UNIVERSITY "ST. KLIMENT OHRIDSKI" - BITOLA
FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGIES - BITOLA
REPUBLIC OF NORTH MACEDONIA

Proceedings of the 13th International Conference on Applied Internet and Information Technologies

AIIT 2023

13 October, 2023, Bitola, Republic of North Macedonia

AIIT
International Conference

UNIVERSITY "ST. KLIMENT OHRIDSKI" - BITOLA
FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGIES - BITOLA
REPUBLIC OF NORTH MACEDONIA

13TH INTERNATIONAL CONFERENCE ON APPLIED INTERNET AND INFORMATION TECHNOLOGIES

AIIT 2023 PROCEEDINGS

Bitola, 2023

Proceedings publisher and organizer of the conference:
University “St. Kliment Ohridski” - Bitola, Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia

For publisher:
Blagoj Ristevski, PhD
Faculty of Information and Communication Technologies - Bitola
University “St. Kliment Ohridski” - Bitola, North Macedonia

Proceedings editors:
Kostandina Veljanovska, PhD
Eleonora Brtka, PhD

Conference Chairmans:
Kostandina Veljanovska, University “St. Kliment Ohridski”, Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia - chair
Eleonora Brtka, University of Novi Sad, Technical Faculty “Mihajlo Pupin”, Zrenjanin, Serbia – co-chair

Technical preparation of the proceedings:

Andrijana Bocevska, PhD

Kostandina Veljanovska, PhD

Cover design:

Andrijana Bocevska, PhD

e-Proceedings
ISBN 978-608-5003-03-7

CIP - Каталогизација во публикација
Национална и универзитетска библиотека "Св. Климент Охридски", Скопје

004-049.8(062)

INTERNATIONAL conference on applied internet and information technologies AIIT 2023 (13 ; 2023 ; Bitola, Republic of North Macedonia)
Proceedings / 13th International conference on applied internet and information technologies AIIT 2023, 13 October, 2023, Bitola, Republic of North Macedonia ; [editors Kostandina Veljanovska, Eleonora Brtka]. - Bitola : University "St. Kliment Ohridski", Bitola Faculty of information and communication technologies, 2023. - 411 стр. : илустр. ; 30 см
Библиографија кон трудовите
ISBN 978-608-5003-03-7

а) Информатичка технологија -- Примена -- Собири
COBISS.MK-ID 62559493

Introduction

The International conference on Applied Internet and Information Technologies is a traditional meeting held every year, that sprouts out of collaboration between the University of Novi Sad, Technical Faculty “Mihajlo Pupin”, Zrenjanin, Serbia and the University “St. Kliment Ohridski”, Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia. The XIII AIIT2023 was held in Bitola, Macedonia on which besides the participants from Serbia and Macedonia there were researchers from Croatia, Bosnia and Herzegovina, Hungary, Finland, Russia, Turkey, Egypt, India and Australia whose contribution was either as authors or as reviewers of the papers.

At the Conference were presented innovative findings in the field of information systems, communications and computer networks, software engineering and applications, data science and big data technologies, artificial intelligence, intelligent systems, business intelligence and IT support to decision-making, data and system security, distributed systems, Internet of Things and smart systems, embedded systems, computer graphics, IT management, e-commerce, e-government, e-education, Internet marketing, and IT practice and experience.

The Conference chairs would like to express gratitude to the authors for their contributions and to express special gratitude to the reviewers for their tremendous work done for selecting the papers with their valuable comments and suggestions that contributed to improve the quality of the papers. Out of more than 60 submitted papers, 51 were selected, presented at the Conference and are published in this proceedings.

The work during the conference was organized in nine sessions: plenary session, five in-person oral sessions, one video session and two poster sessions. During the conference, a round table with participants from academic organizations and IT industry was successfully organized. The theme of the discussions at the round table was "Strengthening the capacities of Faculty of ICT for the realization of strategic cooperation with companies from the IT industry".

AIIT 2023 was very successful conference with fruitful exchange of experiences among the participants reviving the hope of further strengthening a friendly environment after the pandemic crisis. We hope that we will continue with the contribution to the further deepening the development of Internet and information technologies research.

Conference chairs:

Kostandina Veljanovska, University “St. Kliment Ohridski”, Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia (chair)
Eleonora Brtka, University of Novi Sad, Technical Faculty “Mihajlo Pupin”, Zrenjanin, Serbia (co-chair)

Conference Chairs

Kostandina Veljanovska, University “St. Kliment Ohridski”, Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia

Kostandina Veljanovska, Ph.D. finished BSc in Computer Science at the University "Sts. Kiril i Metodi", Skopje. Her first MASc in Applied Engineering she received at the University of Toronto, Toronto, Canada. Her second MSc and also her PhD in Technical Sciences she received at the University “St. Kliment Ohridski” - Bitola, R. Macedonia. Her postdoctoral studies in Artificial Intelligence she attended at the Laboratory of Informatics, Robotics and Microelectronics at the University of Montpellier, Montpellier, France. She worked as a Research assistant at the Faculty of Applied Science, University of Toronto, Canada. She also, worked as a researcher in research team for Constraints, Learning and Agents at LIRMM, University of Montpellier. Since 2008, she works as a Full Professor in Information Systems and Networks, Artificial Intelligence and Systems and Data Processing at the Faculty of Information and Communication Technologies, University “St. Kliment Ohridski” - Bitola, Rebublic of North Macedonia. Her research work is focused on artificial intelligence, machine learning techniques and intelligent systems. She has published numerous scientific papers in the area of interest, as well as several monographic items. She is a reviewing referee for well-known publishing house, journals with significant impact factor in science and also, member of editorial board of several international conferences.

