Explainable Recommendations of Drugs for Diabetic Patients

Priscila Valdiviezo-Diaz

Department of Computer Science, Universidad Técnica Particular de Loja, Loja, Ecuador

Keywords: Diabetes, Drug, Collaborative Filtering, Explainable Recommendation, Recommender System.

Abstract: Currently, recommender systems are widely used for different purposes, for example, to recommend resources, products, and services. In the health domain, recommender systems are being used to recommend drugs, treatments, food plans, and healthcare services in general. Collaborative filtering is the most popular technique in the recommender system area. This technique can be of two types: memory-based collaborative filtering and based-model collaborative filtering. One of the problems of recommender systems is that most of them focus on enhancing the precision of the recommendation and do not provide a justification for the suggestions given to the user. Therefore, it is important to provide explainable recommendations so that the user understands why an item is recommended. To address this problem, in this paper the use of a Bayesian method for explainable drug recommendations for diabetic patients is presented. Several experiments are carried out using a dataset with information on diabetic patients with three collaborative filtering approaches: the memory-based approach IbCF, and two model-based approaches: item-based NBCF, and Hybrid NBCF. The experimental results present good results for the Hybrid NBCF approach compared to the other approaches tested. Moreover, it is observed a better quality of prediction and an increase in recommendation precision with Hybrid NBCF.

1 INTRODUCTION

The advancement of technology has allowed health institutions to have a large amount of information about patients, which can be analyzed and used to help doctors prescribe treatment and medication properly. This information, often available in medical databases, refers to laboratory test results, treatments, diagnoses, and prescribed medications. According to (Wang et al., 2022), it is scientifically important to use drugs to improve their effectiveness in disease treatment.

In this sense, recommender systems (RS) are being widely applied to the health domain to support medical suggestions and provide personalized attention to the patient (Tran et al., 2019). Recent research on RS on the health has focused on disease prediction and recommending the precautions (Rustam et al., 2022), content recommendations to patients with diabetes (Nagaraj and Deepalakshmi, 2022), recommendations for national fitness items (Li and Yang, 2022), and recommendations for healthcare services (Meng et al., 2022). Nowadays, recommender systems developed within the health care setting, have been used for disease management programs, for example, hypertension (Sajde et al., 2022), and diabetes (Kamath et al., 2022), providing a personalized user experience.

In the health domain, it is also important to have explainable machine learning models that help health professionals in decision-making, and internal actions, thus, by explaining the results of a prediction, the trust of doctors is gained making it possible to apply the predictive model in practical situations (Yang, 2022). These models are being used in recommender systems to provide justification for the suggestions that the system provides.

On the other hand, one of the diseases that has attracted a lot of attention from health researchers is diabetes. Diabetes is a chronic disease caused by a lack of physical activity and unhealthy eating habits (Bankhele et al., 2017). In this context, (Ali et al., 2018) develop a recommender system to suggest physical activity and diet plans to help patients control this disease and avoid future complications.

These systems also help health professionals to provide medical recommendations on treatments or medications to prescribe to the patient. According to (Calero Valdez et al., 2016), these systems are expected to minimize time and effort in the healthcare decision-making process. Although some works re-
lated to recommender systems have been developed in the health domain, most focus on the recommendation of a specific item or service and enhancing the recommendation’s precision. Still, they do not provide an explanation of why a particular drug is recommended.

Traditional recommendation systems are based on machine learning techniques such as matrix factorization (MF) (Azri et al., 2023), KNN (Nagaraj et al., 2022), and neural networks (Chaithra et al., 2023). However, models based on MF and neural networks are difficult to explain the recommendations. According to (Ammar and Shaban-Nejad, 2020), explainable machine learning models promote credibility and trust in critical areas, such as medicine, by combining machine learning techniques that explicitly show why a recommendation is made. Therefore, unlike the works reviewed in the state-of-the-art, in this paper, in addition to providing a list of medications for diabetic patients, an explanation is provided for these recommendations. Previous this, an analyzing the performance of three collaborative filtering recommendation methods which facilitate the explanation of recommendations is carried out, and then the approach that better results present is selected.

