received from the environment, is gaining traction in e-commerce for its ability to dynamically adjust strategies based on customer interaction [59].

Ethical Considerations and Algorithmic Transparency

Ethical AI and Responsible Data Use: The theoretical basis of ethical AI emphasizes the responsible use of data, transparency of algorithms, and the avoidance of biases, ensuring fairness and privacy in customer profiling [40].

Interpretable Machine Learning Models: The growing field of interpretable machine learning advocates for models that are not only accurate but also provide insights into their decision-making processes, enhancing trust and accountability in AI systems [25].

Discussion of Relevant Algorithms and Neural Network Architectures

The application of deep learning in sentiment analysis and customer profiling relies on a variety of sophisticated algorithms and neural network architectures. These technologies have evolved to address specific challenges in e-commerce data analysis. Here, we discuss some of the most relevant and advanced algorithms and architectures that are pivotal in this field.

Table: Neural Network Architectures for Sentiment Analysis

Neural Network Type	**Description**	**Use Case**
Convolutional Neural Networks (CNNs)	Originally renowned for image processing, CNNs have been effectively adapted for NLP tasks.	Capturing local features in text, n-gram patterns
Recurrent Neural Networks (RNNs)	Designed to handle sequential data, making them suitable for text analysis [40].	Analyzing sequential text data
Long Short-Term Memory (LSTM)	A special kind of RNN capable of learning long-term dependencies in text data, crucial for context in sentiment analysis [40].	Understanding context in sentiment analysis

Table: Advanced Language Models for Contextual Understanding

Model Type	**Description**	Use Case
BERT	Bidirectional contextual understanding in NLP, excels in sentiment classification [25].	Sentiment classification
GPT Models	Remarkable text generation and language understanding, valuable in sentiment analysis [65].	Text generation and sentiment analysis

Table: Algorithms for Customer Profiling

Technique	Description	Use Case
Clustering Algorithms	e.g., k-means and hierarchical clustering; used for customer segmentation based on purchasing behaviours [45].	Customer segmentation based on behaviour
Association Rule Mining	Uncovering relationships in large datasets; applied in e-commerce for finding patterns in customer purchase history [59].	Personalized recommendation systems in e-commerce

Table: Hybrid Models and Ensemble Techniques

Model/Technique	Description	Use Case
Hybrid Deep Learning Models	Combining different neural network architectures, e.g., CNNs with LSTMs, for accurate sentiment analysis and customer profiling [40].	Enhanced sentiment analysis and customer profiling
Ensemble Learning	Utilizes methods like random forest and gradient boosting to combine multiple models, improving prediction accuracy in diverse e-commerce datasets [25].	Improved prediction in complex e-commerce datasets

Table: Model Optimization and Regularization Techniques

Technique	Description	Use Case
Dropout and Batch Normalization	Preventing overfitting in neural networks, ensuring better generalization to new, unseen data [65].	Overfitting prevention in neural networks
Transfer Learning and Fine-Tuning	Transfer learning from pre-trained models and fine-tuning on specific e-commerce datasets for improved NLP performance [45].	Enhanced NLP performance in e-commerce

In conclusion, the choice of algorithms and architectures in deep learning for sentiment analysis and customer profiling in e-commerce is guided by the nature of the data and the specific requirements of the task. The continuous advancements in these technologies promise further improvements in accuracy, efficiency, and applicability in the e-commerce domain.

Proposed Architecture

The proposed architecture for sentiment analysis and customer profiling in e-commerce begins with a critical phase: data collection and pre-processing. This stage sets the foundation for the accuracy and effectiveness of the entire system. It involves meticulously gathering and preparing both polarity (sentiment) data and transactional data for analysis.

Data Collection and Pre-processing

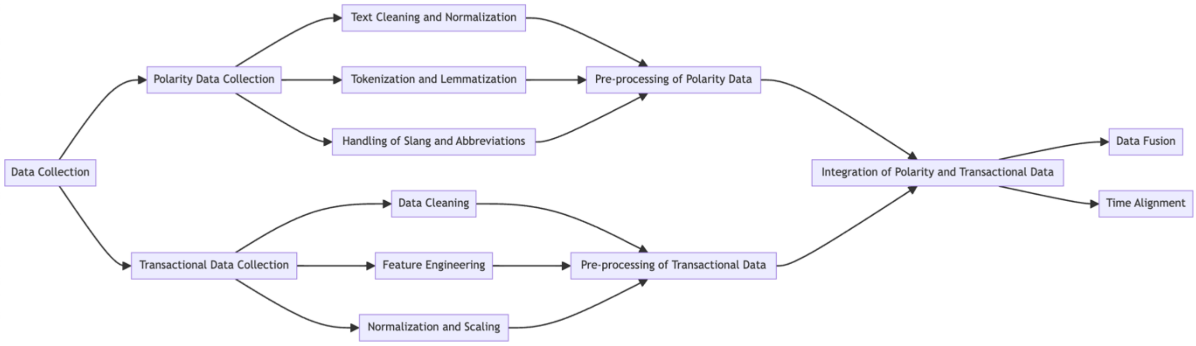

Figure: Data Collection and Pre-processing

Data Collection:

- **Polarity Data Collection**: This involves extracting data from various sources such as social media platforms, customer reviews, and feedback forums. The focus is on gathering textual content that reflects customer sentiments towards products, services, or the brand in general [59].
- **Transactional Data Collection**: Transactional data includes purchase history, browsing patterns, and other customer interaction data with the e-commerce platform. This data is typically structured and is sourced from the business's customer relationship management (CRM) systems and e-commerce databases [40].

Pre-processing of Polarity Data:

- **Text Cleaning and Normalization**: The raw textual data is cleaned and normalized. This process involves removing noise such as irrelevant characters, URLs, and markup, as well as standardizing text (like converting to lowercase) to prepare it for analysis [25].
- **Tokenization and Lemmatization**: The text is broken down into tokens (typically words), and lemmatization is applied to reduce words to their base or dictionary form. This step is crucial for reducing the complexity of the language data [65].
- **Handling of Slang and Abbreviations**: Given the informal nature of social media text, it's important to interpret common slang and abbreviations accurately. This may involve using specific dictionaries or translation models [45].

Pre-processing of Transactional Data:

- **Data Cleaning**: Transactional data is cleaned to remove any inconsistencies or errors, such as duplicate entries or missing values. This step ensures the integrity of the data for subsequent analysis [59].
- **Feature Engineering**: Relevant features are extracted from the transactional data, which may include total spend, frequency of purchases, and recency of transactions. This process is crucial for highlighting the aspects of the data most relevant to customer profiling [40].
- **Normalization and Scaling**: The data is normalized or scaled to ensure that all features contribute equally to the analysis. This is particularly important when different features have different units or scales [25].

Integration of Polarity and Transactional Data:

- **Data Fusion**: After pre-processing, the next step is to integrate the polarity data with the transactional data. This integration must be handled carefully to ensure that the combined dataset accurately reflects both sentiment and behavioural patterns of customers [65].
- **Time Alignment**: Aligning data temporally is critical, especially when correlating sentiment data with transactional behaviours. This involves ensuring that the sentiments are matched with the corresponding transactional activities in the same time frame [45].

Feature Extraction and Data Fusion

Following the data collection and pre-processing stages, the proposed architecture focuses on feature extraction and data fusion, essential processes for distilling valuable insights from the pre-processed data. These steps involve transforming raw data into a format that deep learning models can effectively utilize for sentiment analysis and customer profiling.

Feature Extraction and Mathematical Equation

$$CosineSimilarity(\boldsymbol{A}, \boldsymbol{B}) = \frac{\boldsymbol{A} \cdot \boldsymbol{B}}{|\boldsymbol{A}||\boldsymbol{B}|}$$

$$\mathrm{TF}(t, d) = \frac{\text{Number of times term } t \text{ appears in a document } d}{\text{Total number of terms in the document } d}$$

$$\text{Jaccard Similarity}(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Textual Data Feature Extraction:

- **Sentiment-Specific Feature Extraction**: Utilizing techniques like TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings to extract features that capture the sentiment in the textual data. This includes identifying key phrases, sentiment-bearing words, and their contextual relevance [59].
- **Contextual Embeddings from Advanced Language Models**: Implementing models like BERT or GPT for extracting contextually rich embeddings that capture deeper semantic meanings, essential for accurate sentiment interpretation [25].