Eleonora Brtka, University of Novi Sad, Technical Faculty “Mihajlo Pupin”, Zrenjanin, Serbia

Eleonora Brtka, Ph.D. is an associate professor at the Information Technology department at the University of Novi Sad, Technical Faculty „Mihajlo Pupin“, Zrenjanin. She has received her PhD in Information technology in 2015. Current research interests include: data science techniques and methods, programming languages, education, artificial intelligence and intelligent agent technologies. She is the author or co-author of several articles published in international journals and in the proceedings of international conferences. She participated in several national funded projects.

Organizing Committee

Chairs

Kostandina Veljanovska (chair), Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia
Eleonora Brtka, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia (co-chair)
Zoran Kotevski, Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia (co-chair)

Members

Željko Stojanov, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia
Blagoj Ristevski, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Višnja Ognjenović, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia
Dalibor Dobrilovic, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia
Monika Markovska, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Dragica Radosav, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Biljana Radulovic, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia
Božidar Milenkovski, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Ivana Berković, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Ljubica Kazi, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Snežana Savoska, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Vladimir Brtka, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Zoltan Kazi, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Siniša Mihajlović, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Velibor Premčevski, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Nikola Rendevski , Faculty of Information and Communication Technologies - Bitola, North Macedonia
Aleksandra Stojkov, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Maja Gaborov, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Andrijana Bocevska, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Milica Mazalica, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Igor Vecštejn, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Tome Dimovski, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Marko Blažić, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Vuk Amižić, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Natasa Blazeska-Tabakovska, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Jovana Borovina, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Dalibor Šeljmeši, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Mimoza Bogdanoska-Jovanovska, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Vladimir Šinik, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Nadežda Ljubojev, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Marina Blažeković Toshevski, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Vladimir Karuović, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Evgeny Cherkashin, Institute of System Dynamic and Control Theory SB RAS, Russia
Anastasia Popova, Matrosov Institute for System Dynamics and Control Theory of the Siberian Branch of the Russian Academy of Sciences, Russia
Filip Tsvetanov, South-west University "Neophyte Rilsky", Faculty of Engineering, Blagoevgrad, Bulgaria

Program Committee

Blagoj Ristevski (president), Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Eleonora Brtka, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia (vice-president)
Željko Stojanov, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Višnja Ognjenović, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Dalibor Dobrilović, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Ljubica Kazi, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Dragica Radosav, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Dragana Glušac, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Borislav Odadžić, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Miodrag Ivković, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Biljana Radulović, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Ivana Berković, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Vladimir Brtka, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Zoltan Kazi, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Jelena Stojanov, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Vesna Makitan, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Nadežda Ljubojev, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Vladimir Šinik, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Igor Nedelkovski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Aleksandar Markoski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Violeta Manevska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Pece Mitrevski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Ilija Jolevski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Dragan Gruevski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Kostandina Veljanovska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Monika Markovska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Snežana Savoska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Ramona Markoska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Sonja Mančevska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Mimoza Bogdanoska-Jovanovska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Nataša Blažeska Tabakovska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Božidar Milenkovski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Zoran Kotevski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Nikola Rendevski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Andrijana Bocevska, Faculty of Information and Communication Technologies, University "St. Kliment

Dana Petcu, West University of Timisoara, Romania
Marius Marcu, "Politehnica" University of Timisoara, Romania
Aleksej Stevanov, South-west University "Neophyte Rilsky", Faculty of Engineering, Blagoevgrad, Bulgaria
Petar Apostolov, South-west University "Neophyte Rilsky", Faculty of Engineering, Blagoevgrad, Bulgaria
Filip Tsvetanov, South-west University "Neophyte Rilsky", Faculty of Engineering, Blagoevgrad, Bulgaria
Francesco Flammini, School of Innovation, Design and Engineering, Division of Product Realisation, Mälardalen University, Eskilstuna, Sweden
Deepak Chahal, Jagan Institute of Management Studies (JIMS, Rohini Sector-5), New Delhi, India
Abdel-Badeeh M. Salem, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt
Dragan Peraković, University of Zagreb, Faculty of Transport and Traffic Sciences, Zagreb, Croatia
Gordana Jotanović, University of East Sarajevo, Faculty of Transport and Traffic Engineering, Doboj, Bosnia and Herzegovina
Goran Jauševac, University of East Sarajevo, Faculty of Transport and Traffic Engineering, Doboj, Bosnia and Herzegovina
Dinu Dragan, Faculty of technical Sciences, University of Novi Sad, Serbia
Gururaj Harinahalli Lokesh, Department of IT, Manipal Institute of Technology, Bengaluru, India
Ertuğrul AKBAŞ, Esenyurt University, Istanbul, Turkiye

C O N T E N T

Invited Speakers

Regular Papers

Machine Learning Algorithms for Heart Disease Prognosis using IoMT Devices

Anita Petreska[1], Blagoj Ristevski[1] Daniela Slavkovska[2,] Saso Nikolovski[1], Pero Spirov[2], Nikola Rendevski[1], Snezhana Savoska[1]

[1] *Faculty of Information and Communication Technologies, Bitola, University "St. Kliment Ohridski"- Bitola, Republic of Macedonia*
[2] *UTMS Skopje, Faculty of Computer Science, Republic of North Macedonia,*

petreska.anita@uklo.edu.mk, blagoj.ristevski@uklo.edu.mk, dslavkovska@yahoo.com, saso.nikolovski@ uklo.edu.mk, spirov.pero@gmail.com, nikola.rendevski@uklo.edu.mk, snezana.savoska@uklo.edu.mk

