Drug Recommendations Using a Reviews and Sentiment Analysis by RNN

Pokkuluri Kiran Sree[1](✉), SSSN Usha Devi N[2], Phaneendra Varma Chintalapati[1], Gurujukota Ramesh Babu[1], and PBV Raja Rao[1]

[1] Department of C.S.E, Shri Vishnu Engineering College for Women (A), Bhimavaram, AP, India
drkiransree@gmail.com

[2] Department of C.S.E, University College of Engineering, JNTU Kakinada, A.P, India

Abstract. Sentiment analysis plays a crucial role in understanding the opinions and attitudes expressed in textual data. This paper explores the utilization of two distinct approaches, Recurrent Neural Networks (RNNs) and Cellular Automata (CA), for recommending drugs based on sentiment analysis of user reviews.

Recurrent Neural Networks (RNNs) have emerged as a powerful tool for analyz ing sequential data. In the context of sentiment analysis, RNNs excel at capturing contextual information and dependencies between words within a sentence. By training an RNN on a labeled dataset of drug reviews, sentiment patterns can be learned, enabling the model to predict the sentiment associated with unseen reviews.

Cellular Automata (CA) offer an alternative approach to sentiment analysis. CA are discrete systems where cells transition between states based on local interactions with neighboring cells. Applying CA to sentiment analysis involves repre senting each word or phrase in a review as a cell, and defining rules that govern sentiment state transitions based on neighboring cells' sentiments. By iteratively updating the cellular automaton over multiple time steps, sentiment dynamics within the text corpus can be modeled.

RNNs are particularly adept at capturing long-term dependencies and contextual nuances within a text sequence. Conversely, CA provide a spatially extended framework that can capture spatial dependencies between words. We propose a hybrid method RNN-CA-DR using both of these methods for developing a robust and accurate classifier for drug recommendation. The developed classifier has reported an accuracy of 91.23% and outperformed few base line models when tested with various parameters F1 Score, precision and recall.

Keywords: RNN (Recurrent Neural Network) · CA (Cellular Automata) · Sentiment Analysis

P. Pareek et al. (Eds.): IC4S 2023, LNICST 536, pp. 135–141, 2024.
https://doi.org/10.1007/978-3-031-48888-7_11

1 Introduction

Sentiment analysis of user reviews plays a crucial role in drug recommendation systems. This abstract focuses on the application of Recurrent Neural Networks (RNNs) for sentiment analysis to recommend drugs based on user reviews. RNNs have proven to be effective in capturing sequential dependencies and contextual information in text data, making them well-suited for sentiment analysis tasks.

The proposed approach involves training an RNN model on a labeled dataset of drug reviews, where each review is associated with a sentiment label (positive, negative, or neutral). The RNN leverages its recurrent nature to process the reviews as sequences, allowing it to capture the temporal dynamics and dependencies between words or phrases in the text. To represent the text data, word embeddings such as Word2Vec or GloVe can be utilized.

Cellular automata is an interesting approach for recommending drugs and performing sentiment analysis on reviews. Cellular automata are mathematical models that consist of a grid of cells, each of which can be in a specific state. The state of each cell evolves over time based on a set of predefined rules and the states of its neighboring cells.

2 Literature Survey on RNN and CA for Drug Recommendation.

2.1 RNN (Recurrent Neural Network)

Wen Zhang, et al. [2] has explored the application of RNNs for drug-target interaction prediction. It demonstrates the effectiveness of using RNNs to capture sequential dependencies in drug-target interaction data and achieve accurate predictions. The research showcases the potential of RNNs in drug recommendation by leveraging their ability to model complex relationships between drugs and their molecular targets.

The authors [4] propose a drug recommendation system using RNNs. They leverage the sequential nature of prescription data to capture temporal dependencies and generate personalized recommendations. The study demonstrates the advantages of RNNs in handling temporal data for drug recommendation tasks and provides insights into the implementation and evaluation of such systems.

This review paper [6] discusses the application of neural network models, including RNNs, in drug discovery and recommendation. It highlights the potential of RNNs in analyzing various data sources, such as chemical structures, genomics, and clinical data, for drug discovery and personalized medicine. The review provides an overview of different RNN architectures and their use in drug recommendation tasks.

The paper [2] presents a comprehensive study on the application of deep convolutional and recurrent neural networks for drug-target interaction prediction. It explores different architectures combining CNNs and RNNs to capture spatial and sequential depend encies in drug-target interaction data. The research showcases the potential of these models in drug recommendation by accurately predicting drug-target interactions.

