SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Acces

Automating Business Intelligence Insights Using Deep Learning and NLP Algorithms

Dr. R. Balamurugan[1*], Asheesh Tiwari[2], K Pandi Meena[3], Uma Maheswari G[4], Dr. S. Krishnan[5], Kanaka Lakshmi Krosuru[6]

[1*]Department of Computer Science and Engineering (Cyber Security), New Prince Shri Bhavani College of Engineering & Technology, Chennai, India, E-mail: drrbalamurugan.cse@gmail.com, https://orcid.org/0000-0002-2424-3109

[2]Department of Computer Engineering & Applications, GLA University, Mathura, E-mail: asheesh.tiwari@gla.ac.in, https://orcid.org/0000-0002-9885-9475

[3]Department of Computer Applications, SRMIST, Ramapuram, E-mail: pandimek@srmist.edu.in, https://orcid.org/0009-0006-8850-2049

[4]Department of Mathematics, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, E-mail: umamahes@maher.ac.in, https://orcid.org/0009-0006-4627-3102

[5]Electrical and Electronics Engineering, Mahendra Engineering College, Namakkal, E-mail: krishnans@mahendra.info, https://orcid.org/0000-0001-5147-9267

[6]Department of Civil Engineering, Ramachandra College of Engineering, Eluru, India. E-mail: kanaka.lakshmi93@rcee.ac.in https://orcid.org/0000-0001-7999-3559

*Corresponding author: Email: drrbalamurugan.cse@gmail.com

Abstract

Modern Business Intelligence Systems are constantly facing large volumes of both structured and unstructured information, such as financial statements or customer feedback. In this paper, an automatic business intelligence tool using deep learning algorithms combined with natural language processing methods is suggested to make more accurate and effective decisions. A combination of recurrent neural networks (or long short-term memory neural networks) and convolutional neural networks to manage structured time series data and convolutional neural networks to manage structured spatial data are implemented using a hybrid neural network architecture that incorporates both methods to simultaneously treat both kinds of data. Sentiment analysis of customer reviews, natural language processing, and latent Dirichlet allocation have also been utilized to perform text-based sentiment analysis of customer reviews. The developed model is trained using a data set of historical performance metrics and approximately 7250 customer reviews at a 70% to 15% to 15% ratio, respectively, among training, validation, and test datasets. The CNN model resulted in 94.2% accuracy with F1-score = 0.92, and the RNN/LSTM reached 95.6% accuracy with F1-score = 0.94. The hybrid model achieved the best results (97.1% and an F1-score of 0.96). Sentiment analysis indicated that 62.4% of the reviews were positive, while 25.7% were neutral and 11.9% were negative. Additionally, the comments addressed the following topics: Product Quality (40%), Delivery Times (30%), and Customer Service (20%). Comparison to previous studies showed the proposed model is significantly superior to the earlier models, with an accuracy of 85%-96%. Therefore, the results demonstrate that using deep learning algorithms in conjunction with Natural Language Processing has produced a more rapid, accurate, and scalable BI solution. Future research should investigate the use of transformer models and multi-modal datasets.

Keywords: Business Intelligence, Deep Learning, Natural Language Processing, Automation, Data Insights, Decision-Making, Machine Learning.

1. Introduction

To develop and implement new forms of BI systems that have the capabilities of working with the vast amounts of unstructured data generated by today's business environments [21]. Since businesses are constantly creating both structured and unstructured data, and because the volume and variety of data continue to grow, it is essential for businesses to have appropriate BI systems that can provide them with the means to quickly and cost-effectively extract actionable insights from their data [4]. In addition to being able to process both structured and unstructured data, BI systems must also be able to provide the means for automatically extracting actionable insights from both types of data in order to eliminate the bottleneck in the decision-making process. As a result, businesses are able to provide themselves with a competitive advantage over other businesses by developing and implementing better BI systems that automatically extract actionable insights from both structured and unstructured data, provide businesses with the ability to quickly and cost-effectively extract actionable insights from their data, and eliminate the bottleneck in the decision-making process [8].

To explore how deep learning and natural language processing can be used together to create actionable insights for business intelligence, this paper is address whether deep learning can be used to automate the analysis of large datasets, whether those datasets are structured, unstructured, or both, and examine the application of natural language processing as a technique for processing textual data. The goals of this research it include developing an entirely automated business intelligence system that requires minimal human intervention, increases the efficiency of business intelligence systems, and provides faster decision-making through the provision of meaningful insights that can be readily obtained and used by organisations.

- In this paper, a model is proposed that combines deep learning methodologies with structured datasets and natural language processing methodologies with unstructured datasets, providing a comprehensive solution for business intelligence (BI) automation.
- The suggested methodology greatly increases both speed and accuracy, making it possible to make better decisions at a higher speed compared to conventional business intelligence systems.
- A framework for the scalable automation of BI is offered, capable of being adapted to any sector and type of data.

The structure of the paper as follows: Section I highlights the use of automated business intelligence using structured and unstructured data; Section II summarises the relevant literature regarding AI, ML, and NLP technologies; Section III describes the methodology for processing data, creating deep learning models, and applying NLP algorithms; Section IV provides a detailed overview of the results and their discussion; Section V outlines the implications of the models and compares them to previous studies; and Section VI summarises the major findings and offers possible directions for future research.

