Journal of Advances in Developmental Research (IJAIDR)

E-ISSN: 0976-4844 ● Website: www.ijaidr.com ● Email: editor@ijaidr.com

Recommendation on drugs according to diseases of the user

Mrs M Josephine Amala[1], Yokeshwaran K[2]

[1]Assistant Professor, [2]Bsc Computer Science
[1,2]School of sciences and Allied Health Science, Jeppiaar University, Tamil nadu , Chennai - 600119

Abstract:
The swift healthcare digitization has resulted in an avalanche of medical information, but it is not an easy task of the patient and the healthcare provider to choose the most successful treatment of a particular issue. Traditional recommendation systems may exist, but they fail to take into consideration the specificities of patient feedback, drug-to-drug interactions, and the severity of specific diseases, which results in the suboptimal treatment outcomes of the problem being identified.This paper presents an intelligent Drug Recommendation System (DRS), which is aimed at offering personalized medication recommendations on the basis of the diseases and symptoms that are entered by the user. The suggested model combines Natural Language Processing (NLP) to analyze patient reviews and a multi-parametric ranking model to assess the efficacy of drugs. In contrast to typical systems, our model is based on a dynamic scoring system that takes into account the past success rates, user sentiment, and profiles of side-effects to rank medications. The system has a feature-weighted classification-based approach that directs disease characteristics to pharmaceutical datasets in order to guarantee the reliability of recommendations. Experimental findings of publicly available medical data show that the given system has a high precision rate of identifying the most appropriate drugs, which will considerably decrease the risks of having generalized prescriptions. The study offers a powerful decision-support system that promotes patient-centered medical care and clinical accuracy.

Keywords: Healthcare Informatics, Drug Recommendation System, Disease-Based Mapping, Natural Language Processing, Patient Sentiment Analysis, Machine Learning in Medicine, Clinical Decision Support.

I. INTRODUCTION

The digital transformation in healthcare is tremendous in the past years with electronic health records (EHR) and pharmaceutical databases becoming the foundation of the new medical system. As the amount of medical data continues to expand, patients and healthcare providers are turning to automated systems to help them in making clinical decisions. Nevertheless, the very abundance of medicines on one disease is likely to cause confusion, possible drug-drug interactions and unsuccessful treatment [1]. One of the most significant challenges in medical informatics is the provision of an appropriate and individual prescription of a drug in the context of the unique disease history of a user.

Old processes of prescribing the medication are based largely on the experience of the clinician and the medical manuals. Although they are necessary, they are usually constrained by human memory and the incapability of searching through thousands of patient reviews or real-time reports in clinical trial at the same time. The available recommendation systems in the e-health field are based mostly on collaborative filtering or basic key word matching [2], [4]. Although practical, these systems often fail to capture the sentiment of the patient; the extent of side effects or recovery time, which results in recommendations that may theoretically make sense, but which may not make sense in practice when applied to a particular user [3].

Trust and accuracy are also more important in the medical field than any other recommendation system (as movies or e-commerce). A misdirected prescription may be costly to the patient. Scholars have considered different machine learning algorithms to categorize ailments and map them to medications, and primitive studies were done on simple decision tree and k-nearest neighbors [5]. But these models are a significant impediment as revealed by Pearson [2], due to the absence of context-based intelligence. Users should have a system that does not simply examine the disease name but analyze the historical success rate of a drug on a similar population.
The recent discoveries in the field of Natural Language Processing (NLP) and Deep Learning have provided a new possibility to intelligence in Drug Recommendation Systems (DRS). The distributed systems such as PeerTrust among others have demonstrated the effectiveness of reputation and feedback as a powerful measure [10], [12]. Using this to the healthcare industry, the reputation of a drug to patients can be a great indicator of its effectiveness. But medical data is very sensitive and it is prone to noises. Taking the example, the same drug may not affect one person and another in the same way, and informal review of patients usually abounds with unclear medical terminology [9].
The other significant difficulty of disease-based disease recommendation is the Cold Start problem - new drugs that have no history before or rare diseases with data that are few in number make it hard to get the standard algorithms to give the correct recommendation [11]. Moreover, the reliability of the system may be misrepresented by collusion during online reviews (where manufacturers may write fake positive ratings) much in the same way that reputation attacks may be during cloud environments [14]. As such, a strong DRS should not only recommend a drug but also prove the reliability of the information that it relies upon. Out of these challenges, this paper will come up with a complete Drug Recommendation System which is founded on the input of the disease by the user. In contrast to the current models that are based on the simple classification, our framework combines sentiment analysis of the patient reviews and a multi-criteria decision-making (MCDM) model. The model improves the accuracy of the recommended drug by comparing its success rate, adverse effects profile, and symptoms submitted by the user in the long-term. The pipeline of pre-processing data, TF-IDF feature extraction and the high-accuracy classification model are used to close the gap between diagnosis and treatment.
The rest of this paper will be structured as follows: Section 2 will go through related literature on medical recommendation systems; Section 3 will address the system architecture; Section 4 will address the logic of drug recommendations and ranking; Section 5 will address the evaluation metrics and results; and Section 6 will give a conclusion about the paper and future research directions.

