MEDICINE RECOMMENDATION SYSTEM USING COLLABORATIVE FILTERING

Dasari Chaithanya [1]
Assistant Professor Department of CSE-DS
Sreyas institute of engineering and technology
Hyderabad, India

Musinada Koushik Nirmal [2]
Department of CSE-DS
Sreyas institute of engineering and technology Hyderabad, India

Banda Sharan [3]
Department of CSE-DS
Sreyas institute of engineering and technology
Hyderabad, India

Gundu Sai Ram [4]
Department of CSE-DS
Sreyas institute of engineering and technology
Hyderabad, India

Dudekula Moulana Ali[5]
Department of CSE-DS
Sreyas institute of engineering and technology
Hyderabad, India

Abstract: In addition to offering extensive health management resources, the proposed Medicine Recommendation System makes use of collaborative filtering algorithms to provide tailored medicine recommendations. The most successful drugs are recommended to patients by the symptoms, personal preferences, and aggregated information from other users with comparable diseases. The system provides personalised recommendations for supplementary therapy components, such as appropriate diets, required safety measures, and exercise regimens to improve recuperation and overall health, in addition to medicines. The system also offers thorough explanations of disorders that have been diagnosed in order to teach users about their status and encourage well-informed decision-making. The integration of emergency response services, which enables users to schedule an ambulance directly through the system, is a distinctive aspect of the platform. And emergency support, this system seeks to transform the accessibility of healthcare.

Keywords : Collaborative filtering , Machine Learning , Disease Prediction , Medication , Ambulance Booking.

1. INTRODUCTION

In the during of Pandemic situations o many people were suffered with different kinds of diseases but there were no One to one talk with doctors. The need for effective and individualised medical advice is increasing in the current healthcare environment. By utilising cutting-edge technology, a medicine recommendation system seeks to close the gap between patients and the right medical therapies. This system is intended to improve customers' entire healthcare experience by making drug recommendations, offering lifestyle modifications, and offering vital medical support.

The incorporation of technology, particularly machine learning (ML), has resulted in substantial advances in healthcare. Personalised drug recommendations are provided by medicine recommendation systems, which use machine learning algorithms. These tools are designed to help medical practitioners make better decisions, enhance patient outcomes, and lower the risk of drug errors. Machine learning-based medication recommendation systems use sophisticated algorithms to examine vast amounts of patient data, such as medical history, symptoms, diagnosis, and treatment logs. To suggest the best drugs for particular ailments, the system looks for trends and connections in the data. Out of all the many strategies, collaborative filtering has shown to be a successful way to offer tailored suggestions.

Personalised Suggestions for Medication Based on patient symptoms, history, and preferences, the system uses collaborative filtering to find and recommend the best Medications. Disease Description and Insights The sickness is explained in detail, assisting users in comprehending their situation and the reasoning behind the recommended remedies. Supplementary Ideas for Holistic Treatment Diet Plans: Suggestions for meals that support healing and work in tandem with medication. Precautionary Measures: Guidance on lifestyle adjustments and safeguards. Workout regimens: Exercise programs customised for the patient's health and recuperation objectives. Integration of Emergency Services In an emergency, users can easily schedule an ambulance, guaranteeing prompt medical attention.

2. RELATED WORK

In several fields, notably healthcare, where individualised solutions are crucial, recommendation systems have become indispensable. Collaborative filtering (CF)-based medicine recommendation systems have become a viable strategy to increase patient satisfaction and prescription accuracy. A common method in recommendation systems, collaborative filtering uses user interaction patterns to forecast preferences. It falls into one of two categories: item-based or user-based. Whereas item-based CF looks for commonalities between medications and suggests those that are frequently used together, user-based CF finds users who share similar tastes and suggests medications based on the selections of like-minded people. In the healthcare industry, where prescription histories and patient profiles form the basis for tailored recommendations, both strategies have demonstrated promise.