This study proposes a drug recommendation model that incorporates both temporal information and tag information using RNNs. The model takes into account the temporal order of drug prescription records as well as the semantic information conveyed by

drug tags. The research demonstrates that integrating temporal and tag information into RNN-based models improves the accuracy of drug recommendations.

This work focuses on drug-drug interaction prediction using RNNs. It explores the abil ity of RNNs to capture sequential dependencies in drug interaction data and predict potential interactions between drugs. The research provides insights into the use of RNNs for drug recommendation by identifying potential drug-drug interactions that may influence the effectiveness and safety of drug combinations.

2.2 Cellular Automata

Cellular automata (CA) have gained significant attention as a versatile computational modeling paradigm with a wide range of applications. This survey presents an overview of the diverse and evolving applications of cellular automata [9] in various fields. Starting with an introduction to the fundamental concepts of cellular automata, including their structure, rules, and behavior, the survey explores their applications across multiple domains [3, 10, 11].

In the realm of physics and engineering, cellular automata have been employed to model physical systems, such as fluid dynamics, lattice gases, and magnetism. These applications have provided valuable insights into complex phenomena and the emergence of collective behavior [1].

In the field of computer science and artificial intelligence, cellular automata have found use in image processing, pattern recognition, and cryptography. They have been utilized for tasks such as image filtering, object detection, and encryption algorithms, showcas ing their ability to handle complex spatial and temporal patterns [5, 12].

Cellular automata have also made significant contributions in urban planning, where they have been employed to model urban growth, simulate traffic flow, and optimize land-use patterns. By capturing the dynamics of urban systems, cellular automata offer a powerful tool for decision-making and policy analysis in urban environments [6].

In the realm of biology and bioinformatics, cellular automata have been utilized to simulate biological processes, model ecological systems, and study genetic phenomena.

They enable researchers [7] to explore the emergence of complex behaviors and patterns in biological systems, aiding in understanding natural processes and designing effective interventions [8].

3 Design of RNN-CA-DR

3.1 RNN (Recurrent Neural Network)

The proposed approach involves training an RNN model on a labeled dataset of drug reviews, where each review is associated with a sentiment label (positive, negative, or neutral). The RNN leverages its recurrent nature to process the reviews as sequences, allowing it to capture the temporal dynamics and dependencies between words or phrases in the text. To represent the text data, word embeddings such as Word2Vec is used as shown in Fig. 1.

During the training phase, the RNN learns to understand the sentiment expressed in the reviews and predicts the sentiment of new, unseen reviews. The model's parameters

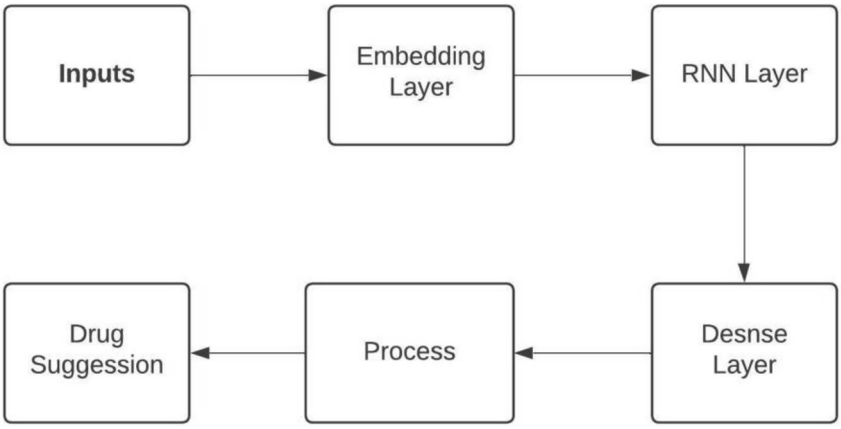

Fig. 1. Design of RNN-CA-DR

are optimized using techniques like backpropagation through time (BPTT) or variants such as Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRUs). These optimization methods enable the RNN to minimize prediction errors and improve its sentiment analysis capabilities.

Once trained, the RNN-based sentiment analysis model can be integrated into a drug recommendation system. Given a new review, the model can classify the sentiment expressed in the text as positive, negative, or neutral. This sentiment information can then be utilized to recommend drugs that align with the user's desired sentiment.