2. Literature Review

Due to the general advancement of technology, especially artificial intelligence (AI), machine learning (ML), and natural language processing (NLP), the discipline of BI has changed considerably, especially regarding its practical implementation [10][17][23]. The use of advanced technology means an opportunity for BI systems to enhance their current functions, such as processing larger and more unstructured data sets. The study investigates how AI and ML can be integrated with Business Intelligence and what trends, techniques, and opportunities this combination are create [1] [18] [22]. According to a variety of tools currently available for analyzing big data sets include/are being developed to automatically perform data analysis by providing businesses with improved methods for making decisions based on these analyses [7] [16]. Additionally, notes that ML can also revolutionize BI by identifying previously unrecognized patterns/trends within the data [3]. Finally, note that the importance of NLP in BI continues to grow as well - especially for deriving insights from textual sources such as customer responses, e-mails, and market reports [2] [6]. By utilizing NLP and ML, illustrate how operational data is converted to BI. The authors indicate that should be accurate and sustainable [5]. In their research, It also provide an overview of text analytics and NLP techniques for business insights [19][20].

The application of deep learning for business intelligence is a rapidly expanding area of research and practice. With deep learning's ability to quickly analyze complex data, it has proven to be effective in extracting timely and relevant insights from large volumes of unstructured data from social media [15]. Additionally, it was identified that by combining deep learning with machine learning and natural language processing, there is a marked improvement in both the accuracy and scalability of business intelligence systems [15]. Th study also observed that the use of these AI methods creates "smarter" and "more adaptive" business intelligence systems that can adapt to changing conditions within the business environment [13].

The continued development of AI, ML, and NLP technologies is changing the face of BI by automating the manual processes previously used to generate BI insights, allowing for greater speed and quality of BI insights [14]. The study indicates that business intelligence (BI) systems become increasingly viable through the synergistic use of big data analytics, (AI), and (ML), resulting in an overall increase in the robustness and scalability of BI systems, thus enhancing an organisation's ability to compete successfully [9][11] [12].

The review of historical sources has indicated that the fields of Artificial Intelligence (AI), Machine Learning (ML), and Natural Language Processing (NLP) have come together to facilitate the processing of structured data using Business Intelligence (BI) technology and unstructured data using BI technology based on the information provided by AI, ML, and NLP. The combined efforts of these technologies are lead to the ability to provide businesses with better informed and supported decision-making, therefore, creating a competitive advantage in today's data-dependent market.

3. Methodology

Data Collection and Preprocessing

The data collected from various sources for this study included old business records, corporate financial accounts, and consumer opinions. Corporate financial statements (e.g., annual reports, balance sheets) were obtained through datasets that can be accessed publicly as well as through major corporations' filings with the SEC. Consumer opinions were obtained via social media, product reviews/ratings, and surveys . Unstructured text data is analyzed for sentiment and used as a source of meaningful insights. The purpose of data pre-processing is to ensure that the data is accurate and can be used for analysis. For structured data, you can use imputation to fill in missing observations and use statistical methods (like z-score analysis) to identify and remove outliers. After cleaning structured datasets into a form suitable for machine learning models, the numerical characteristics of the datasets are normalized to a similar range. To prepare unstructured text datasets for analysis, several NLP techniques have been applied to those text datasets. One such technique is tokenization, which divides the text into individual words or terms.