Abstract:
Artificial intelligence (AI) and Machine learning (ML) algorithms have improved the capabilities of Internet of Medical Things (IoMT) systems in the direction of better quality of life and efficient healthcare. The paper analyzes various IoMT technologies and the classification, architecture, and communication of IoMT. The paper presents the application of machine learning algorithms for the detection of anomalies crucial in the process of prognosis of heart diseases. Cardiovascular disease is the primary global cause of death, prompting increased interest in leveraging artificial intelligence to analyze data obtained from wearable devices. The methodology for predicting heart disease risk using IoMT data includes data collection, pre-processing, and application of machine learning algorithms. A comparative evaluation of five machine learning models: logistic regression, Support Vector Machine (SVM), Decision Tree, Random Forest, and k-nearest neighbors (KNN) was conducted. The purpose of this paper is to emphasize the role of ML in the field of cardiology and the critical importance of data quality, as well as the selection of an appropriate algorithm in order to improve cardiovascular risk assessment. The introduction of ML in the prognosis of heart disease is a significant step towards the realization of predictive, preventive, and personalized health care and the reduction of health care costs.

Keywords:
Internet of Medical Things, machine learning, artificial intelligence, medical Big Data, exploratory data analysis

1. Introduction

The Internet of Medical Things are devices that continuously collect and transmit health data in real-time, enabling early disease detection and personalized patient care. Heart disease is a global health concern and its accurate prediction is of paramount importance. This paper explores the application of machine learning algorithms to cardiac disease prognosis using data collected with the IoMT.

The second part looks at the role of IoMT in the healthcare sector, the architecture, communication, and management of IoMT devices, and defines the technologies used and their security.

Section 3 presents the motivation for introducing ML in healthcare with special reference to commonly used ML algorithms for detecting patterns and anomalies crucial in the process of heart disease prognosis modeling.

In Section 4, five machine learning algorithms are analyzed to detect patterns and anomalies that are key to detecting and diagnosing heart disease risk. The methodology used, data collection, data pre-processing, exploratory data analysis [9], and metrics for evaluation and comparison of results are presented. Section 5 discusses the limitations faced by ML algorithms used to detect heart disease. Section 6 discusses the conclusions and future challenges of ML algorithms for heart disease prediction.

2. The role of IoMT in healthcare, architectural framework, technologies, and security

The IoMT is a network of medical devices, sensors, and software applications connected over the Internet. This network collects and delivers real-time data related to patient health, providing healthcare professionals with invaluable real-time insights into patient conditions enabling personalized healthcare interventions [1]. IoMT encompasses different categories of devices: wearable, implantable, ingestible, stationary, diagnostic, therapeutic, assistive, smart, industrial, emergency, telemedicine, and prosthetic [2]. IoMT devices facilitate continuous, remote patient monitoring, and dynamic data-driven insights, enabling healthcare professionals to detect subtle changes in a patient's condition and adjust interventions accordingly. A core feature of the transformative potential of IoMT is the integration of multimodal data, providing healthcare providers with a comprehensive perspective on patient health and thereby facilitating more effective diagnostic and treatment strategies.

The IoMT architecture consists of a synergy of sensors and the IoMT Gateway, which work together to enable the collection, transmission, analysis, and storage of vital health data. Wireless communication serves as the key, with technologies such as Bluetooth, Wi-Fi, ZigBee, and cellular networks offering mobility and flexibility, while near-field communication (NFC) and radio frequency identification (RFID) facilitate contactless communication [3]. Standardized communication protocols such as Health Level 7 (HL7) and Digital Imaging and Communications in Medicine (DICOM) ensure seamless interoperability and data exchange between IoMT devices and Electronic Health Records (EHR). IoMT management is segmented across different layers, each with a distinct function.

In the field of technology and security [4], the IoMT uses a comprehensive range of tools and measures: cloud computing, blockchain technology, big data analytics, interoperability standards, edge computing, artificial intelligence, machine learning, blockchain technology, and virtual and augmented reality. , medical device integration (MDI) and security measures underpin the IoMT landscape [5], [6]. Cloud computing assumes a key role, offering scalable storage and computing resources [7]. Blockchain technology protects the security and privacy of data, creating a decentralized database [8]. Artificial intelligence identifies patterns and improves patient care. Virtual and augmented reality technologies contribute to professional training, remote consultation with patients, and monitoring of treatment progress.

Security measures [9] and [10] are imperative due to the sensitivity of medical data. Encryption, authentication, access control, physical security, vulnerability assessments, and timely remediation are necessary to protect patient data and system integrity [11], [12].

3. Motivation for using ML to detect heart disease

The motivation for using ML in cardiac disease prognosis within the IoMT is multifaceted. First, heart disease is a leading cause of morbidity and mortality worldwide, requiring early diagnosis and intervention. ML algorithms offer the potential to analyze vast datasets, spanning diverse clinical and physiological attributes to uncover complex patterns and dependencies that may not be apparent through conventional methods.

Heart disease encompasses a range of conditions affecting the heart and blood vessels that can lead to a variety of health problems, including narrowing of the blood vessels, chest pain, stroke, and heart attack. The main causes of heart disease are diabetes, obesity, unhealthy diet, increased weight, excessive alcohol use, and physical inactivity. Early prediction of heart disease is paramount for patients and healthcare providers. Early identification allows health professionals to implement preventive measures, effective diagnosis and treatment, and patients valuable insights into their health. Machine learning plays a key role in identifying and predicting heart disease. Machine learning algorithms applied to relevant medical data serve as powerful tools for identifying patients at risk of heart disease before symptoms become apparent.