Explaining the recommendations to the patient can improve users’ trust and thus increase acceptance of the system by health personnel. In accordance with (Tran et al., 2021), trust is even more critical for RS to convince patients to follow health-related recommendations. This aspect can be enhanced by providing explanations for recommendations (Ammar and Shaban-Nejad, 2020).

The rest of the paper is structured as follows: Section II presents related work. Section III encloses the context of the present work. Section IV includes the material and method used for the experiments. Section V shows the experimental results and the process for the explanation of recommendations. Section VI encloses the conclusions and future work.

2 RELATED WORK

In this section, studies focusing on the development of recommender systems for diabetes, medicine recommendation, and works including explaining recommendations are presented.

2.1 Recommender Systems for Diabetes Patients

Several studies on recommender systems have been developed for the treatment and control of diabetes. For example, (Zeng et al., 2017) use information retrieval approaches in recommending diabetic patient education materials based on diabetic questions posted on the TuDiabetes forum. In (Bankhele et al., 2017) propose an android application based on a user-based collaborative filtering approach to suggest probable medication, diet, and exercise to help people manage their diabetes well. This application can also remind users to carry out the recommendations which are provided by the system. Authors in (Rehman et al., 2017) present a cloud-based food recommendation system, for dietary recommendations based on users’ pathological reports. An ant colony algorithm is used to generate an optimal food list and recommends suitable foods according to the values of pathological reports. Likewise, in (Bhat and Ansari, 2021) a machine learning technique is used for diagnosis of diabetes and recommend proper diet for diabetic patient.

(Nagaraj and Deepalakshmi, 2022) propose an intelligent fuzzy inference rule-based predictive diabetes diagnosis model, providing content recommendations to patients with diabetes. The model predicts the risk of diabetes disease using fuzzy inference based on Mamdani’s technique, then the recommendations for a normal life, nutrition, exercise, and medications are given to patients.

In (Almulla, 2020) propose an expert system to diagnoses diabetes and recommends the right medication depending on the location where the patient lives and on the symptoms of the patient and other effective factors. The system outputs a list of names of locally available brand names of medications that suit the diabetes type.

2.2 Medicine Recommender Systems

In recent years research related to medication recommendations based on machine learning has been developed, for example, a medicine recommendation model based on the incorporation of graphs to recommend appropriate drugs for patients is proposed in (Wang et al., 2022). This model generates the recommended drug list by calculating the cosine similarity between disease combination representations and drug combination representations. A recommendation algorithm called LEAP is presented in (Zhang et al., 2017), which uses records of current patient visits and drug-drug interactions to predict a list of med-
ications. This algorithm uses a recurrent decoder to model label dependencies and content-based attention is used to capture label instance mapping. (Wedagu et al., 2020) propose a recommendation method called DIMERS, which combines a prior medical knowledge of doctors with bidirectional Long Short-Term Memory (BiLSTM). In this study, authors use a weighted block with prior medical knowledge to enhance the learning of deep neural networks.

A drug recommendation model based on message propagation neural network is proposed in (Ren et al., 2022), in this paper, the Drug-Drug Interaction (DDI) knowledge is introduced into the model to reduce the DDI rate of recommended results. A two-stage personalized medication recommender system is presented by (Bhoi et al., 2021). Authors use various weights in the system to compute the contributions from the information sources for the recommended medications. The system models the drug interaction from an external drug database and the drug co-occurrence from the electronic health records as graphs.