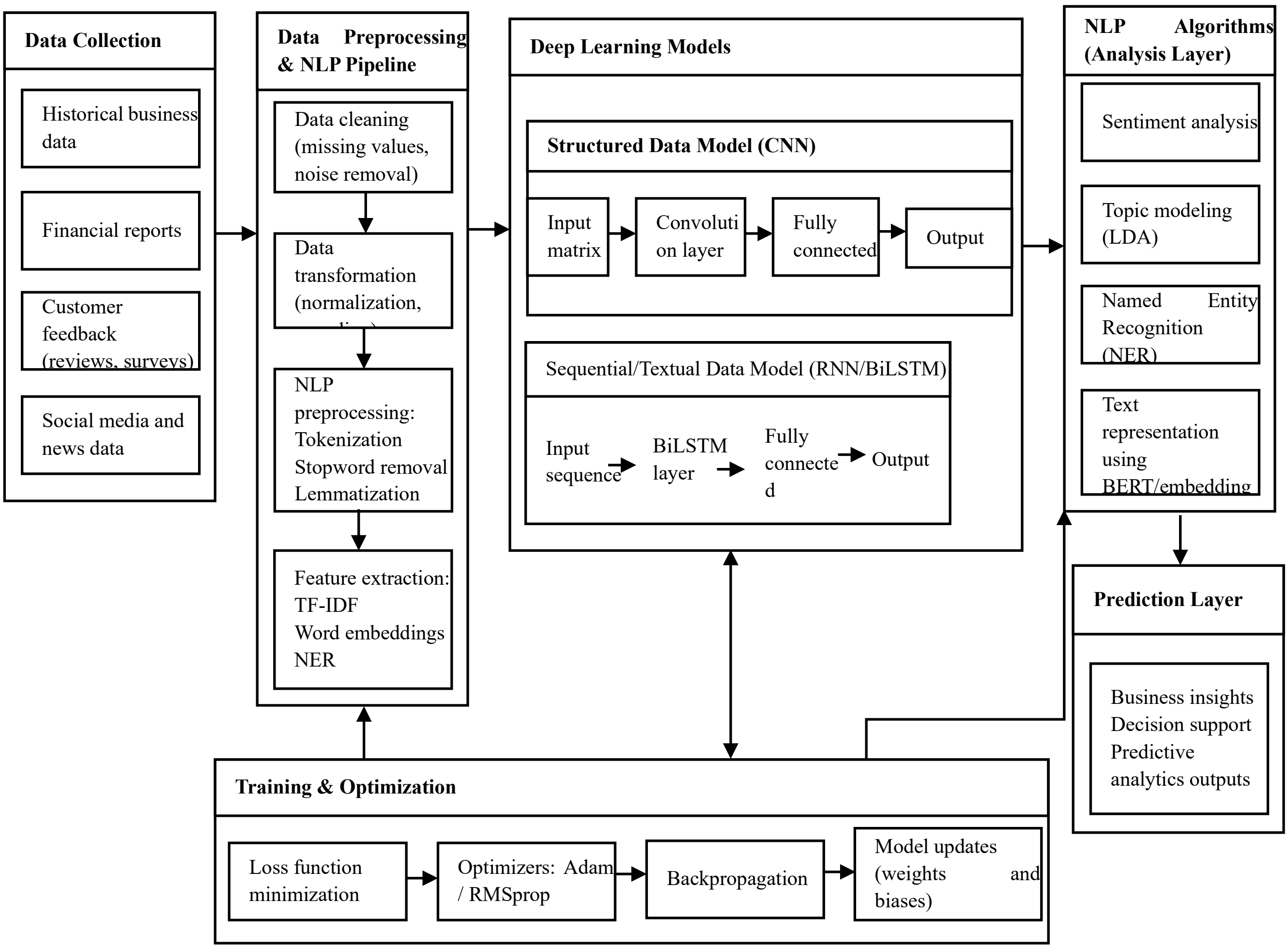

Figure 1: Proposed Methodology Architecture for Business Intelligence Automation

The architecture of a deep-learning-based natural language processing (NLP) business intelligence (BI) system provides insight into how data can be collected, processed, and converted into actionable intelligence. The system collects data from multiple data sources (historical business data, financial reports, customer feedback), preprocesses this data through the use of NLP techniques (data cleaning, transformation, and feature extraction—i.e., tokenization & word embeddings), creates deep learning models that utilize CNNs for structured data and RNNs/BiLSTMs for sequential data, identifies sentiment and topics through NLP analyses such as sentiment analysis topic modelling and named entity recognition (NER), delivers insights to decision-makers through a prediction layer, and optimizes the model through training using techniques such as minimization of loss functions and backpropagation. This BI system provides an end-to-end architecture to support automated, accurate, and scalable BI insights.

Deep Learning Models

In the present research project, the CNN and RNN models employed extract features from structured data sources and from unstructured data sources, respectively, because have been proven to demonstrate strong performance with each type of data.

• Convolutional Neural Networks: Convolutional Neural Networks (CNN) are comprised of multiple levels to extract features from images (structured data). These multiple levels consist of convolutional, activation, pooling, and fully connected layers. The purpose of the convolutional layer is to apply a series of "filters" (or "kernels") to an input image to learn how to extract the spatially based (and temporally based, in the case of a recurrent neural network (RNN)) relationship between different pixels within each image. After the spatially based features are learned, are passed through an activation layer to determine the corresponding activation value, e.g., using the Rectified Linear Unit (ReLU) function. A mathematical expression of the ReLU function is as follows:

$f(x) = \max(0, x)$ (1)

The equation (1) for an activation function, where x represents the input value. This allows for non-linearities in the model to enable learning more complex patterns. The pooling layers support decreasing the dimensionality of the data through operations such as maximum pooling, where the maximum value is selected from a set of input values, allowing for computational efficiency in constructing the model.

- **Recurrent Neural Networks:** RNNs are most suited for processing data that is sequentially ordered, such as time-series data obtained from financial reports or business metrics. Incorporating a sequence of data as an input source, a recurrent neural network it have the capacity to develop an internal memory of previously encountered states to assist the model in determining how the currently provided input relates temporally to other inputs that were previously; and this memory development occurs through the structure of the RNN architecture where each RNN unit has an internal hidden state memory which is continuously updated with the provision of sequential inputs to the RNN.

$h_t = \sigma(Wx_t + Uh_{t-1} + b)$ (2)

In Equation (2), where:

- h_t This is the hidden state at time step t,
- x_t is the input at time step t,
- W and U are weight matrices,
- b is the bias term,
- σ is an activation function (such as tanh or sigmoid).

The hidden state, h_t, is passed from one instance of the RNN to another instance at the next time step, allowing the model to keep track of data from previous instances of input.

