




ORIGINAL

## Recommender System for E-Health

### Sistema de recomendación para la sanidad electrónica

Ahmad Abdullah Aljabr<sup>1</sup>, Kailash Kumar<sup>1</sup> 

<sup>1</sup>College of Computing and Informatics, Saudi Electronic University, Riyadh-11673, Kingdom of Saudi Arabia.

Cite as: Abdullah Aljabr A, Kailash Kumar K. Recommender System for E-Health. Data and Metadata. 2024; 3:.370. <https://doi.org/10.56294/dm2024.370>

Submitted: 29-01-2024

Revised: 21-04-2024

Accepted: 31-08-2024

Published: 01-09-2024

Editor: Adrián Alejandro Vitón-Castillo 

Corresponding author: Kailash Kumar 

#### ABSTRACT

**Introduction:** e-healthcare management services can be significantly enhanced through the implementation of recommender systems, as highlighted in various research papers. These systems, such as Healthcare Recommender Systems (HRS) and Health Care Recommender Systems (HCRS), utilize advanced algorithms and machine learning techniques to provide personalized health recommendations based on user input and medical data.

**Objective:** recommend healthcare services based on patient's state. Model healthcare information network for efficient service recommendation.

**Method:** recommends healthcare services based on patient's critical situation and requirements. Offers re-configurable healthcare workflows to medical staff. Machine learning method classification is applied using decision tree and its result is presented which reflects 70 to 75 % accuracy in predictive models which ensure that health recommender system is a full proof system.

**Result:** hospital recommender systems represent a significant advancement in healthcare, providing personalized and data-driven recommendations to patients.

**Conclusion:** the integration of recommender systems in e-healthcare management services holds great potential in improving personalized patient care, promoting health awareness, and optimizing the quality of healthcare recommendations. In this paper author analyzed and estimated the level of accuracy of recommendation systems in healthcare for personalized medical treatment. It surveys current applications, challenges, and future directions in this field.

**Keywords:** Health Care Recommender System; Machine Learning; Predictive Modeling; Classification Method; Decision Tree; Confusion Metrics; Classification Performance.

#### RESUMEN

**Introducción:** los servicios de gestión de la sanidad electrónica pueden mejorarse notablemente mediante la implantación de sistemas de recomendación, como se ha puesto de relieve en diversos trabajos de investigación. Estos sistemas, como los Healthcare Recommender Systems (HRS) y los Health Care Recommender Systems (HCRS), utilizan algoritmos avanzados y técnicas de aprendizaje automático para ofrecer recomendaciones sanitarias personalizadas basadas en las aportaciones del usuario y en datos médicos.

**Objetivo:** recomendar servicios sanitarios en función del estado del paciente. Modelar una red de información sanitaria para recomendar servicios de forma eficaz.

**Método:** recomienda servicios sanitarios basados en la situación crítica y los requisitos del paciente. Ofrece flujos de trabajo sanitarios reconfigurables al personal médico. Se aplica el método de aprendizaje automático de clasificación mediante árbol de decisión y se presentan los resultados, que reflejan una precisión del 70 al 75 % en los modelos predictivos, lo que garantiza que el sistema de recomendación sanitaria es un sistema

a toda prueba.

**Resultado:** los sistemas hospitalarios de recomendación representan un avance significativo en la atención sanitaria, ya que proporcionan recomendaciones personalizadas y basadas en datos a los pacientes.

**Conclusión:** la integración de sistemas de recomendación en los servicios de gestión de la sanidad electrónica encierra un gran potencial para mejorar la atención personalizada al paciente, fomentar la concienciación sobre la salud y optimizar la calidad de las recomendaciones sanitarias. En este artículo, el autor analiza y estima el nivel de precisión de los sistemas de recomendación en sanidad para el tratamiento médico personalizado. Analiza las aplicaciones actuales, los retos y las orientaciones futuras en este campo.