2.3 Explainable Recommender Systems

In the health domain, a low number of papers have been found to focus on providing explainable recommendations, some of them are oriented to recommend food and explain these recommendations, for example, (Padhiar et al., 2021) include explanations to users for food-related suggestions. These authors model food recommendations, using concepts from the explanation domain to create responses to user questions about food recommendations. Likewise, in (Pecune et al., 2022) a conversational system that recommends recipes aligned with its users’ eating habits and current preferences is presented. This system is also able to justify its recipes recommendation by explaining the trade-off between them.

Authors in (Zoppis et al., 2019) present a computational model for promoting targeted communication and supplying social explainable recommendations, in order to support the formation of communities of patients and health services. In (Gutiérrez et al., 2022) present the design and implementation of a recommender engine and a mobile application designed to support call recommendations and explain these recommendations.

In the work of (Cai et al., 2022), a model of explainable recommendation on account of knowledge graph as well as many-objective evolutionary algorithm is proposed, which combines recommendation and explanation. Likewise, (Chicaiza and Valdiviezo-diaz, 2022) present a research related to explainable recommender systems. In this research, the authors describe two scenarios based on the TripAdvisor dataset to generate restaurant explainable recommendations. Explainable recommendations are evaluated using Fidelity and Transparency metrics.

3 CONTEXT

A previous work on drug recommendations for diabetes patients was presented without considering explaining the recommendation and using other prediction techniques. This previous work is based on collaborative filtering (CF) and clustering techniques for recommending drugs to diabetes patients (Morales et al., 2022). This system realizes suggestions according to drug information and the characteristics of patients. The clustering technique is applied to group patients with similar characteristics, and the collaborative filtering technique is applied to represent the patient’s explicit data, then based on the group to which the patient belongs, the recommendation is made considering the drugs with the highest prediction value.

Currently, Diabetes is a chronic metabolic disease that generates a great impact on the world population. According to (Saeedi et al., 2019) it disease is among the top 10 causes of death in adults. Diabetes can be treated through physical activity, healthy eating, medication, and regular checkups to prevent complications.

In (Association, 2020) different types of medications recommended by the American Diabetes Association (ADA) related to or used in the treatment of this condition are presented. It is possible to determine that there are many different types of drugs that can work to lower your blood sugar. Sometimes one medication is enough, but the doctor may prescribe a combination of medicines in other cases.

Authors in (Al-Sofiani et al., 2021) propose a medication algorithm scheme for the treatment of people with Diabetes. Authors put special emphasis on medication cost and medication adherence as determining factors in the choice of diabetes medications recommended.

The aim is to use the information on essential medicines and doses prescribed for diabetes patients to make drug explainable recommendations for treating this disease. Therefore, in the present work, a step forward from the drug recommendation is taken, considering the justify the recommendation given to the patient based on the diabetic patient’s explicit data (patient-dose-drug), and using probabilistic algorithms that allow understanding of the recommenda-
tion. As a result, three collaborative filtering algorithms were tested using a dataset with diabetic patient information.

In the work presented by (Valdiviezo-Díaz et al., 2019) a bayesian hybrid approach facilitates understanding and explaining the recommendation of the user. This CF approach called Naive Bayes Collaborative Filtering (NBCF) recommends items by using similar users’ and items’ information, respectively. This novelty approach has been considered to explain recommendations taking into consideration the information of patients and the drugs specified in the dataset used, and instead of the rating, the dose of the drugs is considered. Moreover, this approach has been selected because it provides successful results in the quality of recommendations.

4 MATERIAL AND METHOD

Many machine learning techniques have been designed for diabetes diagnosis, the prediction of this disease, and providing useful analysis of medical data. In our work, a probabilistic machine learning model for explainable drug recommendation for diabetic patients is used.

For the recommendation, the collaborative filtering recommendation approach will be use, since, according to (Wang et al., 2020) this approach is one of the most applied methods in the health recommendation systems. The collaborative filtering method recommends to the active user items that other users with similar preferences have liked in the past (Ricci et al., 2015). CF can be of two types: memory-based CF and model-based CF (Yang et al., 2016). In the memory-based CF method, the RS uses the ratings to find neighbors for the target user or item and computes the predicted value for the unknown rating. Model-based CF uses a model to predict the value for the unknown ratings based on the rating matrix.