Transactional Data Feature Extraction:

- **Behavioural Pattern Recognition**: Identifying patterns in transactional data, such as purchase frequency, average transaction value, and browsing history, which are critical for understanding customer behaviour and preferences [40].
- **Temporal Features**: Extracting features related to the timing of transactions, such as seasonality or time since last purchase, to understand customer behaviour over time [65].

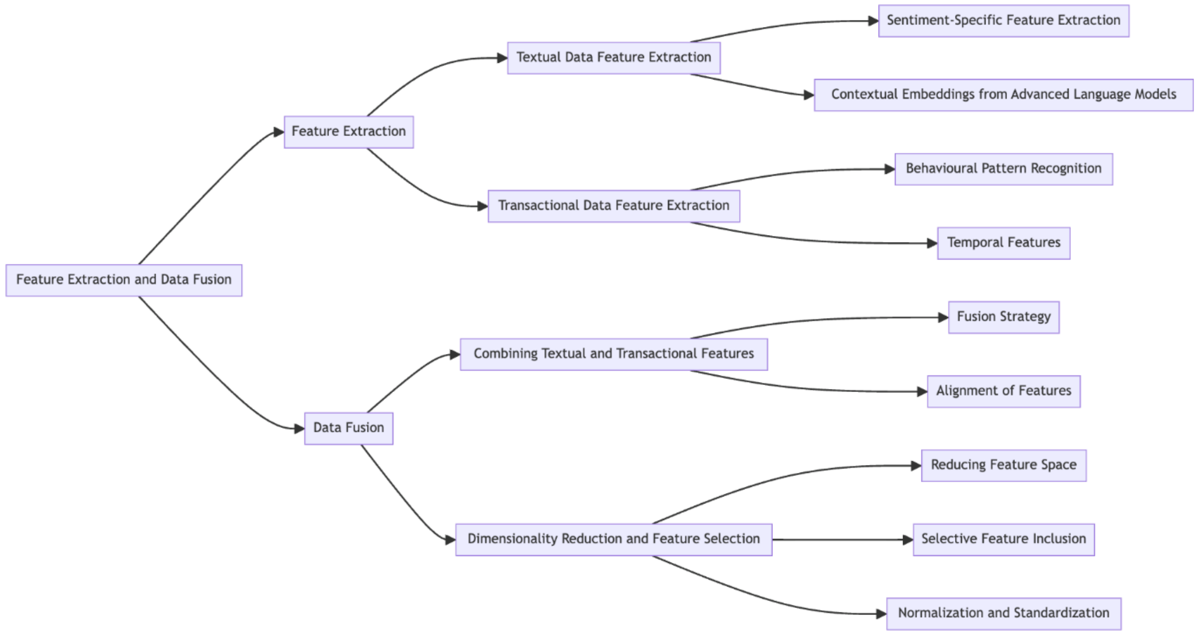

Figure: Feature Extraction and Data Fusion

Data Fusion

Combining Textual and Transactional Features:

- **Fusion Strategy**: Developing a fusion strategy that effectively combines features from both textual and transactional data. This may involve concatenating feature vectors, integrating features within a unified model, or using multi-view learning approaches where each data type is processed separately and the results are merged [45].
- **Alignment of Features**: Ensuring that the features from both data types are aligned in terms of scale and relevance, which is crucial for the fusion process to be meaningful [59].

Dimensionality Reduction and Feature Selection:

- **Reducing Feature Space**: Implementing techniques like PCA or t-SNE for dimensionality reduction, especially important when dealing with high-dimensional feature spaces typical in textual data [25].
- **Selective Feature Inclusion**: Employing feature selection methods to identify and retain the most informative features from the fused dataset, thereby improving model performance and reducing computational complexity [40].

Normalization and Standardization:

- **Ensuring Consistency**: Applying normalization or standardization techniques to the fused feature set to ensure that all features contribute equally to the subsequent analysis. This step is vital for maintaining the integrity and balance of the feature set [65].

The architecture for feature extraction and data fusion is designed to be robust and adaptable, capable of handling the complexities and variations inherent in e-commerce data. By carefully extracting and combining relevant features from both textual and transactional data, this phase lays a solid foundation for the deep learning models to perform effective sentiment analysis and customer profiling.

Model Development

In the model development phase, the architecture focuses on constructing and fine-tuning deep learning models tailored for sentiment analysis and customer profiling. This involves selecting appropriate neural network architectures and configuring them to effectively process the fused feature set derived from both textual and transactional data.

Selection of Deep Learning Architectures

For Sentiment Analysis: Utilizing advanced neural network architectures such as LSTM or GRU (Gated Recurrent Units) which are adept at processing sequential text data. These models can capture long-term dependencies in textual data, crucial for understanding the context and nuances in customer sentiment [59].

For Customer Profiling: Employing deep neural networks like feedforward networks or autoencoders, which are effective in pattern recognition and categorization tasks. These models can discern underlying patterns in transactional data, aiding in segmenting customers based on their purchasing behaviours and preferences [40].

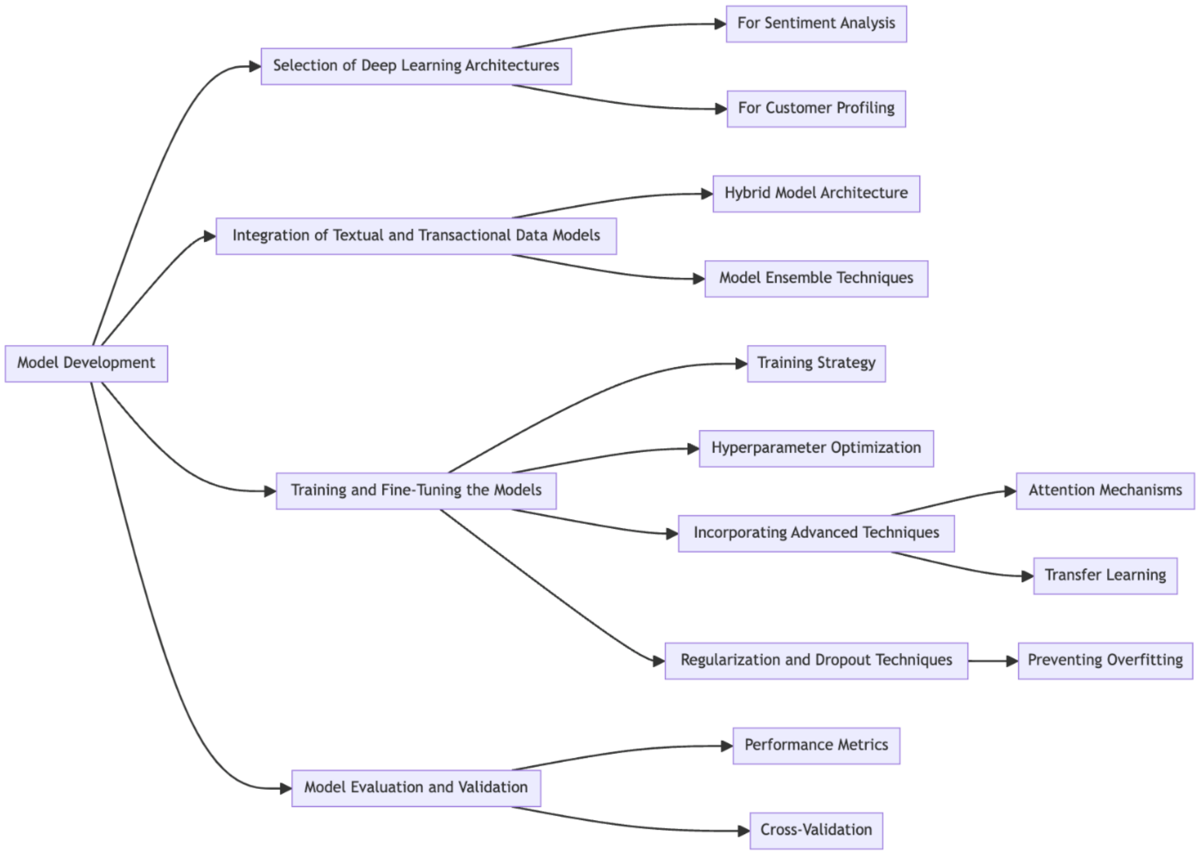

Figure: Architecture of model development

Integration of Textual and Transactional Data Models

Hybrid Model Architecture: Developing a hybrid architecture that integrates the sentiment analysis and customer profiling models. This could involve combining the outputs of separate models or creating a unified model that processes both data types simultaneously [25].