**Palabras clave:** Sistema de Recomendación en Atención Sanitaria; Aprendizaje Automático; Modelado Predictivo; Método de Clasificación; Árbol de Decisión; Métricas de Confusión; Rendimiento de la Clasificación.

## INTRODUCTION

E-healthcare management services have seen a significant rise in popularity due to the integration of recommender systems. These systems are designed to predict user preferences based on various factors such as user profiles, past evaluations, and additional learning.<sup>(1)</sup> The utilization of recommender systems in the healthcare sector has proven to be effective in enhancing health outcomes and preventing diseases. These systems leverage environmental information to support health improvement and disease prevention.<sup>(1,2)</sup> The application of recommender frameworks in e-health services aims to provide comprehensive guidance to patients anytime, anywhere, and for any health-related queries. Health information systems play a crucial role in delivering healthcare services, and recommender systems act as complementary tools in the decision-making process within healthcare services and administrations. Research in this area focuses on understanding the trends and applications of recommender systems in healthcare services, offering insights to healthcare professionals and researchers for future directions.<sup>(3,4)</sup> By analyzing current developments, challenges, and opportunities in e-health with a focus on health recommender systems, researchers aim to evaluate the accuracy using machine learning methods to develop innovative methodologies to further enhance the efficiency and effectiveness of e-healthcare management services.<sup>(5,6)</sup> The application of recommender systems in healthcare services, highlighting their ability to predict user preferences based on past evaluations and user profiles. The research aims to provide a comprehensive overview of recommender system applications in e-health services, offering insights into current developments, challenges, and opportunities in the field. By analyzing trends and applications of recommender systems in healthcare, the paper aims to guide healthcare professionals and researchers in understanding the role of these systems in decision-making processes and improving healthcare services.<sup>(7)</sup>

### Objectives of the Paper

- The paper explores the application of recommender systems in healthcare services, focusing on providing a comprehensive overview of their use in e-health service information dissemination to patients.
- By examining current developments, challenges, and opportunities in e-health, particularly in the context of health recommender systems, the paper tried to explore and estimate accuracy of predictive aspect of recommender system to identify areas for improvement and innovation in e-healthcare management services.
- To evaluate the performance of the Recommender System on real-world medical dataset particularly level of accuracy in respect of predictive outcome.
- The research contributes to enhancing the decision-making process in healthcare services and administrations by highlighting the role of health recommender systems as complementary tools for personalized recommendations based on user preferences and environmental information.

### Literature Survey

The paper by Md Alimul Haque et al. (2022) discusses the significant role of the Internet of Things (IoT) in combating the COVID-19 pandemic. IoT technologies have been pivotal in enhancing healthcare responses, enabling real-time monitoring of patients, and facilitating contact tracing through smart devices. This integration has improved data collection and analysis, leading to more informed decision-making in public health strategies. Moreover, the research highlights the effectiveness of IoT in managing resources, such as ventilators and hospital beds, which is crucial during surges in COVID-19 cases. However, challenges such as data privacy concerns and the need for robust cybersecurity measures are also noted, as these can hinder the full potential of IoT applications in healthcare. Overall, while IoT presents transformative opportunities for pandemic response, addressing its limitations is essential for maximizing its benefits in future health crises.<sup>(8)</sup>

The research paper builds upon existing literature to develop a framework for evaluating recommender systems based on user experience. It mentions that prediction algorithms play a crucial role in enhancing the accuracy of recommendation systems, which aligns with previous studies emphasizing the importance of algorithms in recommendation systems. The paper highlights the significant impact of recommendation systems on user experience, a concept supported by prior research that focuses on user satisfaction and system effectiveness in recommender systems. It integrates subjective experiences of users, such as perceived quality and recommendation variation, into the evaluation framework, drawing from literature that emphasizes the role of user perceptions in shaping their experiences with recommendation systems. The study also explores the relationship between objective system aspects and user behavior, a topic that has been addressed in previous research on user interactions with recommendation systems and the influence of system features on user choices and satisfaction levels. Additionally, the paper discusses the correlation between system aspects and user characteristics like feedback preferences, aligning with existing literature that explores how user traits and behaviors interact with system functionalities in recommender systems.<sup>(9)</sup>