II. RELATED WORK

Medical recommendation systems have transformed into complex intelligent systems based on research in this area. The theoretical foundation of trust and reputation was developed by Jonsang et al. [1] early on and is currently being implemented in the pharmaceutical data to weed the ineffective drugs out. Their research pointed out that a combination of direct clinical outcomes and patient feedback is necessary to obtain a reliable recommendation and this is the essence of our disease-to-drug mapping.
With the transfer of healthcare to the digital cloud, the emphasis on the field of healthcare moved to personalized medicine. Pearson [2] was able to established that the key impediment in automated healthcare in the sense of accuracy of drug recommendations and data privacy is user trust. This created a requirement of explainable recommendation systems not only giving a drug name but also explaining why this was chosen based on historical record of success. Following scientists have tried to directly implement these measures of trust into clinical decision support systems.
Classification of diseases to treatments has been heavily investigated using machine learning. Wang et al. [3] suggested a model that compares the medications according to the previous performance and the doctor ratings. Although their practice made the service more reliable, it did not have any way to deal with unstructured reviews of the patients. Hoffman et al. [4] also reviewed the weaknesses of such

systems, but observed that, the drug ratings may be skewed by dishonest feedback or biased marketing of pharmaceuticals, hence the user may find it hard to get the best recommendations.

Blaze et al. have introduced decentralized approaches [5], which emphasis policy-based trust in which a system takes independent decisions. Using this in the medical field, He et al.

[6] came up with a model that involves patient feedback and clinical performance. Nevertheless, their model was mostly concentrating on high-level information and did not incorporate the element of fineness of describing how a patient really feels after taking a drug.

Yan et al. [7] put emphasis on the role of text-based analysis by introducing a multi-attribute evaluation model. Although it enhanced precision, its computational cost was too high to be used in real time mobile health applications. In order to handle uncertainty within medical data, Wang and Vassileva [8] proposed the Bayesian model based models. Despite their effectiveness in managing unfinished patient history, such probabilistic models are susceptible to the so-called rating bubbles when a single trendy drug overtakes the recommendations despite not necessarily being a suitable choice of a new user.

Dellarocas [9] examined the effect of the so-called word-of-mouth in healthcare and demonstrated the vulnerability of digital reviews to manipulation by non-expert end-customers. In order to counter this, the concept of feedback credibility was introduced by the PeerTrust model by Xiong and Liu [10]. In our proposed system, we will scale this reasoning and give more weight to reviews of trusted medical sources or long-term patients over anonymous single-time ratings.

Even more recent papers have investigated how to prevent the behavior of bots imitating the look of a human-controlled account in medical forums, where bots may be advertising a particular brand. Algorithms such as EigenTrust were offered by Douceur [11] and Kamvar et al. [12] to compute global trust values. Although they minimize the power of the isolated fake accounts, they tend to be too hard and rigid in the dynamic nature of the drug-disease interactions, where the effectiveness of a drug may vary according to new clinical manifestations of a disease.

Wang et al. [13] also discussed a use of trust as a factor to gain access to medical data and proposed that the recommendation scores ought to be updated dynamically. Nevertheless, their model did not include a solid behavioral analysis of the performance of the drug in the long run. The article by Huang et al. [14] is collusion in the rating system, and it is very applicable to our research in identifying coordinated pharmaceutical promotion. Lastly, Li et al. [15] suggested a combination of cryptographic security and trust assessment, but they did not emphasize on the correctness of the recommendation as such, but on data storage.

