Original Paper

Development of an Online Health Care Assessment for Preventive Medicine: A Machine Learning Approach

Cheng-Sheng Yu[1,2], PhD; Yu-Jiun Lin[1,2], MD; Chang-Hsien Lin[1,2], MD; Shiyng-Yu Lin[1,2], MD; Jenny L Wu[1,2], BSc; Shy-Shin Chang[1,2], MD, PhD

[1]Department of Family Medicine, Taipei Medical University Hospital, Taipei, Taiwan

[2]Department of Family Medicine, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

Corresponding Author:
Shy-Shin Chang, MD, PhD
Department of Family Medicine
School of Medicine
College of Medicine, Taipei Medical University
250 Wuxing St
Taipei, 11031
Taiwan
Phone: 886 2 23565926
Email: sschang0529@gmail.com

Related Article:
This is a corrected version. See correction statement in: https://www.jmir.org/2020/7/e21753/

Abstract

Background: In the era of information explosion, the use of the internet to assist with clinical practice and diagnosis has become a cutting-edge area of research. The application of medical informatics allows patients to be aware of their clinical conditions, which may contribute toward the prevention of several chronic diseases and disorders.

Objective: In this study, we applied machine learning techniques to construct a medical database system from electronic medical records (EMRs) of subjects who have undergone health examination. This system aims to provide online self-health evaluation to clinicians and patients worldwide, enabling personalized health and preventive health.

Methods: We built a medical database system based on the literature, and data preprocessing and cleaning were performed for the database. We utilized both supervised and unsupervised machine learning technology to analyze the EMR data to establish prediction models. The models with EMR databases were then applied to the internet platform.

Results: The validation data were used to validate the online diagnosis prediction system. The accuracy of the prediction model for metabolic syndrome reached 91%, and the area under the receiver operating characteristic (ROC) curve was 0.904 in this system. For chronic kidney disease, the prediction accuracy of the model reached 94.7%, and the area under the ROC curve (AUC) was 0.982. In addition, the system also provided disease diagnosis visualization via clustering, allowing users to check their outcome compared with those in the medical database, enabling increased awareness for a healthier lifestyle.

Conclusions: Our web-based health care machine learning system allowed users to access online diagnosis predictions and provided a health examination report. Users could understand and review their health status accordingly. In the future, we aim to connect hospitals worldwide with our platform, so that health care practitioners can make diagnoses or provide patient education to remote patients. This platform can increase the value of preventive medicine and telemedicine.

KEYWORDS

machine learning; online healthcare assessment; medical informatics; preventive medicine

Introduction

In the ever-changing technological era, the internet can provide rapid and convenient medical services in the form of health care, preventive medicine, and telemedicine. Medical informatics is a multidisciplinary field that comprises medicine and computer science. As computer technology continues to advance, medical informatics can be used to develop various applications such as electronic medical records (EMRs), medical image processing, clinical diagnosis decision systems, hospital information management systems, telemedicine, and internet and health information systems [1-4].

To construct a health care information system, several factors must be considered: the hospital information system, including both clinical management and diagnosis services; the storage and processing of patient information, such as EMRs and electronic health records; decision support systems, such as expert diagnosis systems; and the artificial intelligence (AI) algorithms that need to be applied to those factors (eg, data mining in EMRs and decision-making in clinical diagnosis) [5-8].

The mass application of EMRs and the digitalization of medical equipment and instruments have led to the continuous expansion of information capacity in hospital databases. Therefore, informatics research should focus on basic electronic medical database construction, data collection and analysis, medical decision support, and automatic knowledge acquisition. Furthermore, the use of machine learning (ML) technology in AI to extract the most important information has led to cutting-edge research in medicine [9-13]. The goal of AI is to construct an intelligent machine that imitates the natural intelligence of humans. Computers, robots, and software that are made with such technology will have human-like thinking processes, but with the ability to utilize superhuman speed and power effectively. Knowledge engineering is an essential part of AI research, especially ML, because AI operations require a significant amount of real-world data.

ML is defined as a "machine that is capable of self-learning without any guidance." Therefore, the main purpose of ML is to make computers self-learning and auto-correcting when analyzing data. The core technology of ML must identify specific patterns and information hidden within very large data sets using statistical analysis and prediction automatically [14-17].

