Atherosclerosis: Open Access

Expert Review Open Access

A Cutting-Edge Medical Diagnosis Aid for Identifying Patients who Would Develop Atherosclerotic Diseases

Oumaima Terrada*

Signaux, Systèmes Distribués et Intelligence Artificielle (SSDIA) Laboratory, ENSET of Mohammedia, Hassan II University of Casablanca, B.P 159, Mohammedia, Morocco

Abstract

The diagnosis of atherosclerosis is a difficult cognitive procedure. In medical diagnosis support systems, artificial intelligence techniques like machine learning algorithms have demonstrated their effectiveness (MDSS). In this study, we created a brand-new machine learning MDSS to improve cardiovascular disease diagnosis. 835 patient medical records with atherosclerosis, which is typically brought on by coronary artery disorders (CAD), were used in our study. These records were gathered from three databases. Several input variables based on three databases, including the Cleveland heart, are included in the system input layer. databases on illness, Hungarian, and Z-Alizadeh Sani. Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Decision Tree (DT), Nave Bayes (NB), Classification Ensemble (CE), and Discriminant Analysis (DA) algorithms are used to evaluate the system. Through a number of performance indicators, the robustness of the suggested methodologies was assessed. The findings demonstrated that the suggested MDSS achieved an accuracy of (98%), which is greater accuracy than the current techniques. These findings mark a positive development in the field of widespread clinical atherosclerosis disease diagnosis. This section includes a review of the literature on a few chosen works on automated heart disease detection that made use of the same existing datasets and that we will later take into consideration for the performance comparison.

In the authors used a neural network ensemble approach to combine the projected values from earlier models to construct new models. 89.01% more accuracy was attained than with the machine learning approach. The authors suggested a clinical decision support system (CDSS) employing weighted fuzzy rules (WFR) for cardiac disease prediction in another paper that was published. Two evaluation scenarios were employed; the first automates the method for producing WFRs, and the second creates a fuzzy rule-based CDSS. Using the Cleveland Heart Disease database, they evaluated their CDSS. The best accuracy score this method achieves in comparison to a neural network-based system is 62.35%.

Keywords: Atherosclerosis; Machine learning techniques; Cardiovascular disease (CVD); Classification; Prediction

Introduction

The term "cardiovascular disease" (CVD) covers a wide range of illnesses and ailments related to the heart. Other CVDs exist, including coronary artery disease (CAD), also referred to as atherosclerosis. Heart disease, especially atherosclerosis, affects a lot of people. The World Health Organization (WHO) reports that this illness is the main cause of death in the majority of industrialized nations. It is challenging to manually pull relevant information from patient records for medical diagnosis of this condition. This highlights the significance of creating a Medical Diagnostic Support System (MDSS) to automate patient classification and cardiovascular disease prediction. However, to produce the greatest therapeutic decisions, medical diagnosis research has to be more precise and effective decisions. Classic MDSS have demonstrated their ability to address the majority of diagnostic problems, but they have a lower accuracy rate and cannot make the proper diagnosis. Artificial intelligence (AI) and machine learning (ML)-based medical diagnosis and treatment systems have drawn more attention in recent years. As a result, these study areas have had an effect on science in areas like biology, applied sciences, economics, and medical applications. In order to build MDSS to forecast or categorize individuals with heart disease and enhance healthcare, a number of works have been proposed [1-5].

In this context, we suggest a new MDSS in this study that makes use of a few particular machine learning techniques. By examining heart disease databases, the major objective is to categorize and forecast patient health based on the key features that have been selected. Both controllable and uncontrollable risk variables are referred to as risk factors. Age, family history, and gender are risk factors for uncontrolled atherosclerosis. Nevertheless, it does not necessarily follow that the individual would develop atherosclerosis, in part because an early diagnosis might lessen the impact of genetics. On the other hand, there are numerous managed risk factors for atherosclerosis, including diabetes, obesity, hypertension, systolic and diastolic blood pressure, cholesterol, and smoking. By altering their lifestyle, patients could lower these risk factors.