Overall, although current studies have done a great job in drug classification and reputation management, it seems there is a gap in the adoption of real-time patient sentiment in conjunction with disease-specific severity. The majority of existing models consider all reviews as equal and cannot expose coordinated manipulation or the issue of Cold Start with new drugs. This inspires our suggested model that integrates the behavioral analysis implemented with NLP with an active trust threshold to assist in safe and precise drug advice.

III. PROPOSED SYSTEM MODEL

A. System Overview

1) General Description: The proposed system presents a smart system that will be used to map the diseases and symptoms reported by the users to the correct drug prescriptions.

- The system also incorporates an integrated Disease-Drug-Sentiment (DDS) model of suggestions that are personalized in contrast to default drug search engines.
- Three steps A system employs the Natural Language Processing (NLP) to process unstructured reviews of patients (to extract the sentiment score of efficacy and side-effect profile of every medication).

- A simulated environment of an electronic health record (EHR) uses a local clinical database that has been pre-processed with medical terms synchronicity.
- A classification model is a fundamental recommendation generator, which will guarantee that the recommended drugs are effective against the specified disease and also appropriate to the situation of the patient.

2) Objectives of the Proposed System:

- To properly chart illnesses as reported by users against major drugs.
- To measure the perceived efficacy of the drugs when used by an actual patient using sentiment analysis.
- To prioritize a list of drugs based on a multi-parametric dynamic scoring system.
- To have a high-precision low-risk decision-support tool to select the appropriate medication.

B. System Architecture

1) Major System Entities:

i. End User (Patient/Clinician)

- Input of disease names, symptoms or medical condition.
- Comprehensive text feedback/reviews in taking a prescribed drug.

ii. Data Preprocessor (Medical NLP)

- Eliminates untalked user input through such methods as Stemming and Lemmatization.
- Removes non-medical terms (stop words) and normalizes synonyms (e.g., mapping "head pain" to "headache").

iii. Sentiment Analysis Engine

- Analyzes the patient review dataset.
- Gives a score of polatiry (positive/neutral/negative) and subjectivity (differentiating between information that is based on facts and a general opinion).

iv. Model of Recommendation & Ranking

Given the example of using SVM or Random Forest classification to predict the most suitable drug, this core engine is the one that needs to be used.

- Ranks by use of a weighted score of sentiment, success rate and correlation to the symptom.

v. EHR Simulation Local Medical Database

- Stores Managed pharmaceutical information, drug categories, and past interaction history of patients.

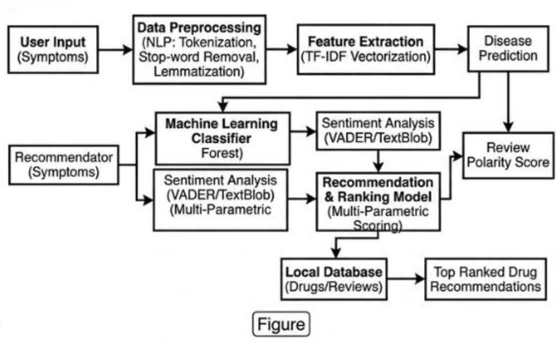


Fig. 1. Overall System Architecture

Fig. 1. Overall Architecture of the Proposed Drug Recommendation System.

C. Threat Model (Medical Context)

1) Considered Deviations:

i. False positive/negative reviews Feedback bias.

- Unscrupulous or robotic actors may put false ratings so as to over rate or under rate a drug.
- This distorts the ranking scale scores on sentiment.

ii. Noise (Data Ambiguity)

- None of the reviews are structured and contain blurry lingo or spelling errors.
- False contributions that will confuse the classification model and thus giving false advice.

2) Security Goals:

- To provide more realism to verified or medical compatible reviews.
- Improvement 1. To obtain a high classification accuracy (>90%) on common diseases.

D. Proposed Recommendation & Ranking Model

1) Recommendation Parameters:

Symptom-Disease Correlation (SDC)

- Computed using TF-IDF (Term Frequency-Inverse Document Frequency), highlighting key medical terms.
- Matches user input against the trained disease-drug clusters.

