



A heterogeneous network embedded medicine recommendation system based on LSTM

Imran Ahmed ^a, Misbah Ahmad ^{b,c}, Abdellah Chehri ^d, Gwanggil Jeon ^{e,f,*}

^a School of Computing and Information Sciences, Anglia Ruskin University Cambridge, Cambridge, UK

^b Faculty of Health and Applied Sciences, University of the West of England, Bristol, UK

^c Department of Animals and Agriculture, Hartpury University, Gloucester, UK

^d Department of Mathematics and Computer Science, Royal Military College of Canada, Ontario, Canada

^e School of Electronic Engineering, Xidian University, Xi'an, 710071, China

^f Department of Embedded Systems Engineering, Incheon National University, Incheon, 22012, Korea

ARTICLE INFO

Article history:

Received 4 February 2023

Received in revised form 20 June 2023

Accepted 1 July 2023

Available online 11 July 2023

Keywords:

Inter of Things (IoT)

Heterogeneous information networks

Deep learning

LSTM

Recommendation system

Healthcare

ABSTRACT

In the healthcare sector, patient data plays a crucial role in medical diagnoses and treatment plans. However, existing techniques for finding similar patients based on Electronic Health Record (EHR) data face challenges due to high-dimensional and sparse vectors. To overcome this challenge, the paper proposes developing a novel heterogeneous network-embedded drug recommendation system. The system focuses on classifying the sentiment of drug users based on their reviews and other relevant features such as their medical condition, drug rating, and usage date. The overall framework of the system follows a step-by-step approach, starting with data exploration and preprocessing, followed by the development of a classification model based on Long-Short-Term Memory (LSTM) networks. During the data exploration phase, various visualization and statistical techniques are employed to analyze the different data types. This process helps in understanding the characteristics of the data, identifying patterns, and preparing the data to align with the research objective. Furthermore, additional variables are considered suitable for the LSTM model, a recurrent neural network (RNN) type designed to handle sequence data and long-term prediction problems. Unlike other models that process individual data points, LSTM incorporates feedback connections to process complete data sequences. This approach enhances the effectiveness of recommendation systems and enables the prediction of new drug user ratings based on existing user ratings. The developed system demonstrates promising results, achieving a classification accuracy of 92%. This indicates its ability to accurately predict the sentiment of drug users based on their reviews and other associated features. The novelty of this research lies in the integration of a heterogeneous network-embedded approach with LSTM-based classification, providing a more comprehensive and accurate drug recommendation system compared to existing methods.

© 2023 Elsevier B.V. All rights reserved.

1. Introduction

The healthcare sector generates a vast amount of data pertaining to patient health issues, demographic information, and required medications [1]. This data is of great interest to both data scientists and medical specialists [2]. One crucial challenge in leveraging artificial intelligence (AI) and cloud computing in medicine is identifying patients who share similarities [3]. This task plays a significant role in medical diagnosis by accurately predicting comparable diseases and offering personalized treatment plans. Existing techniques for patient similarity search typically extract relevant medical events from Electronic Health Record (EHR) data and map them to vectors. These vectors

represent the similarity or dissimilarity between patients based on their medical occurrences. However, the resulting vectors often have high dimensions and sparse representations, making it difficult to precisely determine patient similarities. Furthermore, most current approaches lack access to the time data within EHRs, which hinders the study of how time influences patient similarity searches.

Heterogeneous information networks (HINs) are complex networks that consist of multiple types of nodes and connections. They are widely observed in various domains, such as biology, bibliographic systems, and social networks. HINs capture intricate interactions and rich semantics present in real-world scenarios, thanks to their diverse node and link types. Consequently, analyzing HINs has emerged as a promising approach for numerous data mining applications. In recent years, there has been a

* Corresponding author.

E-mail address: gjeon@inu.ac.kr (G. Jeon).

shift towards HIN embedding, driven by advancements in network embedding techniques. HIN embedding aims to project the nodes of a network into a lower-dimensional space while preserving the structural and semantic characteristics of the HIN. This learned low-dimensional embedding has proven to be an effective solution for various tasks, including node classification, link prediction, and recommendation systems [4,5].

With the significant growth of the web and the web-based enterprise, product ratings, and reviews have become crucial and essential elements for developing, selling, and purchasing products worldwide. People all around the globe have gotten used to reading scores/reviews/ratings and browsing websites before deciding to buy something. The area of healthcare or therapies has only occasionally been explored in this regard. Mostly previous research focuses on rating expectations and ideas in the E-Commerce sphere. The number of people looking for a diagnosis online and concerned about their health has grown. According to a Pew American Research Center poll manipulated in 2013 [6], almost 60% of adults explored online health-related topics, and about 35% of users sought to analyze health conditions on the Internet.

A drug recommendation system is essential to help doctors and patients in order to expand their knowledge about medicines for different medical issues. A usual system known as a recommendation framework suggests a product to the consumer based on their benefit and need. These frameworks use consumer surveys to analyze the responses and provide recommendations based on the customers' precise needs. Product ratings and reviews have become crucial in the web-based enterprise as the Internet has experienced significant growth. They play a vital role in developing, selling, and purchasing products worldwide. People across the globe now rely on reading scores, reviews, and ratings, as well as browsing websites, before making purchase decisions. However, the application of ratings and reviews in the healthcare and therapy sector has been limited. Previous research has primarily focused on rating expectations and ideas in the field of e-commerce. Nevertheless, there has been a noticeable increase in the number of individuals seeking online diagnoses and expressing concerns about their health. According to a 2013 Pew American Research Center poll, nearly 60% of adults explored health-related topics online, and approximately 35% of users searched for information about specific health conditions on the Internet. To address this growing need, a drug recommendation system is crucial in helping doctors and patients expand their knowledge about medicines for various medical issues. Typically, a recommendation framework is used to suggest products to consumers based on their specific needs and benefits. These frameworks utilize consumer surveys to analyze responses and provide tailored recommendations based on precise customer requirements.

The medicine recommendation system described in [2,7] leverages sentiment analysis and feature engineering to suggest medications based on patient evaluations [2,7]. Sentiment analysis involves techniques for extracting emotional information from text, such as opinions, thoughts, and perspectives [6]. In contrast, adding features through engineering increases the number of characteristics already present in a model, enhancing performance. One of the common machine learning issues that a deep learning algorithm may resolve is sentiment analysis. Sequential data should be evaluated while processing text using cutting-edge neural network architecture. LSTM, a type of artificial recurrent neural network (RNN) architecture used in deep learning, is the appropriate one [8,9]. This study aims to look at sentiment analysis of patient drug reviews. LSTM was used to analyze the dataset and build an effective model to predict the rating based on the provided reviews. A dataset used to develop

the model was taken from,¹ a popular pharmaceutical website that offers drug-related information. Each example in the data set includes user ratings and reviews on a particular medicine and a related condition. Rating labels are divided into three sentiment groups (positive, negative, and neutral). The main contributions of this research include employing LSTM for sentiment analysis of patient drug reviews, building a predictive model using the dataset from www.Drugs.com, and utilizing feature engineering techniques to improve the model's performance. The novelty lies in the combination of sentiment investigation, feature engineering, and LSTM architecture to create a robust medicine recommendation system based on patient evaluations. The main objectives of the work is provided as follows:

- To conduct the appropriate data analysis for investigating the data set's insights. This analysis involves extracting meaningful information from user ratings and reviews of medications and related conditions.
- To develop a heterogeneous network-based recommendation system that uses various review ratings for medicine selection and user-reported sickness conditions to identify the optimal treatment for illness. By incorporating different sources of information, the recommendation system aims to provide more accurate and personalized medication suggestions.
- To address the task of sentiment analysis using deep learning algorithms, specifically focusing on sequential data processing with the help of LSTM (Long Short-Term Memory) neural network architecture. The performance of the developed system is enhanced by incorporating additional features through engineering.
- To assess the model's performance through comparisons to state-of-the-art models and accuracy outcomes. This comparison aims to demonstrate the effectiveness of the developed system in predicting ratings based on provided reviews.

The contributions of the above work include developing a medicine recommendation system that incorporates sentiment analysis and feature engineering, the utilization of LSTM for analyzing patient drug reviews, and integrating the system with a heterogeneous network. The research work presented in this paper is organized as follows: Section 2 provides a summary of heterogeneous network-based recommendation systems. Section 3 briefly describes the data sets used in the research. Section 4 discusses the developed recommendation system for medical analysis and prediction, including detailed experimental results. Finally, Section 5 presents the conclusion of the work and provides recommendations for future research.