The remainder of this paper is organized as follows: We explore some pertinent literature in Section of this article. We outline and elucidate our suggested system procedure in Section 3. We specifically give the global flowchart of the suggested MDSS and a few machine learning methods, together with the CAD datasets that were employed. The evaluation measures used to assess and contrast our MDSS

***Corresponding author:** Oumaima Terrada, Signaux, Systèmes Distribués et Intelligence Artificielle (SSDIA) Laboratory, ENSET of Mohammedia, Hassan II University of Casablanca, B.P 159, Mohammedia, Morocco, E-mail: oaimaterda@gmail.com

Received: 01-Jul-2022, Manuscript No. asoa-22-76099; **Editor assigned:** 04-Jul-2022, PreQC No. asoa-22-76099 (PQ); **Reviewed:** 17-Jul-2022, QC No. asoa-22-76099; **Revised:** 19-Jul-2022, Manuscript No. asoa-22-76099(R); **Published:** 30-Jul-2022, DOI: 10.4172/asoa.1000178

Citation: Terrada O (2022) A Cutting-Edge Medical Diagnosis Aid for Identifying Patients who Would Develop Atherosclerotic Diseases. Atheroscler Open Access 7: 178.

performance with that of comparable systems are described in Section We described the implementation in Section and provided the findings and discussion. This work is concluded in Section which also offers some suggested viewpoints.

Subjective Heading

In the authors used a neural network ensemble approach to combine the projected values from earlier models to construct new models. 89.01% more accuracy was attained than with the machine learning approach. The authors suggested a clinical decision support system (CDSS) employing weighted fuzzy rules (WFR) for cardiac disease prediction in another paper that was published .Two evaluation scenarios were employed; the first automates the method for producing WFRs, and the second creates a fuzzy rule-based CDSS. Using the Cleveland Heart Disease database, they evaluated their CDSS. The best accuracy score this method achieves in comparison to a neural network-based system is 62.35%.

The Fast Decision Tree (FDT) and pruned C4.5 tree approaches were employed by the authors in. This strategy tries to incorporate the outcomes of the machine learning investigation into several CAD databases. Results revealed that the classification accuracy was 78.06%, exceeding the 75.48% average classification accuracy of distinct datasets. A hybrid neural network-genetic (HNNG) was developed by the authors of in 2017 as a way to enhance a neural network's initial weights using a genetic method. Using the Cleveland Heart Disease database with the Z-Alizadeh Sani data set, the maximum accuracy percentage is 93.85% [6].

The problem of heart disease medical diagnosis has been addressed by other strategies. The performance of a decision support system for predicting the risk of heart failure is shown by the authors in Reference. An artificial neural network (ANN) and fuzzy-AHP are the foundations of this system. The proposed method, when compared to the traditional ANN method, could achieve an average prediction accuracy of 91.10%, according to the findings that were actually achieved. In 2018, the authors talked about designing and putting in place an expert system for heart disease. The Fuzzy-AHP and Fuzzy Inference System were used by the authors to create this system (FIS). The results of the developed approach have demonstrated the risk of acquiring heart disease. This system demonstrated that AI and ML techniques in the medical field had, in accordance with the Experimental results.

Three classifiers-Naive Bayes, and Support Vector Machine-were employed by the authors in (SVM). These techniques were used to diagnose coronary artery disease (CAD) in the Z-Alizadeh Sani database based on the stenosis of three coronary arteries: the left circumflex (LCX), left anterior descending (LAD), and right coronary artery (RCA). As a result of this study, the maximum accuracy for CAD identification was 96.40%. Using two datasets, Cleveland and Z Alizadeh Sani, the authors presented an ensemble of the nu-Support Vector Classification (NEnu SVM) model in 2019. This model integrates different machine learning techniques and ensemble learning approaches to predict CAD. The suggested model's maximum accuracy for predicting CAD was 94.66% when using the Z-Alizadeh Sani dataset and 98.60% when using the Cleveland CAD dataset. A unique technique known as Hybrid Feature Selection (2HFS) was recently created by the authors in Ref using classifiers from Random Forest (RF), Gaussian Naive Bayes (GNB), Decision Tree (DT), and XGBoost. The Nasarian CAD database was used by the authors of this study. Additionally, they tried this strategy on databases from Hungary, Long Beach VA, and Z-Alizadeh Sani, achieving accuracy rates of 83.94%, 81.58%, and 92.58%, respectively[7,8].

The goal of the current work is to suggest an unique MDSS for the diagnosis of patients with atherosclerosis-related heart disorders. The proposed strategy in this work was based on seven hand-picked machine learning algorithms, including DT, Discriminant Analysis (DA), Discriminant Analysis (DA), Naive Bayes (NB), K-Nearest Neighbor (KNN), Artificial Neural Network (ANN), and Discriminant Analysis. To evaluate model performance, the study simulates the execution of several algorithm configurations. From there, it chooses the best model while applying performance evaluation techniques to enhance each one. The actual research advances our past findings. The main steps to establish our MDSS The goal of the current work is to suggest an unique MDSS for the diagnosis of patients with atherosclerosis-related heart disorders. The proposed strategy in this work was based on seven hand-picked machine learning algorithms, including DT, Discriminant Analysis (DA), Discriminant Analysis (DA), Naive Bayes (NB), K-Nearest Neighbor (KNN), Artificial Neural Network (ANN), and Discriminant Analysis. To evaluate model performance, the study simulates the execution of several algorithm configurations. From there, it chooses the best model while applying performance evaluation techniques to enhance each one. The actual research advances our past findings [9].