The research paper presents an innovative recommender system designed to match health tourist preferences with health and tourism providers. It focuses on offering complete health tourism products by aligning user profiles with attributes of service providers, ensuring users receive their desired treatment in the right location, period, and cost. The recommender system incorporates a database of cases, including medical, wellness, and tourism service providers, each described by various attributes such as service category, cost, infrastructure, and accreditations. By applying the facility location method, the system provides a final list of recommended cases that are both similar to the user query and diverse from each other, enhancing the user's choices and overall experience.<sup>(10)</sup>

Recommender systems (RS) have become essential in various domains like e-commerce and social networks to provide personalized recommendations to users based on their preferences and behaviors. RS can use collaborative filtering, content filtering, or a hybrid approach to generate recommendations for users. Health recommender systems (HRS) are particularly valuable in healthcare services, aiding in effective decision-making and complementing traditional healthcare systems. The use of big data analytics platforms like MapReduce Framework and Hadoop technology can enhance the scalability and efficiency of HRS by processing large volumes of healthcare data. The paper discusses the importance of RS in saving time for users by reducing the number of actions needed to find relevant information, ultimately aiding in effective decision-making. By leveraging predictive analysis, RS can present targeted information to users, helping them arrive at desired conclusions efficiently.<sup>(11)</sup>

The research paper focuses on the development of a health application recommender system using partially penalized regression. It highlights the importance of behavioral intervention technologies in mobile applications for enhancing patient care and improving efficacy. The paper emphasizes the need for individualized health application recommender systems to optimize the utility of mobile health applications. The proposed approach formalizes the recommender system as a policy that maps individual subject information to a recommended app, aiming to maximize expected utility. The method introduced in the paper, PRO-aLasso, utilizes partial regularization via orthogonality and the adaptive Lasso to estimate the optimal policy for recommending health applications. The study demonstrates the convergence rate of the estimated policy to the true optimal policy and shows that the PRO-aLasso estimators exhibit oracle properties similar to the adaptive Lasso. Simulation studies presented in the paper indicate that the PRO-aLasso method produces simpler, more stable policies with improved results compared to other existing methods. The performance of the proposed method is illustrated through an example using IntelliCare mobile health applications, showcasing its effectiveness in real-world scenarios.<sup>(12)</sup>

The research paper focuses on the development of a Health Recommender System (HRS) for Cervical Cancer prognosis in women, utilizing a Multi Objective Genetic Algorithm (MOGA) for feature selection. Prior studies have shown the significance of utilizing recommender systems in the healthcare domain to aid in disease prediction and decision-making for patients. Existing literature highlights the importance of leveraging digital data in the health domain to enhance patient-oriented decision-making and improve disease prognosis. Studies have emphasized the role of feature selection methods, such as genetic algorithms, in improving the accuracy and effectiveness of prediction models in healthcare applications. The paper contributes to the existing body of literature by proposing a novel approach that combines feature selection techniques with prediction models to provide personalized information and support for women in understanding their clinical status related to Cervical Cancer.<sup>(12)</sup>

Bernd Blobel et al. discussed a context-aware recommender system focusing on improving daily health activities by considering user information, interests, time, location, and regular exercise.<sup>(13)</sup> Ekoyot Kaur et al. presented a user context model based on usability attributes for implementing mobile applications in the healthcare domain, specifically evaluating users undergoing peritoneal dialysis (PD). Andre C.M. Costa et al. introduced CORES, a context-aware ontology-based recommender system for service recommendations, emphasizing personalized and proactive services in an infra-aware context. Agusti Solanas et al. explored the

concept of smart health in content-aware recommendation systems for mobile health in smart cities, highlighting intelligent health representation, challenges, and research opportunities. Sergio Ilarri et al. discussed the use of various sensors to enhance recommendation accuracy in healthcare, e-learning, and mobile computing, addressing open issues in context-aware recommendation systems when utilizing sensors.<sup>(14)</sup>

**METHOD**

For estimating the accuracy in predictive model for health care recommender system, a real health data set is collected with the help of kaggle dataset, and machine learning method classification is applied using decision tree and its result is presented which reflects 70 to 75 % accuracy in predictive models which ensure that health recommender system is a full proof system. The whole process is described as below.

