Recommending Relevant Services in Electronic and Mobile Health Platforms

Gjorgji Noveski[1] and Jakob Valič[1]
E-mail: gjorgji.noveski@ijs.si, jakob.valic@ijs.si
[1]Department of Intelligent Systems, Jožef Stefan Institute, Ljubljana, Slovenia

Keywords: Electronic and mobile health, recommendation systems, embeddings

Received: May 6, 2022

Electronic and mobile health (EMH) is becoming an integrated part of healthcare as we move in the future. The opportunity in bringing closer healthcare services with the advent of the internet is growing larger. This is why it is important to adequately provide those services to the people that need them and to also further improve them. Regarding electronic and mobile healthcare systems, it is fairly easy for users to get lost while searching for some information due to the vast amount of data that is present for different illnesses, healthcare institutions and healthcare services. In this paper we present a platform that provides various healthcare services to people, namely the Insieme platform (ISE-EMH). Knowing the difficulty of finding relevant information on platforms and that user preferences vary to a great extent, we additionally give an overview of an implementation of the recommendation system that is part of the Insieme platform which helps users pick services that might be relevant to them.

Povzetek: Razvita ja pratforma ISE-EMH za Italijo in Slovenijo kot jedrnata verzija "dr. Googla".

1 Introduction

With the increasing popularity of digital interventions and digital solutions, electronic and mobile health applications are becoming more popular throughout many sectors of the public healthcare. The idea started since the beginning of the era of smartphones, where the convenience of accessing the internet from the comfort of our homes, moved to conveniently accessing the internet from the palm of our hands. This shift provided fertile ground for a vast amount of new solutions and applications to emerge. Naturally, the healthcare domain was also affected by this and with the collaboration of healthcare professional and people from the domain of information and communication technologies, solutions are built that benefit the people.

In the digital era, more things are becoming available online and with this, general quality of life of people is increasing. Many companies, institutions and various stakeholders shift towards online solutions regarding their data availability. This is beneficial to both the companies and the people using their services for various reasons. Firstly, it allows remote access to employees regardless of location, giving them the option to continue working even when they are not on-site. Furthermore, it can allow commonly used services to be available to the general public. Such examples are online booking for doctor's appointment or online overview of waiting queues. Advances and increase in usage of wearable devices give opportunities for tracking many aspects of a patient's vital signs, heart rate, breathing, blood pressure, etc. This information can be integrated into a system that builds up an electronic health record from the user and provides health checks, monitoring and suggestions. Clearly we can see that there are many opportunities to be explored related to using digital technologies and data on person's health.

During design and implementation of healthcare platforms, ease-of-use and user experience need to be considered. This is because the platforms are going to be used by a variety of individuals, the healthy and the ill. Because of this, the process in which they obtain the required information from the healthcare platform must be as easy as possible. Many different digital solutions have been developed which provide healthcare assistance like embedded [1], and decision support systems [3]. In recent time advances in artificial intelligence pave way to more complex solutions that bring a bigger array of services and benefits in the field of healthcare. These benefits are from areas such as predictive analysis of diseases, patient care, patient monitoring, etc. [6].

Like previously mentioned, the time needed to reach desired information on a healthcare platform should be as low as possible. In order to achieve this, a certain system must be able to learn the patterns of user's interaction with the platform. This way, the system provides information to the user while also building a model representation for that user. This learned context gives the system an ability to tailor which information is shown to which user, thus providing more specialized experience. Such functionalities are typically implemented using recommendation systems.

In our paper, we present the methodology for recommending services on the Insieme platform, i.e., an Electronic and Mobile Health platform that we developed within the ISE-EMH project [4]. We give an overview of the platform as well as the recommendation system used to provide helpful service suggestions. Additionally, analysis is performed on gathered user-service interaction data in order to get some insight on site usage.

The rest of the paper is organized as follows. Section 2 provides examples of related work in the field of Electronic and Mobile Health as well as similar applications. In Section 3 the implementation of the recommendation system is explained while Section 4 gives brief explanation on the EMH platform that the system is used on. Interaction analysis from user data is given in Section 5 and lastly a conclusion is presented in Section 6.

2 Related work

Taking a look at related systems, we can find examples of different platforms and applications that work towards bringing ease of use to the user. The actors in the healthcare system are the patients and the clinicians. Consequently, solutions that benefit either the patients or the clinicians benefit the whole healthcare system. We will present some solutions that are aimed at both clinicians and patients.