Numerous AI algorithms and models use the data generated by wearable sensor devices. In the domain of establishing diagnostic and predictive models using data from wearable devices, classical machine learning and deep learning techniques are of paramount importance. The development of appropriate algorithms and models tailored to specific categories of heart disease remains imperative.

By using machine learning algorithms, healthcare providers and patients can work together to detect heart disease in its early stages, enabling timely interventions and informed decisions.

Deep learning is a powerful tool in the field of heart disease prediction that uses multilayer neural network architectures to automatically learn and extract complex patterns and features from complex medical data. Convolutional Neural Networks (CNNs) are used to automatically detect anomalies and irregularities in ECG signals, helping in the early detection of cardiac conditions such as arrhythmias [13]. Convolutional neural networks (CNNs) are applied to medical imaging data, including heart scans (echocardiograms, MRIs, CT scans). They help in the automatic interpretation of the images, helping in the diagnosis of structural problems of the heart. Recurrent neural networks (RNNs) and their specialized variant [14], long-short-term memory (LSTM) networks excel at modeling temporal sequences [15]. They can capture patterns and dependencies in time series data, making them invaluable in predicting heart attacks and strokes.

4. Machine learning algorithms for prognosis of heart diseases

Machine learning uses different algorithms and models to predict heart disease. These algorithms analyze medical data, such as patient records, medical images, and diagnostic tests, to generate predictive insights. Some of the most common machine learning algorithms for heart disease prediction include logistic regression, random forests, support vector machines, neural networks, gradient boosting, and K-nearest neighbors, among others. These algorithms take into account multiple factors, including a patient's medical history, genetics, lifestyle, and environmental factors to estimate the risk of developing heart disease.

Efficient machine learning models have been developed for the prognosis of heart disease. Five machine learning algorithms were used in our study: logistic regression, SVM (Support Vector Machines), Decision Tree, Random Forest, and KNN [16].

Logistic regression is used to predict the probability of heart disease based on a combination of clinical and physiological factors, as input characteristics and the binary outcome of the presence (1) or absence (0) of heart disease.

Support Vector Machine is a powerful machine learning algorithm that recognizes complex patterns in medical data, particularly when classifying individuals into different risk categories for heart disease. Support vector machine is also efficient for high-dimensional data and non-linear relationships between them, making it excellent for capturing complex dependencies in medical data.

The Decision Tree algorithm builds on a tree-like hierarchical structure that efficiently evaluates data, aiding early detection and personalized management of heart disease.

Random Forest combines the predictions of many decision trees, providing a robust and accurate way to estimate heart disease risk. Random Forest excels in dealing with complex and high-dimensional medical data by being able to capture complex relationships in this data, which cannot be done with simple methods. Random Forest not only increases the prediction accuracy and improves the generalization of the model, but also can reduce overfitting.

The K-Nearest Neighbors (KNN) algorithm is an intuitive method for classifying individuals into risk categories based on their similarity to other patients. KNN although simple and flexible can detect complex relationships in data, making it a good choice for heart disease risk prediction.

4.1. Methodology

The purpose of this study is to determine whether a patient will experience a heart attack based on the data collected by the IoMT device. For this purpose, machine learning methods have been proposed that would be used by doctors in order to diagnose heart diseases more easily.

The methodology includes a series of procedures that take place in 3 phases. In the first phase, data from IoMT devices is collected and prepared for processing. In the second stage, preprocessing is performed, which includes analysis of missing values, cleaning, and standardization. The third stage involves applying a classifier to build an appropriate machine learning model. The proposed model uses 5 machine learning algorithms: logistic regression, SVM, Decision Tree, Random Forest and KNN.

The process of heart disease risk prediction is shown in Figure 1. Data are loaded from Kaggle database [17], cleaned, significant features are extracted. The database is divided into training and testing sets with the ratio (80% and 20%). The data from the five considered machine learning models are trained and tested and finally the results are evaluated and compared in order to evaluate the efficiency of the proposed methodology by evaluating the models and determining the accuracy, cross-validation result, ROC_AUC result.

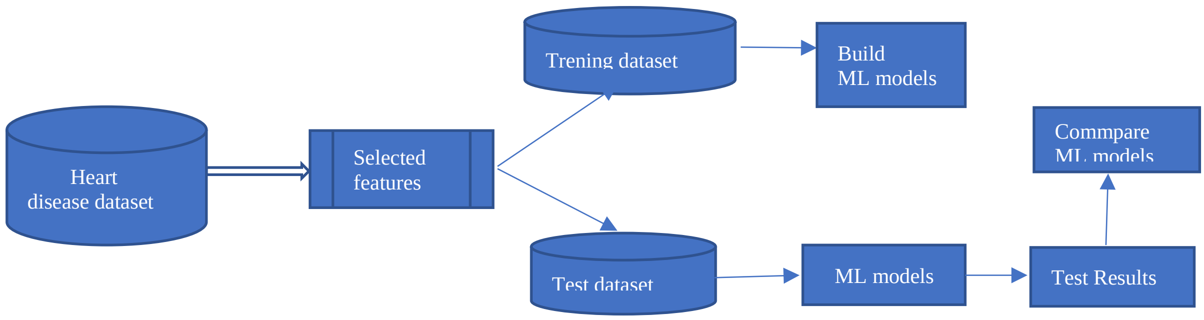


Figure 1: Flowchart of ML algorithms for heart disease prediction

4.2. Data collection

In our study, a database downloaded from Kaggle is used, which consists of 303 queues containing 14 attributes of which 13 are input attributes and one target attribute indicating the presence or absence of heart disease. All 14 attributes shown in Table 1 contain demographic data and clinical parameters that cover a different aspect of an individual's health profile.