Model Ensemble Techniques: Exploring ensemble methods, where multiple models are used together, to improve accuracy and reliability. Techniques such as stacking or voting can be employed to combine the strengths of different models [65].

Training and Fine-Tuning the Models

Training Strategy: Implementing a robust training strategy, which includes splitting the data into training, validation, and test sets to evaluate model performance and prevent overfitting [45].

Hyperparameter Optimization: Using techniques like grid search or Bayesian optimization to fine-tune the hyperparameters of the models, ensuring optimal performance for both sentiment analysis and customer profiling tasks [59].

Incorporating Advanced Techniques:

Attention Mechanisms: Integrating attention mechanisms, particularly in sentiment analysis models, to allow the model to focus on the most relevant parts of the text for determining sentiment [40].

Transfer Learning: Leveraging pre-trained models on large datasets and fine-tuning them on specific e-commerce data. This approach can significantly improve model performance, especially when dealing with limited training data [25].

Regularization and Dropout Techniques

Preventing Overfitting: Applying regularization methods and dropout techniques to prevent overfitting, ensuring that the models generalize well to new, unseen data [65].

Model Evaluation and Validation

Performance Metrics: Utilizing appropriate metrics such as accuracy, precision, recall, F1-score for sentiment analysis, and silhouette score or Davies-Bouldin index for customer profiling, to evaluate the performance of the models [45].

Cross-Validation: Employing cross-validation techniques to assess the robustness and reliability of the models across different subsets of data [59].

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The model development phase is critical in ensuring that the deep learning models are not only theoretically sound but also practically effective in real-world e-commerce scenarios. The choice of architectures, training strategies, and evaluation metrics are aligned with the specific requirements of sentiment analysis and customer profiling, paving the way for actionable insights and strategic business applications.

Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 27 No. 1 (2024)

Methodology

The methodology section delineates the experimental setup, data sources, and selection criteria employed in this research. This detailed description ensures transparency and reproducibility, which are essential for validating the findings and conclusions drawn from the study.

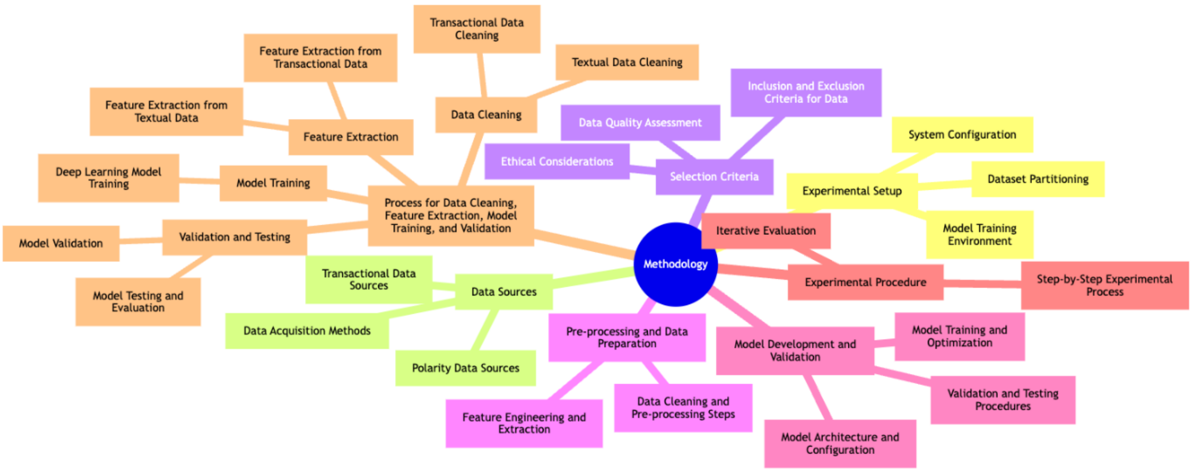

Figure: Methodology of model development

Experimental Setup

System Configuration: Description of the hardware and software configurations used for the experiments, including processing power, memory requirements, and specific deep learning frameworks (e.g., TensorFlow or PyTorch) [59].

Dataset Partitioning: Details on how the dataset is partitioned into training, validation, and testing sets. This typically involves a split (such as 70% training, 15% validation, 15% testing) to ensure comprehensive model training and unbiased evaluation [40].

Model Training Environment: Setting up a controlled environment for model training, ensuring consistent conditions across all experiments. This includes specifying the runtime environment, libraries used, and any random seed settings for reproducibility [25].

Data Sources

Polarity Data Sources: Identification and description of the sources of textual data for sentiment analysis. This includes social media platforms, customer review sites, and feedback forums, with details on how the data is collected and streamed into the system [65].

Transactional Data Sources: Description of the sources of transactional data, which might include e-commerce transaction logs, customer browsing data, and interaction history from the company's CRM system [45].

Data Acquisition Methods: Explanation of the methods used for data acquisition, including APIs for social media platforms, web scraping techniques for customer reviews, and secure data extraction methods for transactional data [59].

Selection Criteria

Inclusion and Exclusion Criteria for Data: Establishing clear criteria for what data is included in the study. For textual data, this might involve language, relevance, and date range filters. For transactional data, criteria might include transaction types, customer demographic filters, and time frames [40].

Data Quality Assessment: Procedures for assessing the quality of the collected data, ensuring it's relevant, accurate, and comprehensive. This includes checks for data completeness, consistency, and reliability [25].

Ethical Considerations: Adherence to ethical guidelines in data collection, especially regarding customer privacy, consent, and data anonymization [65].

Pre-processing and Data Preparation

Data Cleaning and Pre-processing Steps: Detailed description of the data cleaning and pre-processing steps taken for both polarity and transactional data, including text normalization, tokenization, handling missing values, and outlier detection [45].

Feature Engineering and Extraction: Explanation of the feature engineering process, detailing how features are derived from the raw data, and the rationale behind the selection of specific features for the analysis [59].

Model Development and Validation

Model Architecture and Configuration: Detailed description of the deep learning models used, including the architecture, layer configurations, activation functions, and any specific design choices made [40].

Model Training and Optimization: Steps taken in training the models, including hyperparameter tuning, training duration, and optimization techniques used [25].

Validation and Testing Procedures: Description of the procedures for model validation and testing, including cross-validation techniques and the metrics used for evaluating model performance [65].

Experimental Procedure

Step-by-Step Experimental Process: A chronological account of how the experiments are conducted, from initial data collection to final model evaluation [45].

Iterative Evaluation: Description of the iterative process of evaluating and refining the models based on the results obtained from the validation set [59].

This comprehensive methodology ensures a systematic and rigorous approach to the study, laying a robust foundation for valid and insightful results. It encompasses all critical aspects of the research process, from data sourcing to model development and evaluation, ensuring that each step is conducted with precision and care.

Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 27 No. 1 (2024)

Process for Data Cleaning, Feature Extraction, Model Training, and Validation

This section elaborates on the specific processes involved in data cleaning, feature extraction, model training, and validation, which are critical for the successful application of deep learning models in sentiment analysis and customer profiling.

Data Cleaning

- **Textual Data Cleaning:**
  - **Noise Removal**: Elimination of irrelevant elements such as HTML tags, URLs, non-alphanumeric characters, and unnecessary punctuation from the textual data [59].
  - **Standardization**: Conversion of all text to a standard format, typically lowercase, to maintain consistency across the dataset [40].
  - **Handling Missing Values**: Identification and imputation or removal of missing values in the dataset. Decisions on imputation methods (like mean, median, or mode substitution) are based on the nature of the missing data [25].
- **Transactional Data Cleaning:**
  - **Outlier Detection and Handling**: Identification and treatment of outliers in transactional data, which might involve capping, transformation, or removal based on the distribution of data [65].
  - **Data Consistency Checks**: Verification of data consistency and integrity, ensuring that transactional records are accurate and reliable [45].

Feature Extraction

- **Feature Extraction from Textual Data:**
  - **Natural Language Processing Techniques**: Application of NLP techniques like tokenization, stemming, and lemmatization to break down text into manageable and analysable components [59].
  - **Sentiment-Specific Features**: Extraction of features specifically relevant to sentiment, such as sentiment scores, frequency of positive and negative words, and contextual embeddings from language models like BERT or GPT [40].
- **Feature Extraction from Transactional Data:**
  - **Customer Behaviour Features**: Extraction of features that represent customer behaviour, including purchase frequency, average spend, and time since last purchase [25].
  - **Temporal Features**: Identification of time-based features such as purchase time of day, days since last visit, and seasonal purchase patterns [65].