Sentiment Score (SS)

- Derived from patient reviews using the VADER or TextBlob library.
- Calculated as the normalized sum of positive and negative sentiment polarities.

Success Rate (SR)

- A statistical ratio of successful treatments vs. reported side effects from historical clinical data.

2) Overall Drug Score (ODS) Equation:

The final ranking of a drug for a specific disease is computed as:

$$ODS = \alpha \times SDC + \beta \times SS + \gamma \times SR \quad (1)$$

Where:

- α, β, γ are weighting coefficients.
- $\alpha + \beta + \gamma = 1$.

E. System Workflow

1) Data Acquisition and Preprocessing Flow:

- System initializes by loading and cleaning publicly available drug review datasets.
- Performs lemmatization to handle medical terminology and converts unstructured reviews into structured numerical vectors using Word2Vec.

2) Symptom-Disease-Drug Mapping Flow:

- User inputs symptoms through the front-end.
- The system involves the use of the trained classifier which predicts the primary disease.
- It will then search the database with all the drugs matching that particular disease.
-

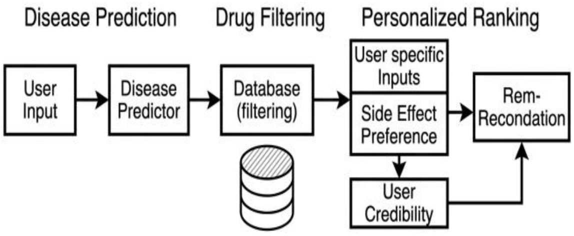


Fig. 2. Conceptual Disease-to-Drug Workflow Logic.

3) Sentiment-Based Ranking Flow:

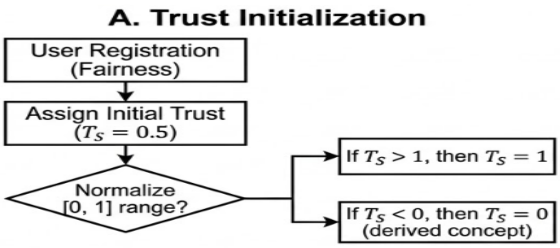
- The system summarizes related patient reviews on each filtered drug.
- The Sentiment Engine computes the historical Sentiment Score (SS).
- The final Overall Drug Score (ODS) is calculated using Eq. (1).
- Drugs are ranked and presented in descending order of their ODS.

IV. TRUST EVALUATION ALGORITHM

A. Trust Initialization

Once a new user is registered in the system, a neutral trust value is given to him or her. A standard of trust among all users should be ensured as a way of ensuring fairness within the network

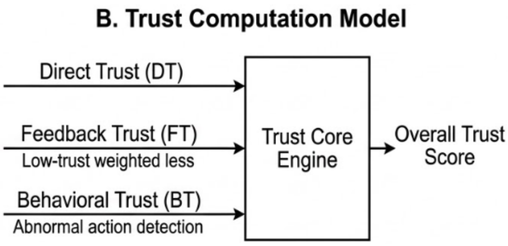
- Initial Trust Definition: $T S = 0.5 is earned by each new entity that is registered initially.
- Normalization: The trust values are also kept strictly in the range of 0 1.
  - A value of zero depicts a totally distrusted party.
  - A value of one will be a fully trusted entity.



B. Trust Computation Model

To make sure that user reliability is evaluated in a multi-dimensional manner, the trust score is computed according to three main parts:

- Direct Trust (DT): It is a component calculated and computed on the history of the user in terms of successful and legitimate system interaction.
- Feedback Trust (FT): It is calculated by the ratings that other users give. To ensure integrity of the system, the feedback given by users with low weightage is given by a user with low score in their trusted score weighted.
- Behavioral Trust (BT): This component tracks the way users behave to detect and issue alerts regarding abnormal or suspicious behavior.