2. Related work

Integration of machine learning and deep learning techniques [10,11] into recommendation systems has gained significant attention due to the rise of heterogeneous information networks and advancements in artificial intelligence [12,13]. Recommendation frameworks are widely used across various industries, such as restaurants, movies, music, education, television programs, books, real estate, business partnerships, websites, conferences, tourism, sports, products, financial services, and learning materials. However, there needs to be more research in the domain of pharmaceutical recommendation systems utilizing sentiment analysis, primarily due to the presence of clinical terminologies and specific characteristics in medicine evaluations [14,15].

¹ www.Drugs.com.

Jin et al. [16] proposed a system that utilizes a heterogeneous network to identify potential drug candidates for COVID-19. This research aimed to assist experts in their search for medication information. To address this need, Doulaverakis et al. [17] introduced GalenOWL, an online platform empowered by semantic technology. GalenOWL integrates clinical terminology and data accurately, leveraging internationally recognized standards such as ICD-10 and UNII to transform clinical information into ontological terms. By doing so, GalenOWL is able to provide personalized medication recommendations based on a patient's illness, sensitivities, and potential interactions with other medicines.

In one study, [18] researchers focused on multilingual sentiment analysis and compared the performance of Naive Bayes, and Recurrent Neural Network (RNN) approaches. They translated multilingual tweets into English using the Google Translator API. The results revealed that the RNN method outperformed Naive Bayes, achieving an accuracy of 95.34% compared to 77.21%. [19] conducted research based on the belief that prescription medicine should consider the patient's individual capabilities. They proposed a risk level categorization approach to assess the patient's immunity. Over 60 risk factors, such as hypertension and alcoholism, were considered to determine the patient's ability to protect themselves against infections. The study aimed to develop a web-based prototype system, utilizing a decision support system, to assist clinicians in selecting appropriate first-line medications.

Authors in [20] presented a drug recommendation system based on machine learning in their work. The authors focused on developing a system that utilizes machine learning techniques to recommend suitable medications. Leilei et al. [21] conducted a study aiming to identify the optimal treatment recommendation for patients by analyzing extensive treatment records. They used a robust semantic clustering algorithm to uncover similarities in treatment data. Additionally, the authors developed a model to evaluate the suitability of suggested treatment plans. This framework can provide tailored therapy regimens based on patients' demographics and medical conditions. Multiple hospitals tested the approach using electronic medical records (EMRs) of patients, and the results showed an improvement in the cure rate.

Xiaohong Jiang et al. [22] conducted a study using treatment data to test three additional algorithms for a drug recommendation system: the decision tree method, support vector machine (SVM), and backpropagation neural network. They evaluated these algorithms based on model accuracy, ability, and adaptability. The authors also proposed implementing a mistake check system to ensure analysis, accuracy, and administration quality. Mohammad Mehedi Hassan et al. [23] proposed CADRE, a cloud-assisted medication recommendation system that suggests medications based on adverse effects experienced by patients. Initially, CADRE employed collective filtering techniques to group medicines into clusters using functional characterization data. However, the model was later enhanced by integrating cloud-based tensor decomposition to address issues like computational complexity and sparse information. Similarly, Sadeghi et al. [24] developed an integrated heterogeneous and graph neural network model for Multi-Labeled drugs in medicine recommendation.

Jiugang Li et al. [25] developed a hashtag-based recommendation system using skip-gram and convolutional neural networks (CNN). They created semantic correction vectors to incorporate hashtag significance in sentiment analysis. These vectors were utilized to categorize hashtags using LSTM RNN. The results demonstrated that this model outperformed traditional approaches such as SVM and Standard RNN. The unique features of this model are lost when subjected to standard AI methods like SVM and collaborative filtering, which is a crucial consideration for realistic expectations. Another study by Hananto et al.

cites hananto2021analysis focused on machine learning and proposed a clustering technique for drug data mining. Additionally, Suhartono et al. [26] proposed a drug product sentiment analysis method using deep neural networks and word embeddings. Furthermore, Doma et al. [27] presented an automated deep learning-based drug recommendation system.

Integrating machine learning and deep learning techniques in recommendation systems has gained attention in various industries, but there is a lack of research in pharmaceutical recommendation systems utilizing sentiment analysis. Existing studies focus on drug candidate identification for COVID-19, personalized medication recommendations, multilingual sentiment analysis, risk level categorization for medication selection, machine learning-based drug recommendation systems, treatment recommendations based on semantic clustering, and cloud-assisted medication recommendation. These studies utilize techniques such as heterogeneous networks, semantic technology, sentiment analysis, decision support systems, machine learning algorithms, and deep neural networks [28,29]. However, there is a gap in the literature regarding pharmaceutical recommendation systems that incorporate sentiment analysis and clinical terminologies effectively.

From the above discussion, it can be concluded that limited research has been done in medicine recommendation system development. Very few studies use heterogeneous networks in the literature with limited data sets. Researchers applied different deep learning models [30]. [12,31–38], as black-box, while in this work, we applied in-depth data analytics techniques and developed a robust model using the heterogeneous network, with feature engineering and LSTM model.

3. Methodology

The method described above involves the use of a heterogeneous information network to connect patients, illnesses, and medications. The network schemas for patients, diseases, and drugs are represented in Fig. 1. The framework consists of multiple interconnected components, including data exploration, data preprocessing, sentiment analysis, and LSTM-based classification. Data exploration involves analyzing user reviews, drug ratings, usage dates, and other relevant features. The preprocessed data is then fed into the LSTM model, which leverages its sequential learning capabilities to predict user sentiment. The system integrates heterogeneous network embedding techniques to capture complex relationships among drugs, conditions, and user preferences.

These embeddings enhance recommendation accuracy by considering contextual information and personalized factors. The system provides robust and effective drug recommendations, leveraging the convergence of artificial intelligence and heterogeneous networks. This illustration demonstrates the comprehensive architecture and workflow of the proposed framework, facilitating accurate and personalized medicine recommendations for improved patient outcomes. In a heterogeneous network, it is possible to have multiple drug types that can treat a single disease, and a single drug may also be effective for various diseases. However, this can lead to inaccurate information. To address this issue and gain more insights from the data, the authors propose utilizing reviews and ratings from individuals with the same illness who are using the same medication. This additional information can be used for data analysis. The authors apply descriptive data analysis techniques to explore the dataset attributes and identify key characteristics. To prepare the data for machine learning modeling, they convert the categorical values of drug names and conditions into numerical values using a technique called Label Encoder. Considering the large number of medications (3667) in the dataset, the authors choose to use

