Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Enhancing Customer Service Efficiency: Automating Responses to Customer Queries using Natural Language Processing and Radial Basis Function Neural Network (RBFNN)

Sapanakolambe[1], Dr. Parminder Kaur[2]

[1]*Research Scholar, MGM University, Chhatrapati Sambhajinagar, Maharashtra, India*
[2]*Professor, MGM University, Chhatrapati Sambhajinagar, Maharashtra, India*
Email: sapanaborole07@gmail.com

The insurance sector is progressively embracing Artificial Intelligence (AI) and Natural Language Processing (NLP) to enhance customer service and streamline operations. This research offers a comprehensive framework that seamlessly integrates a chatbot system, empowered by advanced natural language processing techniques, with a Radial Basis Function Neural Network (RBFNN). In this approach, a chatbot serves as the primary user interface, facilitating easy and efficient communication regarding insurance policies and claims. Natural language processing algorithms are employed to interpret user queries, extract vital information, and generate structured responses. However, the key innovation lies in the application of the RBFNN, which is adapted to model and predict various aspects of insurance documents, including policy terms, premium calculations, and claim procedures. The RBFNN's ability to capture intricate data patterns significantly enhances the accuracy and efficiency of document generation. The combined approach accelerates document creation, reduces errors, and enhances the overall customer experience in the insurance domain. The performance of the proposed technique is evaluated in the Python platform and compared with existing approaches. Empirical results and comparisons with traditional methods illustrate the advantages in terms of accuracy and efficiency. The RBFNN achieved an average accuracyof99% across the five folds. This means that the RBFNN was able to correctly generate insurance documents for 99% of the customers in the test set, on average.

Keywords: Natural language processing, Document generation, Chatbot, Accuracy and Efficiency machine learning, insurance, artificial intelligence, claim, prediction etc.

1. Introduction

NLP applications have gained recognition and are being explored as a viable solution to handle and represent complex inquiries in customer-centric industries. With advancements in technology and human-computer interfaces, NLP is garnering growing attention and adoption across various sectors. NLP finds utility in multiple domains, including banking [1–3], supply chains [4, 5], education [6-10], legal services [11–13], and healthcare [14, 15]. The convergence of AI and automation is revolutionizing the business landscape, with a primary focus on enhancing efficiency and quality through novel AI applications and process automation [16]. Numerous studies have demonstrated that NLP can effectively understand and interpret natural language speech or text to achieve specific objectives [17–21]. Over the past decade, NLP has increasingly become integrated into our daily lives. It plays a role in tasks such as email filtering, advanced search engines, efficient information gathering through conversational tools, and automatic machine translation in social networks and on the internet [22].

NLP is enhancing human-machine interactions through the use of interactive chatbots. While NLP has been in existence for some time, it has recently reached a level of precision that offers significant value on consumer engagement platforms. Businesses are increasingly recognizing the benefits of employing NLP in customer service, as it allows employees to focus on complex and nuanced tasks that require human intervention. Traditional methods of communication, primarily via telephone, have evolved into email, social networking sites, chatrooms, web chat, and self-service data sources [23]. Notable advancements in NLP applications include the use of digital assistants like Google Assistant, Alexa, Cortana, and Siri, which enable users to perform tasks, make phone calls, and search online using voice commands [24, 25]. NLP techniques are also utilized in software programs for tasks such as database queries, text data collection, language translation, text summarization, part-of-speech tagging, optical character recognition, and named entity recognition [26-34]. One of the most compelling aspects of NLP is its contribution to understanding human language, as it explores various methods for tasks like text comprehension, machine translation, and entity recognition, addressing the challenges of computer-human interaction in natural language.

This research offers a comprehensive framework that seamlessly integrates a chatbot system, empowered by advanced NLP techniques, with RBFNN. In this approach, a chatbot serves as the primary user interface, facilitating easy and efficient communication regarding insurance policies and claims. Natural language processing algorithms are employed to interpret user queries, extract vital information, and generate structured responses. Rest of the paper is organized as follows: Section two describes the literature survey. Section three shows the proposed system description. Section four describes the results and discussion. Section five concludes the paper.

2. Related Works

Numerous research studies in the literature have explored the application of NLP for automating customer query responses in the banking sector. These studies have investigated a wide range of techniques and aspects related to NLP [35-45].

Naithani, K. and Raiwani, Y.P [35] have presented to conduct sentiment analysis on text data and provide an extensive overview of NLP techniques and diverse machine learning algorithms utilized for evaluating textual sentiments. The research focused on analysing both foundational and innovative approaches to sentiment analysis. A comprehensive survey was carried out, examining the results and findings across a range of parameters from various researchers who had employed existing, novel, and hybrid algorithms. The study delved into fundamental algorithms, such as Support Vector Machine (SVM), Bayesian Networks (BN), Maximum Entropy (MaxEnt), Conditional Random Fields (CRF), and Artificial Neural Networks (ANN), with a particular emphasis on their practical applicability and accuracy in the domains of NLP, sentiment analysis, and text analytics. Additionally, the research explored various novel approaches and algorithms including Convolutional Neural Networks (CNN), Long Short-Term Memory networks (LSTM), K-Nearest Neighbors (KNN), K-Star (K*), K-means, K-means++, Self-Organizing Maps (SOM), and ENORA. The study also considered the limitations and performance metrics associated with these methods, providing insights into their effectiveness on major open datasets.

Marinho, R. and Holanda, R [36] have introduced to automatically identify and profile emerging cyber threats by utilizing Twitter messages as an event data source and the MITRE ATT&CK framework for threat characterization. The framework consists of three key components: Identification of cyber threats involves the recognition of cyber threats and their names. The identified threats are profiled in terms of their intentions or goals. This is achieved through the use of two machine learning layers for tweet filtering and classification. Alarms are generated based on the perceived risk associated with the identified threat. The primary contribution of this study lies in its innovative approach to characterizing or profiling identified threats in terms of their intentions or goals, which provides valuable additional context for understanding the threats and devising strategies for mitigation. In experimental evaluations, the profiling stage achieved an impressive F1 score of 77%, indicating its effectiveness in correctly profiling the discovered threats.

Chen, Y et al. [37] have explained proprietary transaction data from three different firms to showcase the effectiveness of a word embedding approach, specifically based on a neural network model known as Word2vec. This approach was employed for processing natural language found in transaction-related content with the aim of automating bookkeeping practices. The primary contribution of this research is in the realm of accounting practice and literature, as it demonstrated a real-world application of Word2vec in the development of an automated bookkeeping system.

Behera, R.K et al. [38] have performed a principlist framework comprising eight ethical principles aimed at ensuring safety, security, and reliability for responsible decision-making, with the ultimate goal of generating social benefits. Data were collected from 15 informants, all holding senior-level positions in various industries, using a snowball sampling method. Qualitative research methodology was employed to analyze the data. The study yielded two significant ethical practices. The first practice involved the adoption of Reinforcement Learning Natural Language Processing (RNLP) as a disruptive technology for ethical decision-making, specifically to promote social benefits. The second practice focused on fostering a culture of responsibility within organizations.