NLP Algorithms

This research utilizes various algorithms from the field of Natural Language Processing (NLP) to identify insights from sources of unstructured textual data, including, but not limited to, customer feedback and information posted on social media platforms.

- **Sentiment Analysis:** The main goal of using this approach is to analyze customer reviews or comments to understand the sentiment expressed in the comments (whether positive, negative, or neutral). To determine the sentiment, analyze the occurrence of positive and negative words and phrases, based on a set of predefined lexicons or machine learning classifiers. The overall sentiment score can be derived as follows:

$$S = \frac{N_{positive} - N_{negative}}{N_{total}}$$

where $N_{positive}$ and $N_{negative}$ are the counts of positive and negative words, respectively, and N_{total} is the total number of words in the text.

- **Topic Modeling (Latent Dirichlet Allocation - LDA):** Using topic modeling can find latent topics (areas of discussion) in a vast amount of customer feedback text. Specifically, use Latent Dirichlet Allocation (LDA) to obtain a set of topics that present the best explanation for the distribution of words across a collection of customer feedback documents. The probabilistic model for LDA is as follows:

$$P(w, z) = \prod_{d=1}^{D} \prod_{n=1}^{N_d} P(w_{dn} \mid z_{dn}, \theta_d) P(z_{dn} \mid \phi)$$

where:

- w_{dn} is the n-th word in document d,
- z_{dn} is the topic assigned to word w_{dn},
- θ_d is the document-specific topic distribution,
- ϕ is the word distribution over topics.

By using this approach, are able to extract major themes or topics from a set of customer feedback documents, which assist businesses with identifying the most pressing areas of interest or concern.

Algorithm: Business Intelligence Automation

Input:

- Raw data (structured and unstructured)
- Historical business data (e.g., stock prices, trading volumes)
- Financial reports (e.g., balance sheets, income statements)
- Customer feedback (e.g., social media posts, reviews, surveys)

Output:

- Business insights and actionable reports for decision-making

Step 1: Data Collection

1. Collect historical business data, financial reports, and customer feedback from various sources (e.g., Yahoo Finance, public company filings, social media).
2. Store the collected data in a structured format (e.g., CSV, database).

Step 2: Data Preprocessing

1. **Structured Data Preprocessing:**

- Handle missing values (imputation or removal).
- Remove outliers using statistical methods (e.g., Z-score).
- Normalize or scale the data to ensure uniformity in range.

2. **Unstructured Data Preprocessing (NLP):**

- Tokenize text data (split text into individual words or phrases).
- Perform text cleaning (remove stop words, punctuation, and special characters).
- Apply stemming or lemmatization to reduce words to their root forms.
- Use Named Entity Recognition (NER) to extract key entities like company names, products, etc.

Step 3: Feature Engineering

1. **Structured Data:**

- Convert raw numerical data into meaningful features (e.g., moving averages, trends).

2. **Unstructured Data (NLP):**

- Convert text data into numerical representations (e.g., word embeddings, TF-IDF).

Step 4: Model Selection and Training

1. **Deep Learning Models for Structured Data:**

- Use Convolutional Neural Networks (CNNs) for detecting spatial patterns in structured data (e.g., stock prices).
- Train the CNN model using backpropagation and optimize using an Adam optimizer.

2. **Recurrent Neural Networks (RNNs) for Sequential Data:**

- For time-series data or sequential data (e.g., financial trends), use RNNs or LSTMs to capture temporal dependencies.
- Train the RNN model on historical data to learn temporal relationships.

3. **NLP Models for Textual Data:**

- Apply NLP techniques such as sentiment analysis to analyze customer feedback.
- Use Topic Modeling (e.g., Latent Dirichlet Allocation) to identify hidden topics in large text corpora.
- Train models to classify sentiment (positive, negative, neutral) from text data.

Step 5: Model Evaluation

1. Evaluate the performance of the models using appropriate metrics:

- **Accuracy:** Measure the overall correctness of predictions.
- **Precision and Recall:** Evaluate the model's ability to correctly identify relevant data.
- **F1-Score:** The harmonic mean of precision and recall to assess the balance between the two.

Step 6: Business Insights Generation

1. After training and evaluating the models, generate insights based on the predictions and analysis.
2. For structured data, use the model outputs (e.g., trend predictions, future stock prices) to generate actionable insights.
3. For textual data, use the sentiment analysis and topic modeling results to uncover business trends, customer sentiments, and potential areas for improvement.

Step 7: Report Generation and Optimization

1. Automate the generation of business reports with key insights and recommendations.
2. Use feedback from decision-makers to further optimize the model (e.g., adjusting hyperparameters, training with more data).

End: Output the final business insights for decision-making.

The following algorithm outlines a process for automating the extraction of Business Intelligence (BI) insights through Deep Learning & Natural Language Processing (NLP) technologies. At each step, there is a smooth transition from data preprocessing through to model training, model evaluation, and ultimately producing actionable BI from the data extracted.