Table 1:
Heart prediction dataset attributes and information

Feature	Characteristic representation	specific	Range/Value
age	Patient's age in years	[continuous variable]	29-77
sex	0 = female ; 1 = male	[categorical variable]	0,1
cp	Chest pain type 0 = typical angina 1 = atypical angina 2 = non-anginal pain 3 = asymptomatic	[categorical variable]	0,1,2,3
trestbps	Resting blood pressure (mm hg)	[continuous variable]	94-200
chol	Cholesterol (mg/dl)	[continuous variable]	126-564
fbs	1 = True, 0 = False	[categorical variable]	0,1
restecg	Resting electrocardiographic results ~ 0 = Normal, 1 = ST-T wave normality, 2 = Left ventricular hypertrophy	[categorical variable]	0,1,2
thalach	Maximum heart rate achieved	[continuous variable]	71-202
exng	Exercise induced angina ~ 1 = Yes, 0 = No	[categorical variable]	0,1
oldpeak	Previous peak	[continuous variable]	0-6,2
slp	Slope of exercise ST segment 0 = unslope 1 = flat 2 = downslope	[categorical variable]	0,1,2
Ca	No. of major vessels [0-2] colored by fluoroscopy	[continuous variable]	0,1,2
thal	Defect type 1 = fixed defect; 2 = normal; 3 = reversable defect	[categorical variable]	1,2,3
target	Has heart disease or not, 0 = no 1 = yes	[target variable]	0,1

In the data preprocessing process, categorical variables are converted into numerical variables, where each category is represented by a binary (0 or 1) indicator variable.

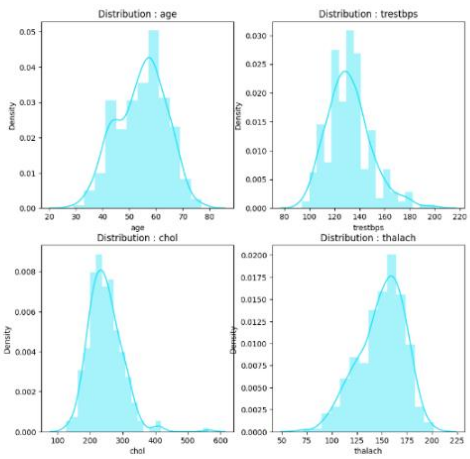
Identifying and addressing missing values is critical. Depending on the extent of missing data, rows with missing data may be discarded or imputation methods may be used that include filling in missing values with the mean, and median, or using more advanced techniques such as interpolation.

It is essential to detect and address outliers appropriately. This may include visualizations (eg, box plots, scatter plots) and statistical methods to identify data points that deviate significantly from the norm.

Scaling is crucial when numerical features have different scales. Standardization involves transforming the data to have a mean of 0 and a standard deviation of 1, while normalization scales the data to a specified range (for example, between 0 and 1).

The listed steps contribute to the preparation of the data for further analysis or machine learning modeling. It is worth noting that the specific techniques and methods used may vary based on the nature of the data, the problem at hand, and the requirements of the chosen analysis or model [18].

Chart 1 and chart 2 present the distribution of numerical characteristics and categorical characteristics respectively.



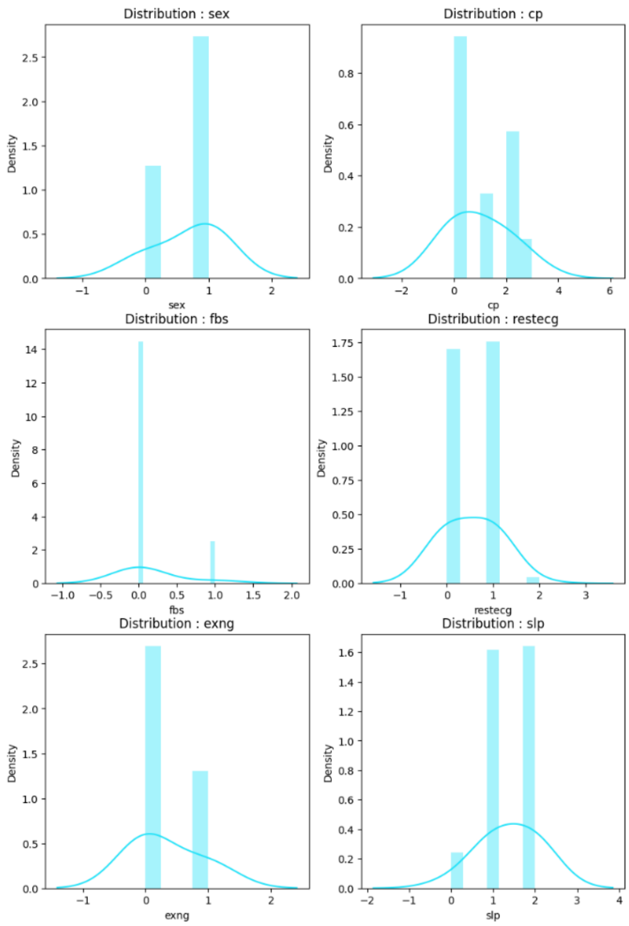


Chart 1: Distribution of numerical features

Chart 2: Distribution of categorical features

Table 2 presents the basic statistical parameters of the database.