In this manuscript, a memory-based CF method and two model-based CF methods that allow recommendation explanation are tested.

For a better demonstration of the explanation of the medications, a recommendation scenario centered on the patient is presented, to whom some medications are suggested by the system with their respective justification.

For the evaluation, the most common metrics will be used to evaluate the performance of recommendation systems. The prediction and recommendation quality are computed using an existing dataset with information on diabetic patients.

4.1 Dataset

For the experiments, a dataset with diabetic patient records available in the UCI Machine Learning Repository is used, which refers to 100,000 observations and 50 features representing patient and hospital outcomes (Dua and Graff, 2017). This dataset contains information related to the personal data of the patients, information about admission, procedures, medications, and diagnostics results (Strack et al., 2014).

The dataset was pre-processed to construct the user-item rating matrix necessary in collaborative filtering. Patients who have been administered at least two medications were selected, leaving a total of 5,148 unique patients. Likewise, a selection of drugs was made considering those that have been prescribed to at least 50% of the selected patients, as a result, there is a total of 10 drugs. Then, to represent the matrix of collaborative filtering (patient-drug-dose), the information on the patient’s dose for each drug is considered, that is, if the dose remains stable, or if the dose is increased.

Table 1 shows a summary of the data to be considered for the experimentation. Each drug is on a scale from 1 to 2 (1: indicates whether the dose is maintained for the patient; 2: if the dose is increased).

Figure 1 shows the drugs selected for the experiments and the percentage of patients using the drug. From figure 1, Insulin is the drug that has been prescribed the most to patients for the treatment of diabetes, on the contrary, Glyburide-metformin and Nateglinide are the drugs that have been prescribed the least.

<table>
<thead>
<tr>
<th>#Patients</th>
<th>5,148</th>
<th>Total patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>#Drugs</td>
<td>10</td>
<td>Total Drugs</td>
</tr>
<tr>
<td>#Ratings (dose)</td>
<td>46,593</td>
<td>1-2 values</td>
</tr>
</tbody>
</table>

4.2 Evaluation Metrics

To assess the accuracy of the prediction of algorithms, Mean Absolute Error (MAE) (Wang and Lu, 2018) is used. Also, Precision and Recall metrics are used to evaluate the quality of the recommendations. Precision represents the percentage of recommended items being relevant, and Recall represents the percentage of relevant items being recommended (Valdiviezo-Díaz and Bobadilla, 2019).

In addition to evaluating the quality of the prediction and the recommendation, this paper presents the
5 RESULTS
This section presents the results of the experiments with the dataset used, and the detail on how to explain the recommendations with the selected method.

The cross-validation method is used to evaluate the performance of the algorithms, where the dataset is split into 80% training set and 20% testing set.

5.1 Performance Comparison
In order to select an algorithm for the explanation of recommendations, we compare the performance of three collaborative filtering methods for recommender systems, for example, model-based CF methods: Item-based NBCF and Hybrid NBCF, and the traditional CF method based on memory IbCF. These methods were selected for experimentation because currently, model-based approaches are achieving better results in accuracy and performance (Valdiviezo-Diaz et al., 2019), and also the Bayesian models-based methods provide a probabilistic interpretation of their results, which facilitates explaining the decision process (Cheng et al., 2017). On the other hand, memory-based approaches are that they are simple to implement and the resulting recommendations are often easy to explain (Charu, 2016).

A more detailed description of the methods selected is presented, join with the relevant hyperparameters set for the experiments.