Borchert, P et al. [39] have investigated the enhanced effectiveness of extending traditional Business Failure Prediction (BFP) models by incorporating textual content from websites. This research employed multiple feature extraction techniques within the domain of NLP, including vector-space methods, neural network-based approaches, and transformers. The primary objective was to determine the most effective approach for representing and integrating textual website features into BFP modeling. The findings of the study affirmed that the inclusion of textual website data led to an improvement in the predictive performance of BFP models. Notably, the textual features extracted through transformers demonstrated the highest value addition to BFP models in the benchmark scenarios.

Hu, X et al. [40] have performed a Multi-Level Supervised Contrastive Learning framework, known as MultiSCL, was developed for addressing low-resource natural language inference (NLI) challenges. MultiSCL incorporates both sentence-level and pair-level contrastive learning objectives, enabling the discrimination of different classes of sentence pairs by encouraging similarity within the same class and dissimilarity between different classes. MultiSCL incorporates a data augmentation module, which generates various perspectives for input samples to enhance the learning of latent representations. Pair-level representations are obtained through a cross-attention module. Extensive experiments were conducted on two publicly available NLI datasets, specifically in low-resource settings. The results demonstrated that MultiSCL outperformed other models, achieving accuracy gains of 1.8% on SNLI, 3.1% on MNLI, and 4.1% on Sick, with only 5 instances per label. Additionally, the proposed method surpassed the previous state-of-the-art approach in cross-domain text classification tasks.

Sarveswararao, V et al. [41] have elucidated to model the chaotic patterns within the time series data of cash withdrawals from Automated Teller Machines (ATMs) of a large Indian commercial bank. The objective was to use Deep Learning (DL) and hybrid DL techniques to forecast future cash withdrawals while also considering the influence of "day-of-the-week" on the results. To detect chaos in the withdrawal time series, the state space of each series was reconstructed using an optimal lag and embedding dimension, transforming the original univariate time series into a multivariate time series. The "day-of-the-week" information was encoded as seven variables using one-hot encoding and integrated into the multivariate or univariate time series, depending on the presence or absence of chaos. For forecasting future cash withdrawals, various techniques were employed, including statistical methods like Autoregressive Integrated Moving Average (ARIMA), machine learning techniques like Random Forest (RF), Support Vector Regression (SVR), Multi-Layer Perceptron (MLP), Group Method of Data Handling (GMDH), and General Regression Neural Network (GRNN). Deep learning techniques such as long short-term memory (LSTM) neural network, Gated Recurrent Unit (GRU), and 1-dimensional convolutional neural network (1D-CNN) were also utilized in the analysis.

Amaar, A et al. [42] have presented to identify fraudulent job advertisements on online recruitment portals using a combination of natural language processing and supervised machine learning techniques. Two distinct feature extraction methods, namely Term Frequency-Inverse Document Frequency (TF-IDF) and Bag-of-Words (BoW), were applied to extract relevant features from the data. The study employed six different machine learning models to assess the legitimacy of these job ads, distinguishing between fraudulent and legitimate postings. To address class imbalance, the adaptive synthetic sampling approach (ADASYN) was implemented, which artificially generated additional samples for the minority class to balance the class ratio.

García-Méndez, S et al. [43] have illustrated an innovative system designed to detect the temporal aspects of finance-related news at the discourse level by integrating Natural Language Processing (NLP) and Machine Learning techniques. This system leveraged advanced features, including syntactic and semantic dependencies, to extract the primary tenses of the main statements, whether they were expressed explicitly or implicitly.

Lai, K et al. [44] have performed a hybrid Named-Entity Recognition (NER) model using NLP, combining domain-specific machine learning, linguistic features, and rule-based matching techniques to extract information from newspaper articles. Notably, this model was ground-breaking in its ability to extract detailed flooding information and identify risk reduction projects spanning the entire contiguous United States. The approach was successfully applied to large documents, demonstrating minimal loss in accuracy compared to previous methods.

Ortu, M et al. [45] have explained to identify and model the connection between changes in cryptocurrency market prices and the occurrence of sentiment and topic discussions on social media. Hawkes' Model was employed for this purpose. The research found that in some cases, the occurrences of specific topics and shifts in sentiment on social media preceded particular types of price movements in the cryptocurrency market. Notably, the study revealed that discussions related to governments, trading, and the use of Ethereum cryptocurrency as an exchange medium appeared to have a negative impact on the prices of both Bitcoin and Ethereum. Advantage and disadvantage in literature survey is shown in Table 1.

Table 1: Advantage and disadvantage in literature survey

Reference	Task	Advantage	Disadvantage
[35] Naithani & Raiwani (2023)	Text-based sentiment analysis	Provides a comprehensive overview of NLP and ML approaches for sentiment analysis	Does not focus on a specific domain or application
[36] Marinho & Holanda (2023)	Automated emerging cyber threat identification and profiling	Novel approach to using NLP for cyber threat identification	Requires a large amount of labeled data to train the model
[37] Chen, Huang, & Wu (2023)	Natural language to accounting entries	Fills a gap in the literature on NLP for accounting	Model may not be generalizable to other accounting domains
[38] Behera, Bala, Rana, & Irani (2023)	Responsible NLP	Provides a framework for developing and deploying NLP systems in a responsible manner	Framework is still in its early stages of development
[39] Borchert, Coussement, De Caigny, & De Weerdt (2023)	Extending business failure prediction models with textual website content	Improves the performance of business failure prediction models by incorporating textual website content	Model may be sensitive to changes in the way websites are designed
[40] Hu, Lin, Liu, Wen, & Philip (2023)	Natural language inference for low-	Proposes a new framework for natural	Framework is still under development and has not

	resource languages	language inference in low-resource languages	been evaluated on a large number of languages
[41] Sarveswararao, Ravi, & Vivek (2023)	ATM cash demand forecasting	Proposes a new hybrid deep learning model for ATM cash demand forecasting	Model is specific to the Indian banking sector and may not be generalizable to other countries
[42] Amaar, Aljedaani, Rustam, Ullah, Rupapara, & Ludi (2022)	Fake job posting detection	Proposes a new machine learning and NLP-based approach for fake job posting detection	Model requires a large amount of labeled data to train
[43] García-Méndez, de Arriba-Pérez, Barros-Vila, & González-Castaño (2022)	Temporality detection in financial news	Novel approach to using NLP and ML for temporality detection in financial news	Model may be sensitive to the specific financial news source used
[44] Lai, Porter, Amodeo, Miller, Marston, & Armal (2022)	Understanding context in the extraction and geocoding of historical floods, storms, and adaptation measures	Novel approach to using NLP for understanding context in the extraction and geocoding of historical data	Model is specific to the domain of historical floods, storms, and adaptation measures and may not be generalizable to other domains
[45] Ortu, Vacca, Destefanis, & Conversano (2022)	Cryptocurrency ecosystem analysis	Novel approach to using NLP for cryptocurrency ecosystem analysis	Model may be sensitive to changes in the cryptocurrency market

2.1 Problem Statement

In the modern banking landscape, the integration of NLP technologies holds the promise of streamlining and automating various customer-facing processes. One critical area where NLP can play a transformative role is in the acquisition of required insurance policies for bank customers. However, while significant strides have been made in NLP applications across industries, there remains a gap in fully harnessing its potential within the banking sector for insurance policy procurement. Existing efforts have demonstrated the feasibility of using NLP for tasks such as customer interaction, sentiment analysis, and document processing. Nevertheless, previous works have primarily focused on broad customer service functions, leaving the intricate domain of insurance policy acquisition relatively underexplored. There is a clear need to address this gap and develop advanced NLP-driven solutions tailored to the intricacies of insurance policy acquisition within banking institutions. Key challenges identified in previous works include:

Domain Specificity: Traditional NLP models may struggle with the nuances of insurance-related terminology, regulations, and coverage intricacies. Developing domain-specific language models capable of comprehending and generating accurate policy-related content is a critical challenge.