*Dataset:* Heart\_Disease\_Prediction.csv

*Data Source:* Machine Learning Repository of University of California Irvine’s

*Tool used:* Rapid miner version 7.3 for windows

*Details of Dataset:* This is a labeled dataset, it is having 270 records and 14 features of patients having heart disease and other associative diseases.

**Features of the data set includes**

1. Age, 2. Sex (1-M, 0-F), 3. Chest pain Type (1 to 4), 4. BP sys(continuous), 5. Cholesterol (continuous), 6. FBS >120 (1-yes, 0-no), 7. EKG/ECG (0-normal, 1/2 abnormal), 8. Max HR/Pulse (continuous), 9. Exercise Angina (0-N, 1-Y), 10. ST depression (continuous), 11. Slope of ST (1-3), 12. No. of vessels fluro (count of main arteries block) , 13. Thallium (Toxicity level in blood), 14. Heart disease (Binominal Y or N)

For this study 14<sup>th</sup> attribute Heart disease is taken as target variable or dependent variable and rest all attributes (features) are considered as independent variables. The data set is in already well normalized so there is no pre-processing is required.

Table 1. (View of main data set)

Age	Sex	Chest pain type	BP	Cholesterol	FBS over 120	EKG results	Max HR	Exercise angina	ST depression	Slope of ST	Number of vessels fluro	Thallium	Heart Disease
70	1	4	130	322	0	2	109	0	2,4	2	3	3	Presence
67	0	3	115	564	0	2	160	0	1,6	2	0	7	Absence
57	1	2	124	261	0	0	141	0	0,3	1	0	7	Presence
64	1	4	128	263	0	0	105	1	0,2	2	1	7	Absence
74	0	2	120	269	0	2	121	1	0,2	1	1	3	Absence
65	1	4	120	177	0	0	140	0	0,4	1	0	7	Absence
56	1	3	130	256	1	2	142	1	0,6	2	1	6	Presence
59	1	4	110	239	0	2	142	1	1,2	2	1	7	Presence
60	1	4	140	293	0	2	170	0	1,2	2	2	7	Presence
63	0	4	150	407	0	2	154	0	4	2	3	7	Presence
59	1	4	135	234	0	0	161	0	0,5	2	0	7	Absence
53	1	4	142	226	0	2	111	1	0	1	0	7	Absence
44	1	3	140	235	0	2	180	0	0	1	0	3	Absence
61	1	1	134	234	0	0	145	0	2,6	2	2	3	Presence
57	0	4	128	303	0	2	159	0	0	1	1	3	Absence
71	0	4	112	149	0	0	125	0	1,6	2	0	3	Absence
46	1	4	140	311	0	0	120	1	1,8	2	2	7	Presence
53	1	4	140	203	1	2	155	1	3,1	3	0	7	Presence
64	1	1	110	211	0	2	144	1	1,8	2	0	3	Absence

**Process steps**

Step 1: Imported the data set into Rapid miner design space and selected all 270 records and 14 columns.

Step 2: Created the process SET ROLE and specified the attribute (Heart disease) as target variable (Label) and connected the output port of data set into the input port of this process SET ROLE.

Step 3: Created the process SPLIT DATA and bifurcated the data set into two partitions, first partition specified as training set with 70 % data records and second partition as test set with 30 % data records and connected the output port of process SET ROLE into the input port of this process SPLIT DATA.