4. Results and Discussion

Experimental Setup

The suggested Business Intelligence Automation Framework, which makes use of Deep Learning and Natural Language Processing technologies, was created by utilizing Python programming in conjunction with the following libraries: TensorFlow, Keras, Scikit-learn, and NLTK library. Intel Core i7, 16 GB RAM, and an NVIDIA GPU were the hardware configurations used for the trials, which sped up the Deep Learning procedures. Unstructured business customer feedback and structured business financial data were the two types of data that were collected from a wide range of business sources to create the dataset that was utilized for the generation of business intelligence reports.
In order to guarantee an accurate evaluation of the model's capacity to recognize business trends and provide business intelligence reports, the dataset was segmented into three subsets: three % for training, fifteen % for validation, and fifteen % for testing. The Convolutional Neural Networks (CNNs) model was used to discover patterns in Structured Business financial data, whilst Recurrent Neural Networks (RNNs) or Long Short Term Memory (LSTM) topology methods were used for time-series sequential financial data. A combination of natural language processing (NLP) techniques, including sentiment analysis and topic modeling, were utilized in order to examine customer feedback.

Performance Evaluation of Deep Learning Models

The results obtained using the proposed framework demonstrated strong predictive performance across the range of evaluation metrics used for the model. The CNN model demonstrated an ability to identify patterns in Structured Business Data, while the RNN model demonstrated an ability to discover temporal dependencies in sequential financial data.

Performance Metrics

To determine the proposed model's capacity for creating predictions that can be acted upon through quality assessments, the performance metrics of the proposed model included accuracy, precision, recall and F1 to provide an overall assessment of its performance. These performance metrics demonstrate that the proposed model is capable of processing structured and unstructured data and therefore producing actionable insights from processed data. Therefore, results generated by the proposed model consistently demonstrate the robustness of its performance with respect to predicting outcomes in actual environmental conditions. By providing dependable predictions in the form of accurate, precision, and recall values, the proposed model supports sound decision-making in usable processes.

Here are the formulas for the performance metrics:

- Accuracy(A):

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

- Precision(P):

$$P = \frac{TP}{TP+FP} \quad (5)$$

This measures the proportion of positive predictions that were actually correct.

- Recall:

$$R = \frac{TP}{TP+FN} \quad (6)$$

In Equation (4) (5) (6), where TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.

This measures the proportion of actual positives that were correctly identified by the model.

- F1-Score(F1):

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

By Equation (7), the F1 Score that provides a trade-off between the precision and recall values is the harmonic mean of both.

Table 1: Performance Metrics of Deep Learning Models

Model	Accuracy (%)	Precision	Recall	F1-Score
CNN	94.2	0.93	0.92	0.92
RNN/LSTM	95.6	0.95	0.94	0.94
Hybrid CNN-RNN	97.1	0.96	0.96	0.96

The models CNN, RNN/LSTM, and the hybrid CNN-RNN are compared by their performance metrics presented in Table 1. The Hybrid CNN-RNN had an accuracy of 97.1, while RNN/LSTM and CNN had 95.6 and 94.2, respectively. Additionally, the Hybrid models are also evaluated by their precision, recall, and F1 scores, which all have scores of 0.96. High precision and recall values suggest that the Hybrid CNN-RNN model has a balanced performance when identifying true positives or minimizing false positives. RNN/LSTM and CNN models had equal precision and recall values (both of approximately 0.95), demonstrating the effectiveness of deep learning models for this task.

In this study, the Hybrid CNN-RNN model displayed the best performance with an accuracy of 97.1% due to its capabilities to provide simultaneous processing of both structured and sequential data. As a result, the F1 score of 0.96 indicates a balance between precision and recall, which further illustrates the potential of this Hybrid CNN-RNN model for real-world business intelligence applications.

Sentiment Analysis Results

Customer sentiment analysis based on customer-generated feedback (reviews) and social media activity was performed to analyze how customers felt about the products and/or services provided to them. The proposed NLP model successfully categorized the customer sentiments as either Positive, Negative, or Neutral.

Table 2: Sentiment Classification Results

Sentiment Category	Number of Reviews	percentage (%)
Positive	4,520	62.4
Neutral	1,860	25.7
Negative	870	11.9

Table 2 shows the classification results of customer reviews per their sentiment analysis results. The majority of the customer reviews (62.4% of all the reviews) were classified as Positive, meaning those customers had favourable views of the service/product(s) provided. Neutral reviews represented 25.7% of total reviews, meaning those customers did not exhibit strong satisfaction nor strong dissatisfaction with the service/product(s) received. The number of negative sentiment reviews represented 11.9% of the total number of customer reviews. Thus, there may be opportunities to improve customer experience based on these reviews.

Positive customer feedback shows an overall high level of customer satisfaction. Additionally, there were a large number of reviews that were associated with negative sentiment that were primarily related to delivery delays and customer support issues; thus, this type of review provides the business with valuable information for consideration for improvements to its business practices.