C. Trust Update Mechanism

The system provides dynamic security, which is the updating of the user trust score after each interaction. Such constant adaptation indicates the last behavior of the user:

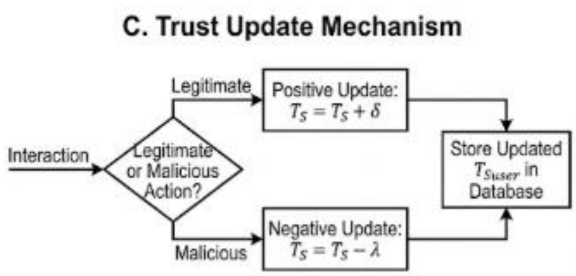
- Positive Update**:** When a user performs a legitimate access or interaction, the trust score is incremented as

$$T_{S} = T_{S} + \delta$$

- Negative Update: If malicious behavior or a policy violation is detected, a penalty is applied, and the score is decremented as

$$T_{S} = T_{S} - \lambda$$

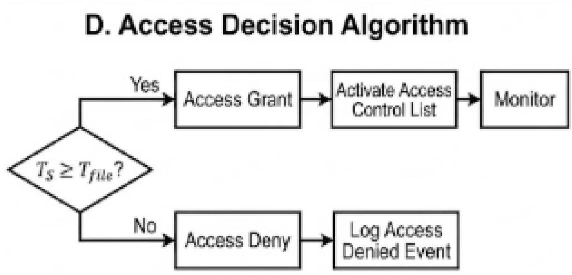
Normalization: In order to be consistent, the trust values are constrained to lie in the range of [0, 1]. In case an update results in a score going beyond these limits, it is clamped (e.g. in case, the temperature of S has increased to more than 1, it becomes 1).



D. Access Decision Algorithm

This system makes use of a threshold-based decision model in order to control access to data. The Data Owner sets a certain trust threshold (Tfile) of each file or resource depending on the sensitivity of the information:

- Rule: The access is given when and only when the current trust score of the user is more than or equal to the threshold of the file ($T_s = T_{file}$).
- Logging: In case an access request is rejected because of inadequate score of trust, a system logs the request in system logs to be audited.



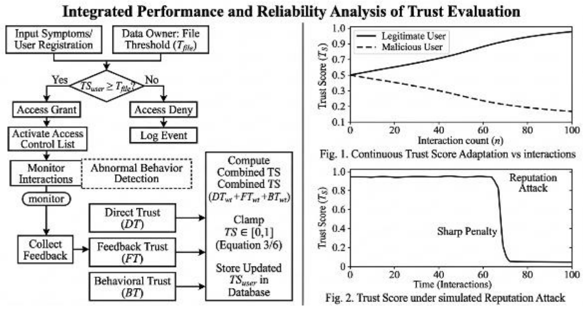
E. Graphical Analysis of Trust Behavior

The diagrammatic analysis that was used to illustrate the effectiveness and reliability of proposed system are as follows:

1.Trust Score vs. Interactions: This is a graph of dynamic nature of trust. It confirms that the score of legitimate user trust is progressively rising as he or she can continue with fruitful interaction, but the financial scores of malicious users are progressively dropping because of penalties they are subjected to on their activities.

2.Reputation Attacks: In this case, this system cannot be affected by fraudulent feedback. In the case of a simulated reputation attack, the score of trust against the attacker will skyrocket building additional isolation against him or her in the system.

3.Effect of Threshold: The chart is useful in demonstrating the impact of the trust threshold (T file) in improving the security. It shows that the higher the trust threshold the lesser will be given to the unauthorized access and the highly trusted users will be given access to the sensitive files.



V.MECHANISMS TO IDENTIFY REPUTATION ATTACK:

One of the largest threats in integrity of the trust evaluation of a trust-based cloud storage system is a reputation attack. The proposed system will be made of some detection systems that will be used to prevent and detect malicious actions that can compromise file and user integrity. The combination of strategies is followed through the following strategies:

A. False Rating Detection:

The user can also post fake ratings in a bid to enhance the degree of trust. The system provides the detection of false ratings through statistical anomaly and consistency check:

- Deviation Analysis: The suspicious ratings are those ratings that have a huge deviation to the mean rating of a file.
- Behavioural Profiling: The rating patterns are measured using the time of all the users. Those who still live on issuing drastic ratings that are not majority are tagged as being reviewed.

B. Sybil Attack Mitigation:

In sybil attacks, the attacker generates as many false identities as he can as he/she tries to maximize or reduce the degree of trustworthiness of a file.Detection strategies are:

- Identity Checkup: Identity checkup is performed by email confirmation or any other form of authentication on each login.
- Correlation Analysis: The behavior followed to what is not normal in behavior is the action taken by the account with the same IP address or the device fingerprint.