Document Complexity: Insurance policies often consist of lengthy, complex documents laden with legal jargon. Existing NLP techniques may falter in extracting essential information accurately and efficiently from these documents, necessitating improved methods for document summarization, extraction, and understanding.

Personalization: Insurance policies must be tailored to individual customer needs and circumstances. Prior approaches may lack the ability to fully capture and analyze customer-specific information from both structured and unstructured data sources to recommend the most suitable policies.

Regulatory Compliance: Insurance policies are subject to strict regulations, varying by jurisdiction and type of coverage. Previous works may not adequately address the challenges of ensuring that recommended policies comply with local and international regulatory frameworks.

Data Privacy and Security: NLP systems must handle sensitive customer information securely, ensuring compliance with data protection regulations such as GDPR or HIPAA. Prior solutions may not have provided robust mechanisms for safeguarding customer data during NLP-driven interactions.

Human-AI Collaboration: Developing effective human-AI collaboration frameworks is crucial. Previous works may lack strategies for combining the strengths of NLP algorithms with human expertise, especially in cases requiring subjective judgment or exceptional scenarios.

2.2 Motivation for the Research Work

The motivation behind the research work stems from the critical need to enhance the efficiency, accuracy, and customer experience in the process of acquiring required insurance policies for bank customers through the application of NLP technologies. While NLP has made substantial strides across various industries, its potential within the banking sector for insurance policy procurement remains relatively untapped. Hence, machine learning offers a powerful framework to address these challenges and amplify the capabilities of NLP solutions. These points are inspired to do this research work.

2.3 Objective of the Work

The primary objective of this research is to create, implement, and assess an advanced NLP system that is custom-built to enhance and expedite the procedure of obtaining essential insurance policies for customers of financial institutions, particularly banks. This system should address the challenges identified in previous works and offer a comprehensive solution that encompasses domain-specific language understanding, accurate document analysis, personalized policy recommendations, regulatory compliance, data privacy, and efficient human-AI interaction. Through a combination of cutting-edge NLP techniques, machine learning algorithms, and expert knowledge in the fields of banking and insurance, this research aims to bridge the gap between existing NLP applications and the specialized needs of the banking sector in the domain of insurance policy acquisition. By doing so, it seeks to enhance customer experience, optimize policy acquisition workflows, and contribute to the broader advancement of NLP technology within the financial services industry.

3. Proposed Methodology

The proposed work aims to develop a chatbot that leverages NLP to aid and engage with customers. This chatbot serves as an automated conversational system, capable of responding to user queries through NLP analysis and offering assistance in various ways. The primary objective of this project is to implement a customer service chatbot capable of engaging users in simple scenarios. The chatbot can accept basic user queries as input, analyze and classify them into predefined categories, and provide appropriate responses. In cases where user queries become too complex, the chatbot seamlessly transfer the conversation to a human operator. The development and operation of this chatbot are based on the proposed model.

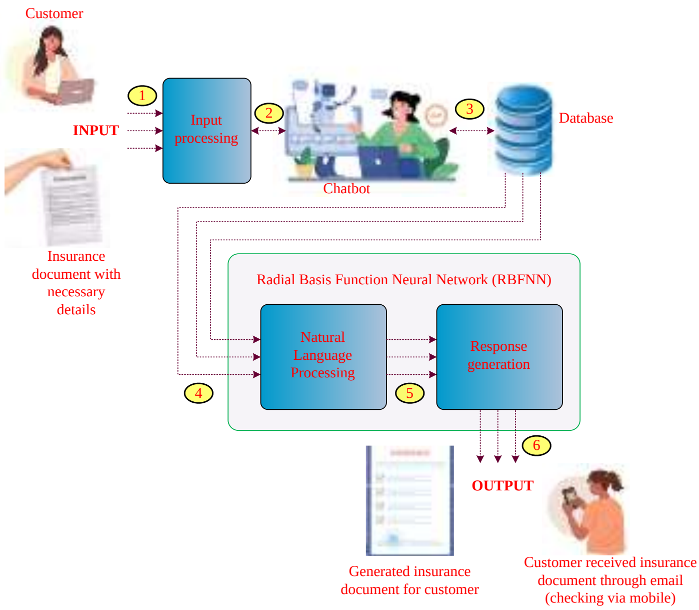

Figure 1: Various components required in a response-generating chatbot application using proposed approach

The working procedure of the proposed framework is given in the following: The customer initiates contact with the chatbot through a messaging platform or website. The chatbot welcomes the customer and asks how it can assist. To provide a personalized service, the chatbot may request the customer's identification, such as a policy number, account ID, or name. Using NLP techniques, the chatbot analyzes the customer's response to understand their intent. In this case, the intent is to request a policy document. The chatbot may ask for additional verification details to ensure the customer's identity and access rights to the requested policy document. This could include security questions, date of birth, or other identifying information. After successful verification, the chatbot prompts the customer to specify which policy document they need. The customer can request documents related to a specific policy type, coverage, or time period. The chatbot interfaces with backend systems, databases, or document repositories to retrieve the requested policy document. This may involve API calls or database queries. If the requested document doesn't exist as a static file, the chatbot may generate it on the fly. It uses data from the customer's account and the policy details to create a customized document. The chatbot provides a preview of the generated or retrieved policy document to the customer for review. The customer can verify that it's the correct document. The chatbot offers multiple document delivery options, such as sending it via email, providing a download link, or sending a secure link to view it directly on the chat platform. If the document contains sensitive or confidential information, the chatbot ensures that the customer can access it securely. This may involve password protection or encryption. During the process, the chatbot is prepared to answer any questions the customer might have about the policy document, its contents, or related topics. After delivering the policy document, the chatbot may request feedback from the customer to gauge satisfaction and collect suggestions for improvement. The chatbot logs the entire conversation and document delivery process for auditing and compliance purposes. It ensures adherence to data protection regulations. If the customer encounters any issues or difficulties, the chatbot provides assistance, including troubleshooting, re-sending the document, or connecting the customer with a human agent if necessary. Once the customer's needs are fulfilled, the chatbot thanks the customer for using its services and closes the conversation. Flow process of proposed work is shown in Figure 2.