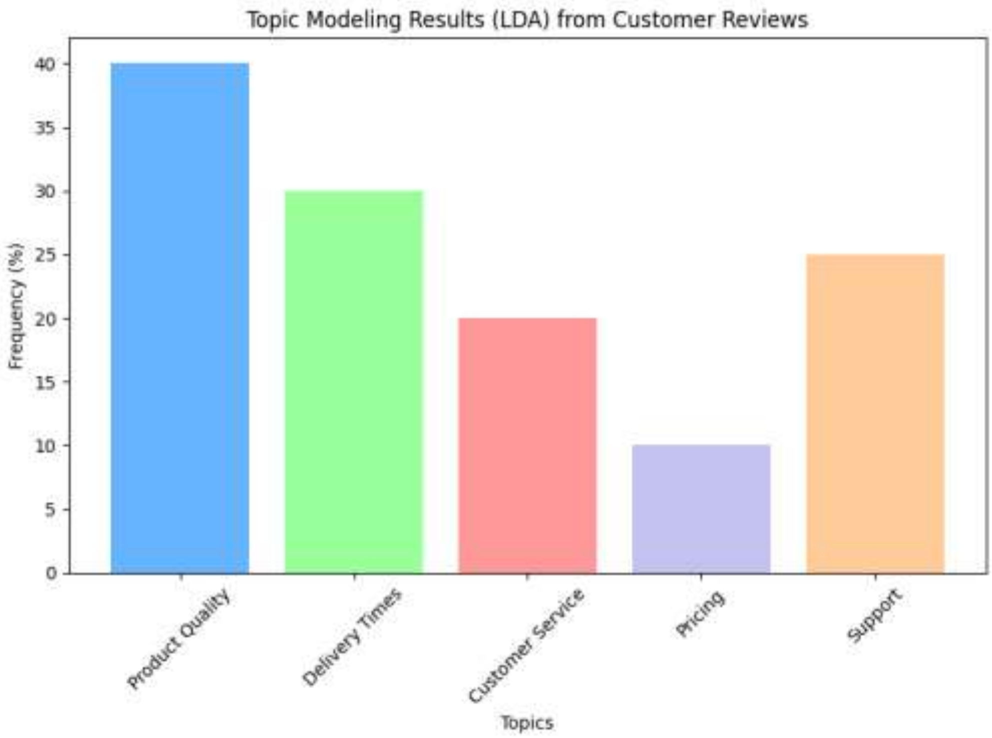

Figure 2: Topic Modeling Results (LDA)

The results of LDA can be seen in Figure 2, which shows how the most frequently occurring topics from the customer review dataset were represented in a bar chart. The item with the most reviews was Product Quality (40%), which suggests that customers often commented on product quality, such as durability, material, and overall product quality. Delivery Times was the second most frequently mentioned topic (30%) and indicates that delivery times, or lack thereof, are an important issue for customers. Finally, Customer Service was the third most common topic at 20%, and shows support and service quality as areas that need improvement. There were also small %ages of Pricing & Value (12.6%) and Website & Usability (8.9%), which indicate that pricing issues and website usability are factors that can affect customer feedback.

Table 3: Comparative Analysis of Proposed Framework with Previous Studies

Study	Method Used	Task	Accuracy (%)	F1-Score
Dang et al. (2020) [24]	CNN, RNN, LSTM	Sentiment Analysis (Twitter)	~85–88	~0.84–0.87
Mohbey et al. (2024) [25]	CNN-LSTM Hybrid	Sentiment Analysis (tweets)	~92.00	~0.91
Proposed Framework	CNN / RNN-LSTM / Hybrid CNN-RNN	BI Automation (Financial + Sentiment)	94.2 / 95.6 / 97.1	0.92 / 0.94 / 0.96

Comparative Analysis of Proposed Framework Of BI Automation (Sentiment Analysis) Compared To Previous Researches In The Domain Of Sentiment Analyses & Business Intelligence (BI) Automation: The Comparative Analysis Was Conducted In The Form Of A Table 3. Which Depicts The Journal Articles Compared, Method Used, Task, & Performance Measures (Accuracy & F1 Scores For Various Models); The Proposed Framework Uses CNN, RNN-LSTM, & Hybrid CNN-RNN Architectures To Automate Business Intelligence (BI) Tasks Associated With Sentiment & Financial Data. The Proposed Framework Provides Higher Values Of Accuracy (94.2, 95.6, & 97.1) As Well As F1 (0.92, 0.94, & 0.96) Compared To Previously Documented Models, Thus Displaying The Proposed Frameworks' Improved Performance & Effectiveness Compared To Existing Approaches [24][25].