Showcasing the importance and benefits of using recommendation systems in the healthcare domain is meticulously presented in [9] where an overview of the existing research in the EMH domain is provided. The analysed scientific papers were also carefully selected and filtered by their criteria which included: selecting papers published no earlier than year 2000, papers referenced with 15 sources and more, containing a detailed discussion on recommendation techniques, etc. After discarding papers that did not meet these criteria, 98 papers remained to be further studied. Examples of the reviewed systems are systems that can recommend food, drugs, healthcare professionals and, etc. The authors listed three main aspects that need to be considered when designing a recommendation system: users, items, and usage context. In our case, users are the patients that will use the system and items are the entities which will be recommended to them. Usage context refers to the set of factors that might influence the selection of the items to be recommended to the user. For example, if the user's health record is integrated into the recommendation system then different users can get different items recommended to them just because of their dietary preference, allergies, etc.

De Croon et al. [2] reviewed a total of 73 published studies that had reported the implementation and evaluation of recommendation systems in the healthcare domain. The authors showed that the most prominent categories of recommended items are about lifestyle, nutrition, general health information, and specific health condition related information. In our work, the recommendation system recommends relevant services to the users, which are related to the specific health condition related information. In literature the most frequently used type of recommendation system is a hybrid one which is a combination of collaborative filtering and content-based filtering (see Section 3. for details).

Diaz Ochoa et al. [7] applied neural networks to create a recommendation system that achieves recall of 98 % and accuracy of 64 %. It takes into account patient data alongside with specific treatments encoded into treatment keys. This information is further processed by clustering in order to lower the dimensionlity of the problem and finally a deep learning model is trained. The advantages of the developed solution are faster training time due to its fewer number of parameters and transparency due to its use of multi-criteria decision operators. The Logic-Operator neural network they use simulates cognitive processes with fuzzy logic. The logical operators are or and and gates which represent the last layers of the network. Furthermore mean squared error is used as a loss function, ReLU as an activation function and the ADAM method as optimization. The recommendation system consists of two separate systems. One of them is trained on the entire patient dataset, while the other is trained only on patients which had positive outcome of their treatments. This provides two predictions and by comparing the two predictions a recommendation is made with low or high confidence. If the predicted treatments match then high confidence is assumed. Likewise if the predictions of the two systems do not match, a treatment is still recommended but with low confidence.

Punn et al. [8] developed a recommendation system in the healthcare domain, which is based on collaborative filtering and recommends remedies by taking as input the patient's symptoms. Since there is a limited amount of data linking remedies to various diseases that is suitable for creating recommendation systems, the authors also provided a dataset for this purpose. This dataset consists of around 1,100 diseases and close to 300 symptoms. Their recommendation system uses singular value decomposition and cosine similarity in order to assess which symptoms give more rise to a certain disease. The system was evaluated with symptoms that were related to one or more diseases and in both cases it recommended one or more remedies respectively.

3 Recommendation systems

Recommendation systems aim at providing reliable recommendations for the domains they are built for, e.g., songs, movies, products at an online shop, etc. With the term "items" we associate all the possible types of recommendations. Because of the ability of the recommendation systems to adapt suggestions based on specific users, several online services that include recommendation systems have greatly improved user experience.

The recommendation system takes as input a search query vector and computes which items are most similar to it. There are several options to compute the similarity between different vectors. Several recommendation systems use one or more of the following metrics:

1. Cosine
2. Dot product
3. Euclidean distance

These metrics calculates a score of how much the current choices of the user are similar to all the other items. The items which relate the highest are taken as recommendation to the users. Our system uses the cosine similarity metric.

There exist three types of recommendation systems: collaborative filtering, content-based filtering, and hybrid ones which are a combination of the first two. In our platform, we use the LightFM recommendation system [5], which is a hybrid one. It works as a hybrid matrix factorization model where it represents the users u and items i as embeddings (latent vectors). These are defined as linear combination of the features that describe each user or item. When these features are present in the model, it defines the recommendation by calculating the dot product of u and i representations, adjusted by the respective biases b_u and b_i. When no item or user features are available, the LightFM model falls back to pure collaborative filtering. These types of recommendation systems are further explained in the sections below.