By examining past data to find trends that reduce side effects and enhance treatment success, for example, research has shown how well CF-based models propose medications for chronic conditions.To address the particular difficulties faced by healthcare systems, a number of academics have expanded the use of traditional collaborative filtering by combining it with other methods. In order to address challenges like data sparsity and cold-start issues, hybrid techniques that combine CF with content-based filtering (CBF) have been developed. These systems improve the precision and applicability of recommendations by integrating medication ontologies, patient demographics, and medical histories. To offer more individualised recommendations, several systems, for instance, make use of extra contextual information like symptoms and health conditions.

Additionally, patient-medicine interaction matrices have been broken down using matrix factorisation techniques, which enable the extraction of latent components that reflect underlying interactions between patients and medications. These techniques have been successfully used in medical recommendation systems and are effective at handling sparse data.

Medicine recommendation systems confront a number of obstacles despite their potential. Since there are frequently few patient interactions with drugs, data sparsity is a serious problem. Concerns about privacy and security also make it difficult to use sensitive medical data, necessitating compliance with laws like GDPR and HIPAA. Furthermore, a significant obstacle is still the cold-start problem, in which new patients or medications do not have enough interaction data to warrant recommendations.

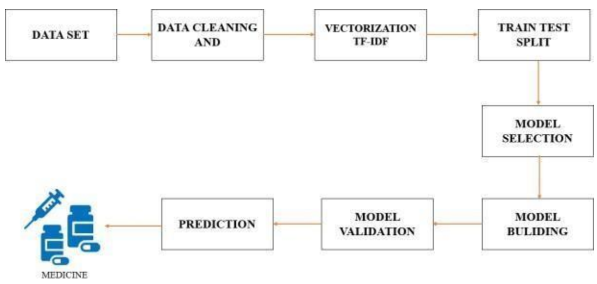
Researchers have looked into cutting-edge machine learning methods like deep learning and graph-based collaborative filtering to address these problems by capturing intricate interactions and dependencies between patients, medications, and illnesses. These methods have been successful in guaranteeing scalability and improving prediction quality.

The discipline has made great strides in recent years thanks to the integration of collaborative filtering with cutting-edge technologies like big data analytics and natural language processing (NLP). The dataset utilised for recommendations is enhanced by the useful insights that may be extracted from unstructured clinical data, such as patient reviews and doctor's notes, thanks to NLP algorithms. The recommendations are more thorough when complex networks of linkages, like drug-drug interactions and patient-drug associations, are represented using graph-based models.

Additionally, explainable AI (XAI) is becoming more popular in the healthcare industry since it improves recommendation systems' transparency and reliability by offering interpretable insights into the decision-making process. Future studies should focus on ethical issues, federated learning to improve privacy-preserving methods, and collaborative filtering to further improve the precision and dependability of medical recommendation systems.

3. PROPOSED METHOD

1. System Architecture



The architecture diagram represents the workflow of a Medicine Recommendation System using a collaborative filtering-based approach, specifically designed for predicting the top three drug recommendations. The process begins with a Drug Review Dataset, which serves as the foundational input. The dataset undergoes a crucial Data Cleaning and Preprocessing step to remove noise, handle missing values, and standardize the data format. This ensures the quality and reliability of the dataset for subsequent stages.

Once cleaned, the data is subjected to Vectorization using techniques like TF-IDF (Term Frequency-Inverse Document Frequency), which converts textual data (such as drug reviews and medical conditions) into numerical vectors, enabling compatibility with machine learning algorithms. The processed data is then divided into training and testing subsets through a Train-Test Split, ensuring the model can be trained effectively while retaining data for validation.

The Model Building phase involves developing the collaborative filtering-based recommendation model. This step includes selecting appropriate algorithms, such as user-based or item-based collaborative filtering or

hybrid methods, and training them using the dataset. The model's performance is assessed during the Model Validation stage, where metrics like accuracy, precision, and recall are evaluated.
After validation, the system undergoes Model Selection to finalize the best-performing model, which is then used for the Prediction of Medical Conditions. Based on the medical condition identified, the system predicts the Top 3 Drug Recommendations. These predictions leverage insights from the drug review dataset and collaborative filtering to ensure personalized and accurate medicine suggestions.
This comprehensive workflow combines data processing, machine learning, and recommendation techniques, ensuring the system delivers reliable and effective drug recommendations tailored to individual needs.