Disease and disability are influenced by several factors: environmental factors, genetic predisposition, pathogens, and lifestyle choices. Some conditions are a dynamic process that can affect an individual before they are aware of any problem [18-20]. The core of preventive medicine is to prevent chronic diseases among people who are at risk of certain diseases. In some cases, it can also be used to reverse their condition, returning them to a good health status. In the past, due to information asymmetry, doctors and hospitals led the medical environment, and patients did not have access to any appropriate methods or information to implement real-time self-management. Patients who failed to obtain an early diagnosis would have to pay higher health care costs. Therefore,

the spirit of prevention medicine is that "an ounce of prevention is worth a pound of cure" [21-23].

Metabolic syndrome (MetS) is a cluster of conditions comprising high blood sugar, high blood pressure, abnormal blood lipid levels, abdominal obesity, and other metabolic risk factors. It is a warning sign of potential future chronic disease. People with MetS have an increased risk of subsequent development of type II diabetes, hypertension, hyperlipidemia, heart disease, and stroke compared with healthy people [24-28].

Chronic kidney disease (CKD) is defined as kidney function that is impaired for longer than 3 months, leading to irreversible damage. The National Kidney Foundation Kidney Disease Outcome Quality Initiative guideline classifies CKD into 5 stages according to the estimated glomerular filtration rate (eGFR) and using the recommended Modification of Diet in Renal Disease (MDRD) equation [29]. There are many causes of CKD, such as congenital anomalies of the kidney, urinary tract obstruction, urinary tract infection, and glomerulopathy. In addition, hypertension, diabetes, and gout are common chronic diseases that cause CKD if undertreated [30,31].

Telemedicine uses information and telecommunication technology to deliver medical information and physicians' diagnoses to patients without the limitations of time and space. It combines information and communication technologies with medical expertise to provide various services: remote consultation and conferencing for doctors; comprehensive medical care for residents in remote and outlying islands; and teaching and training opportunities for medical staff. The internet can be used to assist with the popularization of telemedicine to achieve a two-way communication channel between patients and medical practitioners [32,33]. Therefore, this study aims to construct an online ML-driven medical database system from EMRs of subjects who have undergone health examination, and provide online self-health evaluation for MetS and CKD.

Methods

Setting

The study was conducted at the Health Management Center (HMC) of Taipei Medical University Hospital (TMUH). Electronic medical records (EMRs) were obtained and reviewed from the HMC, which receives approximately 60 to 70 visits per month.

Ethics

The study was approved by the Institutional Review Board (IRB) of TMUH prior to data collection (TMUH TMU-JIRB number N202003088), in accordance with the original and amended Declaration of Helsinki. The IRB waived the need for informed consent because of the retrospective nature of this study.

EMR Database and System

The databases and the selected predicting variables (Table 1) were derived from previous publication on MetS and CKD [34-36]. Figure 1 shows an overview of the system and the main functions. Briefly, using a series of complicated procedures,

the two databases (MetS and CKD) were connected to an internet platform to construct one integrated system. This web-based system was embedded with ML models to provide various medical evaluations and analyses. The online system was constructed on a server as a web-based environment. The frontend implementation included the programming language JavaScript (Oracle Corp), the framework VueJS (Vue), and the styling Syntactically Awesome Style Sheets (Sass). The backend implementation used Java and R as the programming languages, and all ML calculations and evaluations were conducted using the statistical program R (version 3.6.1, R Foundation for Statistical Computing). The back web framework was Spring Boot (Pivotal Software), connecting the MySQL (Oracle Corp) database as the storage system.

Study Populations

Figure 2 [37-39] shows an overview of the main study population and the validation populations. Briefly, the starting study population included 48,628 EMRs of Taiwanese adults aged over 18 years who underwent a self-paid health examination at TMUH from July 2015 to December 2019. All the study participants completed a self-questionnaire on demographics, existing medical conditions, and the use of medications.

XSL•FO
RenderX

Table 1. The list of predicting variables in the electronic health care records.