- **Item-based NBCF** recommends items to the user according to the ratings received by each item (Valdiviezo-Diaz et al., 2019). The hyperparameter set for this method is \( \alpha = 0.01 \).
- **Hybrid NBCF** combines user-based NBCF and item-based NBCF approaches to complement each other and improve the accuracy of the model (Valdiviezo-Diaz et al., 2019). The hyperparameter set for this method is \( \alpha = 0.01 \). This approach first computes the prior distributions and the likelihood for user-based NBCF and item-based NBCF approaches, then both approaches are combined using a weighted product.
- **Item-based CF (IbCF)** recommends similar items to the item the active user has already preferred in the past (Kant and Mahara, 2018). The cosine correlation coefficient is used as a similarity measure.

The hyperparameter values were selected in order to maximize the accuracy of algorithms for quality measures used.

In NBCF, the explanation is realized considering the evidence set of the bayesian method. On the other hand, the IbCF algorithm can explain the recommendations provide, considering the similarity between the items.

Table 2 shows the results of the comparative analysis of the CF algorithms based on the performance evaluation metrics on the dataset used.

From the experimental results, it is observed that: NBCF methods present a better performance in contrast to IbCF. The results show that Hybrid NBCF outperforms all the other methods in terms of MAE. A smaller value of this metric means better performance. Moreover, it indicates that the predicted and actual values are closer. Therefore, we can conclude that the NBCF model is making a good prediction of the drug dose.

Likewise, the results show more accurate values concerning Precision for the probabilistic methods (NBCF) in comparison to IbCF. From the table 2 we can see that 74% of the recommended drugs are relevant or adequate for the patient. However, Recall results show that IbCF is better than the other two methods tested. Therefore, analyzing the performance of the three CF methods, Hybrid NBCF is selected for the explanation of recommendations because it presents better performance in most of the comparison process (Cheng et al., 2017). On the other hand, memory-based approaches are that they are simple to implement and the resulting recommendations are often easy to explain (Charu, 2016).

A more detailed description of the methods selected is presented, join with the relevant hyperparameters set for the experiments.

- **Item-based NBCF** recommends items to the user according to the ratings received by each item (Valdiviezo-Diaz et al., 2019). The hyperparameter set for this method is \( \alpha = 0.01 \).
- **Hybrid NBCF** combines user-based NBCF and item-based NBCF approaches to complement each other and improve the accuracy of the model (Valdiviezo-Diaz et al., 2019). The hyperparameter set for this method is \( \alpha = 0.01 \). This approach first computes the prior distributions and the likelihood for user-based NBCF and item-based NBCF approaches, then both approaches are combined using a weighted product.
- **Item-based CF (IbCF)** recommends similar items to the item the active user has already preferred in the past (Kant and Mahara, 2018). The cosine correlation coefficient is used as a similarity measure.

The hyperparameter values were selected in order to maximize the accuracy of algorithms for quality measures used.

In NBCF, the explanation is realized considering the evidence set of the bayesian method. On the other hand, the IbCF algorithm can explain the recommendations provide, considering the similarity between the items.

Table 2 shows the results of the comparative analysis of the CF algorithms based on the performance evaluation metrics on the dataset used.

From the experimental results, it is observed that: NBCF methods present a better performance in contrast to IbCF. The results show that Hybrid NBCF outperforms all the other methods in terms of MAE. A smaller value of this metric means better performance. Moreover, it indicates that the predicted and actual values are closer. Therefore, we can conclude that the NBCF model is making a good prediction of the drug dose.

Likewise, the results show more accurate values concerning Precision for the probabilistic methods (NBCF) in comparison to IbCF. From the table 2 we can see that 74% of the recommended drugs are relevant or adequate for the patient. However, Recall results show that IbCF is better than the other two methods tested. Therefore, analyzing the performance of the three CF methods, Hybrid NBCF is selected for the explanation of recommendations because it presents better performance in most of the comparison process (Cheng et al., 2017). On the other hand, memory-based approaches are that they are simple to implement and the resulting recommendations are often easy to explain (Charu, 2016).