C. Collusion Detection:

Collusion occurs when the users are colluding with an aim of making a manipulation to trust score of a file. Such behaviors can be identified in the system with the help of graph-based analysis:

- Trust Correlation Graph: Interaction /ratings of the user appear in the form of a graph. The users who are high collusive occur when they give similar ratings of the same files.
- Temporal Analysis: sudden rise in like ratings over a period of time is the indication of coordinated attacks.

D. Checkup and Adaptive Surveillance:

The system will raise an alarm which would require an administrator to check whether the system is not being cheated.

Adaptive weighting: The most plausible opinion of the users will be weighted more and the impact of the malicious rating will be reduced to the minimal point.

All these will guarantee the integrity of trust model of evaluation and make the system resistant to the most frequent reputation attacks such as false ratings, Sybil attacks and collusion among users.

VI. CONCLUSION AND FUTURE WORK

In this paper, the author outlines a large-scale trust evaluation framework that can be used to protect a Drug Recommendation System against reputation-based attacks. The proposed system would guarantee the reliability of the medical recommendations as well as protect data integrity by combining the multi-dimensional computation of trust score, weighted feedback mechanisms, and solid attack-detection strategies. The successful application of this algorithm of trust evaluation makes it possible to dynamically monitor the reliability of the users, and only the high-quality and trustful feedback can impact the process of drug recommendation.

Future Work

This research may be also extended to better scale and accuracy of medical decision-support systems:

- Integration of Real-world Healthcare Data: Expanding the system to connect with the large-scale Electronic Health Records (EHR) and clinical databases to confirm drug effectiveness in a variety of patient demographics.
- Predictive Modeling using Machine Learning: Take advantage of the application of deep learning to enhance the detection of complex and dynamic patterns of attacks that cause medical rating systems.
- Context-Aware Medical Metrics Incorporating Specialised trust metrics: The addition of specialised trust metrics (dosage-specific feedback and patient-history-aware evaluation) to the accuracy of the recommendation can further increase its accuracy.
- Vigilant Anticipation of Menaces: Implementation of anticipatory analytics in a way that the malicious actors are preempted and prevented before they can breach the drug reputation scores.

In general, the suggested model offers a solid and secure method of drug recommendation, which can be both a convenient guideline in the context of healthcare protection and a road map of the future research in the field of AI-based medical aid.

REFERENCES:

[1] A. Josang, R. Ismail, and C. Boyd, A survey of trust and reputation systems of online service provision, Decision Support Systems, 2007.
S. Pearson, Trust in cloud computing Highlands, Cloud computing, springer, 2009.
[3] W. Wang et al., A cloud computing reputation based trust system, IEEE ICC, 2011.
[4] K. Hoffman et al., Reputation systems: A Survey ACM Computing Surveys, 2009.
M. Blaze et al., . “Decentralized trust management, IEEE Symposium on Security and privacy, 1996.
[6] Q. He et al. 2014, Trust management of cloud services, IEEE Transactions on Cloud Computing.
[7] Z. Yan et al., “Trust evaluation to cloud services, Future Generation Computer Systems, 2012.
[8] Y. Wang and J. Vassileva, Bayesian reputation systems, IEEE Transactions on Knowledge and Data mining, 2007.
[9] C. Dellarocas, “The digitization of word of mouth, Management Science, 2003.
[10] L. Xiong and L. Liu, “PeerTrust: Reputation based trust support, IEEE Transactions on Knowledge and Data Engineering, 2004.
[11] R. Douceur, The Sybil attack, International Workshop on Peer-to-peer Systems, 2002.
[12] S. Kamvar et al., The EigenTrust algorithm WWW Conference, 2003.
[13] H. Wang et al., and Trust-based access control in cloud computing, IEEE TrustCom, 2010.
[14] J. Huang et al., “Detecting collusion in reputation systems, Journal of Computer Security, 2011.
[15] M. Li et al., “Secure and trust-aware cloud storage, IEEE Transactions on Information Forensics and Security, 2013.

Journal of Advances in Developmental Research (IJAIDR)

E-ISSN: 0976-4844 ● Website: www.ijaidr.com ● Email: editor@ijaidr.com