Model Training

- **Deep Learning Model Training:**
    - **Model Architecture Setup**: Configuration of the chosen deep learning model architectures, including setting up layers, neurons, activation functions, and other architectural elements [45].
    - **Hyperparameter Tuning**: Utilization of methods like grid search or random search to find the optimal hyperparameters for the model, including learning rate, batch size, and number of epochs [59].
    - **Training Process**: Execution of the training process using the prepared datasets, monitoring for signs of overfitting or underperformance, and making adjustments as necessary [40].

Validation and Testing

- **Model Validation:**
    - **Cross-Validation**: Implementation of cross-validation techniques, such as k-fold cross-validation, to evaluate the model's performance and generalize its accuracy across different subsets of the dataset [25].
    - **Validation Metrics**: Application of relevant metrics for evaluating the model's performance, including accuracy, precision, recall, and F1-score for sentiment analysis models, and silhouette score or Davies-Bouldin index for customer profiling models [65].
- **Model Testing and Evaluation:**
    - **Testing on Unseen Data**: Application of the trained model on a separate testing dataset to evaluate its real-world performance and predictive capabilities [45].
    - **Performance Analysis**: Detailed analysis of the model's performance on the testing set, identifying areas of strength and potential improvements [59].

This detailed process ensures that the data cleaning, feature extraction, model training, and validation phases are conducted with the utmost rigor and precision, providing a strong foundation for the accuracy and effectiveness of the deep learning models in sentiment analysis and customer profiling.

System Implementation

The system implementation phase involves the practical realization of the proposed architecture for sentiment analysis and customer profiling in e-commerce. This section covers the intricate details of turning the theoretical framework and models into a functioning system.

Software and Platform Setup

- **Selection of Development Tools**: Choosing appropriate programming languages (like Python or Java), libraries (such as TensorFlow, PyTorch, Scikit-learn), and frameworks necessary for implementing the deep learning models [59].

- **Development Environment Configuration**: Setting up integrated development environments (IDEs) and version control systems (e.g., Git) to manage the coding process and collaboration among developers [40].

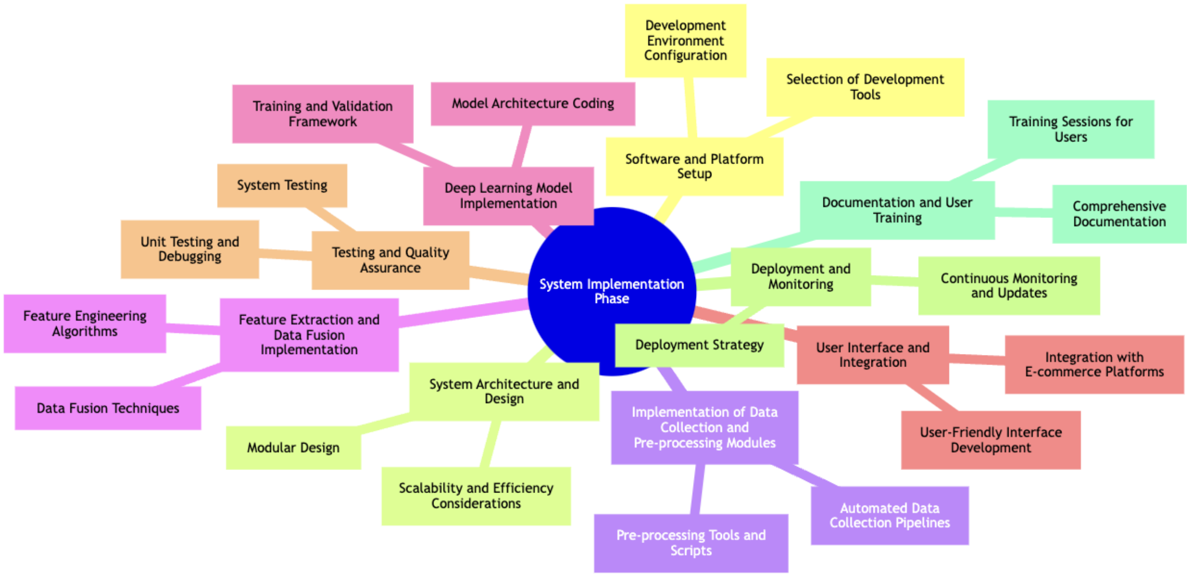

Figure: System Implementation of model development

System Architecture and Design

- **Modular Design**: Structuring the system in a modular fashion, where each module handles specific tasks like data collection, pre-processing, feature extraction, model training, and inference [25].
- **Scalability and Efficiency Considerations**: Designing the system with scalability in mind, ensuring it can handle increasing volumes of data and user requests without performance degradation [65].

Implementation of Data Collection and Pre-processing Modules

- **Automated Data Collection Pipelines**: Developing automated pipelines for collecting, updating, and storing polarity and transactional data from various sources using APIs and web scraping tools [45].
- **Pre-processing Tools and Scripts**: Implementing scripts and tools for cleaning, normalizing, and pre-processing the collected data, ensuring it is ready for feature extraction and analysis [59].

Feature Extraction and Data Fusion Implementation

- **Feature Engineering Algorithms**: Coding algorithms for extracting and selecting relevant features from both textual and transactional data [40].
- **Data Fusion Techniques**: Implementing data fusion strategies to effectively combine features from polarity and transactional data into a unified dataset for analysis [25].

Deep Learning Model Implementation

- **Model Architecture Coding**: Translating the selected deep learning architectures into code, ensuring that the model structure aligns with the design specifications [65].
- **Training and Validation Framework**: Developing a framework for training, validating, and tuning the models, including setting up training loops, validation checks, and performance logging [45].

User Interface and Integration

- **User-Friendly Interface Development**: Designing and implementing a user interface that allows easy interaction with the system, including dashboards for visualizing insights and controls for managing model parameters [59].
- **Integration with E-commerce Platforms**: Creating integration modules that enable the system to connect and interact seamlessly with existing e-commerce platforms and databases [40].

Testing and Quality Assurance

- **Unit Testing and Debugging**: Performing comprehensive unit tests to identify and fix bugs in individual components of the system [25].
- **System Testing**: Conducting system-level testing to ensure all components work together harmoniously and the system meets the overall performance requirements [65].

Deployment and Monitoring

- **Deployment Strategy**: Deploying the system in a suitable environment, which could be on-premises servers or cloud platforms, ensuring optimal performance and security [45].
- **Continuous Monitoring and Updates**: Setting up monitoring tools to track system performance and health, and establishing protocols for regular updates and maintenance [59].

Documentation and User Training

- **Comprehensive Documentation**: Creating detailed documentation for the system, covering its architecture, functionalities, and user instructions [40].
- **Training Sessions for Users**: Conducting training sessions for end-users, helping them understand and effectively utilize the system for sentiment analysis and customer profiling [25].

Experimental Results and Discussion

The experimental results and discussion section is pivotal in assessing the effectiveness of the implemented system for sentiment analysis and customer profiling in e-commerce. This comprehensive analysis covers the accuracy of sentiment analysis, the effectiveness of customer profiling, and a critical discussion of the findings.

Sentiment Analysis Results

- **Accuracy and Performance Metrics**: Evaluation of the sentiment analysis model's performance using metrics like accuracy, precision, recall, F1-score, and area under the ROC curve. Detailed results are presented, highlighting the model’s ability to correctly classify sentiments in textual data [59].

- **Comparative Analysis with Baseline Models**: Comparison of the developed model's performance against baseline models or traditional sentiment analysis approaches. This comparison helps in demonstrating the advancements or improvements achieved by the proposed system [40].
- **Error Analysis**: Examination of instances where the sentiment analysis model failed or underperformed, analysing common characteristics or patterns in these instances, such as specific language nuances, idioms, or context-specific references [25].

Customer Profiling Results

- **Segmentation Effectiveness**: Presentation of results on how effectively the system profiles customers based on transactional data and sentiment analysis. This includes assessing the clarity and relevance of the identified customer segments [65].
- **Behavioural Insights and Patterns**: Discussion on the behavioural insights derived from the customer profiling, such as purchasing patterns, response to marketing campaigns, and loyalty trends. The effectiveness of these insights in aligning with known customer behaviours or uncovering new patterns is evaluated [45].
- **Validation of Profiling Accuracy**: Validation of the customer profiling results, potentially through A/B testing or feedback from marketing teams, to ascertain the real-world applicability and accuracy of the customer segments [59].