3.1 Content-based filtering

This is the simplest type of recommendation system. In essence it uses the similarity between items to recommend items similar to what the user has chosen in the past. From the constructed user-item interaction matrix, item features are extracted and a similarity metric is calculated. This approach has several advantages and disadvantages, for instance it is scalable to a large number of users because it considers only one user at a time. In addition, it can provide recommendations to very niche items, by learning from the user's previous interactions with the whole system. Downside of such a system is that it needs domain knowledge to some extent for the creation of item features. These item features for example can represent the categories of a certain movie (horror, comedy, action, etc.). Further downside is that it can not expand the user's taste by suggesting items which are not closely similar to what the user has chosen in the past.

3.2 Collaborative filtering

To address some of the concerns and drawbacks of the content-based filtering method, collaborative filtering was developed. Collaborative filtering takes into account the similarities between users and items simultaneously in order to provide recommendations. The benefit of this is that while the system looks for similar items, it also looks for similar users to the target user, which means new items can be suggested to him/her. Consequently, the users are able to discover new interests.

The recommendation system can operate in two modes, depending on how it interprets the user-item interaction matrix. These two modes are:

1. Implicit feedback
2. Explicit feedback

In the implicit feedback mode, the model regards the absence of data in some fields as negative feedback. This means that for the items the user did not supply a rating, it considers them as if the user gave a negative rating. The idea behind this is that the user consciously made a decision about which items he/she will rate, so the system regards the items that aren't rated, as items that the user does not like. On the other hand, an explicit feedback model is the opposite. A rating for an item has to be supplied in order for the system to make any judgements. The items which do not have a rating are considered as data that is unknown. Which one to use is typically dependant on a use-case basis.

3.3 Embeddings

An embedding represents a relatively low dimensional space which captures some semantic meaning from a higher dimensional input. When working with categorical data, as is in our case, having many features quickly brings the size of the input to a large number. This results in longer training times of the models and more data that needs to be processed. Embeddings can transform the data from a categorical one, to a real-valued representation in fewer dimensions. On each of these dimensions the system learns to represent some concept, for example the N-th dimension might represent the fact whether the item belongs to a class of horror movies or action movies.

In our recommendation system, embeddings are learned internally and used for predicting services. We also use them in order to acquire a list of most similar tags to be suggested to the user. To create meaningfull embeddings, LightFM gives the option to add additional features that describe the items and features that describe the users. Consequently, better embeddings are learned by the system and used for generating better recommendations.

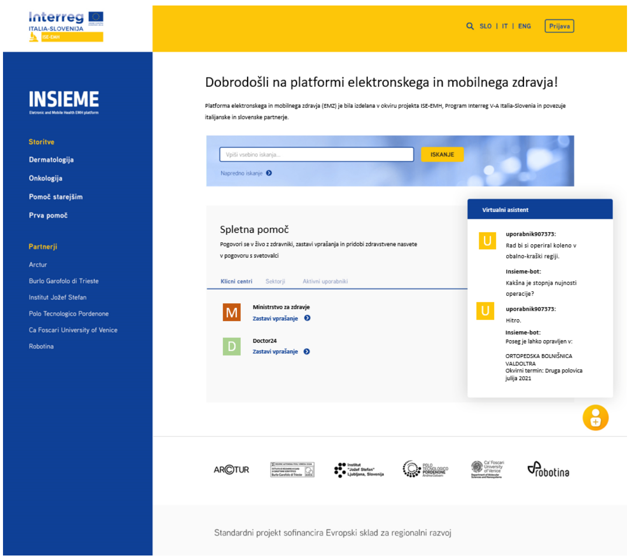

Figure 1: The Insieme platform.

4 Insieme platform

The ISE-EMH platform connects various partners, medical institutions, and patients from both Italy and Slovenia. The medical institutions and partners provide information about their services through the platform so the users can have one location where that information is easily accessible. The platform also facilitates communication between a specialist and a patient. Furthermore, a chatbot is developed that integrates various services and question answering. For example, it can list the waiting queues for different medical institutions and provide helpful information about diseases. When getting information about various diseases, the user is also provided with information about medical institutions that deal with them alongside the symptoms associated with the disease. In Figure 1 we can see the graphical user interface of the platform.