Table 2:
Basic statistical parameters

	count	mean	std	min	25%	50%	75%	max
age	303.00	54.37	9.08	29.00	47.50	55.00	61.00	77.00
sex	303.00	0.68	0.47	0.00	0.00	1.00	1.00	1.00
cp	303.00	0.97	1.03	0.00	0.00	1.00	2.00	3.00
trestbps	303.00	131.62	17.54	94.00	120.00	130.00	140.00	200.00
chol	303.00	246.26	51.83	126.00	211.00	240.00	274.50	564.00
fbs	303.00	0.15	0.36	0.00	0.00	0.00	0.00	1.00
restecg	303.00	0.53	0.53	0.00	0.00	1.00	1.00	2.00
thalach	303.00	149.65	22.91	71.00	133.50	153.00	166.00	202.00
exng	303.00	0.33	0.47	0.00	0.00	0.00	1.00	1.00
oldpeak	303.00	1.04	1.16	0.00	0.00	0.80	1.60	6.20
slp	303.00	1.40	0.62	0.00	1.00	1.00	2.00	2.00
ca	303.00	0.73	1.02	0.00	0.00	0.00	1.00	4.00
thal	303.00	2.31	0.61	0.00	2.00	2.00	3.00	3.00
target	303.00	0.54	0.50	0.00	0.00	1.00	1.00	1.00

4.3. Exploratory data analysis

Exploratory Data Analysis (EDA) is a fundamental step in predicting heart disease using machine learning and aims to explore and understand essential insights and patterns in a database. EDA is performed with a systematic approach, includes visual and statistical examinations, provides a deep understanding of the data, and creates a basis for the development of accurate and relevant models [19]. Visual displays such as histograms, scatterplots, and ROC curves provide a comprehensive understanding of data distributions, feature importance, and model performance metrics. Data visualization is an indispensable tool in the field of predicting heart disease using machine learning models, helping to understand the complex relationships between characteristics such as age, cholesterol levels, and blood pressure and their impact on heart disease risk. Metrics such as ROC curves, confusion matrices, and precise regression curves provide insight into the accuracy and ability of the model to distinguish patients with and without heart disease [20]. Figure 2 shows the mean values of all attributes for cases with and without heart disease risk.

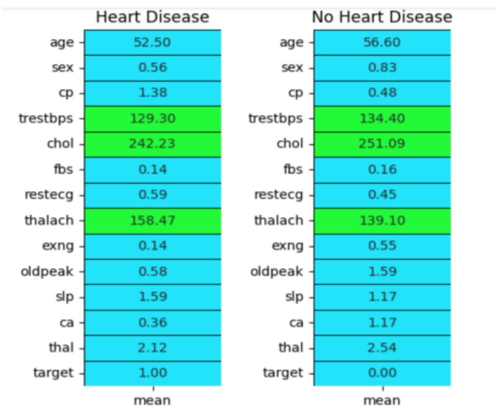
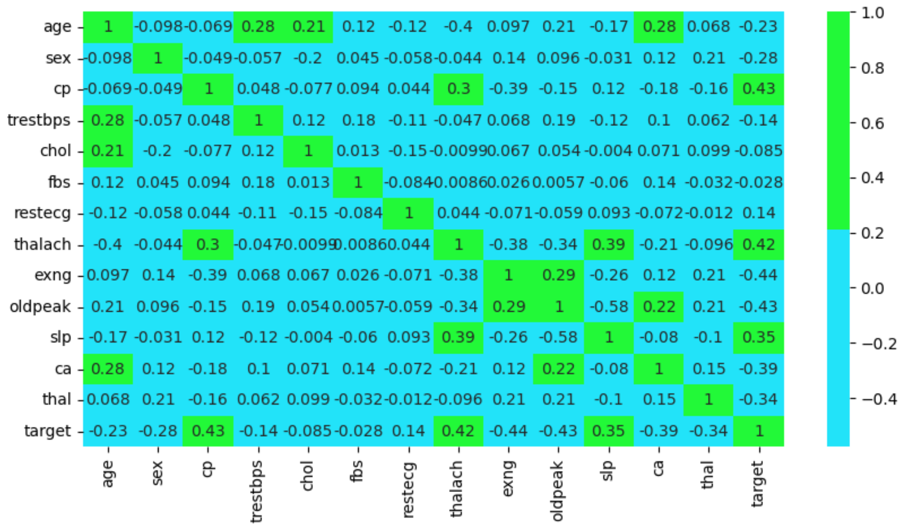


Figure 2: Mean values of all characteristics there is and there is no risk of heart disease

Correlation was performed to reveal potential relationships between variables in order to highlight interactions and dependencies in the dataset. This analysis provides valuable insights into the potential influence between variables as well as their joint influence on heart disease risk. From the correlation matrix, insight is obtained into which characteristics are positively or negatively correlated with each other and with the target variable. According to Table 3, the target variable is a dependent variable that is negatively correlated with some of the factors such as thali, and positively correlated with some of the parameters such as that, ca, etc. Visualization of the correlation matrix provides insight into which characteristics are positively or negatively correlated with the target variable.

Table 3: Correlation matrix



4.4. Evaluation metrics

Evaluation metrics are crucial for evaluating the performance of machine learning models [21]. Accuracy is a metric that measures the percentage of correctly classified cases out of all cases and represents the ability to correctly identify true positives while minimizing false positives, which is critical to patient safety.

Cross-validation divides the dataset into multiple subsets so that training and testing are performed multiple times, allowing each subset of data to be part of both the training and testing set. Cross-validation helps to identify potential problems such as overfitting and underfitting [22].

Receiver Operating Characteristics – AreaUnder Curve (ROC-AUC) is used to quantify the ability of the predictive model to distinguish between positive and negative cases. In binary classification tasks. The ROC-AUC metric helps assess the ability of a model to maintain a balance between true positives and false positives at different classification thresholds [23].