Step 4: Created the process DECISION TREE and by keeping the parameter as default, connected the label data (70 %) output port of process SPLIT DATA into the training input port of this process DECISION TREE.

Step 5: Created the test process APPLY MODEL and kept the parameters as default, then connected the

model output port of the process DECISION TREE into the model input port of this process and also connected an label data (30 %) output port of the process SPLIT DATA , into an label input port of this process APPLY MODEL, as because this process needs two input one from Decision Tree and another from Split Data.

Step 6: In this step we have created a model evaluation process that is CLASSIFCATION PERFORMANCE to measure the accuracy of predictive model using performance metrics which is called confusion metrics. Then connected the label output port of APPLY MODEL process into the label input port of this process.

Step 7: Finally, in this last step we have to connect required outputs into result port. So we have connected the performance output port and example output port of the process CLASSIFCATION PERFORMANCE, both into the result port. And also we have connected the model output port of the process APPLY MODEL into the result port.

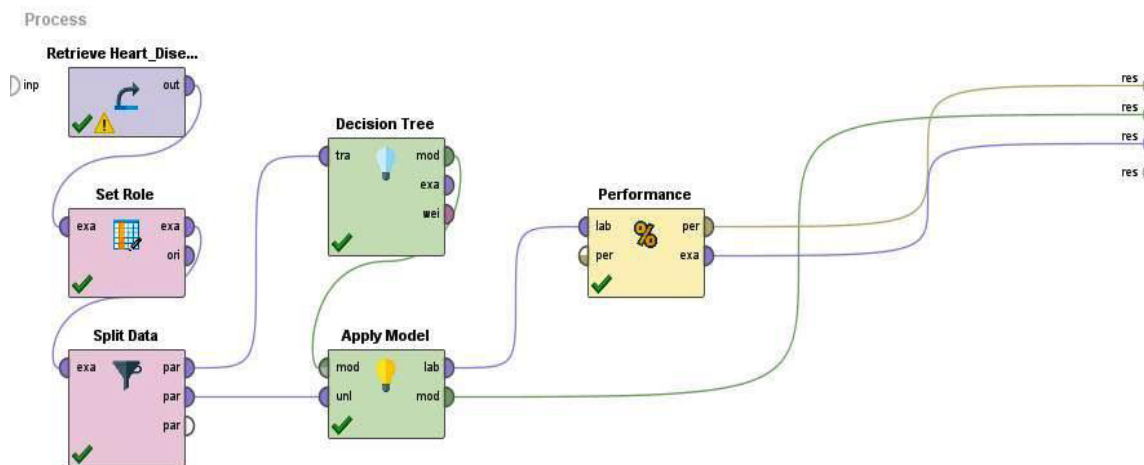


Figure 1. View of the Processes design in Rapid miner

Therefore after creations and configuration of all above processes we have RUN the overall integrated process and received three types of results which are, the model (Tree), the predicted example and the performance, all these three are presented below.

**RESULTS AND DISCUSSION**

First result is Decision Tree; this is the predictive model generated by classification algorithm based on the training data set (70 %) provided. In this tree leaf nodes in blue represent presence of heart disease and red leaf nodes represent absence of heart disease. Root is generated based on gain ratio. However, there are some nodes which are partially blue and red which are misplaced nodes.