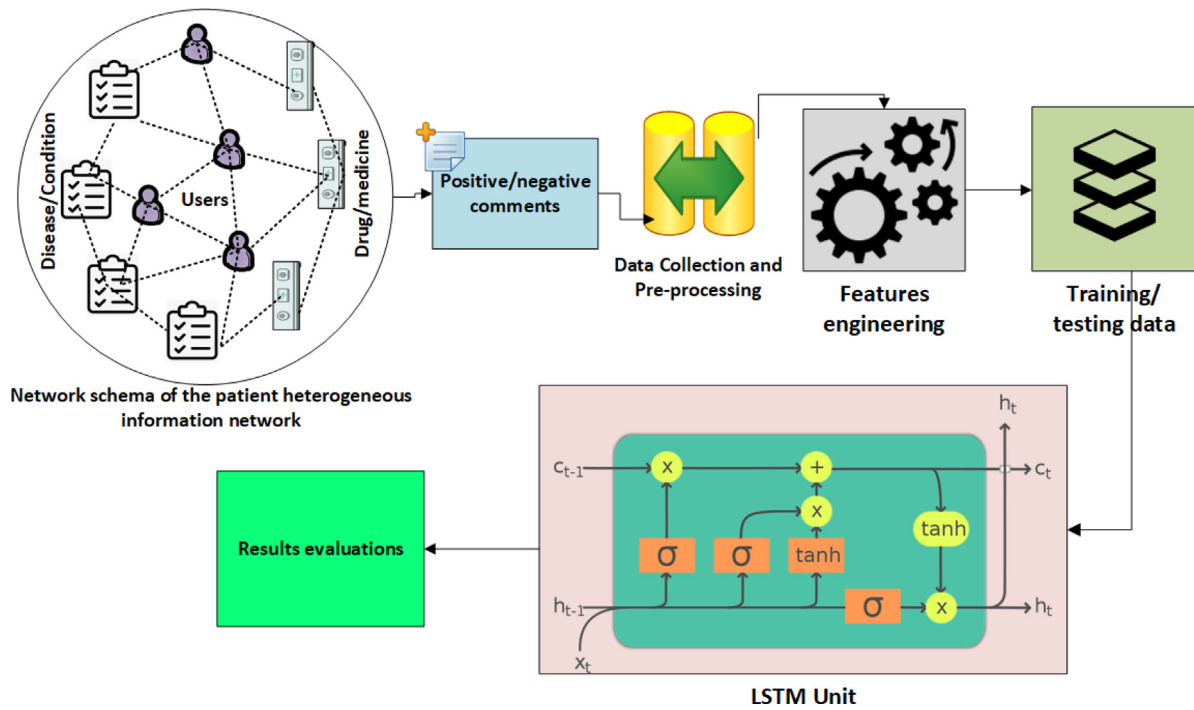


Fig. 1. The illustration depicts a heterogeneous network embedded medicine recommendation system based on LSTM, encompassing data exploration, preprocessing, sentiment analysis, and LSTM-based classification. The framework leverages heterogeneous network embedding techniques and sequential learning to provide accurate and personalized drug recommendations, integrating AI and heterogeneous networks for improved patient outcomes.

Table 1

Dataset description contains patient feedback on medications, including information on associated ailments and patient ratings.

Data	#Train	#Test	#conditions	#drugs	length	rating	label	%
Drugs.com								
Overall Rating	161297	53766	836	3654	458.32 (240.76)	rating ≤ 4 4 < rating < 7 rating ≥ 7	-1 0 1	25 9 66
Side Effects (Annotated)	-	400	141	243	500.385 (209.42)	No Side Effects Mild / Moderate Side Effects Severe / Extremely Severe Side Effects	0 1 2	32 28 40
Druglib.com								
Overall Rating	3107	1036	1808	541	277.57 (283.21)	rating ≤ 4 4 < rating < 7 rating ≥ 7	-1 0 1	21 10 69
Benefits (Effectiveness)	3107	1036	1808	541	212.87 (198.51)	Ineffective Marginally / Moderately Effective Considerably / Highly Effective	0 1 2	8 19 73
Side Effects	3107	1036	1808	541	177.36 (197.93)	No Side Effects Mild / Moderate Side Effects Severe / Extremely Severe Side Effects	0 1 2	30 53 17

a one-hot encoder, which would result in 3667 new features. However, this approach can be computationally costly. To aid in categorization and identify important features, feature engineering is applied to analyze the key characteristics of the data. The dataset is then split into training and testing samples for further analysis.

For classification purposes, the authors employ the LSTM (Long Short-Term Memory) model, which is a type of recurrent neural network known for its ability to handle sequential data. More details about the LSTM model are discussed in a subsection (not provided). After the classification process, the predicted results are evaluated based on accuracy, which measures how well the model predicts the correct classifications.

The novelty of this approach lies in the utilization of a heterogeneous information network to connect patients, diseases, and drugs. Additionally, incorporating reviews and ratings from patients with similar illnesses and medications provides extra information for data analysis. The presented work highlights the use

of descriptive data analysis, feature engineering, and the LSTM model for classification, which contribute to a comprehensive and detailed exploration of the dataset.

3.1. Dataset description and data analysis

The dataset referred to as [39] consists of patient feedback related to specific medications. The dataset includes information about the ailments associated with the drugs and a patient rating, which represents the overall satisfaction level on a scale of 1 to 10. The data was collected by extracting information from internet drug review websites. The dataset is stored in two.tsv (tab-separated-values) files, one designated for training purposes, which accounts for 75% of the data, and another file for testing purposes, containing the remaining 25% of the data. The dataset description can be found in Table 1.

In this section, we will employ various data analysis and visualization tools to gain insights from the data. Our objective is

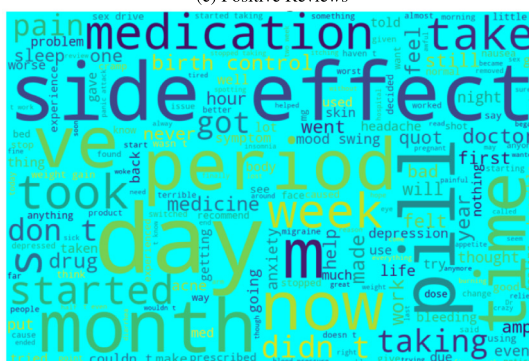
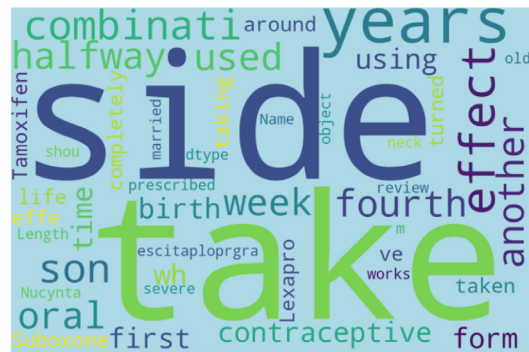


Fig. 2. Wordcloud analysis of dataset.

to present the subject matter, preprocess the data to align with our goals and generate different variables that are compatible with our model. Specifically, we have utilized a WordCloud visualization tool to showcase the dataset used in this study, as illustrated in Fig. 2. The WordCloud technique emphasizes the

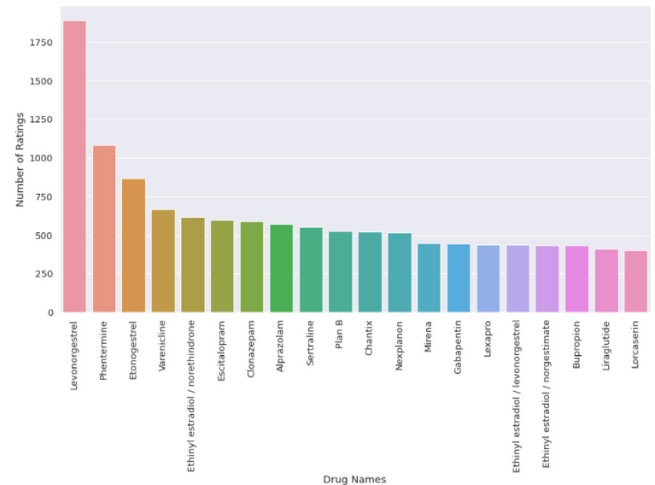


Fig. 3. Top 20 medicines/drugs with 10/10 rating score.

significance and frequency of different terms within the dataset. The size of each word in WordCloud indicates its frequency of occurrence. By examining the sample figures, we can observe that the most frequently mentioned medicines are 'Valsartan', 'Lybrel', and 'Guanfacine'. Additionally, when analyzing the review comments, we identified that the most prominent topics revolved around the side effects experienced by individuals and the act of taking medication itself.

Fig. 3 presents a bar graph illustrating the top 20 medications in the dataset that have received a perfect score of 10 out of 10. Among these medications, 'Levonorgestrel' stands out with approximately 1883 ratings, making it the drug with the highest number of 10/10 ratings. On the other hand, 'Lorcasrin' has the lowest number of 10/10 ratings among the top 20 medications. The graph visually represents the frequency or count of medications using the text size used for displaying the names of the medications. The size of each word corresponds to the number of 10/10 ratings received by the respective medication. By examining the graph, we can observe that 'Levonorgestrel' indicates the highest number of 10/10 ratings. In contrast, 'Lorcasrin' shows the lowest number of 10/10 ratings among the top 20 medications.

The data analysis reveals that Levonorgestrel stands out among the top 20 medications in the dataset with the highest percentage of perfect ratings. This indicates a significant level of user satisfaction with this medication. On the other hand, Lorcasrin receives the lowest number of 10/10 ratings among the top 20 medications, suggesting a comparatively lower level of satisfaction among users. The findings provide valuable insights into the satisfaction levels of users regarding different medications. Levonorgestrel emerges as the most highly rated medication, implying that many users have expressed satisfaction with its effectiveness or other aspects. This high rating could be attributed to various factors such as the medication's efficacy, tolerability, ease of use, or minimal side effects. In contrast, Lorcasrin receives the least number of perfect ratings among the top 20 medications. This suggests that a smaller percentage of users have reported complete satisfaction with this medication. The reasons for this lower satisfaction level could be diverse, including inadequate efficacy, unpleasant side effects, difficulties in administration, or other factors that have led users to provide lower ratings.