Disease and predicting variable	Unit
Metabolic syndrome	
Sex	Male/Female
Age	years
Body mass index	kg/m^2
Waist circumference	cm
Glutamic-oxaloacetic transaminase	IU/L
Glutamate pyruvate transaminase	IU/L
γ-Glutamyl transpeptidase	U/L
Total bilirubin	mg/dL
Alkaline phosphatase	IU/L
Blood urea nitrogen	mg/dL
Creatinine	mg/dL
Uric acid	mg/dL
Albumin	g/dL
Cholesterol	mg/dL
High-density lipoprotein	mg/dL
Low-density lipoprotein	mg/dL
Hemoglobin A_{1c}	%
Glucose AC	mg/dL
Triglycerides	mg/dL
Systolic blood pressure	mm Hg
Diastolic blood pressure	mm Hg
Elastic modulus (E) score	kPa
Controlled attenuation parameter (CAP) score	dB/m
Chronic kidney disease	
Sex	Male/Female
Age	years
Body mass index	kg/m^2
Waist circumference	cm
Glutamic-oxaloacetic transaminase	IU/L
Glutamate pyruvate transaminase	IU/L
γ-Glutamyl transpeptidase	U/L
Total bilirubin	mg/dL
Alkaline phosphatase	IU/L
Blood urea nitrogen	mg/dL
Creatinine	mg/dL
Uric acid	mg/dL
Albumin	g/dL
Cholesterol	mg/dL
High-density lipoprotein	mg/dL
Low-density lipoprotein	mg/dL