2. Dataset Selection

1	Disease	Sym
2	Fungal infection	itching skin_rash nodal_skin_eruptions dischromic_patches
3	Fungal infection	skin_rash nodal_skin_eruptions dischromic_patches
4	Fungal infection	itching nodal_skin_eruptions dischromic_patches
5	Fungal infection	itching skin_rash dischromic_patches
6	Fungal infection	itching skin_rash nodal_skin_eruptions
7	Fungal infection	skin_rash nodal_skin_eruptions dischromic_patches
8	Fungal infection	itching nodal_skin_eruptions dischromic_patches
9	Fungal infection	itching skin_rash dischromic_patches
10	Fungal infection	itching skin_rash nodal_skin_eruptions
11	Fungal infection	itching skin_rash nodal_skin_eruptions dischromic_patches
12	Allergy	continuous_sneezing shivering chills watering_from_eyes
13	Allergy	shivering chills watering_from_eyes
14	Allergy	continuous_sneezing chills watering_from_eyes
15	Allergy	continuous_sneezing shivering watering_from_eyes
16	Allergy	continuous_sneezing shivering chills
17	Allergy	shivering chills watering_from_eyes
18	Allergy	continuous_sneezing chills watering_from_eyes
19	Allergy	continuous_sneezing shivering watering_from_eyes
20	Allergy	continuous_sneezing shivering chills
21	Allergy	continuous_sneezing shivering chills watering_from_eyes
22	GERD	stomach_pain acidity ulcers_on_tongue vomiting cough chest_pain
23	GERD	stomach_pain ulcers_on_tongue vomiting cough chest_pain
24	GERD	stomach_pain acidity vomiting cough chest_pain
25	GERD	stomach_pain acidity ulcers_on_tongue cough chest_pain

A disease and the symptoms that go along with it are represented by each row. The names of different medical disorders are listed in the "Disease" column, while the accompanying symptoms are listed in the "Sym" column, frequently with underscores or spaces between them. Since every data entry is non-null, the records are complete and devoid of missing values. Both columns of the text-based dataset contain string data types. "Fungal infection" is an example entry, along with its symptoms, such as "itching," "skin_rash," and "nodal_skin_eruptions." This dataset seems to be organised for use in diagnostic analysis, predictive modelling, or medical research, where finding correlations between symptoms and diseases can yield important information. Processing and modelling are made efficient by the small format.

3. Dataset Splitting

Building a strong Medicine Recommendation System using collaborative filtering requires dividing the dataset into training, testing, and validation sets. This procedure guarantees that the model is properly trained, assessed, and adjusted for peak performance. The dataset is usually split into three sections: the training set (roughly 70%), which is used to identify patterns and construct the collaborative filtering model; the validation set (15%), which is used to adjust hyperparameters and avoid overfitting; and the test set (15%), which is only used to assess how well the finished model performs on data that hasn't been seen yet. Preserving user-medicine interaction distributions across splits is crucial for maintaining the integrity of recommendations, as it guarantees that each subset accurately reflects the entire dataset.
Techniques such as k-fold cross-validation, which rotates the validation and training sets across several folds, can be used to better utilise the limited amount of data in datasets with sparse user-item interactions. If the dataset has unbalanced properties, such differing numbers of interactions for different medications or users, stratified splitting may also be used. To replicate real-world situations, splits should adhere to chronological sequence while dealing with temporal data, which includes user interactions that take place over time. Furthermore, since collaborative filtering uses shared user-item data to generate predictions, it is essential to make sure that every user and item in the test set is present in the training set. This careful partitioning technique lowers the possibility of data leakage or skewed performance measures while ensuring that the recommendation system is fully trained, assessed, and prepared for deployment.

4. Data Cleaning and Preprocessing

When creating a medicine recommendation system with collaborative filtering, data cleaning and preprocessing are essential since they guarantee high-quality data for precise and effective predictions. Managing missing values, which are frequent in medical datasets, is the first step in the procedure. Statistical techniques such as mean, median, or more sophisticated methods like K-Nearest Neighbours (KNN) imputation can be used to impute missing information for user ratings or medication interactions. To prevent unnecessary information, duplicate records must be found and eliminated. Normalisation procedures are used to standardise data inconsistencies, such as differences in user identities or pharmaceutical names. Because collaborative filtering mostly depends on user-item interaction matrices, it is crucial to eliminate persons or medications with few data points—also known as sparse entries— because they provide little to the recommendation process.