Fig. 4 displays a bar graph depicting the top 20 medications from our dataset. The ratings scale ranges from 1 to 10, with 1 being the lowest rating. In this graph, the medications are ranked based on the number of 1/10 ratings they have received. The medication 'Miconazole' stands out with the highest number

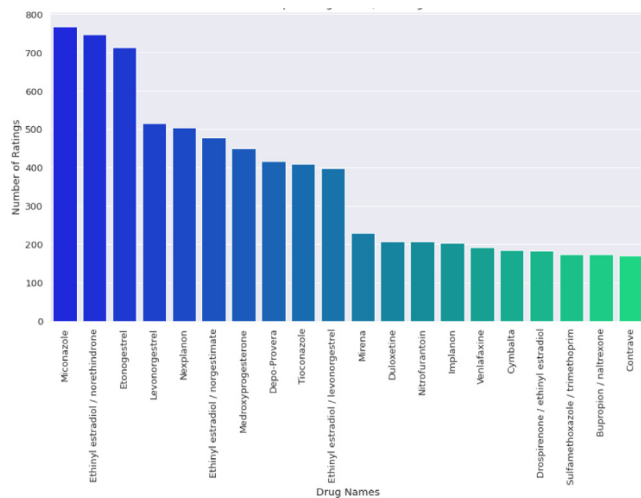


Fig. 4. Top 20 medicines/drugs with 1/10 rating score.

of 1/10 ratings, totaling approximately 767. This indicates that 'Miconazole' received the most negative feedback among the top 20 medications in our dataset.

In Fig. 5, we have depicted the typical illness and disease conditions. It is worth noting that the number of evaluations for birth control exceeds that of anthrax. This observation can be explained by considering the widespread use and relevance of birth control measures in comparison to the relatively rare occurrence of anthrax cases. To elaborate further, Fig. 4 visually represents various illnesses and disease conditions. It is a graphical depiction that provides an overview of the frequency of evaluations conducted for different medical issues. In this context, evaluations refer to the assessment and examination of individuals to diagnose or manage their health conditions. When comparing birth control and anthrax, we find that birth control measures, such as contraceptives and family planning methods, are commonly utilized by a significant portion of the population. Consequently, the demand for evaluations related to birth control is relatively high.

On the other hand, anthrax, a rare and potentially serious bacterial infection, is considerably less prevalent in comparison. Therefore, the number of evaluations associated with anthrax is comparatively lower. The contrasting number of evaluations for birth control and anthrax in Fig. 4 reflects the varying degrees of attention and resources dedicated to these health concerns. It highlights the significance of birth control in terms of public health and its relevance to a larger population compared to the relatively limited occurrence of anthrax cases.

Comparing the above figures, Fig. 5 displays information related to the common medications recommended for various illness conditions. We conducted an exploration of the dataset and discovered that each of the top eight disorders is associated with approximately 100 different drugs, as illustrated in Fig. 6. It provides a visual representation of the frequency distribution of medications across the top eight disorders. It shows that there is a wide range of medications available for each illness condition, with approximately 100 drugs associated with each disease. The figure helps us understand the diversity and abundance of treatment options available for these specific conditions. Additionally, it is worth noting that within the dataset, there is a peculiar statement associated with the illness condition, which states, "3 users considered this remark useful". However, this appears to be an error resulting from the crawling process and should be disregarded as it does not provide relevant information about the medications or the illness condition.

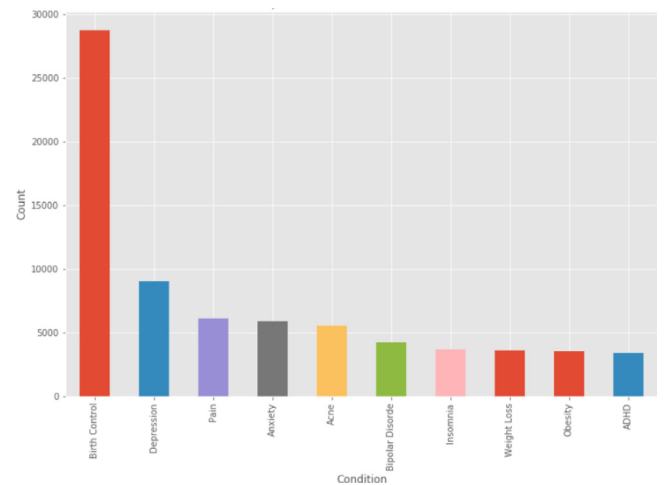


Fig. 5. To 10 most common conditions.

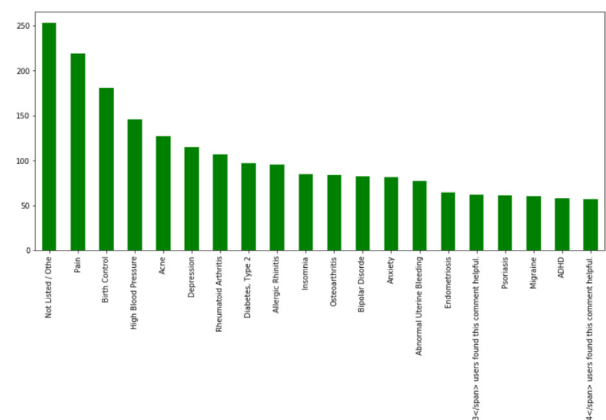


Fig. 6. Top 10 medications for different conditions.

Fig. 7 represents the distribution of ratings within a given dataset. The figure consists of two visualizations side by side. On the left side of Fig. 7a, a bar graph is displayed, which illustrates the frequency or count of each rating value ranging from 1 to 10. The height of each bar corresponds to the number of occurrences for that specific rating. This visualization provides a clear overview of the distribution of ratings within the dataset. On the right side of Fig. 7, a distribution plot is shown. This plot provides a visual representation of how the ratings are distributed across the range of values. It typically consists of a line or curve that represents the density or probability of each rating value occurring. The shape of the plot provides insights into the overall pattern and characteristics of the rating distribution, such as whether it is symmetric, skewed, or multimodal. The combination of the bar graph and the distribution plot in Fig. 7b allows for a comprehensive understanding of the dataset's rating distribution. The bar graph provides a discrete representation of the frequencies, while the distribution plot offers a continuous visualization of the density or probability. Together, these visualizations enable viewers to assess the prevalence and pattern of different ratings in the dataset.

Fig. 8 presents the results of sentiment analysis, indicating the total number of reviews. The analysis revealed that four specific rating values were predominantly chosen by users: 10, 9, 1, and 8. Among these four ratings, 10 had the highest occurrence, followed by 9, 1, and 8. Remarkably, the combined occurrence of 10 and 9 ratings was more than twice as high as the occurrence