Disease and predicting variable	Unit
Hemoglobin A_{1c}	%
Hypertension	Yes/No

Figure 1. The structure of web-based machine learning medical system. API: application programming interface; EMR: electronic medical record; ML: machine learning.
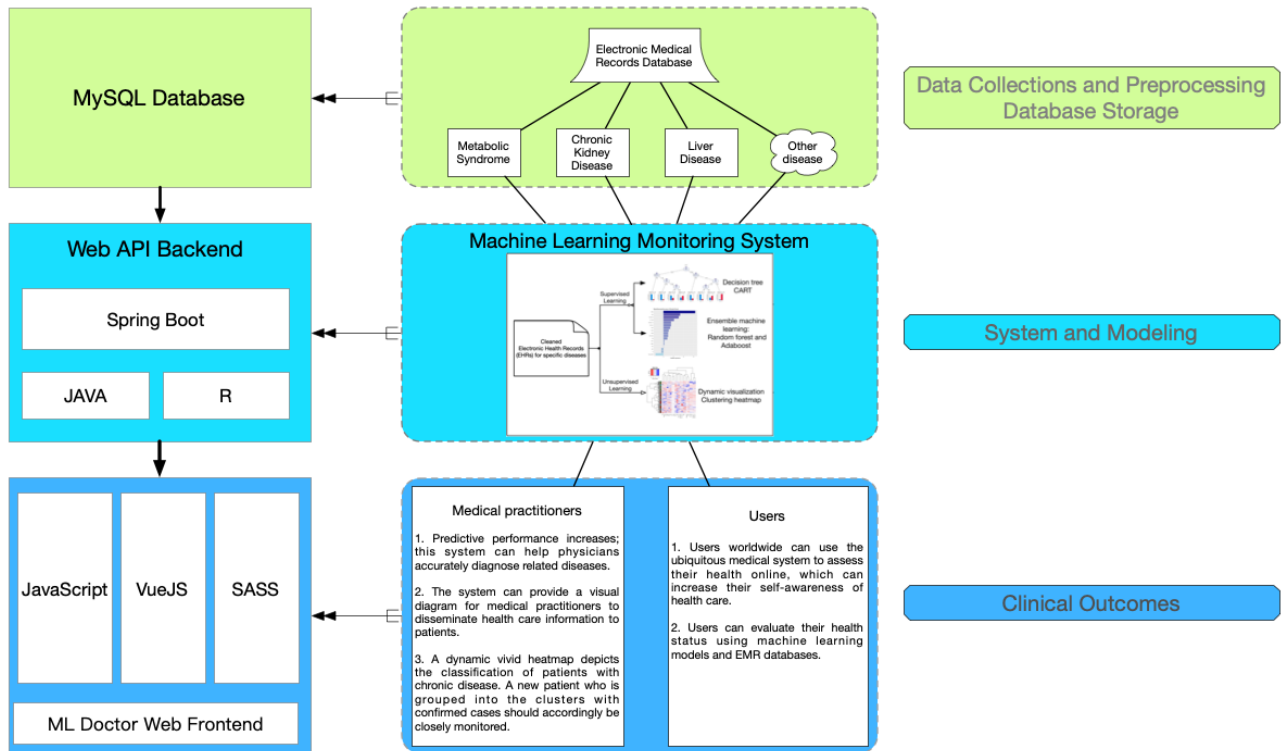
Figure 2. Flowchart of data collection and preprocessing for MetS and CKD data sets including training and validation sets. SAS Enterprise Guide is a software that combines the analytic ability of SAS software with a user-friendly interface. It provides several functions of Structured Query Language (SQL), which includes a text mining technique. ACC: accuracy; AUC: area under the curve; BUN: blood urea nitrogen; CKD: chronic kidney disease; KNN: k-nearest neighbors algorithm; MetS: metabolic syndrome; UA: uric acid. * Centers for Disease Control and Prevention (CDC) and National Center for Health Statistics (NCHS) [37], ** Iimori et al [39], *** De Nicola et al [38].
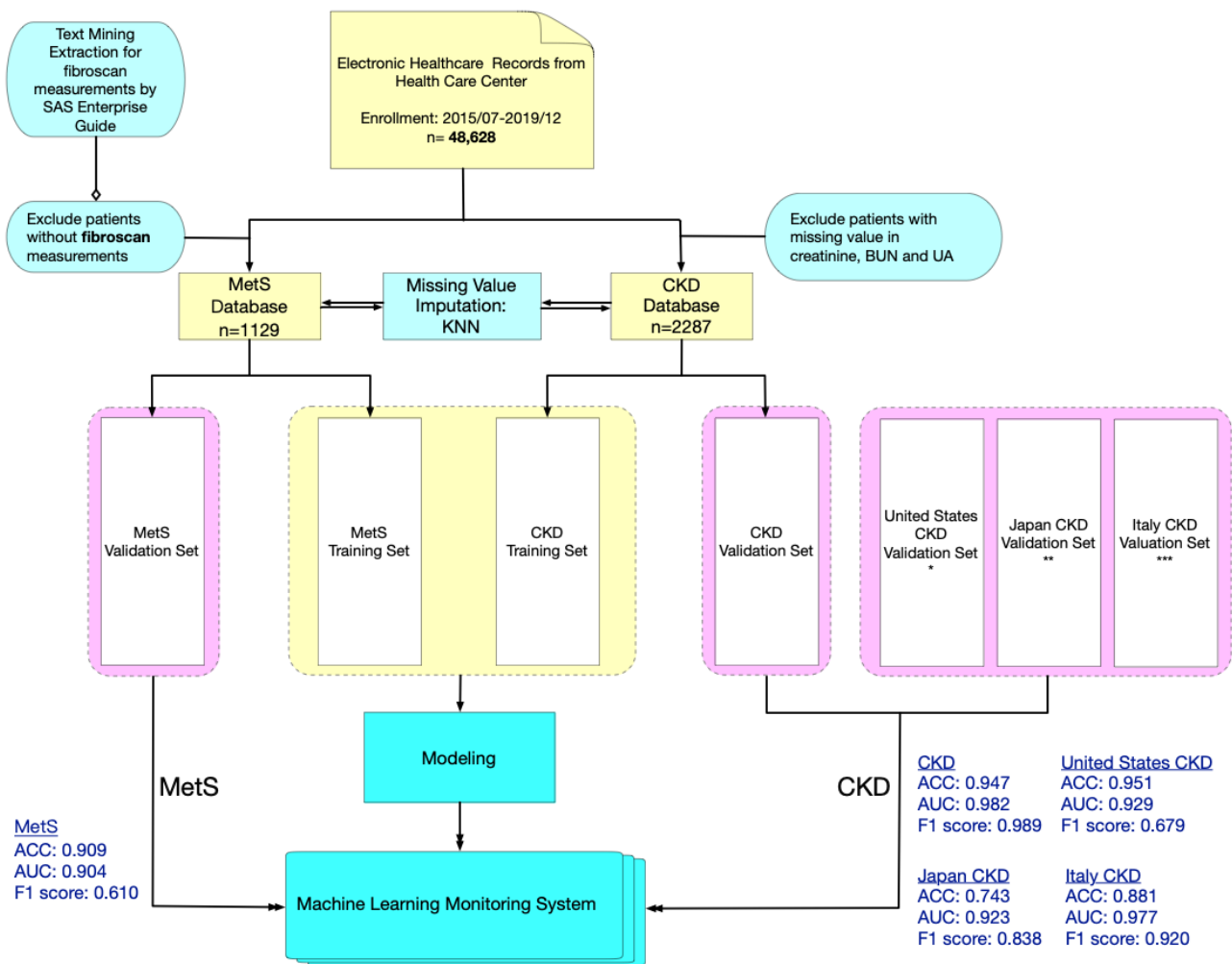
Subsequently, the starting population data underwent data cleaning and preprocessing to form two distinct databases (MetS and CKD) for ML. For the MetS database, there were a total of 1129 participants after the exclusion of participants without FibroScan (Echosens) measurements. For the CKD database, there were a total of 2287 participants after the exclusion of participants without values for creatinine, blood urea nitrogen, and uric acid.