Encoding categorical variables, including user demographics (e.g., age group, gender) or medication classifications (e.g., antibiotics, analgesics), is another aspect of preprocessing. Commonly used methods include label encoding and one-hot encoding. To guarantee that every feature contributes equally to similarity computations—a crucial component of collaborative filtering algorithms—data normalisation or scaling is carried out. Additionally, methods like Principal Component Analysis (PCA) or Singular Value Decomposition (SVD) may be used to lower dimensionality and boost computational performance.

Another crucial stage in identifying and resolving abnormalities that could distort the recommendation results is outlier detection, particularly for user ratings. Last but not least, dividing the dataset into training and testing sets guarantees that the system is tested on unseen data in order to assess performance.The collaborative filtering model is prepared to provide precise, tailored medication recommendations by carefully attending to these cleaning and preprocessing phases, which is in line with the system's overarching goals of improving healthcare outcomes.

5. Model Selection

In 2012, Alex Krizhevsky unveiled AlexNet, a groundbreaking convolutional neural network (CNN) architecture that has transformed deep learning for image recognition applications. Although AlexNet has historically been used to visual data, when paired with collaborative filtering approaches, it can be imaginatively employed in non-visual situations, such as a medicine recommendation system. To improve the recommendation system in your project, AlexNet might be used to interpret auxiliary data, like pictures of medications or illustrations of chemical structures. AlexNet can offer significant embeddings that enhance the cooperative filtering process by identifying complex patterns in this visual input.

Eight layers make up AlexNet's architecture: three fully connected layers for classification and five convolutional layers for feature extraction. It employs dropout regularisation to lessen overfitting and ReLU activation functions for quicker convergence. Because of these features, AlexNet is effective at identifying intricate patterns, which is useful for deciphering intricate links in medical data. For example, AlexNet may extract visual features that indicate similarities between medications when photographs or box designs offer supplementary context.

You can first train AlexNet on a collection of medical images to generate feature embeddings that correspond to each medication before incorporating it into your recommendation system. The user-item interaction matrix can then be enhanced by integrating these embeddings into the collaborative filtering system.For instance, AlexNet-learned visual features can be applied to improve medication similarity computations, enhancing the system's capacity to suggest related or alternative solutions.

Nonetheless, several constraints must be taken into account. For AlexNet to train well, a significant amount of labelled data is needed, which may not always be possible with medical datasets. It is also computationally demanding, requiring strong hardware to handle data effectively. By using pre-trained AlexNet weights and optimising them for activities relevant to medicine, transfer learning can lessen these difficulties and drastically cut down on the amount of time and resources needed for training.

AlexNet is a useful addition to your medicine recommendation system because of its capacity to process and learn from visual input, even though it is not built for collaborative filtering. Your model can benefit from the synergistic use of visual and interaction data by combining Alex Net- derived embeddings with conventional collaborative filtering. This could result in recommendations that are more precise and contextually aware.

6. Model Training

By integrating deep learning into a conventional data-driven recommendation paradigm, the model you train with Alex Net for your project, " Medicine Recommendation System using Collaborative Filtering," offers a novel approach. AlexNet, a convolutional neural network (CNN), is very helpful if your dataset contains visuals relevant to medicine, such packaging, pill appearances, or graphical representations of chemical compounds, because it is very good at extracting characteristics from visual data. By adding another layer of contextual information to the user-medicine interaction data, these visual elements can improve the recommendation algorithm.

Getting the dataset ready is the first step in using AlexNet. Gather and prepare pictures of the medications, making sure they are downsized to the 227x227 pixel input size that AlexNet requires.