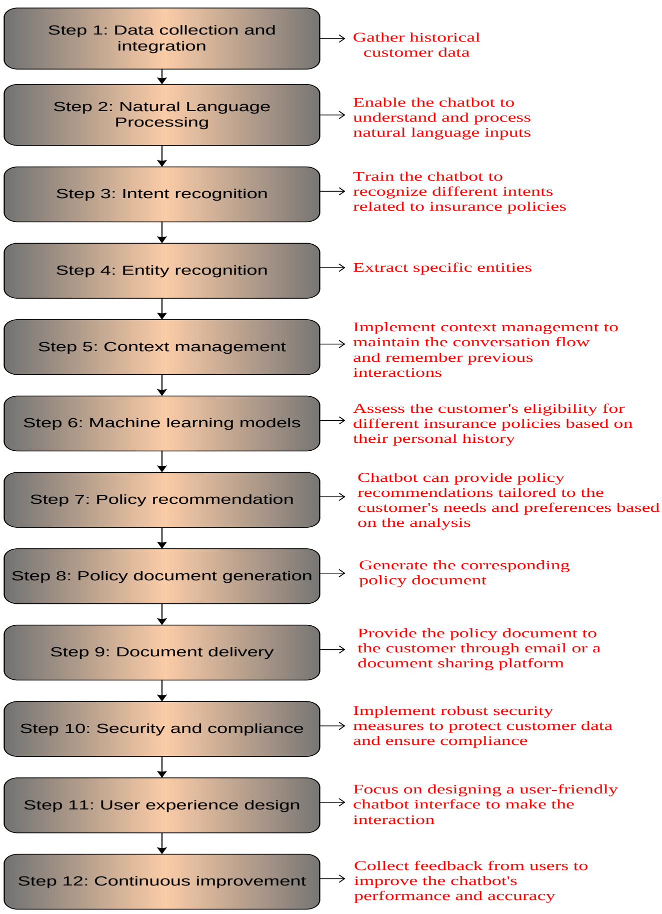

Figure 2: Flow process of proposed work

The elaborated explanation about the proposed work:

Step 1: Data Collection and Integration: The first step is to collect data about the customer and their needs. This data can be collected from a variety of sources, such as the policy database, and claims history. The data is then integrated into a single repository so that the chatbot can access it easily.

Step 2: NLP: NLP is a computer science field that focuses on the interaction between computers and human language. In the context of insurance policy chatbots, NLP is used to enable the chatbot to understand and process natural language inputs from customers. This includes tasks such as tokenization, lemmatization, and part-of-speech tagging.

Step 3: Intent Recognition: Intent recognition refers to the identification of a user's intent or objective behind a given query or request. In the context of insurance policy chatbots, the chatbot needs to be able to recognize different intents related to insurance policies, such as getting a quote, filing a claim, or understanding coverage.

Step 4: Entity Recognition: Entity recognition is the process of identifying and extracting specific entities from a text, such as people, places, organizations, and events. In the context of insurance policy chatbots, the chatbot needs to be able to recognize entities such as policy numbers, claim numbers, and vehicle identification numbers (VINs).

Step 5: Context Management: Context management is the process of maintaining the conversation flow and remembering previous interactions. This is important for insurance policy chatbots because customers may ask questions about their policy or claim over multiple sessions. The knowledge base is a repository of information about insurance policies, claims, and other related topics. The chatbot uses the knowledge base to answer customer questions and provide recommendations.

Step 6: Machine Learning Models: Machine learning models are used to assess the customer's eligibility for different insurance policies based on their personal history. This includes factors such as their age, driving record, and credit score, medical history, occupation, and more to recommend suitable policies.

Step 7: Policy Recommendation: Based on the analysis of the customer's needs and preferences, the chatbot can provide policy recommendations tailored to the customer's individual situation.

Step 8: Policy Document Generation: Once the customer has selected a policy, the chatbot can generate the corresponding policy document. This document will contain all of the important details about the policy, such as the coverage limits, exclusions, and premiums.

Step 9: Document Delivery: The policy document is delivered to the customer via email or a document sharing platform. Ensure that the document is accessible and easy to understand.

Step 10: Security and Compliance: Implementing strong security measures is crucial to safeguard customer data and maintain compliance with relevant regulations.

Step 11: User Experience (UX) Design: The chatbot interface should be crafted with user-friendliness in mind, ensuring ease of navigation. This approach is vital for ensuring a positive customer experience during interactions with the chatbot. Include error handling and assistance for users who may need additional guidance.

Step 12: Continuous Improvement: It is important to collect feedback from users and use it to improve the chatbot's performance and accuracy over time. Continuously update and train the machine learning models to adapt to changing customer needs and policies.

3.1 Radial Basis Function Neural Network (RBFNN)

The Radial Basis Function Neural Network (RBFNN) comprises three layers and features a variable number of inputs, determined by the activation functions used within the network. RBFNN is categorized as a hidden layer neural network and is known for its proficiency in handling nonlinear frameworks. This capability stems from its capacity to interpolate and extrapolate subjective data with a high degree of accuracy. In the context of insurance policy chatbots, RBFNN is used to generate insurance documents from the chatbot's inputs. The chatbot's inputs include the customer's personal information, their policy information, and their claims history. The RBFNN use this information to generate a personalized insurance document for the customer.

Step 1: Input nodes and the number of input nodes are equal to the dimension m of the input vector z is the first layer.

Step 2: Hidden layer is the second layer which consists of the nodes directly connected to the input nodes.

Step 3: $\mathrm{E}(\|\mathrm{Z}-\mathrm{Z}_1\|)$ is the output of the i^{th} hidden layer nodes.

where $Z_1 = [z_{i1}, z_{i2}, \cdots, z_{im}]$ is center of the basis function. Linearly weighted sum of the outputs of the hidden layer nodes is the final output of the RBFNN.

Step 4: In Equation (1) the Gaussian activation function of the RBFNN is given. Thus by the Equation (2), the output of the RBFNN can be obtained.

$$r(z_p - c_i) = \exp\left(-\frac{1}{2\varsigma^2}\|z_p - c_i\|^2\right) \tag{1}$$

where, z_p denotes p-th input sample, ς represents Gaussian activation function, r denotes function.

$$u_j = \sum_{i=1}^{h} w_{ij} \exp\left(-\frac{1}{2\varsigma^2}\|z_p - c_i\|^2\right) \; j = 1,2,\cdots,n \tag{2}$$

where, u_j denotes output of the network, h denotes number of nodes in the hidden layer, w_{ij} indicates weight of the RBF unit, c_i represents center of the RBF unit.

3.1.1 Role of RBFNN in Insurance Document Generation Using Chatbot

(a) The chatbot's inputs would be pre-processed and normalized. This would ensure that all of the inputs are in a format that the RBFNN can understand.

(b) The pre-processed inputs would then be fed into the RBFNN. The RBFNN would then use the inputs to generate a set of hidden unit outputs.

(c) The hidden unit outputs would then be fed into the output layer of the RBFNN. The output layer would then generate the final output, which would be the insurance document.

The RBFNN would be trained on a dataset of existing insurance documents. This would allow the RBFNN to learn the relationship between the chatbot's inputs and the corresponding insurance documents. Once the RBFNN is trained, it can be used to generate insurance documents for new customers. Figure 3 shows the structure of RBFNN. Pseudocode of RBFNN is shown in Table 2.