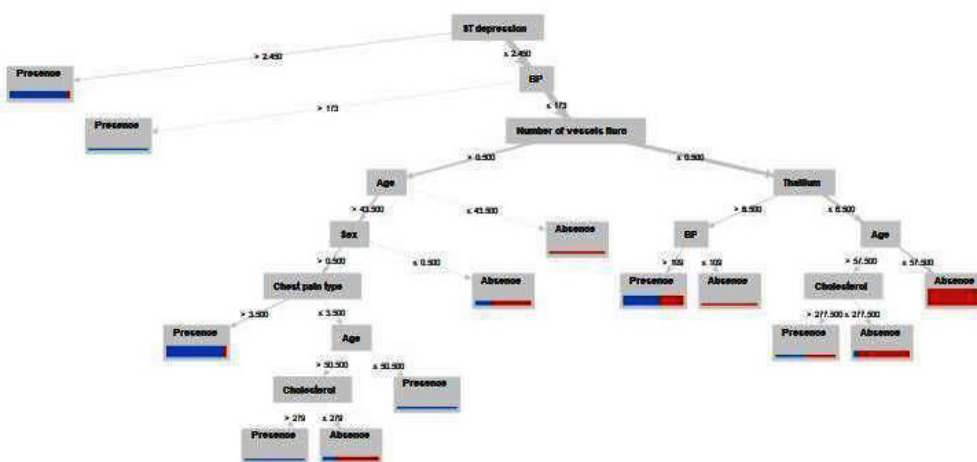


Figure 2. Decision Tree representing the predictive model

The second result is the example sheet which represents the predicted values against actual values. In this



output predicted values are grouped under confidence of Presence and confidence of Absence, it is decided on the basis of standard thresh hold value 0,5, if predicted value for Presence is greater than 0,5 then it is labeled as Heart diseases (presence), similarly if the predicted value for Absence is greater than 0,5 then it is labeled as Heart diseases (absence).

Row No.	Heart Disease	prediction(Heart Disease)	confidence(Presence)	confidence(Absence)	Age	Sex	Chest pain L..
2	Presence	Absence	0.222	0.778	56	1	3
3	Presence	Presence	0.941	0.059	60	1	4
4	Presence	Presence	0.952	0.048	63	0	4
5	Absence	Absence	0.059	0.941	64	1	1
6	Absence	Absence	0	1	47	1	4
7	Absence	Absence	0.273	0.727	54	0	2
8	Absence	Absence	0	1	51	0	3
9	Absence	Absence	0.273	0.727	71	0	3
10	Absence	Absence	0.059	0.941	66	1	4
11	Absence	Absence	0	1	37	0	3
12	Presence	Presence	0.952	0.048	59	1	4
13	Presence	Presence	0.941	0.059	61	1	4
14	Presence	Absence	0.059	0.941	59	1	1

Figure 3. Predicted sample values against actual values

Below is the bar graph generated on the basis of Age Vs Confidence(presence), which is showing chances of heart diseases grows with the age of patients. X - axis is confidence(presence) and Y-axis is Age.