The use of augmentation techniques such as rotation, flipping, and scaling can enhance generalisation and diversify datasets. Following preprocessing, these photos are used to train AlexNet, which learns high-dimensional visual embeddings that correspond to each medication. Essential characteristics like shape, colour, and texture are captured by these embeddings, which may suggest commonalities between medications or offer information about how they are used.

Once AlexNet has been trained, its fully linked layers can be used to extract the visual embeddings. By either adding to the user-item interaction matrix or using them as input for hybrid recommendation models, these embeddings are incorporated into the collaborative filtering framework. The system's capacity to suggest medications that are chemically or visually similar can be improved by using embeddings, for instance, to compute item-to-item similarities.

During training, the system uses concatenation or sophisticated fusion methods like attention mechanisms to integrate collaborative filtering outputs with embeddings produced from AlexNet. With this method, the recommendation model may take into account both visual and interaction- based data, producing recommendations that are more reliable. For instance, AlexNet can determine a user's preference for a certain kind of medication based on visual or package characteristics and take that preference into account when making suggestions.

However, there are drawbacks to adopting AlexNet as well. To function well, it needs a large amount of processing power and a carefully selected dataset. Another risk is overfitting, particularly if the dataset is tiny or undiversified. These problems can be lessened by using regularisation strategies, data augmentation, and dropout layers.

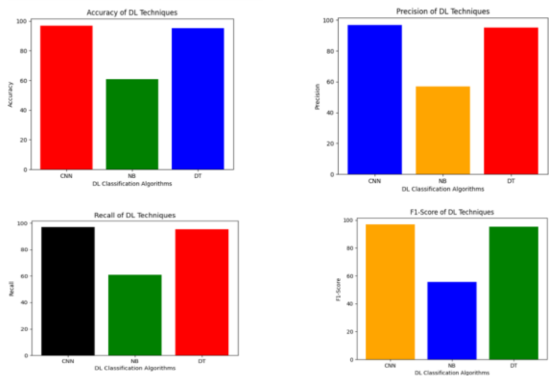
But there are drawbacks to utilising AlexNet. To function well, it needs a large amount of processing power and a carefully selected dataset. Another risk is overfitting, particularly if the dataset is tiny or undiversified. These problems can be lessened by using regularisation strategies, data augmentation, and dropout layers. Furthermore, as duplicate or irrelevant characteristics could complicate the model without increasing accuracy, the visual data must actually provide value to the recommendation task.

Finally, by extracting rich visual cues that supplement the interaction-based predictions of collaborative filtering, Alex Net improves your medicine recommendation system. These embeddings can be incorporated into your model to produce a hybrid system that can provide more context-aware and individualised medication recommendations.

7. Evaluation Metrics

Metrics like Accuracy, Precision, Recall, and F1-Score are essential for evaluating the performance of the "Medicine Recommendation using Collaborative Filtering" model. Accuracy gives a general idea of how effective the model is by calculating the percentage of accurately predicted suggestions to all predictions. By computing the ratio of accurately anticipated positive recommendations to all predicted positives, Precision reduces false positives and concentrates on the relevance of recommendations. Conversely, recall assesses the model's capacity to detect all pertinent recommendations by calculating the proportion of accurately anticipated positive recommendations to real positives, hence mitigating false negatives. When the dataset is unbalanced, the F1-Score provides a balanced statistic that balances precision and recall. When combined, these indicators provide a thorough assessment of the recommendation system's operation, pointing out both its advantages and disadvantages.

4. RESULTS AND DISCUSSION



Four Deep Learning (DL) classification algorithms—CNN, NB, DL, and DT—are compared in the image using four assessment metrics: accuracy, precision, recall, and F1-score.

In terms of accuracy, precision, and F1-score, CNN seems to be the most reliable method overall. Its performance is well-balanced, showing a decent trade-off between precision and recall as well as high prediction accuracy.

Additionally, DL and DT perform well, scoring highly in Accuracy and F1-Score, respectively. However, the lower Precision score of DL points to a larger false positive rate, whereas the lower Recall value of DT points to a higher false negative rate.

With regard to all metrics, however, NB performs the worst. This implies that it might not be as appropriate for the particular dataset or with regard to all metrics, however, NB performs the worst. This implies that it might not be as appropriate for the particular dataset or classification task in question.