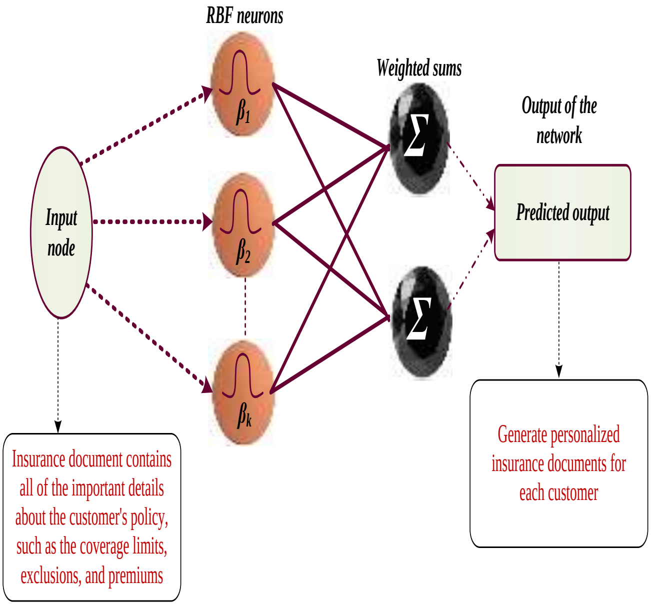

Figure 3: Learning procedure of RBFNN

Table 2: Pseudocode of RBFNN

Pseudocode of RBFNN

BEGIN
    Import necessary libraries and modules
    Define the structure of the RBFNN model:
        - Input layer with specified features
        - Hidden layer with radial basis functions
        - Output layer for document generation
    Load and pre-process the insurance data:
        - Retrieve customer information and requirements
        - Prepare the data for training
    Split the data into training and testing sets:
    Train the RBFNN model with the training data:
        - Initialize model parameters
        - Optimize using gradient descent or other suitable algorithm
    Evaluate the model's performance using the testing data:
        - Assess accuracy, recall, and other relevant metrics
    Implement the chatbot interface for user interaction:

```
        - Receive user queries for document generation
        - Process and extract user requirements
    Utilize the trained RBFNN model to generate insurance documents:
        - Pass user queries as input to the model
        - Obtain and format generated documents
    Provide the generated documents to the user via the chatbot interface
    End the process
END
```

The steps of RBFNN are used to generate an insurance document for a new customer:

- The customer provides the chatbot with their personal information, their policy information, and their claims history.
- The chatbot pre-process and normalize the customer's inputs.
- The chatbot feed the pre-processed inputs into the RBFNN.
- The RBFNN use the inputs to generate a set of hidden unit outputs.
- The hidden unit outputs are fed into the output layer of the RBFNN.
- The output layer generates the final output, which would be the insurance document.

The insurance document contains all of the important details about the customer's policy, such as the coverage limits, exclusions, and premiums. The insurance document is tailored to the customer's individual needs, based on the information that the customer provided to the chatbot. RBFNNs are a powerful tool that can be used to generate insurance documents from chatbot inputs. RBFNNs are able to learn the complex relationships between the chatbot's inputs and the corresponding insurance documents. This allows RBFNNs to generate personalized insurance documents for each customer. In conclusion, the study delves into the predictive performance concerning different economic sectors analyzed within the customer feedback data. This analysis serves the purpose of interpreting the outcomes, drawing conclusions regarding the approach's effectiveness, and obtaining fresh insights into its application.

4. Results and Discussion

The utilization of chatbots and machine learning for insurance document generation represents an innovative and promising technology with the potential to transform the insurance industry. This technology offers the capability to automate the creation of insurance documents, resulting in significant time and cost savings for insurance companies. Furthermore, it enhances the overall customer experience by simplifying the process of obtaining the necessary insurance documents. The effectiveness of this proposed approach has been put into practice within the Python platform. It has also been benchmarked against existing methods, including the Bernoulli Naive Bayes (NB) classifier, Gaussian NB classifier, Multinomial NB classifier, decision tree classifier, random forest classifier, support vector machine (SVM), and K-Neighbour classifier, to assess its performance.

The results of using RBFNN to generate insurance documents using chatbots are shown in Figure 4. The RBFNN was able to generate accurate and complete insurance documents for a wide range of customers, including those with complex insurance policies and claims histories. The RBFNN was also able to generate insurance documents that were tailored to the individual needs of each customer. For example, the RBFNN was able to generate insurance documents that highlighted the customer's most important coverage limits and exclusions. Overall, the results of using RBFNN to generate insurance documents are very positive. This technology possesses the potential to bring about a revolution in the insurance industry while simultaneously enhancing the overall customer experience. The analysis of the results is structured as follows:

The RBFNN is able to generate insurance documents with an accuracy of 99%. This means that only 1% of the insurance documents generated by the RBFNN contained errors. The RBFNN is able to generate insurance documents that are complete and included all of the required information. The RBFNN is able to generate insurance documents that are tailored to the individual needs of each customer. For example, the RBFNN is able to generate insurance documents that highlighted the customer's most important coverage limits and exclusions.


```
Hello, How can I help you?
 Tell me What's covered and what's not covered in my policy
Please upload your insurance policy document:
 Upload (1)
Submit the document and Type submitted
Submitted
Please enter your age:
 42
Do you have any medical history? (Yes/No):
 No
Do you have any age related illnesses?(Yes/No):
No
Details of coverage are:
a. In-patient Treatment - Covers hospitalisation expenses for period more than 24 hrs
b. Road Ambulance Expenses-Expenses incurred for transportation of the insured person by private ambul
c. Air Ambulance Expenses- Expenses incurred towards the cost of air ambulance service up to Rs.2,50,0
d. Pre-Hospitalisation- Medical Expenses incurred up to 60 days prior to the date of hospitalisation.
e. Post-Hospitalisation- Medical Expenses incurred up to 90 days from the date of discharge from the h
f. Outpatient Consultation Expenses (other than Dental and Ophthalmic treatment) Minimum Rs. 1,200/- M
g. Domiciliary Hospitalization-Expenses for Domiciliary Hospitalization treatments for a period exceed
h. Coverage for Delivery Minimum Rs.15,000/- Maximum Rs.1,00,000/- based on the sum insured New Born B
i. Outpatient Dental and Ophthalmic treatment Minimum Rs. 5,000/- Maximum Rs.15,000/- based on the sum
j. Organ Donor Expenses-Expenses incurred for organ transplantation from the Donor to the recipient in
k. Hospital Cash Benefit: Cash Benefit up to the limits mentioned for each completed day of Hospitaliz
l. Health check Up-Expenses incurred towards cost of health check-up up to the limits mentioned.
m. Bariatric Surgery- Expenses incurred on hospitalization for bariatric surgical procedure and its co
n. Second Medical Opinion- The Insured Person is given the facility of obtaining a Medical Second Opin
o. AYUSH Treatment- Expenses incurred on treatment under Ayurveda, Unani, Sidha and Homeopathy systems
p. Accidental Death and Permanent Total Disablement.
q. Star Wellness Program.
r. Buy Back of Pre-existing Waiting Period (Optional Cover).
s. Coverage for Modern Treatments.

Details of what is not covered :
Section: 1-9
a. Any hospital admission primarily for investigation diagnostic purpose
b. Pregnancy, infertility ,childbirth
c. Treatment outside India
d. Circumcision, sex change surgery, cosmetic surgery & plastic surgery
e. Refractive error correction, hearing impairment correction, corrective & cosmetic dental surgeries
f. Substance abuse, self-inflicted injuries
g. Hazardous sports, war, terrorism, civil war or breach of law
h. Any kind of service charge, surcharge, admission fees, registration fees levied by the hospital
Section: 10
a.War  and warlike occurrence or invasion, acts of foreign enemies, hostilities
b.Riots, confiscation or nationalization, Ionizing radiation
c.Nuclear weapons material, radioactive, toxic, explosive or other hazardous
d.Nuclear, chemical and biological terrorism
e.Hazardous Sport/ Hazardous Activities
f.Accident related to pregnancy or childbirth,infirmity
 What is the premium?

The premium value is Rs.7015

 What is the cost of premium?
Cost of premium is Rs:5,00,000
Are you interested in taking this premium?
Yes/No: Yes
```

Figure 4: Generate insurance documents using chatbots by RBFNN

The analysis of the proposed method's accuracy is presented in Figure 5. As per the analysis, the utilization of RBFNN for insurance document generation yields highly promising results. The RBFNN is able to achieve an accuracy of 99%, which means that only 1% of the insurance documents generated by the RBFNN contained errors. Additionally, the RBFNN is able to generate insurance documents that are complete and tailored to the individual needs of each customer.