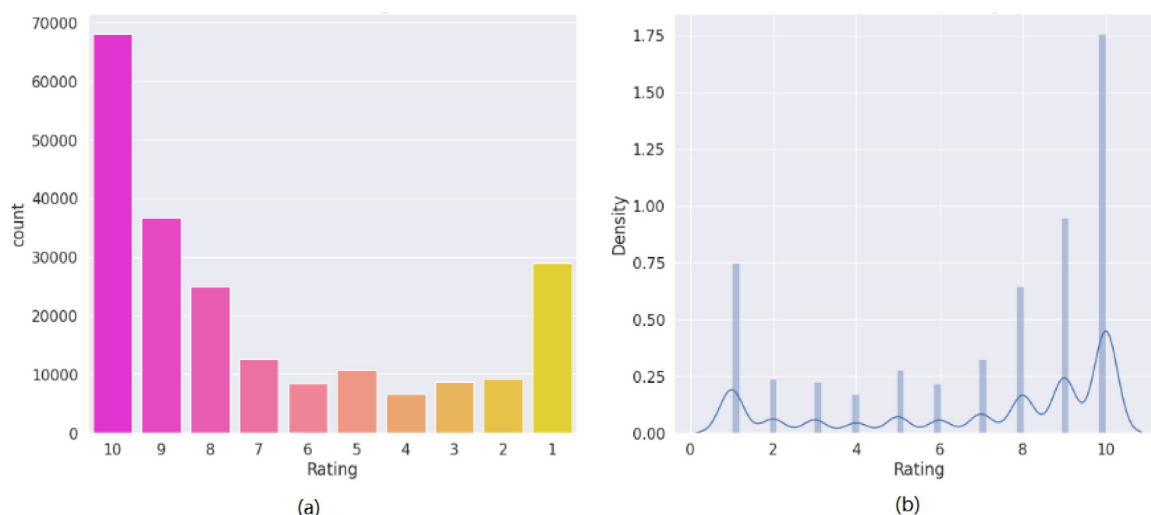
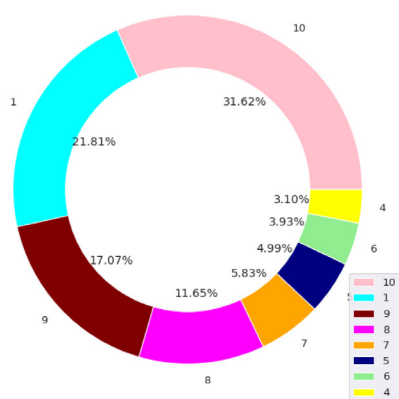
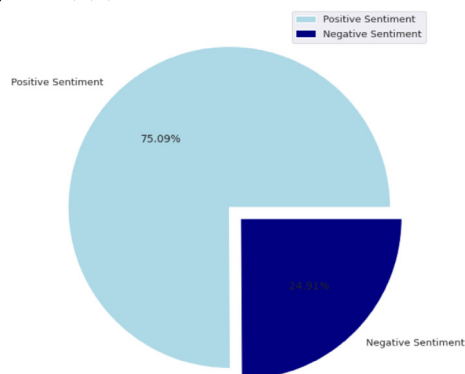


Fig. 7. Rating distribution of all reviews.



(a) Total number of reviews indicating four specific rating values were predominantly chosen by users: 10, 9, 1, and 8.



(b) Overall review analysis (sentiment analysis), indicating the total number of reviews, including positive and negative.

Fig. 8. An.

of the other three ratings combined. This observation suggests that users' responses to the subject under review are highly dramatic. Furthermore, it indicates a significantly larger proportion of positive ratings compared to negative ratings in Fig. 8(b).

The feature significance plot is displayed in Fig. 9. According to the plot, the mean word length is identified as the most crucial factor in determining a certain outcome. The patient's condition

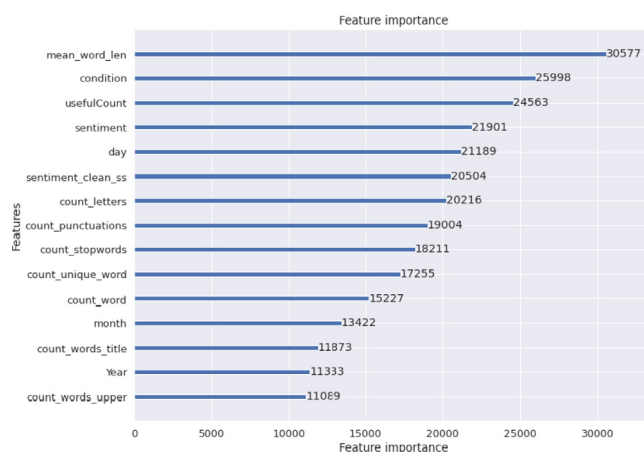


Fig. 9. Feature importance plot, identifying the mean word length which is most crucial factor in determining an outcome.

is ranked second in terms of significance. On the other hand, the number of words in upper case is found to be the least significant aspect among all the factors analyzed in the plot. The plot provides a visual representation of the relative importance of these factors, indicating that mean word length has the highest impact. In contrast, the number of words in upper case has the lowest impact.

In summary, the above data analysis and visualization provide insights of the dataset. A WordCloud showed the most frequently mentioned medications and common topics in the reviews. A bar graph highlighted the top drugs with perfect ratings of 10/10, indicating high user satisfaction for Levonorgestrel and lower satisfaction for Lorcasrin. Another bar graph showed medications with the highest number of negative ratings, with Miconazole standing out. Visual representations depicted the prevalence of evaluations for birth control compared to anthrax and the abundance of medications for different illness conditions. Distribution plots and bar graphs provided an overview of the rating distribution in the dataset. Sentiment analysis showed a higher occurrence of positive ratings (10 and 9) compared to negative ratings. Finally, a feature significance plot ranked mean word length and the patient's condition as the most important factors in determining outcomes.

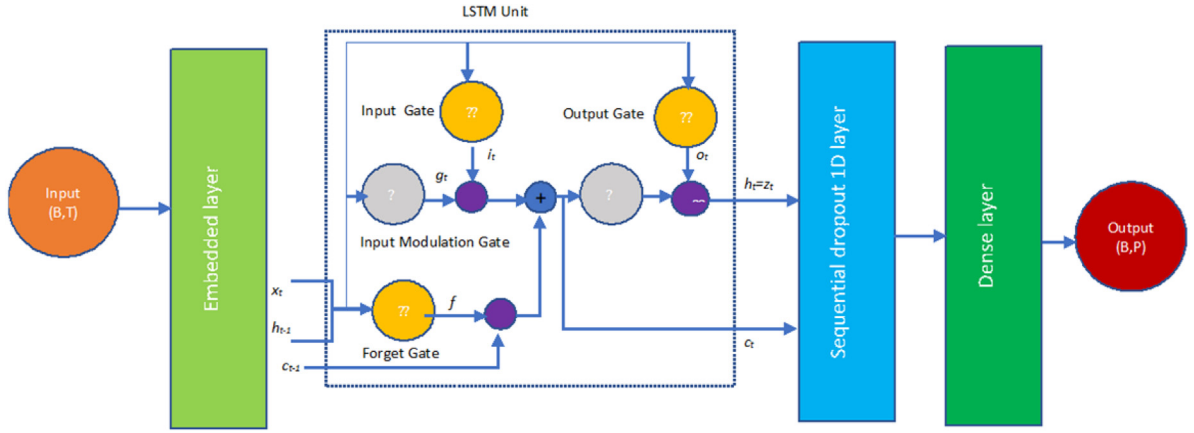


Fig. 10. An illustration of the proposed model. Including the embedding, LSTM, sequential dropout, and dense layers.

3.2. LSTM-based classification model

We employ five layers model in this study, including an Embedding layer, a SpatialDropout1D layer, two LSTM layers, and a Dense layer. The embedding layer is used For text processing. This layer converts the integer that was word-encoded into vectors. The training set uses a total number of (5000) unique words. The second input is the total number of vector dimensions that will be produced. The final one is input length = 200, which indicates that each word is a unique integer number with a size of 200. Similar to a dropout function is the spatial dropout 1D layer. When we enter the process with a rate of 0.1, a portion of the input units are dropped. However, the function is not the same precisely. The overall architecture is shown in Fig. 10. The model's components are shown in Figure: a word embedding and preprocessing layer to prepare and collect the data, an LSTM layer to forecast the date of medication, and a mix of dense and sequential layers to record and track the irregular time-lapse of segments.

Categorical data can be represented more expressively with a one-hot encoding. As a first step, we produced one-hot encodings of a predetermined time period described as B for each input sample. Using Keras padding in the second stage of the preprocessing pipeline ensured that all input samples were the same length. The words retrieved from the textual input were described as dense vectors in the third stage of the preprocessing channel using an embedding layer in Keras as a hidden layer. A vector represents the projection of a word in continuous vector space. The input data must be integer encoded such that a different integer represents each word as a requirement for this embedding layer. The embedding layer is started with random weights. The preprocessing pipeline's output (B, T) is an embedding with dimensionality.