5 Interaction analysis

An integrated part of the platform is logging of the user interactions on every click. When a new user visits the site and he/she click on a service, that interaction is saved in a database. The system keeps track of which users interact with which services by means of internet cookies. We consider the time between the user visiting the platform and leaving it as a “session”. The information from the sessions is crucial if we want to train our recommendation system. Each user has his/her own unique session identification number for as long as he/she uses the platform. It is possible for the same person to use the platform multiple times but on different devices. This way the system will assign a different identification number to the different sessions. During design we presume that each identification number belongs to a different person even though it may not always be the case. This choice was taken since we assume that the number of occurrences of this event will be insignificant. The goal of this analysis is to acquire and present empirical data of the platform usage to acquire a better understanding of user patterns in order to further refine the recommendation system.

The interaction analysis was performed as follows. We analysed the number of times services were visited. As seen in Figure 2, majority of services have been visited between one and thirty times so far. A popularity trend can be also observed, e.g., the top three services which were the most chosen are:

1. Psoriasis
2. Ambient assisted living
3. Breast cancer

Figure 3 shows how many services were visited in each session. This figure shows that the majority of users (more than half at around 500) have only clicked on a single service. Consequently, we can assume that most of the users come to the platform to obtain specific information and know what they are searching for. It should be noted that sessions in which more than 50 services were visited are filtered out from this data. We consider them as outliers and rare occurrences. Nevertheless, a high number of services visited per session might indicate users which do not come to the platform with a specific goal in mind, but just casually educate themselves in resources the Insieme platform provides.

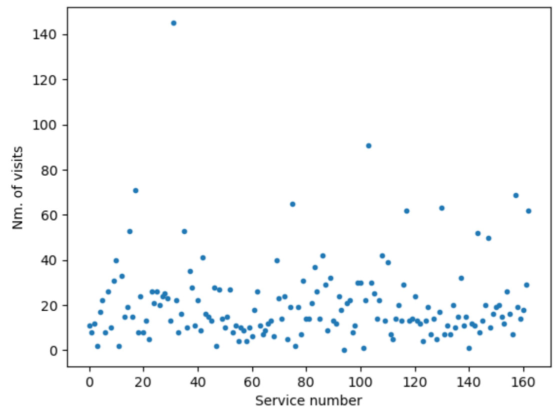

Figure 2: Number of times a certain service was clicked.

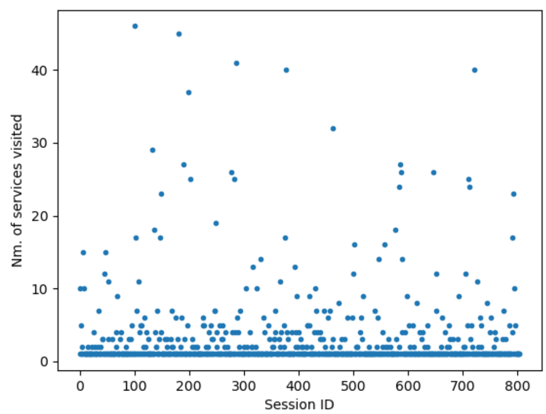

Figure 3: Number of services visited in each session.

The implementation of the recommendation system on the Insieme platform is regularly trained with new data. The time period between each training is one day. This ensures enough time for new data to be generated by users throughout the day whilst at night when site traffic is lower the system is trained and improved. The presented interaction data was collected over the course of several months.

Table 1 shows an example of interaction matrix between (subsets of) user sessions and services. The “+” sign indicates that a service was visited during a session.

Each of the services on the Insieme platform has a set of tags associated with it, for example, organ-liver, duration-

Table 1: The interaction matrix with subsets of user sessions and services.

Services / Sessions	Acne	Dermatology	Heart Diseases	Dementia	Epilepsy	Elderly support	Mental illness
4relftnjnvvzjogwtu3ap2pnx	+	+					
7dmar25i0cqm0a64fajznt4fm			+		+	+	
3sgxsbskwqszpbg2dom2y6k7f				+			
0n5c4ms9jk27khzcbnfvou7qt				+			+
clainomcth543lm7ss6lpisv0	+	+					
8b9k48ced5pbij88nk1ommvkx						+	+
0b8w50rwyel1bxutiz4m2ujg0	+				+		
4di23cq5t5637es2eoa43at4n			+	+	+		

chronic, disease course-progressive, etc. By providing recommendations to users based on these tags, the users are able to easily search for and identify the needed services. In addition, the usage of the tags enables us to learn better embeddings and provide improved suggestions. An additional procedure for suggestion making is to obtain the most similar services embedding-wise. The system can provide recommendations by searching the proximity of the queried service in its lower dimensional embedding space and returning services which are in that proximity. Since embeddings can capture semantics, the obtained services will be those that are closely related to the searched service.