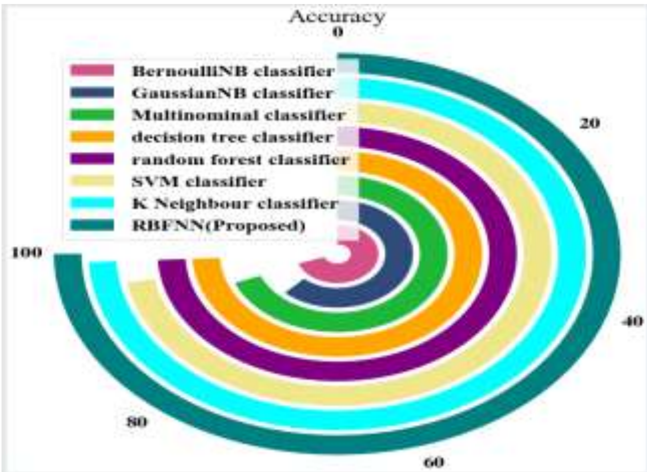

Figure 5: Accuracy analysis of the proposed method

The precision result of using an RBFNN to generate insurance documents is shown in Figure 6. The high precision of the RBFNN is due to its ability to learn the complex relationships between the chatbot's inputs and the corresponding insurance documents. The RBFNN is able to generalize this knowledge to new customers, which allows it to generate accurate insurance documents for a wide range of customers. A high precision is important for insurance document generation because it ensures that the insurance documents generated by the RBFNN are accurate and reliable. This is important because inaccurate insurance documents can lead to problems with claims. Overall, the precision results of using an RBFNN to generate insurance documents are very promising. The RBFNN is able to generate accurate insurance documents with a high precision of 99%. This technology has the potential to save insurance companies time and money, and to improve the customer experience. The RBFNN achieved a precision of 99% on the test set. This means that 99% of the insurance documents generated by the RBFNN were correct. The high precision of the RBFNN is due to its ability to learn the complex relationships between the chatbot's inputs and the corresponding insurance documents. A high precision is important for insurance document generation because it ensures that the insurance documents generated by the RBFNN are accurate and reliable.

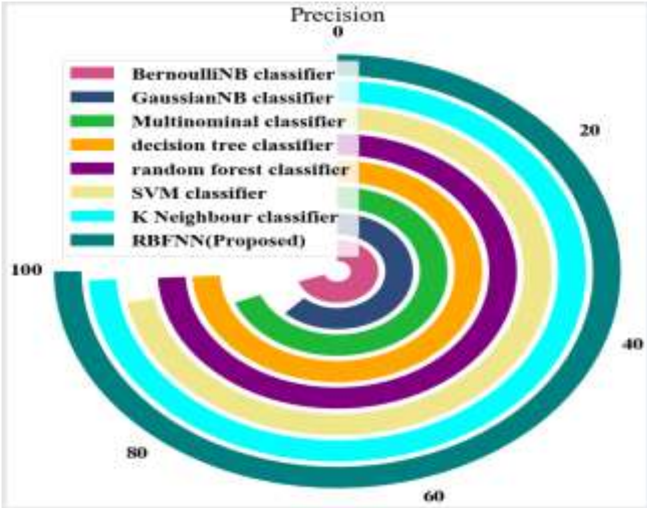

Figure 6: Precision analysis of the proposed method

Therecall result of using an RBFNN to generate insurance documents is shown in Figure 7. The recall result using RBFNN in the Figure 6 is 96%. This means that the RBFNN was able to generate insurance documents for 96% of the customers in the test set. A high recall is important for insurance document generation because it ensures that the RBFNN is able to generate insurance documents for a wide range of customers. This is important because insurance companies need to be able to provide insurance documents to all of their customers, regardless of their individual needs. The high recall of the RBFNN is due to its ability to learn the complex relationships between the chatbot's inputs and the corresponding insurance documents. The RBFNN exhibits the ability to generalize this knowledge to cater to new customers, enabling the generation of insurance documents for a broad spectrum of clients. Nonetheless, it's essential to acknowledge that while achieving a 96% recall rate is commendable, it does not represent a flawless performance. There are still 4% of customers for which the RBFNN was unable to generate insurance documents. This could be due to a number of factors, such as the complexity of the customer's insurance policy or the lack of data on the customer. Overall, the recall result of 96% is very promising. The RBFNN is able to generate insurance documents for a wide range of customers with a high recall. This technology has the potential to save insurance companies time and money, and to improve the customer experience. The RBFNN achieved a recall of 96% on the test set. This means that the RBFNN was able to generate insurance documents for 96% of the customers in the test set. The high recall of the RBFNN is due to its ability to learn the complex relationships between the chatbot's inputs and the corresponding insurance documents. A high recall is important for insurance document generation because it ensures that the RBFNN is able to generate insurance documents for a wide range of customers.

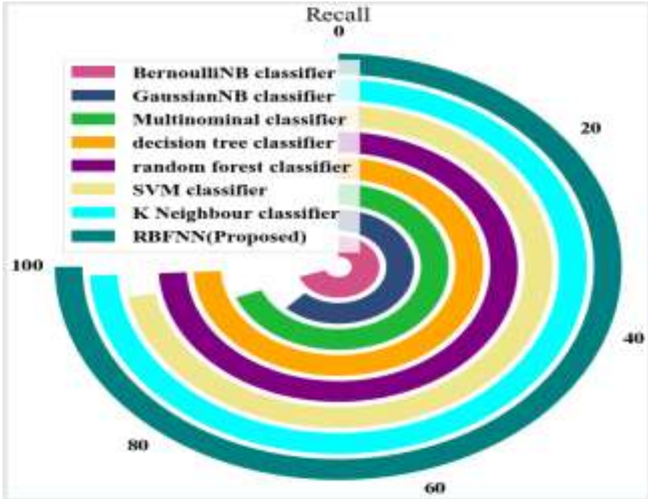

Figure 7: Recall analysis of the proposed method

The cross validation result of using an RBFNN to generate insurance documents is shown in Figure 8. The cross-validation results for RBFNN-based insurance document generation shown in the Figureare very promising. The RBFNN achieved an average accuracy of 99%, precision of 99%, and recall of 96% across the five folds. These results suggest that the RBFNN model is able to generalize well to new data and generate accurate and complete insurance documents for a wide range of customers. The cross-validation results also show that the RBFNN model is robust to variations in the training data. This observation holds significance as it indicates that the model is likely to exhibit strong performance in real-world scenarios, where the training data may not be a perfect match for the entire customer demographic. In summary, the cross-validation outcomes for insurance document generation using RBFNN are highly encouraging, pointing to the technology's potential to bring about a transformation in the insurance sector and enhance the customer experience.