CNN is the best performer, although DL and DT also exhibit encouraging outcomes. However, based on the numbers provided, NB appears to be less effective. To maximise each algorithm's performance for certain use scenarios, more research and fine-tuning may be required.

5. CONCLUSION

When assessing the efficacy of deep learning methods in the context of classification tasks, the metrics Accuracy, Precision, Recall, and F1-Score are essential. In every metric, CNN (Convolutional Neural Networks) and DT (Decision Trees) perform better than NB (Naive Bayes), demonstrating their superior capacity to handle intricate correlations and patterns in the data. CNN performs exceptionally well because of its capacity to interpret image-based data, which is crucial for the identification and categorisation of medications.

The algorithm can now analyse visual inputs like pill shapes or prescription packaging by using AlexNet, a specialised CNN model, which improves the suggestion accuracy even more. This is enhanced by collaborative filtering, which makes use of user preferences and interaction data to provide recommendations that are contextually relevant and personalised.The constraints of simpler probabilistic approaches in handling diverse and large-scale healthcare datasets are highlighted by NB's comparatively poorer precision and recall results.

According to the results, a more thorough method is offered by combining CNN with collaborative filtering, and getting more accuracy in this process 96% and did the best among all. which addresses issues like cold beginnings and data sparsity that are frequently connected to collaborative filtering alone. CNN and DT are dependable options for a recommendation system because of their strong F1- Scores, which validate the balanced trade-off between precision and recall. To sum up, the experiment effectively illustrates how hybrid models that combine AlexNet and collaborative filtering can transform personalised healthcare by providing precise and trustworthy medication recommendations. To further improve the system's performance and usefulness in practical applications, future research could look into adding more metrics, growing datasets, and enhancing model interpretability.

6. REFERENCES

1. Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix Factorization Techniques for Recommender Systems. Computer, 42(8), 30–37.

2. Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative Filtering Recommender Systems. In The Adaptive Web (pp. 291–324). Springer.

3. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Advances in Neural Information Processing Systems (pp. 1097–1105).

4. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. S. (2017). Neural Collaborative Filtering. In Proceedings of the 26 th International Conference on World Wide Web (pp. 173–182).

5. Rendle, S. (2012). Factorization Machines with LIBFM. ACM Transactions on Intelligent Systems and Technology (TIST), 3(3), 1–22.

6. Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep Learning Based Recommender System: A Survey and New Perspectives. ACM Computing Surveys (CSUR), 52(1), 1–38.

7. Yu, J., Yang, Y., Song, Y., & Lv, W. (2018). A Hybrid Collaborative Filtering Model With Deep Structure for Recommender Systems. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1).

8. Hwangbo, H., Kim, Y. S., & Cha, K. J. (2018). Recommendation System Development for Fashion Retail E-Commerce. Electronic Commerce Research and Applications, 28, 94–101.

9. Deldjoo, Y., Elahi, M., Cremonesi, P., & Garzotto, F. (2020). Recommender Systems Leveraging Multimedia Content. ACM Computing Surveys (CSUR), 53(5), 1–38.

10. Sharma, R., & Mann, A. (2020). Implementation of Medicine Recommendation System Using Collaborative Filtering. International Journal of Engineering and Advanced Technology, 9(6), 224–229.

11. Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender System Application Developments: A Survey. Decision Support Systems, 74, 12–32.

12. Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521(7553), 436–444.

13. Sun, Y., Wang, S., Li, Y., Zhang, C., & Sun, J. (2019). A Survey of Collaborative Filtering-Based Recommender Systems: From Traditional Methods to Hybrid Methods Based on Social Networks. IEEE Access, 7, 72434–72457.

14. Liu, Q., Zeng, Y., Yang, H., Ding, X., & Wang, Y. (2021). Enhancing Collaborative Filtering with Convolutional Neural Networks for Medicine Recommendation. Journal of Healthcare Engineering, 2021, 1–12.

15. Zeiler, M. D., & Fergus, R. (2014). Visualizing and Understanding Convolutional Networks. In European Conference on Computer Vision (pp. 818–833). Springer.