3.2 NCA (Non Linear Cellular Automata)

To apply cellular automata for recommending drugs and analyzing sentiment in reviews, you can follow these steps:

1. Data Collection: Gather a dataset of drug reviews, including the text of the reviews and corresponding sentiment labels (e.g., positive, negative, neutral).
2. Preprocessing: Preprocess the reviews by removing noise, such as special characters, punctuation, and stop words. You may also consider stemming or lemmatization to normalize the words.
3. Sentiment Analysis: Perform sentiment analysis on the reviews using established techniques such as lexicon-based approaches, machine learning models (e.g., Naive Bayes, Support Vector Machines), or deep learning models (e.g., recurrent neural networks, transformers). Assign sentiment labels (e.g., positive, negative, neutral) to each review.
4. Cellular Automata Representation: Represent the sentiment labels of the re- views as the states of the cellular automata. For example, you can map positive sentiment to one state (e.g., "1"), negative sentiment to another state (e.g., "0"), and neutral sentiment to a third state (e.g., "2").
5. Cellular Automata Rules: Define the rules for evolving the states of the cellular automata based on the neighboring cells. These rules can be designed to capture

patterns and dependencies in the sentiment labels. For example, you might consider rules that promote the spreading of positive sentiment or rules that dampen the impact of negative sentiment.
6. Simulation: Run the cellular automata simulation for a certain number of time steps. Each time step represents the evolution of the sentiment labels based on the defined rules and the current state of the neighboring cells.
7. Drug Recommendation: Analyze the final state of the cellular automata and extract information about the sentiment distribution. Based on the sentiment patterns observed, you can recommend drugs that have received positive sentiment feedback and avoid drugs associated with negative sentiment.

We have augmented both RNN output and CA output to propose a robust classifier RNN-CA-DR which trained and tested on Winter 2018 Kaggle University Club Hackathon datasets [13].

4 Result Analysis and Comparisons

Here are some key aspects that contribute to the performance of a drug recommendation system:

1. Data quality and coverage: The system should have access to comprehensive and up-to-date drug information, including indications, contraindications, side effects, interactions, and dosage guidelines. The data should be reliable and regularly updated to reflect the latest research and clinical guidelines.
2. Algorithmic approach: Different recommendation algorithms can be used, such as collaborative filtering, content-based filtering, or hybrid approaches. The chosen algorithm should be able to effectively analyze the input data and generate meaningful recommendations based on patient-specific factors, such as medical history, allergies, current medications, and demographic information.
3. Personalization: The system should take into account individual patient characteristics and preferences to provide tailored recommendations. Factors like age, gender, comorbidities, genetic profile, and lifestyle choices can influence the suitability of a particular drug for a patient.
4. Accuracy and relevance: The recommendations provided by the system should be accurate, relevant, and aligned with the specific needs and condition of the patient. The system should consider the latest clinical guidelines, evidence based medicine, and known drug-drug interactions or contraindications.
5. Evaluation metrics: Performance evaluation is crucial to assess the effectiveness of a drug recommendation system. Metrics such as precision, recall, F1 score, and accuracy can be used to measure the system's ability to provide relevant recommendations and avoid false positives or negatives.

The RNN-CA-DR recommendation system is shown in the Fig. 2 as explained above (Table 1).

RNN-CA-DR reports an accuracy of 91.23 as next promising method is Neural Network(NN) in this parameter. The proposed classifier reports an F1 score of 0.952 and next promising work is reported as Regression model. RNN-CA-DR and NN are the top two

Table 1. Comparison of the performance of RNN-CA-DR with Base Line Methods.

Base line methods	Accuracy	F1 Score	Precision	Recall
RNN-CA-DR	91.23	0.952	0.963	0.961
NN(Neural Network)	87.9	0.885	0.953	0.923
Decision Tree	88.3	0.921	0.902	0.936
Regression	78.3	0.896	0.895	0.802

promising methods in the precision parameters reporting 0.963 and 0.953 respectively. RNN-CA-DR reports an recall value as 0.96 and the next promising technique is Decision Tree.