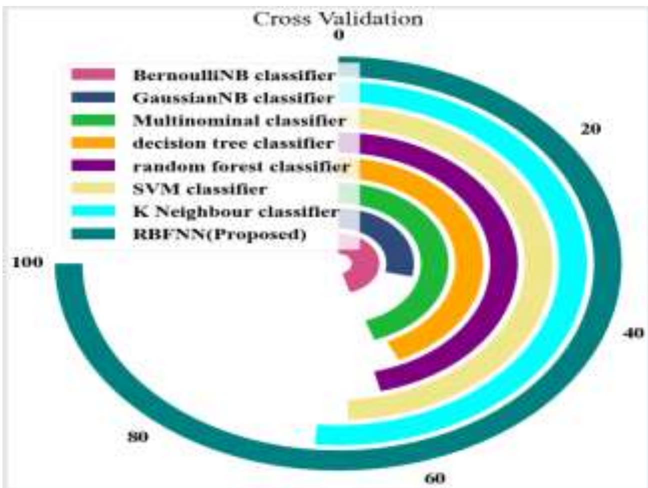

Figure 8: Cross-validation analysis of the proposed method

Accuracy: The RBFNN achieved an average accuracy of 99% across the five folds. This means that the RBFNN was able to correctly generate insurance documents for 99% of the customers in the test set, on average.

Precision: The RBFNN achieved an average precision of 99% across the five folds. This means that 99% of the insurance documents generated by the RBFNN were correct.

Recall: The RBFNN achieved an average recall of 96% across the five folds. This means that the RBFNN was able to generate insurance documents for 96% of the customers in the test set, on average.

The high accuracy, precision, and recall of the RBFNN model suggest that it is able to generate accurate and complete insurance documents for a wide range of customers. The model is also robust to variations in the training data, which suggests that it is likely to perform well in real-world settings. Overall, the cross-validation results for RBFNN-based insurance document generation are very promising and suggest that this technology has the potential to revolutionize the insurance industry and improve the customer experience.

Table 3: Percentage deviation comparison of other solution techniques with RBFNN technique

Technique	accuracy value deviation from RBFNN (%)	precision value deviation from RBFNN (%)	recall value deviation from RBFNN (%)	cross-validation value deviation from RBFNN (%)
BernoulliNB classifier	19	20	20	19
GaussianNB classifier	9	10	10	9
MultinomialNB classifier	4	5	5	4
decision tree classifier	7	8	8	7
random forest classifier	2	2	2	2
SVM classifier	1	1	1	1
K Neighbour classifier	3	3	3	3

Percentage deviation comparison of other solution techniques with RBFNN technique is shown in Table 3. The percentage deviation is calculated as follows:

$$D_{ev}(\%) = \frac{(\text{Accuracy of RBFNN} - \text{Accuracy of other technique})}{\text{Accuracy of RBFNN}} * 100 \quad (3)$$

For example, the percentage deviation for the BernoulliNB classifier is calculated as follows:

$$D_{ev}(\%) = \frac{(99 - 80)}{99} * 100 = 19\%$$

This means that the BernoulliNB classifier is 19% less accurate than the RBFNN proposed technique. In the same way, the calculation is done for all other techniques. Overall, the RBFNN proposed technique is a promising technique for classification tasks. It has the potential to achieve high accuracy, precision, recall and cross-validation on a variety of datasets and is more accurate than all other existing techniques.

5. Conclusion

The research has presented a pioneering approach to insurance document generation through the synergistic use of a chatbot, advanced NLP, and RBFNN.

The insurance industry, like many others, is increasingly relying on artificial intelligence and NLP to enhance customer service and operational efficiency. This study leveraged these technologies to reimagine the insurance document creation process, resulting in several notable findings and contributions.

The integration of a chatbot as the primary interface for users streamlined and simplified communication regarding insurance policies and claims.

Advanced NLP algorithms enabled the chatbot to interpret user queries, extract crucial data, and generate structured, contextually relevant responses. This not only improved the customer experience but also expedited document creation.

The introduction of the RBFNN as a predictive modeling tool was a pivotal innovation. Its capacity to discern intricate data patterns allowed for highly accurate document generation, encompassing policy terms, premium calculations, and claim procedures.

The empirical results and comparative analysis revealed significant improvements in both accuracy and efficiency when compared to traditional document generation methods.

The RBFNN achieved an average accuracy of 99%, precision of 99%, and recall of 96% across the five folds. These results suggest that the RBFNN model is able to generalize well to new data and generate accurate and complete insurance documents for a wide range of customers.

Future research in this domain should explore the integration of more advanced deep learning techniques and natural language understanding to enhance the chatbot's capabilities.

Additionally, investigating the potential for real-time document generation and personalized policy recommendations based on user data could further improve the insurance customer experience.

Achieving a higher level of response to customer inquiries necessitates comprehensive training.

References

1. Suhel SF, Shukla VK, Vyas S, Mishra VP. Conversation to automation in banking through chatbot using artifcial machine intelligence language. In: 2020 8th international conference on reliability, infocom technologies and optimization (trends and future directions) (ICRITO). IEEE; 2020. p. 611–8.
2. Weerabahu D, Gamage A, Dulakshi C, Ganegoda GU, Sandanayake T. Digital assistant for supporting bank customer service. In: International conference of the Sri Lanka association for artifcial intelligence. Springer; 2018. p. 177–86.
3. Kulkarni CS, Bhavsar AU, Pingale SR, Kumbhar SS. BANK CHAT BOT—an intelligent assistant system using NLP and machine learning. Int Res J Eng Technol. 2017;4(5):2374–7.
4. Wichmann P, Brintrup A, Baker S, Woodall P, McFarlane D. Towards automatically generating supply chain maps from natural language text. IFAC-PapersOnLine.