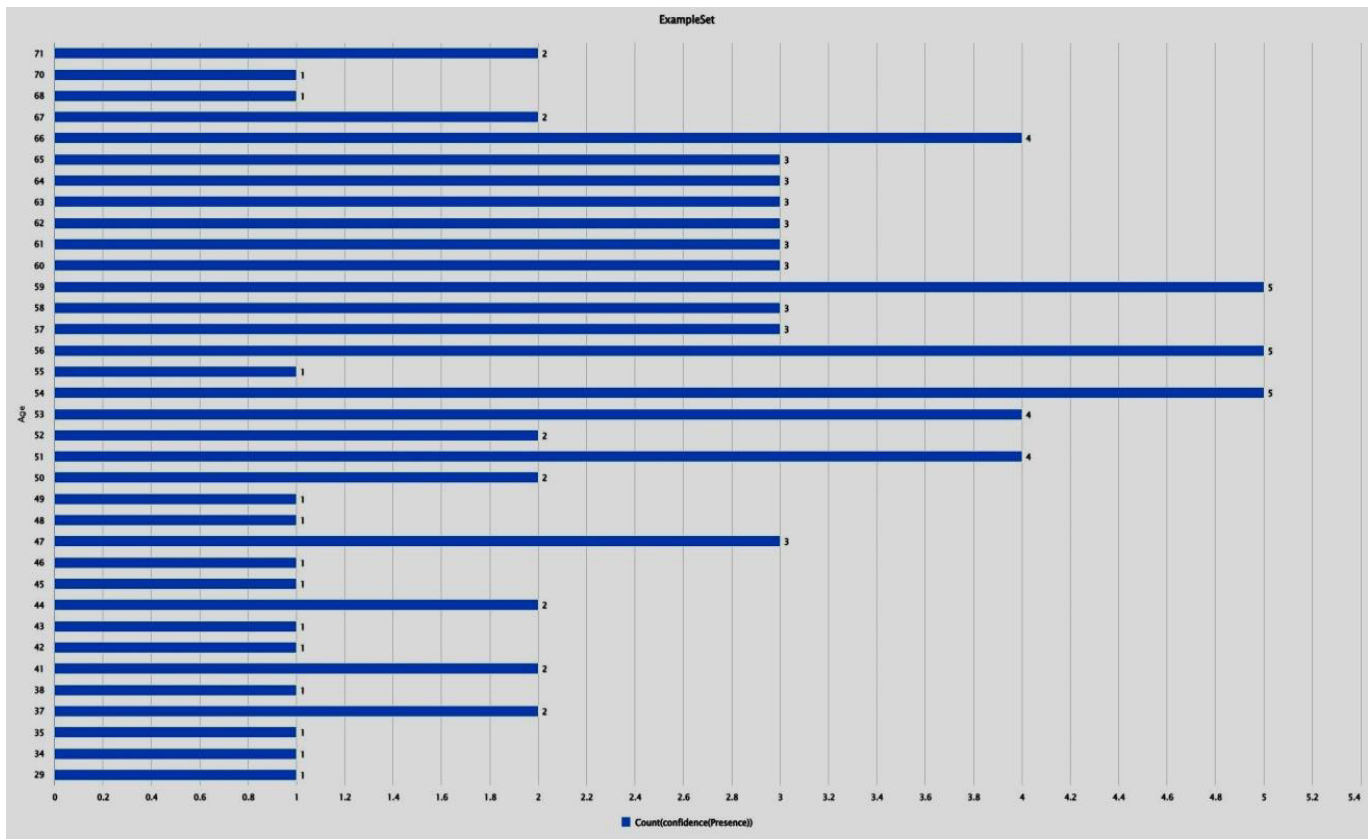


Figure 4. Age vs Confidence (Presence) of heart disease

Third result is performance evaluation of the predictive model. It is measured using classification performance metrics which is also called confusion metrics. The outcome the metrics is presented as below.

**accuracy: 74.07%**

	true Presence	true Absence	class precision
pred. Presence	26	11	70.27%
pred. Absence	10	34	77.27%
class recall	72.22%	75.56%	

**Figure 4.** Performance measured of the model (Confusion metrics)

Above two by two metric is presenting class wise accuracy and overall accuracy as percentage. As we see that True presence is predicted with 72,22 % with 10 records misplaced or wrongly predicted, similarly True absence is predicted with 75,56 % with 11 records misplaced or wrongly predicted. Therefore, the diagonal values 26(TP) and 34(TA) is reflecting accuracy of the model so the overall accuracy 74,07 % is assuring the success of the model.

## CONCLUSION

It is justified that health care predictive applications (Hospital recommender systems) build on the aforesaid predictive model then it will deliver approx. 75 % accurate prediction for the patients whether they have a potential risk of heart disease or not. Hospital recommender systems represent a significant advancement in healthcare, providing personalized and data-driven recommendations to patients. By leveraging sophisticated algorithms and robust system architecture, these systems can enhance patient satisfaction and improve healthcare outcomes. Future research should focus on addressing challenges related to data privacy, quality, and algorithmic bias to further refine and optimize hospital recommender systems.

## ACKNOWLEDGMENT

The authors extend their appreciation to the Deanship of Scientific Research at Saudi Electronic University for funding this research (9521).