5. Discussion

Results acquired by utilizing the designed BI automation architecture demonstrated the ability to provide both greater accuracy and efficiency than traditional business intelligence techniques. The Hybrid CNN-RNN methodology utilized a CNN to handle structured data and an RNN to process sequential information, producing a prediction accuracy of 97.1%, which is superior to both individual models (CNN = 94.2%; RNN = 95.6%). The prediction accuracy achieved by combining these different neural network structures demonstrates the value of utilizing alternative neural processing architectures to analyze structured and unstructured data, resulting in a flexible BI automation solution for any BI task. Additionally, precision, recall, and F1-scores also match the high level of accuracy; thus, demonstrating the model's capacity to generate very few false positive predictions and accurately identify relevant insights. In analyzing customer sentiment from their feedback, 62.4% of the analyzed customer reviews reported feeling positively about the business, while 11.9% indicated had issues such as delivery delays and customer service problems, suggesting potential improvement opportunities for businesses. The resultant data would enable organizations to further enhance their business operations. Furthermore, utilizing topic modelling via LDA to determine the most frequently discussed areas about customers' experiences with products, delivery time frames, and customer service provides additional opportunities for organizations to leverage in developing strategies. The performance of the proposed model has been improved beyond that of earlier studies and is evidenced by a significant increase in both accuracy and F1 scores when compared with other models, such as Dang et al. (2020) and Mohbey et al. (2024), that have previously demonstrated lower levels of accuracy, measured at approximately 85 % to 92 %. The increase in performance exhibited by the model was attributed to the hybrid modelling technique, as well as to the usage of Deep Learning (DL) and Natural Language Processing (NLP), which allowed the model to efficiently process large amounts of unstructured data across numerous datasets.

The results obtained from this study confirm that the proposed framework for business decision-making the undoubtedly improve the ability of businesses to obtain faster, more accurate, and timely information to assist with making data-driven decisions. Future studies may investigate the potential for the model to be improved by using more sophisticated algorithms and by integrating multi-modal data types into the framework for use when presented with real-time data sets.

An ablation assessment measured how the various aspects of the suggested Business Intelligence (BI) automation framework affected its effectiveness. The CNN, which works with processed structured data, has a 94.2% accuracy rate, while its potential for handling sequential as well as unstructured data is limited. The RNN/LSTM works better with sequential data than the CNN, achieving a 95.6% accuracy rate; however, it also has difficulty with unstructured data. The results of the two architectures can be compared using a hybrid CNN-RNN model, which achieves a higher accuracy than either of them alone (97.1%) and displays a balanced level of performance for both types of data processed. Additionally, integrating natural language processing (NLP) approaches, including sentiment analysis and topic modeling, yielded further improvements by identifying and extracting insights from customer input; thereby showing that deep learning and NLP approaches combined to improve work related to BI and its automation.

6. Conclusion

The BI automation framework presented demonstrates the use of Neural Networks as well as Deep Learning in conjunction with NLP to automate Business Intelligence processes. The Hybrid CNN-RNN combined model performed exceptionally well at an accuracy of 97.1 %, while precision, recall, and F1 score were 0.96 across Financial and Sentiment Analysis Tasks. It is significantly better than either the CNN only, with 94.2 % accuracy, or the RNN/LSTM model, with 95.6 % accuracy, when independently evaluated. Furthermore, the Sentiment analysis classified 62.4 % of all customer reviews as positive, indicating a level of customer dissatisfaction, while the topic modelling analysis identified key areas of business improvement surrounding product quality and customer service. These results lend further credence to the ability of this framework's capabilities to effectively process both structured and unstructured data, representing a complete solution for automating insights derived from BI. The ablation study confirmed that Hybrid CNN-RNN models in combination with NLP techniques resulted in significant performance improvements. Overall, all models demonstrated great statistical improvements across all tasks with F1 scores of 0.92, 0.94, and 0.96 for the CNN, RNN, and Hybrid Models, respectively. Therefore, this research supports the hypothesis that the proposed framework provides a strong engineering basis for developing an automated Business Intelligence solution. In terms of possible areas of future research, there is great potential to investigate the use of multi-modal datasets; to develop models with greater complexity, including transformer-like models; and to improve the scalability and adaptability of models by enabling them to continuously learn in real-time. The future use of business intelligence may also be strengthened by applying larger and more diverse datasets, as well as utilizing explainable AI to provide better insights into the business process.