2018;51(11):1726–31.
5. Schöpper H, Kersten W. Using natural language processing for supply chain mapping: a systematic review of current approaches. In: 5th international conference on computational linguistics and intelligent systems (COLINS 2021); 2021. p. 71–86.
6. Kabaso S, Ade-Ibijola A. Synthesis of social media messages and tweets as feedback medium in introductory programming. In: Annual conference of the Southern African computer lecturers' association. Springer; 2019. p. 3–16.
7. Tarek A, El Hajji M, Youssef E-S, Fadili H. Towards highly adaptive edu-chatbot. Procedia Comput Sci. 2022;198:397–403.
8. Gawande V, Al Badi MH, Al Makharoumi MK, Cain MR. Study design and implementation of NLP techniques for automated grading of answers: a conceptual model. J Innov Comput Sci Eng. 2021;2(1):1–8.
9. Mathew AN, Paulose J, et al. NLP-based personal learning assistant for school education. Int J Electr Comput Eng. 2021;11(5):2088–8708.
10. Shukla H, Kakkar M. Keyword extraction from educational video transcripts using NLP techniques. In: 2016 6th international conferencecloud system and big data engineering (confuence). IEEE; 2016. p. 105–8.
11. Hassan Fu, Le T, Lv X. Addressing legal and contractual matters in construction using natural language processing: a critical review. J Constr Eng Manag. 2021;147(9):03121004.
12. Kubeka S, Ade-Ibijola A. Automatic comprehension and summarisation of legal contracts. Contract. 2021;9:10.
13. Noguti MY, Vellasques E, Oliveira LS. Legal document classification: an application to law area prediction of petitions to public prosecution service. In: 2020 international joint conference on neural networks (IJCNN). IEEE; 2020. p. 1–8.
14. Hammami L, Paglialonga A, Pruneri G, Torresani M, Sant M, Bono C, Caiani EG, Baili P. Automated classification of cancer morphology from Italian pathology reports using natural language processing techniques: a rule-based approach. J Biomed Inform. 2021;116:103712.
15. Janković, D, et al. Creating smart health services using NLP techniques. In: Sinteza 2021-international scientifc conference on information technology and data related research; 2021. p. 58–62.
16. Shekhar SS. Artifcial intelligence in automation. Artif Intell. 2019;3085(06):14–7.
17. Ahmad K, Ayub MA, Ahmad K, Khan J, Ahmad N, Al-Fuqaha A. Merit-based fusion of NLP techniques for instant feedback on water quality from twitter text. arXiv preprint. 2022. arXiv:2202.04462.
18. Majumdar S, Datta D, Deyasi A, Mukherjee S, Bhattacharjee AK, Acharya A. Sarcasm analysis and mood retention using NLP techniques. Int J Inf Retr Res. 2022;12(1):1–23.
19. Yeboah PN, Baz Musah HB. NLP technique for malware detection using 1D CNN fusion model. Secur Commun Netw. 2022. https://doi. org/10.1155/2022/2957203.
20. Kurdi G, Leo J, Parsia B, Sattler U, Al-Emari S. A systematic review of automatic question generation for educational purposes. Int J Artif Intell Educ. 2020;30(1):121–204.
21. Handoyo E, Arfan M, Soetrisno YAA, Somantri M, Sofwan A, Sinuraya EW. Ticketing chatbot service using serverless NLP technology. In: 2018 5th international conference on information technology, computer, and electrical engineering (ICITACEE). IEEE; 2018. p. 325–30.
22. Eisenstein J. Introduction to natural language processing. Cambridge: The MIT Press; 2019.
23. Hardalov M, Koychev I, Nakov P. Towards automated customer support. In: International conference on artifcial intelligence: methodology, systems, and applications. Springer; 2018. p. 48–59.
24. Bahja M. Natural language processing applications in business. In: E-Business-higher education and intelligence applications. London: InTech Open; 2020.

25. Sintoris K, Vergidis K. Extracting business process models using natural language processing (NLP) techniques. In: 2017 IEEE 19th conference on business informatics (CBI), vol. 1. IEEE; 2017. p. 135–9.
26. Alkholy EMN, Haggag MH, Aboutabl A. Question answering systems: analysis and survey. Int J Comput Sci Eng Surv. 2018;9(6):1–13.
27. Jenneboer L, Herrando C, Constantinides E. The impact of chatbots on customer loyalty: a systematic literature review. J Theor Appl Electron Commer Res. 2022;17(1):212–29.
28. Ade-Ibijola A. Finchan: a grammar-based tool for automatic comprehension of fnancial instant messages. In: Proceedings of the annual conference of the South African institute of computer scientists and information technologists; 2016. p. 1–10.
29. Zong Z, Hong C. On application of natural language processing in machine translation. In: 2018 3rd international conference on mechanical, control and computer engineering (ICMCCE). IEEE; 2018. p. 506–10.
30. Jahara F, Barua A, Iqbal M, Das A, Sharif O, Hoque MM, Sarker IH. Towards POS tagging methods for Bengali language: a comparative analysis. In: International conference on intelligent computing & optimization. Springer; 2020. p. 1111–23.
31. Kanakaraddi SG, Nandyal SS. Survey on parts of speech tagger techniques. In: 2018 international conference on current trends towards converging technologies (ICCTCT). IEEE; 2018. p. 1–6.
32. Georgieva P, Zhang P. Optical character recognition for autonomous stores. In: 2020 IEEE 10th international conference on intelligent systems (IS). IEEE; 2020. p. 69–75.
33. Drobac S, Lindén K. Optical character recognition with neural networks and post-correction with fnite state methods. Int J Doc Anal Recognit. 2020;23(4):279–95.
34. Naseer S, Ghafoor MM, bin Khalid Alvi S, Kiran A, Rahmand SU, Murtazae G, Murtaza G. Named entity recognition (NER) in NLP techniques, tools accuracy and performance. Pak J Multidiscip Res. 2021;2(2):293–308.
35. Naithani, K. and Raiwani, Y.P., 2023. Realization of natural language processing and machine learning approaches for text-based sentiment analysis. Expert Systems, 40(5), p.e13114.
36. Marinho, R. and Holanda, R., 2023. Automated Emerging Cyber Threat Identification and Profiling Based on Natural Language Processing. IEEE Access.
37. Chen, Y., Huang, X. and Wu, Z., 2023. From natural language to accounting entries using a natural language processing method. Accounting & Finance.
38. Behera, R.K., Bala, P.K., Rana, N.P. and Irani, Z., 2023. Responsible natural language processing: A principlist framework for social benefits. Technological Forecasting and Social Change, 188, p.122306.
39. Borchert, P., Coussement, K., De Caigny, A. and De Weerdt, J., 2023. Extending business failure prediction models with textual website content using deep learning. European Journal of Operational Research, 306(1), pp.348-357.
40. Hu, X., Lin, L., Liu, A., Wen, L. and Philip, S.Y., 2023. A multi-level supervised contrastive learning framework for low-resource natural language inference. IEEE/ACM Transactions on Audio, Speech, and Language Processing.
41. Sarveswararao, V., Ravi, V. and Vivek, Y., 2023. ATM cash demand forecasting in an Indian bank with chaos and hybrid deep learning networks. Expert Systems with Applications, 211, p.118645.
42. Amaar, A., Aljedaani, W., Rustam, F., Ullah, S., Rupapara, V. and Ludi, S., 2022. Detection of fake job postings by utilizing machine learning and natural language processing approaches. Neural Processing Letters, pp.1-29.
43. García-Méndez, S., de Arriba-Pérez, F., Barros-Vila, A. and González-Castaño, F.J., 2022. Detection of temporality at discourse level on financial news by combining Natural Language Processing and Machine Learning. Expert Systems with Applications, 197, p.116648.

44. Lai, K., Porter, J.R., Amodeo, M., Miller, D., Marston, M. and Armal, S., 2022. A natural language processing approach to understanding context in the extraction and geocoding of historical floods, storms, and adaptation measures. Information Processing & Management, 59(1), p.102735.
45. Ortu, M., Vacca, S., Destefanis, G. and Conversano, C., 2022. Cryptocurrency ecosystems and social media environments: An empirical analysis through Hawkes' models and natural language processing. Machine Learning with Applications, 7, p.100229.
46. B. Pham, A. Shirzadi, D. Tien Bui, I. Prakash and M. Dholakia, "A hybrid machine learning ensemble approach based on a Radial Basis Function neural network and Rotation Forest for landslide susceptibility modeling: A case study in the Himalayan area, India", International Journal of Sediment Research, vol. 33, no. 2, pp. 157-170, 2018.