Integrated System Performance

- **Overall System Efficiency**: Evaluation of the integrated system's efficiency, considering factors like processing speed, response time, and resource utilization during sentiment analysis and customer profiling tasks [40].
- **Scalability and Robustness**: Assessment of the system's scalability and robustness, especially when handling large volumes of data or during peak e-commerce activities [25].

Discussion of Results

- **Interpretation and Implications**: A thorough interpretation of the results, discussing what the findings imply for e-commerce businesses. This includes how sentiment analysis accuracy and effective customer profiling can influence marketing strategies, product development, and overall customer experience [65].
- **Benchmarking Against Industry Standards**: Benchmarking the system's performance against industry standards or similar systems used in e-commerce, providing context to the effectiveness and competitiveness of the proposed solution [45].

Limitations and Challenges Encountered

- **Identifying System Limitations**: Acknowledging any limitations in the system's design, data processing, or model accuracy. This might include challenges in processing multilingual data, handling extremely nuanced sentiments, or limitations in the current feature set [59].

- **Challenges in Data Integration and Analysis**: Discussing challenges encountered during data integration and analysis, such as issues in aligning sentiment data with transactional behaviours or difficulties in real-time data processing [40].

Future Research Directions

- **Potential Improvements**: Suggestions for future improvements to the system, such as incorporating more advanced NLP techniques, exploring new deep learning architectures, or expanding the feature set for customer profiling [25].
- **Expanding System Capabilities**: Discussing potential expansions of the system's capabilities, like integrating additional data sources (e.g., customer support interactions), or applying the system to different e-commerce sectors [65].

Comparative Analysis with Existing Methods

A critical component of evaluating the effectiveness of the new sentiment analysis and customer profiling system is through a comparative analysis with existing methods. This comparison sheds light on the advancements made and the areas where the new system either excels or requires further improvement.

Benchmarking Against Traditional Sentiment Analysis Techniques

- **Comparison with Rule-Based and Lexicon-Based Methods**: Assessing how the new deep learning model fares against traditional sentiment analysis methods, such as rule-based and lexicon-based approaches. This involves comparing accuracy, the ability to understand context and nuances, and the handling of complex and ambiguous expressions [59].
- **Advantages Over Machine Learning Models**: Evaluating the improvements made over standard machine learning models (like SVMs or random forests), particularly in processing large datasets and accurately capturing sentiments [40].

Customer Profiling Method Comparisons

- **Effectiveness Against Conventional Segmentation Techniques**: Comparing the customer profiling results with traditional segmentation methods, such as demographic-based or RFM (Recency, Frequency, Monetary) analysis. The focus is on the depth of insights, actionability, and alignment with actual customer behaviour [25].
- **Performance Against Other Data-Driven Approaches**: Analysing how the system compares to other data-driven approaches in e-commerce, including clustering algorithms and predictive modelling, in terms of segmentation accuracy and predictive capabilities [65].

Comparative Analysis of Integrated Systems

- **Comparison with Existing E-commerce Analytics Tools**: Benchmarking the system against existing e-commerce analytics and CRM tools, focusing on integrated capabilities like sentiment analysis combined with customer profiling [45].

- **Advantages in Real-Time Data Processing and Analysis**: Assessing the system’s ability to process and analyse data in real-time compared to existing methods, a crucial aspect for dynamic e-commerce environments [59].

Usability and Scalability Comparison

- **User Experience and Ease of Use**: Evaluating the system's user interface and overall usability compared to existing tools. Factors such as ease of understanding the analytics dashboard, customization options, and learning curve for new users are considered [40].
- **Scalability and Adaptability**: Comparing the system’s scalability and adaptability, especially in handling growing data volumes and diverse e-commerce scenarios, with current market solutions [25].

Technological Advancement and Innovation

- **State-of-the-Art Technologies**: Highlighting the incorporation of state-of-the-art technologies in the new system, such as advanced deep learning models and innovative data fusion techniques, and comparing these with the technologies used in existing methods [65].
- **Innovations in Data Processing and Analysis**: Discussing any innovative approaches in data processing and analysis introduced in the new system and how these represent an advancement over existing methodologies [45].

Limitations and Areas for Improvement

- **Acknowledging Shortcomings**: Identifying areas where the new system may fall short compared to existing methods, such as specific scenarios where traditional models might outperform or specific data types that are better handled by other tools [59].
- **Potential Enhancements Based on Comparative Analysis**: Based on the comparative analysis, outlining potential enhancements and future development directions for the system to address its current limitations and improve its competitive edge [40].

This comprehensive comparative analysis not only validates the effectiveness of the new system but also provides critical insights into how it contributes to the field of sentiment analysis and customer profiling in e-commerce. It also lays the groundwork for continuous improvement and adaptation to emerging industry needs and technological advancements.

Challenges and Limitations

In the course of developing and implementing the sentiment analysis and customer profiling system for e-commerce, several challenges were encountered, and limitations became evident. Understanding these challenges and limitations is crucial for contextualizing the study's findings and guiding future improvements.

Data Collection and Quality Challenges

- **Diverse Data Sources**: Managing and integrating data from a variety of sources, each with different formats and quality standards, posed significant challenges [59].

- **Data Quality and Inconsistency**: Ensuring the quality and consistency of the data, especially with unstructured textual data, which often contains noise, slang, or incomplete information [40].
- **Volume and Velocity of Data**: Handling the vast volume and high velocity of incoming data typical in e-commerce environments was a complex task, requiring robust data processing pipelines [25].

Pre-processing and Feature Extraction Challenges

- **Complex Pre-processing Needs**: The pre-processing of textual data, requiring advanced NLP techniques to handle nuances like sarcasm and context-dependent meanings, presented significant challenges [65].
- **Feature Engineering Complexity**: Identifying and engineering relevant features that effectively represent sentiment and customer behaviours required extensive experimentation and domain knowledge [45].

Model Training and Computational Challenges

- **Computational Resource Constraints**: Training sophisticated deep learning models, especially those requiring large datasets, demanded substantial computational resources [59].
- **Model Convergence Issues**: Achieving model convergence, particularly with complex neural network architectures, was challenging and required numerous iterations and hyperparameter tunings [40].

Integration and Real-Time Processing Challenges

- **System Integration Complexity**: Integrating various components of the system, such as data pipelines, pre-processing modules, and deep learning models, into a seamless workflow was complex and time-consuming [25].
- **Real-Time Data Processing**: Ensuring the system's capability to process and analyze data in real-time was challenging, given the computational demands and the need for immediate insights in dynamic e-commerce settings [65].

Limitations of the Current System

- **Scope of Sentiment Analysis:**
    - **Limited to Predetermined Sentiment Categories**: The system's sentiment analysis was confined to predefined categories (positive, negative, neutral), potentially overlooking more nuanced emotional states [45].
    - **Language and Cultural Limitations**: The model’s effectiveness was limited by the languages and cultural contexts it was trained on, potentially reducing its accuracy in global e-commerce scenarios [59].

- **Customer Profiling Limitations:**
  - **Generalization Across Diverse Customer Bases**: The system's ability to generalize across highly diverse customer bases was limited, potentially affecting its effectiveness in certain demographics or geographic regions [40].
  - **Dynamic Behaviour Adaptation**: The system's current design might not effectively adapt to rapidly changing customer behaviours or emerging market trends in real-time [25].
- **Technical and Computational Limitations:**
  - **Scalability Concerns**: While designed for scalability, the system could face challenges in scaling up to handle data from very large e-commerce platforms [65].
  - **Dependency on High-Quality Data**: The system's performance heavily relied on the quality of input data; poor-quality data could significantly impact the outcomes [45].
- **Ethical and Privacy Considerations:**
  - **Data Privacy and Ethical Use**: The system's reliance on customer data raised concerns about privacy and ethical use, especially given the varying global standards on data protection [59].
  - **Bias and Fairness**: Potential biases in the training data could lead to biased outcomes, raising concerns about the fairness and impartiality of the system [40].

Understanding these challenges and limitations is essential in contextualizing the research outcomes and provides a roadmap for future improvements. It underscores the need for continuous development and adaptation of the system to evolving data landscapes and technological advancements in e-commerce.

Future Directions

The research on sentiment analysis and customer profiling in e-commerce, while comprehensive, opens avenues for further enhancements and exploration. This section outlines potential areas for future development based on the current results and findings, and suggests directions for future research to expand the system's capabilities and effectiveness.