The confusion matrix provides a detailed overview of the classification results of the model, categorizing the predictions into four basic categories: true positives (correctly predicted cases of heart disease), true negatives (correctly predicted cases without heart disease), false positives (incorrectly predicted cases of heart disease), and false negatives (incorrectly predicted cases without heart disease [24].

4.5. Comparison of results

Table 4 shows the Evaluation Scores of the five considered ML algorithms by evaluating their performance in terms of accuracy, Cross Validation Score, and ROC_AUC Score in order to build a model for predicting heart diseases.

In terms of accuracy, SVM achieved the highest accuracy of 90.16%, indicating that it correctly predicted the heart disease status for the majority of cases. Random forest and logistic regression also performed well, with 86.89% accuracy. KNN and Decision Tree scored lower for accuracy, with KNN at 85.25% and Decision Tree at 77.05%.

In terms of the Cross Validation Score metric, Random Forest received the highest cross-validation score of 89.67%, indicating its consistent performance across different data subsets. SVM had a slightly lower but still strong cross-validation score of 87.96%. Logistic regression also performed well in cross-validation, with a score of 88.52%. KNN and Decision Tree had lower cross-validation scores of 84.98% and 79.82%, respectively.

Table 4: Accuracy, Cross Validation Score, and ROC AUC Score for different ML models

model	accuracy	Cross Validation Score	ROC_AUC Score
Logistic regression	86.89%	88.52%	87.34%
SVM	90.16%	87.96%	90.62%
Decision Tree	77.05%	79.82%	76.99%
Randoom Forest	86.89%	89.67%	87.34%
KNN	85.25%	84.98%	85.61%

Analysis of the ROC-AUC results showed that SVM achieved the highest ROC-AUC score of 90.62%, indicating its excellent ability to discriminate between patients with and without heart disease. Random forest and logistic regression share the same ROC-AUC score of 87.34%, indicating their comparable discrimination power. KNN and Decision Tree had slightly lower ROC-AUC scores of 85.61% and 76.99%, respectively, suggesting that they may not perform as well in distinguishing between positive and negative cases.

Of the algorithms analyzed, SVM stands out as the top-performing model, boasting the highest accuracy and ROC-AUC score, making it a strong choice for heart disease prediction. Random Forest and Logistic Regression also provide robust performance across all metrics. KNN and Decision Tree, while achieving reasonably good results, have slightly lower scores in accuracy, cross-validation, and ROC-AUC, indicating slightly lower predictive power in this particular dataset.

The choice of the best model ultimately depends on the specific application requirements and the relative importance of the specified metrics in the context of heart disease prediction. The methodology must be constantly improved and developed, so that other parameters such as echocardiographic data and medical recordings may be included in some future considered models.

5. Limitations of machine learning models for heart disease prognosis

Data quality and quantity play a key role in the performance of ML algorithms for heart disease prognosis. IoMT devices generate huge amounts of data, which often leads to problems with data imbalance, noise, and missing values. The need for comprehensive, standardized data collection protocols is becoming apparent, as the absence of such practices can hinder the accuracy and generalizability of predictive models. The interpretability and transparency of ML algorithms in health applications, especially for the prognosis of critical heart diseases, despite their high prediction accuracy, cannot often provide meaningful insights into the decision-making process, which raises ethical and regulatory concerns, given the fact that doctors and patients demand transparent explanations for the recommendations made by these algorithms. The issue of bias and fairness in ML models for heart disease prognosis is a critical limitation.

Biases in data and algorithmic decisions may disproportionately affect certain demographic groups, potentially leading to unfair results. Providing robust, universally applicable models requires careful consideration of these factors.

Limitations imposed by the hardware and software capabilities of the IoMT devices themselves can hamper the effectiveness of ML algorithms. Connectivity issues, security issues, and device compatibility challenges must be addressed to facilitate seamless integration of these devices with ML algorithms.

Future research should focus on developing more robust and personalized models to improve cardiovascular risk assessment.

Machine learning algorithms integrated with IoMT devices hold great promise for advancing heart disease prediction, enabling early detection, personalized care, remote monitoring, and more efficient allocation of healthcare resources. However, addressing data privacy, data quality, regulatory compliance, and model interpretability are essential to unlocking the full potential of this technology to strengthen patient trust and data security.

6. Conclusions and challenges of machine learning models for heart disease prognosis

IoMT devices improve patient outcomes by enabling early detection of health problems, continuous monitoring, and personalized interventions. IoMT has the potential to reduce healthcare costs by minimizing hospital readmissions, optimizing resource utilization, and streamlining workflows and patient care in rural settings. It empowers individuals to actively participate in their own health management and promotes patient engagement.

The data processing process applies advanced analytics, machine learning algorithms, and artificial intelligence techniques to extract meaningful insights from the data. Choosing an appropriate machine learning algorithm depends on the specific task and the characteristics of the data. These insights help healthcare professionals make the right and personalized decisions. IoMT facilitates remote patient monitoring, allowing healthcare providers to monitor patients' health conditions and intervene in real-time when needed. IoMT enables virtual consultations, remote diagnostics, and telehealth services, and supports chronic disease monitoring, medication management, and preventive care.

The future of machine learning models in the prediction of numerous diseases brings enormous challenges in the direction of applying advanced algorithms such as deep learning, neural networks, hybrid model research, and enabling real-time monitoring. The continued evolution of machine learning has the potential to revolutionize early diagnosis, intervention, and patient care in healthcare.