When there are unknown time delays of undetermined magnitude and length between occurrences, LSTMs are ideally suited to producing predictions. Input gates, forget gates, output gates, and a memory cell make up the basic LSTM. This typical architecture implicitly assumes that data is evenly distributed during a sequence's elapsed duration. Detailed mathematical formulations of the LSTM that was employed are provided in Fig. 10. Our model's intermediate outcomes are produced by the LSTM Layers, which have output dimensions of (B, T, v) , (B, T, L) ,

(B, TL) , and (B, P) . Using $P = 1$ as the output, we may calculate the medicine rating for the positive feedback. The output from one batch of the period is represented by one LSTM matrix represented in the following equations;

$$i_t = \text{sigmoid}(W_{h_i} \times h_{t-1} + W_{x_i} \times x_t + b_i) \quad (1)$$

$$o_t = \text{sigmoid}(W_{h_o} \times h_{t-1} + W_{x_o} \times x_t + b_o) \quad (2)$$

$$f_t = \text{sigmoid}(W_{h_f} \times h_{t-1} + W_{x_f} \times x_t + b_f) \quad (3)$$

$$g_t = \tanh(W_{h_g} \times h_{t-1} + W_{x_g} \times x_t + b_g) \quad (4)$$

$$c_t = (f + t \cdot c_{t-1} + (i_t \cdot g_t)) \quad (5)$$

$$h_t = o_t \cdot \tanh c_t \quad (6)$$

In our approach, we used a sequential dropout layer to enhance the efficiency of the unfolding process. Unfolding refers to processing input data over multiple time steps in recurrent neural networks (RNNs) such as LSTM (Long Short-Term Memory) networks. By applying dropout, a technique that randomly sets a portion of input units to zero during training, we aim to accelerate the unfolding process. Additionally, we incorporated a dense layer into our network architecture. This layer is responsible for integrating the time segments, which are irregularly sampled, into a fully-connected network. Irregularly sampled time frames refer to instances where the intervals between consecutive observations or measurements are not constant. The dense layer helps capture the temporal patterns present in these irregularly sampled time frames.

To generate the output for each LSTM matrix, we produce a sequence with a dimension of 1. This sequence represents the irregular time-lapse segments. For this purpose, we use a sigmoid activation function, which maps the output values to a range between 0 and 1. The sigmoid function is commonly used in binary classification tasks, where the goal is to classify inputs into one of two categories. Overall, our approach leverages a sequential dropout layer to speed up the unfolding process in RNNs and a dense layer to incorporate irregularly sampled time frames into a fully-connected network. Using a sigmoid activation function, the resulting output sequence represents the irregular time-lapse segments. We used a thick layer to merge the time segments into a fully linked network to speed up the unfolding process and detect sporadically sampled time frames. The irregular time-lapse segments for each LSTM matrix are then produced as a sequence (with dimension 1) made up of the numbers 0 and 1.

The novelty of the above model lies in its integration of a five-layer architecture that includes an Embedding layer, SpatialDropout1D layer, two LSTM layers, and a Dense layer, along with the incorporation of sequential dropout and a dense layer to handle irregularly sampled time frames. This combination allows for a more expressive representation of categorical data through

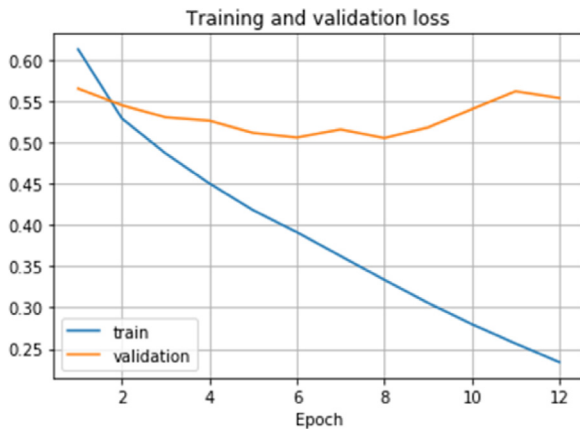


Fig. 11. Training and validation loss.

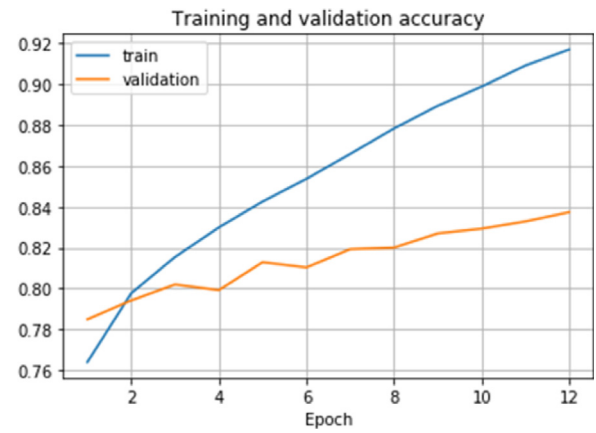


Fig. 12. Training and validation accuracy.

one-hot encoding, effective processing of textual input through word embeddings, and accurate predictions of unknown time delays. The model's ability to capture temporal patterns in irregularly sampled time frames and classify inputs into two categories using a sigmoid activation function further adds novelty. Overall, this comprehensive approach offers an innovative solution for efficient and accurate forecasting in time-series analysis.

4. Results

The training and evaluation results of the discussed model are presented in this section. The model was trained for 112 epochs, and its performance was analyzed using various metrics. Two figures, Figs. 11 and 12, were used to visualize the findings. Fig. 11 depicts the training and validation loss values changes across multiple epochs. The x-axis represents the number of epochs, while the y-axis represents the corresponding loss values. The plot reveals a gradual decrease in the loss values after the tenth epoch. The validation loss, indicated by the curve, reaches its lowest point at 0.55. This value suggests that the model performs relatively well in generalizing to unseen data. A lower validation loss signifies that the model can make accurate predictions on new, unseen examples. On the other hand, the training loss, represented by the curve as well, achieves a minimum value of 0.25. This lower value indicates a closer fit between the model's predictions and the training data. It suggests that the model has learned the patterns and details present in the training dataset quite well.

Additionally, Fig. 12, presents the training and validation accuracy of the model. The accuracy values indicate the model's ability to make correct predictions for the target variable. The training accuracy, measured at 0.92, signifies that the model accurately predicts the target variable for 92%. Comparatively, the validation accuracy is slightly lower at 0.80. This suggests a slight reduction in performance when the model encounters unseen data from the validation set. The lower validation accuracy might indicate that the model struggles to generalize as effectively to new examples as it does with the training data. In summary, the loss and accuracy values provide insights into the performance and generalization ability of the model. The decreasing loss values demonstrate that the model learns and improves over the epochs. The lower validation loss and relatively high training accuracy suggest that the model can make accurate predictions on both the training and validation data, although there seems to be a slight decline in performance on unseen data, as indicated by the lower validation accuracy. These figures provide a visual representation

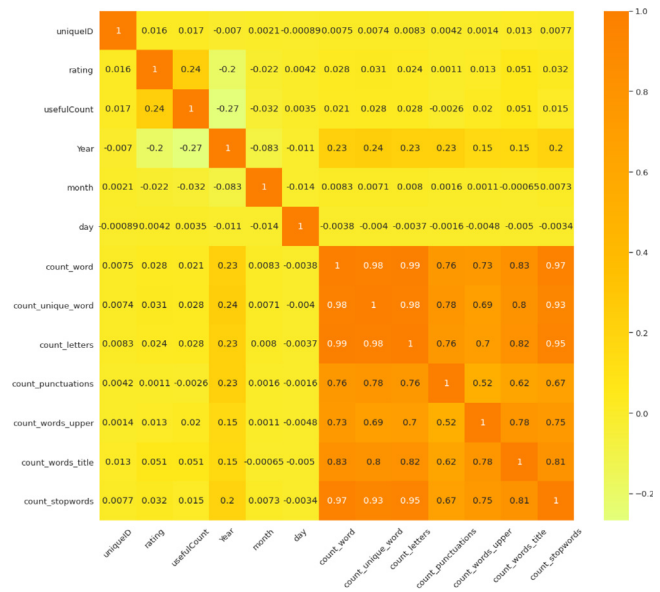


Fig. 13. Confusion matrix.

of the model's training and validation process, offering insights into its performance and generalization capabilities.

The evaluation results using a confusion matrix are shown in Fig. 13 and presented in Table 2. A confusion matrix is a table that allows visualization of the performance of a classification model by displaying the number of correct and incorrect predictions made by the model. Fig. 13 likely represents such a matrix, with rows representing the actual classes and columns representing the predicted classes. Each cell in the matrix shows the count or percentage of samples that belong to a specific combination of true and predicted classes.