In order to extract relevant patient data from atherosclerosis characteristics, this is the processing goal. Please take note that these data records should be split into two datasets. Then, in order to create the models for the prediction method, we produce the training dataset and testing dataset. We applied machine learning techniques in the prediction phases after picking the desired attributes from each database. Therefore, it is necessary to assess the patient's data in this stage for various results. The final stage of the suggested method produces a result that is appropriate for the patient with atherosclerosis. The proposed method's flowchart, which makes use of ML techniques, is show in the text [10].

Discussion

A development of computational learning-related Artificial Intelligence (AI) techniques is machine learning (ML). It makes it possible for the computer to be implicitly and automatically trained while learning. The purpose of machine learning (ML) is to create and design learning algorithms to establish a thorough data analysis. Using the ML approach, algorithms and prediction models in data domains become simple and straightforward to use. As a result, it enables researchers to make wise choices and provide valuable results.

Patients with heart illness at Tehran's Shaheed Rajaei Cardiovascular, Medical and Research Center provided the Z-Alizadeh Sani dataset at random. This dataset, which has 303 samples and 54 features for each patient, was created for the diagnosis of CAD. The selected elements include the primary data on the patient's physical and biological examinations, echocardiograms (ECGs) laboratory tests, demographic characteristics, and symptoms.

Patients were divided into two output classes by: 71% of patients had CAD and 29% were in good health. The stenosis prediction outputs of the LCX, RCA, and LAD coronary arteries are also included in this dataset. According to the atherosclerosis risk factor used in this investigation, 17 elements were manually chosen as the most crucial aspects. The aspects that the suggested system for atherosclerosis has chosen are listed. Andras Janosi gathered the Hungarian data at the Hungarian Institute of Cardiology in Budapest. Ten features are present in this database. 34 samples out of the 294 dataset samples were deleted due to missing values, and 262 records were often used and divided into groups of 62.21% healthy participants and 37.78% heart disease patients.

The term "biochemical analysis" describes a set of processes and assays that allow researchers and medical professionals to examine all the compounds present in organs and chemical reactions. For specialized research and diagnostic labs, the most crucial of these methods is still used. The following situations are provided as examples for the heart disease datasets under consideration: FBS (Fasting Blood Sugar) (mg/dl), TG (Triglyceride) (mg/dl), and LDL.

In medicine, symptoms or signs are declarations of a pathology that the doctor may note during a clinical examination (clinical sign, para-clinical sign, or complementary sign). Different observational levels can produce these indicators. Clinical signs are those that are gathered without the need of bulky equipment, such as a thermometer, blood pressure monitor, stethoscope, etc. There are numerous indications and symptoms, including general indicators like tiredness and fever as well as localised signs like pimples and bubbles as well as functional signs like discomfort, dyspnea, palpitations, and syncope (abnormal stethoscope noise, redness, etc.).

Symptoms are typically interpreted as a positive sign, which is the notable existence of a sign. For instance, a positive indicator for diabetes is a fasting blood sugar level greater than 1.26 g/L; a negative sign is the clear absence of a sign. For instance, fasting blood sugar less than 1.26 g/L is a warning indicator for diabetes.

A subject's clinical or biological status that raises their risk of contracting a disease or experiencing trauma is known as a cardiovascular risk factor. Cardiovascular risk factors can be divided into two categories: uncontrollable risk factors and manageable risk factors. The strongest associations between risk factors and cardiovascular disease are those that are under your control. We give the following examples for the considered datasets: Smoking, arterial hypertension, total cholesterol, low-density lipoproteins (LDL), triglycerides, high-density lipoproteins (HDL), obesity (body mass index, or BMI), diabetes, etc.

Uncontrollable risk factors are aspects that may contribute to the development of heart disease but over which we have no control. Even while we are powerless to change these characteristics to make them disappear, we can still use them to identify at-risk individuals and give them the right preventive and therapeutic interventions. These kinds of characteristics typically relate to demographic characteristics. The quantitative and qualitative study of population dynamics is known as demography. Almost all medical databases contain demographic information. Age, gender, family history, and other factors are taken from heart disease datasets.