A more detailed description of the methods selected is presented, join with the relevant hyperparameters set for the experiments.
metrics calculated, for example: in MAE and Precision.

5.2 Explainable Recommendations for Diabetic Patients

The selected algorithm Hybrid NBCF allows explaining the predictions. For the explanation, what is mentioned in (Valdiviezo-Diaz et al., 2019) is considered, which indicates that to explain a recommendation to the user \( u \) is necessary to consider the case in which the system has recommended the item with an estimated rating, in our case, would be that the system has recommended a drug with an estimated dose value.

Based on the algorithm for the Hybrid NBCF approach, we have the following recommendation for a patient:

The case in which the system has recommended the drug \( i \) with an estimated dose \( \hat{r} = 2 \) is considered. As explained in (Valdiviezo-Diaz et al., 2019) it necessary to obtain the \( P \) and \( Q \) evidences corresponding to the user-based NBCF and item-based NBCF approaches, respectively. So, firstly, all drugs that have been prescribed to patient \( u \) according to their likelihood within the drug \( i \) are sorted from highest to lowest. Secondly, the items from the list whose dose has been increased for the patient are extracted, and add them to the set of \( P \) evidences, in this case: Insulin.

Next, all patients who have been prescribed the drug \( i \) according to their likelihood within the patient \( u \) are sorted from highest to lowest. Then the patients of the list whose dose of medication \( i \) has been increased are extracted and added to the list of \( Q \) evidences, in our case, patients: 2844, 3019, 3286, 3751. Finally, both sets of evidence are combined by adjusting the \( P \) and \( Q \) evidences according to their likelihood within the patient.

Figure 2 presents the explain the recommendation given to a patient.

From figure 2 can be observed that in addition to presenting the recommendation to the user, an explanation of the recommendation is shown. It is hoped that this explanation can help the user understand why the drug is recommended and give the user greater confidence in the system.

5.3 Evaluation of the Explainable Recommendations

(Zhang and Chen, 2018) introduce approaches to evaluate recommendation explanations. The first approach evaluates the percentage of recommended items that can be explained by the explainable recommendation model, regardless of the quality of the explanations; and the second evaluates the quality of the explanations exactly. This paper applies the first approach to evaluate the explicability of the Hybrid NBCF recommendation algorithm used to explainable drug recommendations.

In this section, we present the evaluation of the recommendations using the Fidelity metric. In our case, the recommended items will be the items whose estimated dose is 2, that is, those items whose dose has been increased. The explainable items will be those items that are recommended but that can be explained because there is at least one element in the set of \( P \) and \( Q \) evidences.

Therefore, applying the equation defined in (Peake and Wang, 2018), the fidelity obtained is 0.66. This means that evaluating the quality of the explanations of the NBCF algorithm will depend on the existence of \( P \) and \( Q \) evidences. The more items that can be explained, the higher the fidelity value.

6 CONCLUSIONS

In the health domain, explainability is essential to gain the trust of healthcare professionals and patients and to enhance the transparency of the recommender system.

In this paper, the use of a probabilistic approach to explain recommendations based on the doses of drugs prescribed to diabetic patients was presented. Because this approach is probabilistic, we think that it makes the explanations easily understandable to users. We have also established how would be the explanation of the recommendation made to a user within the system with the Hybrid NBCF approach, and how to evaluate the explanations determining the faithfulness of the explainable recommendation model.

We focused our research on testing CF methods which allow the explanation of recommendations. From tested methods, NBCF shows better results in drug dose prediction accuracy. Moreover, the results of experiments conducted on a real dataset of diabetic patients verify the good recommendation performance and explanatory ability of Hybrid NBCF.
Future work includes: a) testing other collaborative filtering methods for explaining recommendations using the same dataset, and b) evaluating the transparency of the recommender system, providing to the patients an understanding of how the system formulated the recommendation.

REFERENCES