Table 2 presents the parameter evaluation results derived from the confusion matrix. The model's overall accuracy is stated as 91%, which indicates the proportion of correctly classified samples out of the total number of samples. An accuracy of 91% suggests that the model performs well in predicting the correct class labels. Precision is another performance metric that evaluates the model's ability to correctly classify positive instances. With a precision value of 92%, the model demonstrates a high level of accuracy in correctly identifying positive samples. The recall value, also known as sensitivity or true positive rate, measures the proportion of actual positive instances that are correctly identified by the model. A recall value of 83% indicates that

Table 2
Comparison results with other state-of-the-art methods.

S.No	Method	Accuracy	Precision	Recall
1	SVM	90%	94%	80%
2	Random forest	89%	91%	78%
3	KNN	84%	90%	76%
4	ANN	90%	94%	80%
5	Proposed LSTM	91%	98%	83%

the model is successful in capturing a significant portion of the positive samples.

Additionally, the results obtained from the proposed model are compared with those from other state-of-the-art methods, and Table 2 likely includes a comparative analysis. It can be inferred from the table that all of the proposed models achieve excellent results, although the exact details are not provided in the query.

In summary, the training and evaluation results demonstrate the model's performance and generalization capabilities. The figures and metrics provide quantitative assessments of the model's accuracy, precision, and recall. The model shows promising results, achieving high accuracy and demonstrating a relatively close fit to the training data. The comparative analysis in Table 2 suggests that the proposed model performs well compared to other state-of-the-art methods, although detailed information on the comparison is not provided in the query.

5. Conclusion and future work

This paper presents a recommendation system that combines deep learning and heterogeneous networks to propose the best medication for a patient's condition based on reviews. The research was conducted in several stages, including data exploration, preparation, and modeling. The researchers used statistical and visual analysis methods to examine different data types during the data exploration phase. This process allowed them to gain insights and understand the dataset's characteristics. Pre-processing steps were then performed to ensure that the data met the requirements of the model. To overcome the limitations of natural language processing, the researchers employed a deep learning model called Long Short-Term Memory (LSTM). By applying sentiment analysis using LSTM, the model achieved an impressive accuracy of 92% in recommending the most suitable medication for a patient's condition. The novelty of this research lies in integrating deep learning techniques, specifically LSTM, with heterogeneous networks for medication recommendation. This approach combines the power of deep learning in understanding and extracting patterns from textual reviews with the versatility of heterogeneous networks in handling diverse data types. This research addresses the important problem of medication recommendation, which can significantly impact patient outcomes and healthcare efficiency. By leveraging deep learning and heterogeneous networks, the proposed recommendation system can improve the accuracy and effectiveness of medication suggestions.

In terms of future directions, the research could be extended to incorporate more diverse and comprehensive datasets to enhance the recommendation system's performance. This could include incorporating additional sources of patient feedback, such as social media data or electronic health records. The model's interpretability could be improved. While deep learning models often achieve high accuracy, they can be considered black boxes due to their complex internal workings. Developing techniques to interpret and explain the recommendations generated by the model would be valuable, especially in the healthcare domain, where

transparency and trust are crucial. Lastly, conducting user studies and evaluating the system's performance in real-world clinical settings would provide valuable insights into its effectiveness and usability.

CRedit authorship contribution statement

Imran Ahmed: Conceptualization, Methodology. **Misbah Ahmad:** Software, Validation, Data curation, Writing, Visualization, Investigation. **Abdellah Chehri:** Review & editing, Software, Validation. **Gwanggil Jeon:** Visualization, Investigation, Review & editing.

Declaration of competing interest

None Declared.

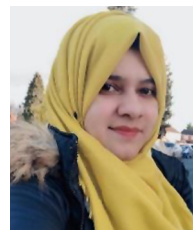
References

- [1] I. Ahmed, G. Jeon, F. Piccialli, A deep-learning-based smart healthcare system for patient's discomfort detection at the edge of internet of things, *IEEE Internet Things J.* 8 (13) (2021) 10318–10326.
- [2] I. Ahmed, M. Ahmad, G. Jeon, F. Piccialli, A framework for pandemic prediction using big data analytics, *Big Data Res.* 25 (2021) 100190.
- [3] A. Razaque, M. Aloqaily, M. Almiani, Y. Jararweh, G. Srivastava, Efficient and reliable forensics using intelligent edge computing, *Future Gener. Comput. Syst.* 118 (2021) 230–239.
- [4] Z. Zhao, X. Zhang, H. Zhou, C. Li, M. Gong, Y. Wang, HetNERec: Heterogeneous network embedding based recommendation, *Knowl.-Based Syst.* 204 (2020) 106218.
- [5] Y. Wang, Y. Yao, H. Tong, F. Xu, J. Lu, A brief review of network embedding, *Big Data Min. Anal.* 2 (1) (2018) 35–47.
- [6] S. Fox, M. Duggan, *Health online 2013*, *Health 2013* (2013) 1–55.
- [7] I. Ahmed, M. Ahmad, G. Jeon, Social distance monitoring framework using deep learning architecture to control infection transmission of COVID-19 pandemic, *Sustainable Cities Soc.* 69 (2021) 102777.
- [8] I. Ahmed, M. Ahmad, A. Chehri, G. Jeon, A smart-anomaly-detection system for industrial machines based on feature autoencoder and deep learning, *Micromachines* 14 (1) (2023) 154.
- [9] I. Ahmed, S. Din, G. Jeon, F. Piccialli, Exploring deep learning models for overhead view multiple object detection, *IEEE Internet Things J.* 7 (7) (2019) 5737–5744.
- [10] I. Ahmed, M. Anisetti, A. Ahmad, G. Jeon, A multilayer deep learning approach for malware classification in 5G-enabled IIoT, *IEEE Trans. Ind. Inform.* 19 (2) (2022) 1495–1503.
- [11] M. Ahmad, I. Ahmed, G. Jeon, An IoT-enabled real-time overhead view person detection system based on cascade-RCNN and transfer learning, *J. Real-Time Image Process.* (2021) <http://dx.doi.org/10.1007/s11554-021-01103-0>.
- [12] I. Ahmed, M. Ahmad, A. Chehri, M.M. Hassan, G. Jeon, IoT enabled deep learning based framework for multiple object detection in remote sensing images, *Remote Sens.* 14 (16) (2022) 4107.
- [13] M. Ahmad, I. Ahmed, K. Ullah, I. Khan, A. Khattak, A. Adnan, Energy efficient camera solution for video surveillance, *Int. J. Adv. Comput. Sci. Appl.* 10 (3) (2019) <http://dx.doi.org/10.14569/IJACSA.2019.0100367>.
- [14] J. Liu, C. Shi, C. Yang, Z. Lu, S.Y. Philip, A survey on heterogeneous information network based recommender systems: Concepts, methods, applications and resources, *AI Open* 3 (2022) 40–57.
- [15] T.N. Tekade, M. Emmanuel, Probabilistic aspect mining approach for interpretation and evaluation of drug reviews, in: 2016 International Conference on Signal Processing, Communication, Power and Embedded System, SCOPES, IEEE, 2016, pp. 1471–1476.
- [16] S. Jin, X. Zeng, W. Huang, F. Xia, C. Jiang, X. Liu, S. Peng, Heterogeneous network-based drug repurposing for COVID-19, 2021, arXiv preprint [arXiv: 2107.09217](https://arxiv.org/abs/2107.09217).
- [17] C. Doulaverakis, G. Nikolaidis, A. Kleontas, I. Kompatsiaris, GalenOWL: Ontology-based drug recommendations discovery, *J. Biomed. Semant.* 3 (1) (2012) 1–9.
- [18] V. Goel, A.K. Gupta, N. Kumar, Sentiment analysis of multilingual twitter data using natural language processing, in: 2018 8th International Conference on Communication Systems and Network Technologies, CSNT, IEEE, 2018, pp. 208–212.