The binary labels "Diagnosis of heart disease" reflecting the patient's real condition is the equivalent outputs utilised for prediction. A patient either has atherosclerosis or is in good health. Here, 0 denotes the absence of atherosclerosis, which corresponds to a diameter narrowing of less than 50%. According to UCI data for Cleveland and Hungarian, the value 1 reflects atherosclerosis disease, which is defined as a diameter narrowing higher than 50%. The outputs for the Z-Alizadeh Sani database are divided into two class labels. As a result, class 0 designates a state of normality that is free of atherosclerotic disease. Class 1 denotes the presence of CAD, or atherosclerosis disease.

Numerous simulations and experiments were run to empirically determine the most effective ML models, which were then used to demonstrate the efficacy of the proposed classifiers and predictors. The three atherosclerosis datasets are utilised in this fashion, and the experimental findings are compiled in the tables using various performance evaluation indicators to gauge the effectiveness of the suggested strategy. It was also done to compare the results with earlier research. We the ANN classification process comparing the testing data with the target data in order to assess the effectiveness of the ANN algorithm on the three databases. The computation of the true positive and true negative values, which represent the actual patient situation, as well as the false-negative and false-positive values, which represent the anticipated patient situation, are all included in the evaluation. Our trained ANN model has demonstrated good accuracy, which means that the projected values converge toward the target values, according to the actual outcomes.

The training dataset, training labels, testing dataset, and initialising the K odd value must all be loaded before the KNN algorithm can run. Selecting the distance type is the next step. In our situation, employing the Hamming distance yields the greatest results. Equation was used by the method to calculate the separation between elements in the training dataset and the testing dataset. The next step is to use equation to determine the majority class label of the K-nearest. This is carried out during the training and testing phases. The optimal K odd values for each of the used databases are provided.

Conclusion

Results from the testing phase are fed into the proposed classifier system to categorise and forecast patients with atherosclerosis. Utilizing the common performance criteria of SS, SP, ACC, FS, and MCC, the results are assessed. Two additional machine learning measures, FS as a binary classification accuracy test and MCC as a binary classification quality measure, are employed to enhance our atherosclerosis prediction method. To evaluate the effectiveness of the system, the FS and MCC metrics should be close to 1. The assessment metrics for the ANN, KNN, CE, NB, DA, SVM, and DT algorithms Table displays the results that were derived from the Cleveland, Hungarian, and Z-Alizadeh Sani databases.

We provided the findings of our study on a few limited ML technique combinations in Table in order to illustrate the interest of the combination of numerous ML approaches in our suggested prediction system before presenting the comparison with the results in related works. In order to use the studied methodologies on the Z-Alizadeh Sani database, we really restricted this study to the accuracy evaluation metric (ACC).

Acknowledgement

I would like to thank my Professor for his support and encouragement.

Conflict of Interest

The authors declare that they are no conflict of interest.

References

1. Ross R (1986) The pathogenesis of atherosclerosis-an update. New Eng J med 314: 488-500.
2. Duval C, Chinetti G, Trottein F, Fruchart JC, Staels B (2002) The role of PPARs in atherosclerosis. Trends Mol Med 8: 422-430.
3. Libby P, Ridker PM, Maseri A (2002) Inflammation and atherosclerosis. Circulation105: 1135-1143.

4. Falk E (2006) Pathogenesis of atherosclerosis. Exp Clin Cardioliology 47: C7-C12.
5. Hansson GK, Hermansson A (2011) The immune system in atherosclerosis. Nat Immunol 12: 204-212.
6. Han TS, Sattar N, Lean M (2006) ABC of obesity: Assessment of obesity and its clinical implications. Bio Med J 333: 695-698.
7. Kiran S, Kumar V, Kumar S, Price RL, Singh UP (2021) Adipocyte, Immune Cells, and mi RNA Crosstalk: A Novel Regulator of Metabolic Dysfunction and Obesity. Cells 24: 10.
8. Lumeng CN, Bodzin JL, Saltiel AR (2007) Obesity induces a phenotypic switch in adipose tissue macrophage polarization. J Clin Invest 117: 175-184.
9. Xu J, Kitada M, Ogura Y, Koya D (2021) Relationship Between Autophagy and Metabolic Syndrome Characteristics in the Pathogenesis of Atherosclerosis. Front Cell Dev Biol 9: 641852.
10. Engin A (2017) The Pathogenesis of Obesity-Associated Adipose Tissue Inflammation. Adv Exp Med Biol 960: 221-245.