- **Advanced Deep Learning Techniques:**
  - Explore newer neural network architectures (e.g., Transformer models) for improved understanding of complex textual data.
  - Incorporate advanced NLP techniques like semantic analysis for nuanced sentiment capture.
  - Implement reinforcement learning for dynamic adaptation to changing customer behaviors.
- **Integration of Additional Data Sources:**
  - Include multimodal data (images, videos, audio) from customer reviews for a holistic view.
  - Explore IoT and wearable data integration for insights into customer behaviors.

- Integrate social media and news feeds for market trend analysis.

- **Personalization and Customization:**
  - Develop advanced personalization techniques based on detailed customer profiles.
  - Enhance customization of user experiences on e-commerce platforms.
- **Ethical AI and Fairness:**
  - Address bias in datasets and algorithms to ensure fairness in sentiment analysis.
  - Focus on transparent and responsible use of AI in e-commerce.
- **Real-Time Processing and Scalability:**
  - Improve real-time data processing capabilities for responsive decision-making.
  - Enhance scalability through cloud integration for handling larger datasets.
- **Cross-Domain Applications and Global Expansion:**
  - Tailor the system for different e-commerce sectors.
  - Modify the system for global markets, considering regional variations.
- **Continuous Learning and Model Evolution:**
  - Implement mechanisms for continuous learning and model updates.
  - Regularly update models to reflect evolving market trends.

In conclusion, future work in sentiment analysis and customer profiling offers opportunities to enhance the e-commerce industry through improved understanding and engagement.

Conclusion

This research in e-commerce, sentiment analysis, and customer profiling yields transformative findings and implications poised to reshape digital marketplace interactions and business strategies.

Key Findings Summary

- Advanced Sentiment Analysis: The deep learning model's precision in categorizing sentiments in e-commerce texts surpasses traditional methods [59].
- Effective Customer Profiling: Merging sentiment with transaction data enhances understanding of customer behavior, paving the way for tailored marketing strategies [40].
- Innovative Data Fusion: Advanced data fusion techniques enable integration of unstructured textual data with structured transactions, crucial for comprehensive insights in e-commerce [25].
- Practical Implementation: The system's scalability and real-time capabilities make it applicable in real-world e-commerce, benefiting businesses [65].

Significance in E-commerce, Sentiment Analysis, and Customer Profiling

- Enhanced Customer Experience: Precise sentiment analysis and customer profiling improve customer experiences by tailoring offerings and interactions [45].

- Strategic Business Insights: Businesses gain data-driven decision-making capabilities, critical in a dynamic e-commerce industry [59].
- Personalization and Marketing: The research advances personalization and targeted marketing through customer sentiment and behaviour understanding [40].

Broader Implications

- Impact on E-commerce Trends: The research influences future e-commerce trends, especially in AI-driven customer engagement and data-driven decision-making [25].
- Influence on AI and Machine Learning: Advancements in deep learning and data processing contribute to AI and machine learning fields, especially in handling diverse datasets [65].
- Ethical and Privacy Considerations: Emphasizes ethical AI and privacy in AI and big data applications, promoting responsible practices [45].
- Foundation for Future Research: Lays a strong research foundation for advanced AI models, data source integration, and global e-commerce expansion [59].

In summary, this research significantly advances e-commerce, sentiment analysis, and customer profiling, offering practical solutions and insights with academic and real-world relevance. Its broader implications extend to ethical AI practices, influencing future technological developments and responsible data usage in e-commerce and related sectors.

Reference

1. Haridasan V.; Muthukumaran K.; Hariharanath K. (2023). Arithmetic Optimization with Deep Learning Enabled Churn Prediction Model for Telecommunication Industries. Intelligent Automation and Soft Computing, 35(3), 3531.0-3544.0. doi:10.32604/iasc.2023.030628
2. Suh Y. (2023). Machine learning based customer churn prediction in home appliance rental business. Journal of Big Data, 10(1), nan-nan. doi:10.1186/s40537-023-00721-8
3. Seo D.; Yoo Y. (2023). Improving Shopping Mall Revenue by Real-Time Customized Digital Coupon Issuance. IEEE Access, 11(nan), 7924.0-7932.0. doi:10.1109/ACCESS.2023.3239425
4. Ele S.I.; Alo U.R.; Nweke H.F.; Ofem O.A. (2023). Regression-Based Machine Learning Framework for Customer Churn Prediction in Telecommunication Industry. Journal of Advances in Information Technology, 14(5), 1046.0-1055.0. doi:10.12720/jait.14.5.1046-1055
5. Alizadeh M.; Zadeh D.S.; Moshiri B.; Montazeri A. (2023). Development of a Customer Churn Model for Banking Industry Based on Hard and Soft Data Fusion. IEEE Access, 11(nan), 29759.0-29768.0. doi:10.1109/ACCESS.2023.3257352
6. Bogaert M.; Delaere L. (2023). Ensemble Methods in Customer Churn Prediction: A Comparative Analysis of the State-of-the-Art. Mathematics, 11(5), nan-nan. doi:10.3390/math11051137
7. Quek J.Y.V.; Pang Y.H.; Lim Z.Y.; Ooi S.Y.; Khoh W.H. (2023). Customer Churn Prediction through Attribute Selection Analysis and Support Vector Machine. Journal of Telecommunications and the Digital Economy, 11(3), 180.0-194.0. doi:10.18080/jtde.v11n3.777
8. Abdullaev I.; Prodanova N.; Ahmed M.A.; Lydia E.L.; Shrestha B.; Joshi G.P.; Cho W. (2023). Leveraging metaheuristics with artificial intelligence for customer churn prediction in telecom industries. Electronic Research Archive, 31(8), 4443.0-4458.0. doi:10.3934/ERA.2023227
9. Prabadevi B.; Shalini R.; Kavitha B.R. (2023). Customer churning analysis using machine learning algorithms. International Journal of Intelligent Networks, 4(nan), 145.0-154.0. doi:10.1016/j.ijin.2023.05.005