- [19] K. Shimada, H. Takada, S. Mitsuyama, H. Ban, H. Matsuo, H. Otake, H. Kunishima, K. Kanemitsu, M. Kaku, Drug-recommendation system for patients with infectious diseases, in: AMIA Annual Symposium Proceedings, Vol. 2005, American Medical Informatics Association, 2005, p. 1112.
- [20] M.N. Uddin, M.F.B. Hafiz, S. Hossain, S.M.M. Islam, Drug sentiment analysis using machine learning classifiers, *Int. J. Adv. Comput. Sci. Appl.* 13 (1) (2022).
- [21] L. Sun, C. Liu, C. Guo, H. Xiong, Y. Xie, Data-driven automatic treatment regimen development and recommendation, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1865–1874.
- [22] Y. Bao, X. Jiang, An intelligent medicine recommender system framework, in: 2016 IEEE 11th Conference on Industrial Electronics and Applications, ICIEA, IEEE, 2016, pp. 1383–1388.
- [23] Y. Zhang, D. Zhang, M.M. Hassan, A. Alamri, L. Peng, CADRE: Cloud-assisted drug recommendation service for online pharmacies, *Mob. Netw. Appl.* 20 (2015) 348–355.
- [24] S. Sadeghi, J. Lu, A. Ngom, An integrative heterogeneous graph neural network-based method for multi-labeled drug repurposing, *Front. Pharmacol.* 13 (2022).
- [25] J. Li, H. Xu, X. He, J. Deng, X. Sun, Tweet modeling with LSTM recurrent neural networks for hashtag recommendation, in: 2016 International Joint Conference on Neural Networks, IJCNN, IEEE, 2016, pp. 1570–1577.
- [26] D. Suhartono, K. Purwandari, N.H. Jeremy, S. Philip, P. Arisaputra, I.H. Parmonangan, Deep neural networks and weighted word embeddings for sentiment analysis of drug product reviews, *Procedia Comput. Sci.* 216 (2023) 664–671.
- [27] V. Doma, S. Singh, N. Arora, G. Ortiz, P.K. Saran, S. Chavarin, M. Pirouz, Automated drug suggestion using machine learning, in: Advances in Information and Communication: Proceedings of the 2020 Future of Information and Communication Conference (FICC), Volume 2, Springer, 2020, pp. 571–589.
- [28] M. Ahmad, I. Ahmed, F.A. Khan, F. Qayum, H. Aljuaid, Convolutional neural network-based person tracking using overhead views, *Int. J. Distrib. Sens. Netw.* 16 (6) (2020) 1550147720934738.
- [29] M. Ahmad, I. Ahmed, G. Jeon, An IoT-enabled real-time overhead view person detection system based on cascade-RCNN and transfer learning, *J. Real-Time Image Process.* (2021) 1–11.
- [30] I. Ahmed, M. Ahmad, M. Nawaz, K. Haseeb, S. Khan, G. Jeon, Efficient topview person detector using point based transformation and lookup table, *Comput. Commun.* (2019) URL <https://www.sciencedirect.com/science/article/abs/pii/S0140366419305092?via%3Dihub>.
- [31] M. Ahmad, I. Ahmed, K. Ullah, I. Khan, A. Khattak, A. Adnan, Person detection from overhead view: A survey, *Int. J. Adv. Comput. Sci. Appl.* 10 (4) (2019) <http://dx.doi.org/10.14569/IJACSA.2019.0100470>.
- [32] I. Ahmed, G. Jeon, A. Chehri, M.M. Hassan, Adapting Gaussian YOLOv3 with transfer learning for overhead view human detection in smart cities and societies, *Sustainable Cities Soc.* 70 (2021) 102908.
- [33] Y. Lin, H. Song, F. Ke, W. Yan, Z. Liu, F. Cai, Optimal caching scheme in D2D networks with multiple robot helpers, *Comput. Commun.* 181 (2022) 132–142.
- [34] X. Deng, E. Liu, S. Li, Y. Duan, M. Xu, Interpretable multi-modal image registration network based on disentangled convolutional sparse coding, *IEEE Trans. Image Process.* 32 (2023) 1079–1091.
- [35] H. Jiang, Z. Xiao, Z. Li, J. Xu, F. Zeng, D. Wang, An energy-efficient framework for internet of things underlying heterogeneous small cell networks, *IEEE Trans. Mob. Comput.* 21 (1) (2022) 31–43.
- [36] X. Liu, J. Zhao, J. Li, B. Cao, Z. Lv, Federated neural architecture search for medical data security, *IEEE Trans. Ind. Inform.* 18 (8) (2022) 5628–5636.
- [37] B. Cheng, D. Zhu, S. Zhao, J. Chen, Situation-aware IoT service coordination using the event-driven SOA paradigm, *IEEE Trans. Veh. Technol.* 13 (2) (2016) 349–361.
- [38] X. Zhang, Y. Wang, M. Yang, G. Geng, Toward concurrent video multi-cast orchestration for caching-assisted mobile networks, *IEEE Trans. Veh. Technol.* 70 (12) (2021) 13205–13220.
- [39] F. Gräßer, S. Kallumadi, H. Malberg, S. Zaunseder, Aspect-based sentiment analysis of drug reviews applying cross-domain and cross-data learning, in: Proceedings of the 2018 International Conference on Digital Health, 2018, pp. 121–125.



Imran Ahmed (Senior Member, IEEE) received a Ph.D. degree in Computer Science from the University of Southampton, Southampton, U.K. Before joining ARU; he worked as an Associate Professor with the Institute of Management Sciences Peshawar, Hayatabad, Peshawar, Pakistan. His research interests include deep learning, machine learning, computer vision, data science, data mining, and IoT. He has also attended several international conferences in these areas. He has published numerous articles in refereed journals and conference proceedings, including IEEE Access, IEEE Transactions on Industrial Informatics, IEEE Internet of Things Journal, International Journal of Machine Learning and Cybernetics, IEEE/CAA Journal of Automatica Sinica, Journal of Ambient Intelligence and Humanized Computing, Multimedia Tools and Applications, Cluster Computing, IEEE Annual Ubiquitous Computing, Computer Communications, Electronics & Mobile Communication Conference (UEMCON), ACM/SIGAPP, Symposium On Applied Computing at Pau, France, and many others. He has also been acting as a reviewer of journals, such as IEEE Transactions on Industrial Electronics, IEEE Access, Journal of Ambient Intelligence, and Elsevier. Because of his excellent research profile, he was also enlisted in the Top 2 scientist list of Stanford University, USA, 2021. He is also an Associate Editor/Guest Editor of well-reputed international journals.



Misbah Ahmad is a Ph.D. student at the University of West of England, Bristol, UK. She is doing her research on animal behavior recognition at Hartpury University, Gloucester, UK. She received her MSCS degree from the Institute of Management Sciences, Hayatabad, Peshawar, Pakistan, in 2019 and her BS-Telecommunication Systems from Islamia College University, Peshawar in 2015. Her research interests include Computer Vision, Image Processing, Machine Learning, Deep Learning, and Data Science.



Abdellah Chehri is a professor at the University of Quebec- UQAC, Canada. Dr. Chehri served as the program chair and workshops chair of the 2014 IEEE International Humanitarian Technology Conference (IEEE IHTC 2014). He served as guest editor and severed as a technical reviewer for several international conferences and journals. Additionally, he is a Senior Member of IEEE Communications Society and IEEE Canadian Humanitarian Initiatives Committee (HIC).



Gwanggil Jeon received the B.S., M.S., and Ph.D. (summa cum laude) degrees from the Department of Electronics and Computer Engineering, Hanyang University, Seoul, Korea, in 2003, 2005, and 2008, respectively. From 2009.09 to 2011.08, he was with the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada, as a Post-Doctoral Fellow. From 2011.09 to 2012.02, he was with the Graduate School of Science and Technology, Niigata University, Niigata, Japan, as an Assistant Professor. From 2014.12 to 2015.02 and 2015.06 to 2015.07, he was a Visiting Scholar at Centre de Mathématiques et Leurs Applications (CMLA), École Normale Supérieure Paris-Saclay (ENS-Cachan), France. From 2019 to 2020, he was a Prestigious Visiting Professor at Dipartimento di Informatica, Università degli Studi di Milano Statale, Italy. He is currently a Full Professor at Incheon National University, Incheon, Korea. He was a Visiting Professor at Sichuan University, China, Universitat Pompeu Fabra, Barcelona, Spain, Xinjiang University, China, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, and University of Burgundy, Dijon, France.

Dr. Jeon is an ACM Member, IEEE Senior Member, an Associate Editor of Sustainable Cities and Society, IEEE Access, Real-Time Image Processing, Journal of System Architecture, and MDPI Remote Sensing. Dr. Jeon was a recipient of the IEEE Chester Sall Award in 2007, ACM's Distinguished Speaker in 2022, the ETRI Journal Paper Award in 2008, and Industry-Academic Merit Award by Ministry of SMEs and Startups of Korea Minister in 2020.