## REFERENCES

1. F. Ricci, L. Rokach, and B. Shapira, "Introduction to recommender systems handbook," in *Recommender systems handbook*, Springer, 2010, pp. 1-35.
2. S. Zeba, M. A. Haque, S. Alhazmi, and S. Haque, "Advanced Topics in Machine Learning," *Mach. Learn. Methods Eng. Appl. Dev.*, p. 197, 2022.
3. M. D. Ekstrand, J. T. Riedl, and J. A. Konstan, "Collaborative filtering recommender systems," *Found. Trends® Human-Computer Interact.*, vol. 4, no. 2, pp. 81-173, 2011.
4. V. Whig, B. Othman, A. Gehlot, M. A. Haque, S. Qamar, and J. Singh, "An Empirical Analysis of Artificial Intelligence (AI) as a Growth Engine for the Healthcare Sector," in *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, IEEE, 2022, pp. 2454-2457.
5. E. Zangerle and C. Bauer, "Evaluating recommender systems: survey and framework," *ACM Comput. Surv.*, vol. 55, no. 8, pp. 1-38, 2022.
6. M. A. Haque, S. Ahmad, D. Sonal, S. Haque, K. Kumar, and M. Rahman, "Analytical Studies on the Effectiveness of IoMT for Healthcare Systems," *Iraqi J. Sci.*, pp. 4719-4728, 2023.
7. P. Kanchanadevi, D. Selvapandian, L. Raja, and R. Dhanapal, "Cloud-based protection and performance improvement in the e-health management framework," in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*, IEEE, 2020, pp. 268-270.
8. M. A. Haque, D. Sonal, S. Haque, K. Kumar, and M. Rahman, "The Role of Internet of Things (IoT) to Fight against Covid-19," in *Proceedings of the International Conference on Data Science, Machine Learning and*

Artificial Intelligence, New York, NY, USA: ACM, Aug. 2021, pp. 140-146. doi: 10.1145/3484824.3484900.

9. Z. Munawar, N. Suryana, Z. B. Sa'aya, and Y. Herdiana, "Framework With An Approach To The User As An Evaluation For The Recommender Systems," in 2020 Fifth International Conference on Informatics and Computing (ICIC), IEEE, 2020, pp. 1-5.

10. A. Panteli, A. Kompothrekas, C. Halkiopoulos, and B. Boutsinas, "An innovative recommender system for health tourism," in Culture and Tourism in a Smart, Globalized, and Sustainable World: 7th International Conference of IACuDiT, Hydra, Greece, 2020, Springer, 2021, pp. 649-658.

11. R. Bateja, S. K. Dubey, and A. Bhatt, "Health recommender system and its applicability with MapReduce framework," in Soft Computing: Theories and Applications: Proceedings of SoCTA 2016, Volume 2, Springer, 2018, pp. 255-266.

12. E. J. Oh, M. Qian, K. Cheung, and D. C. Mohr, "Building health application recommender system using partially penalized regression," Stat. Model. Biomed. Res. Contemp. Top. Voices F., pp. 105-123, 2020.

13. J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," Knowledge-based Syst., vol. 46, pp. 109-132, 2013.

14. M. S. Kumar and J. Prabhu, "Healthcare recommender system based on Smart-health Routes," in International Conference for Phoenixes on Emerging Current Trends in Engineering and Management (PECTEAM 2018), Atlantis Press, 2018, pp. 42-46.

#### **FUNDING**

Research funding by Deanship of Scientific Research at Saudi Electronic University (9521).

#### **AVAILABILITY OF DATA AND MATERIALS**

The datasets used in this research are publicly available and properly cited in our dataset section for transparency and ease of replication.

#### **CONFLICT OF INTEREST**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### **AUTHOR CONTRIBUTIONS**

*Conceptualization:* Kailash Kumar, Ahmad Abdullah Aljabr.

*Investigation:* Kailash Kumar, Ahmad Abdullah Aljabr.

*Methodology:* Kailash Kumar, Ahmad Abdullah Aljabr.

*Writing - original draft:* Kailash Kumar, Ahmad Abdullah Aljabr.

*Writing - review and editing:* Kailash Kumar, Ahmad Abdullah Aljabr.