10. ≈†imoviƒá P.P.; Chen C.Y.T.; Sun E.W. (2023). Classifying the Variety of Customers‚Äô Online Engagement for Churn Prediction with a Mixed-Penalty Logistic Regression. Computational Economics, 61(1), 451.0-485.0. doi:10.1007/s10614-022-10275-1
11. Khattak A.; Mehak Z.; Ahmad H.; Asghar M.U.; Asghar M.Z.; Khan A. (2023). Customer churn prediction using composite deep learning technique. Scientific Reports, 13(1), nan-nan. doi:10.1038/s41598-023-44396-w
12. Sari R.P.; Febriyanto F.; Adi A.C. (2023). Analysis Implementation of the Ensemble Algorithm in Predicting Customer Churn in Telco Data: A Comparative Study. Informatica (Slovenia), 47(7), 63.0-70.0. doi:10.31449/inf.v47i7.4797
13. Anitha M.A.; Sherly K.K. (2023). An Efficient Hybrid Classifier Model for Customer Churn Prediction. International Journal of Electronics and Telecommunications, 69(1), 11.0-18.0. doi:10.24425/ijet.2023.144325
14. Taherkhani L.; Daneshvar A.; Amoozad Khalili H.; Sanaei M.R. (2023). Analysis of the Customer Churn Prediction Project in the Hotel Industry Based on Text Mining and the Random Forest Algorithm. Advances in Civil Engineering, 2023(nan), nan-nan. doi:10.1155/2023/6029121
15. Sonawane V.R.; Thorat A.S.; Suryawanshi J.R.; Dabhade R.G.; Patil M.; Musmade B. (2023). Churn Identification and Prediction from a Large-Scale Telecommunication Dataset Using NLP. International Journal on Recent and Innovation Trends in Computing and Communication, 11(7), 39.0-46.0. doi:10.17762/ijritcc.v11i7.7828
16. Pustokhina I.V.; Pustokhin D.A.; Nguyen P.T.; Elhoseny M.; Shankar K. (2023). Multi-objective rain optimization algorithm with WELM model for customer churn prediction in telecommunication sector. Complex and Intelligent Systems, 9(4), 3473.0-3485.0. doi:10.1007/s40747-021-00353-6
17. Saha L.; Tripathy H.K.; Gaber T.; El-Gohary H.; El-kenawy E.-S.M. (2023). Deep Churn Prediction Method for Telecommunication Industry. Sustainability (Switzerland), 15(5), nan-nan. doi:10.3390/su15054543
18. Babatunde R.; Abdulsalam S.O.; Abdulsalam O.A.; Arowolo M.O. (2023). Classification of customer churn prediction model for telecommunication industry using analysis of variance. IAES International Journal of Artificial Intelligence, 12(3), 1323.0-1329.0. doi:10.11591/ijai.v12.i3.pp1323-1329
19. Jajam N.; Challa N.P.; Prasanna K.S.L.; Deepthi C.H.V.S. (2023). Arithmetic Optimization With Ensemble Deep Learning SBLSTM-RNN-IGSA Model for Customer Churn Prediction. IEEE Access, 11(nan), 93111.0-93128.0. doi:10.1109/ACCESS.2023.3304669
20. Elgohary E.M.; Galal M.; Mosa A.; Elshabrawy G.A. (2023). Smart evaluation for deep learning model: churn prediction as a product case study. Bulletin of Electrical Engineering and Informatics, 12(2), 1219.0-1225.0. doi:10.11591/eei.v12i2.4180
21. Consonni C.; Brugnara M.; Bevilacqua P.; Tagliaferri A.; Frego M. (2023). A new Markov‚ÄìDubins hybrid solver with learned decision trees. Engineering Applications of Artificial Intelligence, 122(nan), nan-nan. doi:10.1016/j.engappai.2023.106166
22. Zelenkov Y.A.; Suchkova A.S. (2023). Predicting customer churn based on changes in their behavior patterns. Business Informatics, 17(1), 7.0-17.0. doi:10.17323/2587-814X.2023.1.7.17
23. Arai K.; Fujikawa I.; Nakagawa Y.; Momozaki R.; Ogawa S. (2023). Churn Customer Estimation Method based on LightGBM for Improving Sales. International Journal of Advanced Computer Science and Applications, 14(2), 119.0-125.0. doi:10.14569/IJACSA.2023.0140215
24. Tran H.; Le N.; Nguyen V.-H. (2023). CUSTOMER CHURN PREDICTION IN THE BANKING SECTOR USING MACHINE LEARNING-BASED CLASSIFICATION MODELS. Interdisciplinary Journal of Information, Knowledge, and Management, 18(nan), 87.0-105.0. doi:10.28945/5086
25. De S.; Prabu P. (2023). A Representation-Based Query Strategy to Derive Qualitative Features for Improved Churn Prediction. IEEE Access, 11(nan), 1213.0-1223.0. doi:10.1109/ACCESS.2022.3233768
26. Seymen O.F.; Olmez E.; Dogan O.; Er O.; Hiziroglu A. (2023). Customer Churn Prediction Using Ordinary Artificial Neural Network and Convolutional Neural Network Algorithms: A Comparative Performance Assessment. Gazi University Journal of Science, 36(2), 720.0-733.0. doi:10.35378/gujs.992738

27. Khoh W.H.; Pang Y.H.; Ooi S.Y.; Wang L.-Y.-K.; Poh Q.W. (2023). Predictive Churn Modeling for Sustainable Business in the Telecommunication Industry: Optimized Weighted Ensemble Machine Learning. Sustainability (Switzerland), 15(11), nan-nan. doi:10.3390/su15118631
28. Karanatsiou D.; Sermpezis P.; Gruda D.; Kafetsios K.; Dimitriadis I.; Vakali A. (2022). My Tweets Bring All the Traits to the Yard: Predicting Personality and Relational Traits in Online Social Networks. ACM Transactions on the Web, 16(2), nan-nan. doi:10.1145/3523749
29. Mirza O.M.; Moses G.J.; Rajender R.; Lydia E.L.; Kadry S.; Me-Ead C.; Thinnukool O. (2022). Optimal Deep Canonically Correlated Autoencoder-Enabled Prediction Model for Customer Churn Prediction. Computers, Materials and Continua, 73(2), 3757.0-3769.0. doi:10.32604/cmc.2022.030428
30. Loukili M.; Messaoudi F.; Ghazi M.E. (2022). Supervised Learning Algorithms for Predicting Customer Churn with Hyperparameter Optimization. International Journal of Advances in Soft Computing and its Applications, 14(3), 49.0-63.0. doi:10.15849/IJASCA.221128.04
31. Saias J.; Rato L.; Gon√ßalves T. (2022). An Approach to Churn Prediction for Cloud Services Recommendation and User Retention. Information (Switzerland), 13(5), nan-nan. doi:10.3390/info13050227
32. Bilal S.F.; Almazroi A.A.; Bashir S.; Khan F.H.; Almazroi A.A. (2022). An ensemble based approach using a combination of clustering and classification algorithms to enhance customer churn prediction in telecom industry. PeerJ Computer Science, 8(nan), nan-nan. doi:10.7717/PEERJ-CS.854
33. Rabbah J.; Ridouani M.; Hassouni L. (2022). A New Churn Prediction Model Based on Deep Insight Features Transformation for Convolution Neural Network Architecture and Stacknet. International Journal of Web-Based Learning and Teaching Technologies, 17(1), nan-nan. doi:10.4018/ijwltt.300342
34. Nhu N.Y.; Van Ly T.; Truong Son D.V. (2022). Churn prediction in telecommunication industry using kernel Support Vector Machines. PLoS ONE, 17(5 May), nan-nan. doi:10.1371/journal.pone.0267935
35. Sana J.K.; Abedin M.Z.; Rahman M.S.; Rahman M.S. (2022). A novel customer churn prediction model for the telecommunication industry using data transformation methods and feature selection. PLoS ONE, 17(12 December), nan-nan. doi:10.1371/journal.pone.0278095
36. Abdulsalam S.O.; Arowolo M.O.; Saheed Y.K.; Afolayan J.O. (2022). Customer Churn Prediction in Telecommunication Industry Using Classification and Regression Trees and Artificial Neural Network Algorithms. Indonesian Journal of Electrical Engineering and Informatics, 10(2), 431.0-440.0. doi:10.52549/ijeei.v10i2.2985
37. Mar√≠n D√≠az G.; Gal√°n J.J.; Carrasco R.A. (2022). XAI for Churn Prediction in B2B Models: A Use Case in an Enterprise Software Company. Mathematics, 10(20), nan-nan. doi:10.3390/math10203896
38. Arshad S.; Iqbal K.; Naz S.; Yasmin S.; Rehman Z. (2022). A Hybrid System for Customer Churn Prediction and Retention Analysis via Supervised Learning. Computers, Materials and Continua, 72(3), 4283.0-4301.0. doi:10.32604/cmc.2022.025442
39. Fridrich M.; Dost√°l P. (2022). User Churn Model in E-Commerce Retail. Scientific Papers of the University of Pardubice, Series D: Faculty of Economics and Administration, 30(1), nan-nan. doi:10.46585/SP30011478
40. Usman-Hamza F.E.; Balogun A.O.; Capretz L.F.; Mojeed H.A.; Mahamad S.; Salihu S.A.; Akintola A.G.; Basri S.; Amosa R.T.; Salahdeen N.K. (2022). Intelligent Decision Forest Models for Customer Churn Prediction. Applied Sciences (Switzerland), 12(16), nan-nan. doi:10.3390/app12168270
41. Mirkovic M.; Lolic T.; Stefanovic D.; Anderla A.; Gracanin D. (2022). Customer Churn Prediction in B2B Non-Contractual Business Settings Using Invoice Data. Applied Sciences (Switzerland), 12(10), nan-nan. doi:10.3390/app12105001
42. Ullah I.; Rios A.; Gala V.; McKeever S. (2022). Explaining deep learning models for tabular data using layer-wise relevance propagation. Applied Sciences (Switzerland), 12(1), nan-nan. doi:10.3390/app12010136
43. Moral R.A.; Chen Z.; Zhang S.; McClean S.; Palma G.R.; Allan B.; Kegel I. (2022). Profiling Television Watching Behavior Using Bayesian Hierarchical Joint Models for Time-to-Event and Count Data. IEEE Access, 10(nan), 113018.0-113027.0. doi:10.1109/ACCESS.2022.3215682