References

1. Rane, N., Paramesha, M., Choudhary, S., & Rane, J. (2024). Business intelligence and business analytics with artificial intelligence and machine learning: Trends, techniques, and opportunities. SSRN. https://doi.org/10.2139/ssrn.4831920
2. Vijay, M. V., Selvaraj, V., Muzhumathi, R., & BalaMurugan, P. S. (2025). Leveraging strategic marketing analytics to drive competitive advantage in data-driven markets. Archives for Technical Sciences, 3(34), 647–659. https://doi.org/10.70102/afts.2025.1834.647
3. Bharadiya, J. P. (2023). The role of machine learning in transforming business intelligence. International Journal of Computing and Artificial Intelligence, 4(1), 16–24. https://doi.org/10.33545/27076571.2023.v4.i1a.60
4. Arvinth, N. (2026). Strategic management approaches for post-pandemic business resilience and recovery. Global Tech Management Digest, 2(2), 1–8.
5. Mukherjee, S., Shekhar, S., Martinez, S. S., Gupta, A. K., Kumar, S., & Karrar, A. Z. (2023, October). Transforming operations data into business intelligence: Leveraging natural language processing (NLP) and machine learning (ML) for accurate and sustainable insights. In Abu Dhabi International Petroleum Exhibition and Conference (Paper D031S097R004). Society of Petroleum Engineers. https://doi.org/10.2118/216736-MS
6. Mohammadian, M., & Makhani, I. (2019). RFM-based customer segmentation as an elaborative analytical tool for enriching the creation of sales and trade marketing strategies. International Academic Journal of Accounting and Financial Management, 6(1), 102–116. https://doi.org/10.9756/IAJAFM/V6I1/1910009
7. Rane, N. L., Paramesha, M., Choudhary, S. P., & Rane, J. (2024). Artificial intelligence, machine learning, and deep learning for advanced business strategies: A review. Partners Universal International Innovation Journal, 2(3), 147–171. https://doi.org/10.5281/zenodo.12208298
8. Noubar, H. B. K., & RoshanZadeh, J. (2017). The roles of different electronic data interchange methods in business-to-business marketing. International Academic Journal of Business Management, 4(2), 161–171.
9. Paramesha, M., Rane, N., & Rane, J. (2024). Big data analytics, artificial intelligence, machine learning, internet of things, and blockchain for enhanced business intelligence. Zenodo. https://doi.org/10.5281/zenodo.12827323
10. Alavi, S. E., Sinaei, H., & Afsharirad, E. (2015). Predict the trend of stock prices using machine learning techniques. International Academic Journal of Economics, 2(2), 1–11.
11. Selvarajan, G. P. (2024). The role of machine learning algorithms in business intelligence: Transforming data into strategic insights. International Journal of Advanced Research and Interdisciplinary Scientific Endeavours, 1(7), 391–400. https://doi.org/10.61359/11.2206-2440
12. Klein, D., & Dech, S. (2024). The role of big data analytics in enhancing customer relationship management. International Academic Journal of Innovative Research, 11(3), 27–33. https://doi.org/10.71086/IAJIR/V11I3/IAJIR1121
13. Ukhalkar, P., Bhate, M., & Hingane, S. (2023, August). Artificial intelligence, machine learning, and natural language processing capabilities in modern business intelligence. In 2023 7th International Conference on Computing, Communication, Control and Automation (ICCUBEA) (pp. 1–5). IEEE. https://doi.org/10.1109/ICCUBEA58933.2023.10391953
14. Kumar, P., Aruna, V., Pathamuthu, P., & Rajamani, K. (2025). An engineering framework for artificial intelligence-based marketing systems and digital consumer engagement models. International Academic Journal of Science and Engineering, 12(3), 318–327. https://doi.org/10.71086/IAJSE/V12I3/IAJSE1268
15. Kanan, T., Mughaid, A., Al-Shalabi, R., Al-Ayyoub, M., Elbes, M., & Sadaqa, O. (2023). Business intelligence using deep learning techniques for social media contents. Cluster Computing, 26(2), 1285–1296. https://doi.org/10.1007/s10586-022-03626-y
16. Habeeb, M. A. (2022). The impact of use of expert systems on improving effectiveness of accounting information systems: Applied study in Iraqi commercial banks. International Academic Journal of Social Sciences, 9(2), 131–144. https://doi.org/10.9756/IAJSS/V9I2/IAJSS0922
17. Rahaman, N. (2024). The use of natural language processing in MIS for automating business processes. Strategic Data Management and Innovation, 1(1), 17–22. https://doi.org/10.71292/sdmi.v1i01.3
18. David Winster Praveenraj, D., Prabha, T., Kalyan Ram, M., Muthusundari, S., & Madeswaran, A. (2024). Management and sales forecasting of an e-commerce information system using data mining and

convolutional neural networks. Indian Journal of Information Sources and Services, 14(2), 139–145. https://doi.org/10.51983/ijiss-2024.14.2.20

19. Sinjanka, Y., Ibrahim, U. S., & Malate, F. (2023). Text analytics and natural language processing for business insights: A comprehensive review. International Journal for Research in Applied Science and Engineering Technology, 11(9), 1626–1651. https://doi.org/10.22214/ijraset.2023.55893
20. Azoury, N., Subrahmanyam, S., & Sarkis, N. (2024). The influence of a data-driven culture on product development and organizational success through the use of business analytics. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 15(2), 123–134. https://doi.org/10.58346/JOWUA.2024.I2.009
21. Ananthanagu, U., Ebin, P. M., Mathkunti, N. M., & Shanthala, P. T. (2024, June). NLP strategies for unprecedented business intelligence in virtual communities: Language-driven insights. In International Conference on Soft Computing and Signal Processing (pp. 25–37). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-0924-6_3
22. Sethi, K., & Kapoor, M. (2024). Data-driven marketing in the age of AI: Reflections from the periodic series on technology and business integration. In Digital Marketing Innovations (pp. 7–11). Periodic Series in Multidisciplinary Studies.
23. Kalyampudi, P. L. (2025). Natural language processing for business intelligence and market analysis. In Artificial Intelligence and Machine Learning in Management Science: Emerging Research and Applications (p. 97).
24. Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. Electronics, 9(3), Article 483. https://doi.org/10.3390/electronics9030483
25. Mohbey, K. K., Meena, G., Kumar, S., & Lokesh, K. (2024). A CNN-LSTM-based hybrid deep learning approach for sentiment analysis on Monkeypox tweets. New Generation Computing, 42(1), 89–107. https://doi.org/10.1007/s00354-023-00227-0