We also trained a secondary model that provides recommendations for tags instead of services. This is a more general model and not as specific as the model used for recommending services. The main advantage of this model is that it requires less training data because we have a smaller number of tags compared to the number of services. Both the service suggestion model and the tag suggestion model can be used side-by-side or independently. Due to the platform requirements, only the service suggestion model is used.

We tested the speed of the Insieme recommendation system and measured the area under the receiver operating characteristic curve (AUC) on predictions obtained from a training and test set. In order to ensure satisfactory functioning of the website, the recommendation system had to make a prediction in less time than it takes for the services database to return a query. The results show that the system complies with this requirement and that its speed does not drastically affect loading times. The interaction data that was collected in the course of several months, was split into training (70 %) and test (30 %) data set. The model that uses services only (without tags) as the input, achieved an AUC score of 0.96 on the training set and an AUC score of 0.7 on the testing set. We expect this metric to improve in due time with further maturing of the website.

6 Conclusion

This paper presented the implementation of the recommendation system in the Insieme electronic and mobile health platform. The experimental results showed the ability of this system to recommend relevant services within given time constraints. Since the recommendation model was trained on data obtained from a time period of several months, constant usage of the platform and further user interactions are expected to increase the relevance of the suggestions. Furthermore, an additional model was trained for suggestion of service tags. In both models it can be observed that semantic meaning is captured through embeddings.

Acknowledgement

The paper was supported by the ISE-EMH project funded by the program Interreg V-A Italy-Slovenia 2014-2020.

References

[1] A. M. Alharbi, N. T. Alharbi, H. M. Alharbi, and D. M. Ibrahim (2019). Patient Assistance System: A Proposed Structure. *10th International Conference on Information and Communication Systems (ICICS)*, pp. 230–233, `https://doi.org/10.1109/iacs.2019.8809136`.

[2] R. D. Croon, L. V. Houdt, N. N. Htun, G. Štiglic, V. V. Abeele, and K. Verbert (2021). Health recommender systems: systematic review. *Journal of Medical Internet Research*, vol. 23, no. 6, article no. 18035, `https://doi.org/10.2196/18035`.

[3] A. Hommersom, P. J.F. Lucas, M. Velikova, G. Dal, J. Bastos, J. Rodriguez, M. Germs, and H. Schwietert (2013). MoSHCA-my mobile and smart health care assistant. *15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013)*, pp. 188–192, `https://doi.org/10.1109/healthcom.2013.6720664`.

[4] ISE-EMH, Interreg Italia-Slovenia project, `https://www.ita-slo.eu/en/ise-emh`. Accessed July 13, 2022.

[5] M. Kula (2015). Metadata embeddings for user and item cold-start recommendations. *Proceedings of the 2nd Workshop on New Trends on Content-Based Recommender Systems co-located with 9th ACM Conference on Recommender Systems (RecSys 2015)*, vol. 1448, pp. 14–21, `https://doi.org/10.1145/2792838.2798718`.

[6] R. Manne and S. C Kantheti (2021). Application of artificial intelligence in healthcare: chances and challenges. *Current Journal of Applied Science and Technology*, vol. 40, no. 6, pp. 78–89, https://doi.org/10.9734/cjast/2021/v40i631320.

[7] J. G. D. Ochoa, O. Csiszár, and T. Schimper (2021). Medical recommender systems based on continuous-valued logic and multi-criteria decision operators, using interpretable neural networks. *BMC medical informatics and decision making*, vol. 21, no. 1, pp. 1–15, https://doi.org/10.1186/s12911-021-01553-3.

[8] Sudhanshu, N. S. Punn, S. K. Sonbhadra, and S. Agarwal (2021). Recommending best course of treatment based on similarities of prognostic markers. *International Conference on Neural Information Processing*, pp. 393–404, https://doi.org/10.1007/978-3-030-92270-2_34.

[9] T. N. T. Tran, A. Felfernig, C. Trattner, and A. Holzinger (2021). Recommender systems in the healthcare domain: state-of-the-art and research issues. *Journal of Intelligent Information Systems*, vol. 57, no. 1, pp. 171–201, https://doi.org/10.1007/s10844-020-00633-6.