44. T√©kouabou S.C.K.; Gherghina »ò.C.; Toulni H.; Mata P.N.; Martins J.M. (2022). Towards Explainable Machine Learning for Bank Churn Prediction Using Data Balancing and Ensemble-Based Methods. Mathematics, 10(14), nan-nan. doi:10.3390/math10142379
45. Zhang T.; Moro S.; Ramos R.F. (2022). A Data-Driven Approach to Improve Customer Churn Prediction Based on Telecom Customer Segmentation. Future Internet, 14(3), nan-nan. doi:10.3390/fi14030094
46. Xiahou X.; Harada Y. (2022). B2C E-Commerce Customer Churn Prediction Based on K-Means and SVM. Journal of Theoretical and Applied Electronic Commerce Research, 17(2), 458.0-475.0. doi:10.3390/jtaer17020024
47. Thakkar H.K.; Desai A.; Ghosh S.; Singh P.; Sharma G. (2022). Clairvoyant: AdaBoost with Cost-Enabled Cost-Sensitive Classifier for Customer Churn Prediction. Computational Intelligence and Neuroscience, 2022(nan), nan-nan. doi:10.1155/2022/9028580
48. Muneer A.; Ali R.F.; Alghamdi A.; Taib S.M.; Almaghthawi A.; Abdullah Ghaleb E.A. (2022). Predicting customers churning in banking industry: A machine learning approach. Indonesian Journal of Electrical Engineering and Computer Science, 26(1), 539.0-549.0. doi:10.11591/ijeecs.v26.i1.pp539-549
49. Wang G.-Y. (2022). Churn Prediction for High-Value Players in Freemium Mobile Games: Using Random Under-Sampling. Statistika, 102(4), 443.0-453.0. doi:10.54694/STAT.2022.18
50. Sina Mirabdolbaghi S.M.; Amiri B. (2022). Model Optimization Analysis of Customer Churn Prediction Using Machine Learning Algorithms with Focus on Feature Reductions. Discrete Dynamics in Nature and Society, 2022(nan), nan-nan. doi:10.1155/2022/5134356
51. Bugajev A.; Kriauziene R.; Vasilecas O.; Chady≈°as V. (2022). The Impact of Churn Labelling Rules on Churn Prediction in Telecommunications. Informatica (Netherlands), 33(2), 247.0-277.0. doi:10.15388/22-INFOR484
52. Ha Thanh N.T.; Vy N.T. (2022). Building a proper churn prediction model for Vietnam's mobile banking service. International Journal of Advanced and Applied Sciences, 9(7), 139.0-149.0. doi:10.21833/ijaas.2022.07.014
53. Khodadadi A.; Hosseini S.A.; Pajouheshgar E.; Mansouri F.; Rabiee H.R. (2022). ChOracle: A Unified Statistical Framework for Churn Prediction. IEEE Transactions on Knowledge and Data Engineering, 34(4), 1656.0-1666.0. doi:10.1109/TKDE.2020.3000456
54. Fareniuk Y.; Zatonatska T.; Dluhopolskyi O.; Kovalenko O. (2022). CUSTOMER CHURN PREDICTION MODEL: A CASE OF THE TELECOMMUNICATION MARKET. ECONOMICS - Innovative and Economics Research Journal, 10(2), 109.0-130.0. doi:10.2478/eoik-2022-0021
55. Mahmood I.N.; Abdullah H.S. (2022). Telecom Churn Prediction based on Deep Learning Approach. Iraqi Journal of Science, 63(6), 2667.0-2675.0. doi:10.24996/ijs.2022.63.6.32
56. Faritha Banu J.; Neelakandan S.; Geetha B.T.; Selvalakshmi V.; Umadevi A.; Martinson E.O. (2022). Artificial Intelligence Based Customer Churn Prediction Model for Business Markets. Computational Intelligence and Neuroscience, 2022(nan), nan-nan. doi:10.1155/2022/1703696
57. AL-Najjar D.; Al-Rousan N.; AL-Najjar H. (2022). Machine Learning to Develop Credit Card Customer Churn Prediction. Journal of Theoretical and Applied Electronic Commerce Research, 17(4), 1529.0-1542.0. doi:10.3390/jtaer17040077
58. Mandiƒá M.; Kraljeviƒá G. (2022). Churn Prediction Model Improvement Using Automated Machine Learning with Social Network Parameters. Revue d'Intelligence Artificielle, 36(3), 373.0-379.0. doi:10.18280/ria.360304
59. Al-Shourbaji I.; Helian N.; Sun Y.; Alshathri S.; Elaziz M.A. (2022). Boosting Ant Colony Optimization with Reptile Search Algorithm for Churn Prediction. Mathematics, 10(7), nan-nan. doi:10.3390/math10071031
60. Kristijansson A.M.; Aegisson T. (2022). Survey on Technique and User Profiling in Unsupervised Machine Learning Method. Journal of Machine and Computing, 2(1), 9.0-16.0. doi:10.53759/7669/jmc202202002
61. √iskarsd√≥ttir M.; G√≠slad√≥ttir K.E.; Stef√°nsson R.; Aleman D.; Sarraute C. (2022). Social networks for enhanced player churn prediction in mobile free-to-play games. Applied Network Science, 7(1), nan-nan. doi:10.1007/s41109-022-00524-5

62. Ibitoye A.O.J.; Onifade O.F.W. (2022). Improved customer churn prediction model using word order contextualized semantics on customers,Äô social opinion. International Journal of Advances in Applied Sciences, 11(2), 107.0-112.0. doi:10.11591/ijaas.v11.i2.pp107-112
63. Elyusufi Y.; Kbir M.A. (2022). Churn Prediction Analysis by Combining Machine Learning Algorithms and Best Features Exploration. International Journal of Advanced Computer Science and Applications, 13(7), 615.0-622.0. doi:10.14569/IJACSA.2022.0130773
64. Sudharsan R.; Ganesh E.N. (2022). A Swish RNN based customer churn prediction for the telecom industry with a novel feature selection strategy. Connection Science, 34(1), 1855.0-1876.0. doi:10.1080/09540091.2022.2083584
65. Dalli A. (2022). Impact of Hyperparameters on Deep Learning Model for Customer Churn Prediction in Telecommunication Sector. Mathematical Problems in Engineering, 2022(nan), nan-nan. doi:10.1155/2022/4720539
66. Subramanian R.S.; Prabha D. (2022). Ensemble variable selection for Naive Bayes to improve customer behaviour analysis. Computer Systems Science and Engineering, 41(1), 339.0-355.0. doi:10.32604/csse.2022.020043
67. Vo N.N.Y.; Liu S.; Li X.; Xu G. (2021). Leveraging unstructured call log data for customer churn prediction. Knowledge-Based Systems, 212(nan), nan-nan. doi:10.1016/j.knosys.2020.106586
68. Sucu M.√á.; Unusan C. (2021). Predictions from an Empirical Study in the Turkish Mobile Telecommunications Market on the Determinants of Mobile Customer Churn. Journal of Telecommunications and the Digital Economy, 9(4), 178.0-189.0. doi:10.18080/jtde.v9n4.463
69. Ahn J.; Hwang J.; Kim D.; Choi H.; Kang S. (2020). A Survey on Churn Analysis in Various Business Domains. IEEE Access, 8(nan), 220816.0-220839.0. doi:10.1109/ACCESS.2020.3042657
70. H√∂ppner S.; Stripling E.; Baesens B.; Broucke S.V.; Verdonck T. (2020). Profit driven decision trees for churn prediction. European Journal of Operational Research, 284(3), 920.0-933.0. doi:10.1016/j.ejor.2018.11.072
71. Seyedan M.; Mafakheri F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. Journal of Big Data, 7(1), nan-nan. doi:10.1186/s40537-020-00329-2
72. Amin A.; Al-Obeidat F.; Shah B.; Adnan A.; Loo J.; Anwar S. (2019). Customer churn prediction in telecommunication industry using data certainty. Journal of Business Research, 94(nan), 290.0-301.0. doi:10.1016/j.jbusres.2018.03.003
73. Parwez M.A.; Abulaish M.; Jahiruddin (2019). Multi-Label Classification of Microblogging Texts Using Convolution Neural Network. IEEE Access, 7(nan), 68678.0-68691.0. doi:10.1109/ACCESS.2019.2919494
74. Ullah I.; Raza B.; Malik A.K.; Imran M.; Islam S.U.; Kim S.W. (2019). A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector. IEEE Access, 7(nan), 60134.0-60149.0. doi:10.1109/ACCESS.2019.2914999
75. Ahmad A.K.; Jafar A.; Aljoumaa K. (2019). Customer churn prediction in telecom using machine learning in big data platform. Journal of Big Data, 6(1), nan-nan. doi:10.1186/s40537-019-0191-6