Data is securely transmitted to cloud-based platforms or local servers, where it can be stored, analyzed, and processed, however, the widespread adoption of IoMT raises security and privacy concerns given that it is sensitive patient data. Interoperability issues between different IoMT devices and systems pose a challenge for seamless integration and data exchange. Healthcare organizations need to ensure their IoMT devices comply with regulatory requirements such as HIPAA and GDPR [25] to avoid legal and financial penalties.

References

[1] Mansour, Romany Fouad, et al. "Artificial intelligence and internet of things enabled disease diagnosis model for smart healthcare systems." *IEEE Access* 9 (2021): 45137-45146.

[2] Dwivedi, Ruby, Divya Mehrotra, and Shaleen Chandra. "Potential of Internet of Medical Things (IoMT) applications in building a smart healthcare system: A systematic review." *Journal of oral biology and craniofacial research* 12.2 (2022): 302-318.

[3] Huang, Chenxi, et al. "Internet of Medical Things: A Systematic Review." *Neurocomputing* (2023): 126719.

[4] Kamalov, Firuz, et al. "Internet of medical things privacy and security: Challenges, solutions, and future trends from a new perspective." *Sustainability* 15.4 (2023): 3317.

[5] Dang, Van Anh, et al. "Intelligent Healthcare: Integration of Emerging Technologies and Internet of Things for Humanity." *Sensors* 23.9 (2023): 4200.

[6] Ameen, Aya Hamid, Mazin Abed Mohammed, and Ahmed Noori Rashid. "Dimensions of artificial intelligence techniques, blockchain, and cyber security in the Internet of medical things: Opportunities, challenges, and future directions." *Journal of Intelligent Systems* 32.1 (2023): 20220267.

[7] Ashfaq, Z., Rafay, A., Mumtaz, R., Zaidi, S.M.H., Saleem, H., Zaidi, S.A.R., Mumtaz, S. and Haque, A., 2022. A review of enabling technologies for the Internet of Medical Things (IoMT) Ecosystem. Ain Shams Engineering Journal, 13(4), p.101660

[8] Elhachmi, J. A. M. A. L., and A. B. D. E. L. L. A. T. I. F. Kobbane. "BLOCKCHAIN-BASED SECURITY MECHANISMS FOR INTERNET OF MEDICAL THINGS (IOMT)." *International Journal of Computer Networks and Communications* (2022): 115-136.

[9] Kamruzzaman, M. M., Ibrahim Alrashdi, and Ali Alqazzaz. "New opportunities, challenges, and applications of edge-AI for connected healthcare in the internet of medical things for smart cities." *Journal of Healthcare Engineering* 2022 (2022).

[10] Savoska, Snezana, et al. "Integration of heterogeneous medical and biological data with electronic personal health records." Medical Informatics, Biometry and Epidemiology (MIBE) 19 (2023): 1-11.

[11] Savoska, Snezana, Blagoj Ristevski, and Vladimir Trajkovik. "Personal Health Record Data-Driven Integration of Heterogeneous Data." *Data-Driven Approach for Bio-medical and Healthcare*. Singapore: Springer Nature Singapore, 2022. 1-21.

[12] Gopalan, Subiksha Srinivasa, Ali Raza, and Wesam Almobaideen. "IoT security in healthcare using AI: A survey." *2020 International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*. IEEE, 2021.

[13] Bao, Xinqi, et al. "Time-frequency distributions of heart sound signals: A comparative study using convolutional neural networks." *Biomedical Engineering Advances* 5 (2023): 100093.

[14] Chitra, S., and V. Jayalakshmi. "Prediction of heart disease and chronic kidney disease based on internet of things using RNN algorithm." *Proceedings of Data Analytics and Management: ICDAM 2021, Volume 1*. Springer Singapore, 2022.

[15] Zhang, Wanruo, et al. "Motion prediction of beating heart using spatio-temporal LSTM." *IEEE Signal Processing Letters* 29 (2022): 787-791.

[16] Uddin, Shahadat, et al. "Comparing different supervised machine learning algorithms for disease prediction." *BMC medical informatics and decision making* 19.1 (2019): 1-16.

[17] https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset

[18] Md, Abdul Quadir, et al. "Enhanced Preprocessing Approach Using Ensemble Machine Learning Algorithms for Detecting Liver Disease." *Biomedicines* 11.2 (2023): 581.

[19] Kandpal, Jyoti, Geetanjali Baluti, and Praful Ranjan. "Integration of Artificial Intelligence with Opto-VLSI for the Biomedical and Healthcare Industries." *Opto-VLSI Devices and Circuits for Biomedical and Healthcare Applications*. CRC Press 161-175.

[20] Senthilkumar, S. A., et al. "Big data in healthcare management: a review of the literature." *American Journal of Theoretical and Applied Business* 4.2 (2018): 57-69.

[21] Hicks, Steven A., et al. "On evaluation metrics for medical applications of artificial intelligence." *Scientific Reports* 12.1 (2022): 5979.

[22] Xiong, Zheng, et al. "Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation." *Computational Materials Science* 171 (2020): 109203.

[23] Habenicht, Richard, et al. "Machine Learning Based Linking of Patient Reported Outcome Measures to WHO International Classification of Functioning, Disability, and Health Activity/Participation Categories." *Journal of Clinical Medicine* 12.17 (2023): 5609.

[24] Vujović, Ž. "Classification model evaluation metrics." *International Journal of Advanced Computer Science and Applications* 12.6 (2021): 599-606.

[25] Shuaib, Mohammed, et al. "Compliance with HIPAA and GDPR in a blockchain-based electronic health record." Materials Today: Proceedings (2021).