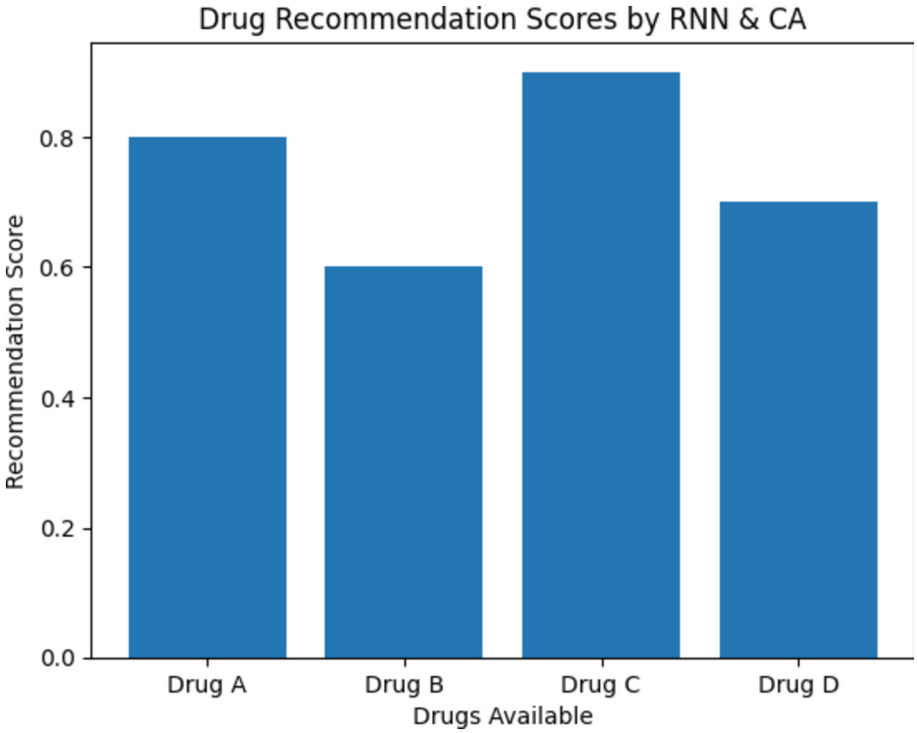

Fig. 2. Sample Drug Recommendation by RNN-CA-DR

5 Conclusion

It is evident that RNNs are widely applied in drug recommendation systems due to their ability to capture sequential dependencies and contextual nuances in drug-related data. Their recurrent nature allows them to learn from sequential patterns and make accurate

predictions or recommendations. On the other hand, while CA offer a different perspec tive by capturing spatial dependencies, they may not be as prevalent in drug recommen dation due to the limitations in capturing long-term dependencies. In conclusion, both RNNs and CA together provide valuable approaches for drug recommendation. RNNs excel in capturing temporal patterns and dependencies within sequential data, making them a popular choice for sentiment analysis or personalized drug recommendation. CA, on the other hand, offer a spatial perspective and can capture spatial interactions but may not be as widely used in drug recommendation due to their limitations in modeling long-term dependencies. We have achieved considerable accuracy in drug recom mendation and this work can be extended to various chronic related recommendations also.

References

1. Pokkuluri, K.S., Nedunuri, S.U.D.: A novel cellular automata classifier for covid-19 prediction. J. Health Sci. **10**(1), 34–38 (2020)
2. Zhang, W., et al.: DRI-RCNN: an approach to deceptive review identification using recurrent convolutional neural network. Inf. Process. Manage. **54**(4), 576–592 (2018)
3. Sree, P.K., Babu, I.R.: Identification of protein coding regions in genomic DNA using unsupervised FMACA based pattern classifier. IJCSNS **8**(1), 305 (2008)
4. Singh, R., Lanchantin, J., Sekhon, A., Qi, Y.: Attend and predict: understanding gene regulation by selective attention on chromatin. Adv. Neural Inf. Process. Syst. **30** (2017)
5. Pokkuluri, K.S., Nedunuri, S.U.D., Devi, U.: Crop disease prediction with Convolution Neural Network (CNN) augmented with cellular automata. Int. Arab J. Inf. Technol. **19**(5), 765–773 (2022)
6. Lo, Y.C., et al.: Reconciling allergy information in the electronic health record after a drug challenge using natural language processing. Front. Allergy **3**, 904923 (2022)
7. Pokkuluri, K.S., Usha, D.N.: A secure cellular automata integrated deep learning mechanism for health informatics. Int. Arab J. Inf. Technol. **18**(6), 782–788 (2021)
8. Wang, Y., Teng, G., Ding, X., Zhang, G., Ling, Y., Wang, G.: A novel method of Chinese electronic medical records entity labeling based on BIC model. J. Softw. **16**(1), 24–38 (2021)
9. Sree, P.K.: Exploring a novel approach for providing software security using soft computing systems. Int. J. Secur. Appl. **2**(2), 51–58 (2008)
10. Su, R., Huang, Y., Zhang, D.G., Xiao, G., Wei, L.: SRDFM: siamese response deep factorization machine to improve anti-cancer drug recommendation. Brief. Bi oinform. **23**(2), bbab534 (2022)
11. Sree, P.K., Babu, I.R., Murty, J.V.R., Ramachandran, R., Devi, N.U.: Power - aware Hybrid Intrusion Detection System (PHIDS) using cellular automata in wireless ad hoc networks. WSEAS Trans. Comput. **7**(11), 1848–1874 (2008)
12. Sree, P.K., Devi, N.U.: Achieving efficient file compression with linear cellular automata pattern classifier. Int. J. Hybrid Inform. Technol. **6**(2), 15–26 (2013)
13. https://www.kaggle.com/datasets/jessicali9530/kuc-hackathon-winter-2018