Due to the inconsistent definition of MetS across the world, the ML performance of the MetS database and the CKD database were validated using different study populations. The ML performance of the CKD database was validated using Taiwanese, Italian, US, and Japanese data sets, but the ML performance of the CKD database was only validated using a Taiwanese data set [37-39]. Since different variables may be unavailable in different validation data sets, unavailable variables were simply excluded in ML performance analysis for a balanced comparison.

ML Techniques

The ML techniques used in this system included supervised learning models, such as classification and regression tree

(CART) and random forest [35,36]. Supervised learning was applied to classify the patients in the training set and predict patients with a specific chronic disease or syndrome in the validation set before the prediction model was available on this system [40,41]. In addition, unsupervised learning (hierarchical clustering using the Ward method and Euclidean distance) was embedded in a heat map, providing classified visualization between new input records and the database. An interactive heat map that could be rearranged or zoomed in and out was applied to this system [42-47].

All outcomes were presented on the web platform after the ML system evaluated the users' EMRs. Although the ML system was developed on a web-based interface, it could be embedded in the Internet of Medical Things (IoMT) environment, for example, as apps or real-time monitoring systems between several medical centers and hospitals [48-50].

Questionnaire Selection

To measure the usability of websites, we invited potential users of the ML system (physicians, medical staff, and potential users) to fill out a system usability scale (SUS) evaluation questionnaire. SUS was chosen as the usability test tool because

previous studies found it to be reliable and quick to answer, and the final score is provided with interpretation based on a well-established reference standard [51,52]. In general, the higher the SUS score, the better the usability of the website. Details about the questionnaire design (the 10 questions), score summary, and results of reliability and validity tests are given in Multimedia Appendix 1.

Results

The web-based health care ML system provides online diagnosis of three diseases (Figure 3), and it is available on the internet [53]. The website provides an assessment of MetS and CKD; the system for noncancer liver disease is still under beta testing. Report pages are provided for online diagnosis of each disease. Therefore, users from all over the world can choose the evaluation provided depending on their requirements. Users input the predicting variables (Table 1) into the website to evaluate their health (Figure 4), and the evaluation results will appear in <5 seconds when there is a single request. Missing predicting variables are allowed, and the missing values will be imputed based on the mean values from the database. However, the users are warned that missing predicting variables may result in poorer prediction accuracy. The details of stress tests with different numbers of requests (100 to 800) can be found in Multimedia Appendix 2. Briefly, a stress test with 800 requests reports a throughput of 4.7 requests per second. To evaluate the usability of the system, we invited 30 volunteers to complete the SUS evaluation questionnaire. The volunteers included 6 physicians, 12 medical staff, and 12 potential users (Multimedia Appendix 1). It was found that the average SUS score is 74, which indicates a good usability rating [54]. In addition, results were found to be reliable and valid by Kaiser-Meyer-Olkin and Bartlett tests. The entire analysis process follows a strict privacy policy, so that none of the patients' private information is ever recorded.

Figure 3. Home page of the machine learning health care system.

Figure 4. Interface of the input page for disease assessment.

The clinical outcomes established by our database are reported on the website when users have finished entering their medical record data on the website (Figure 5). The CART model and ensemble learning model (random forest) are shown in the output interface. A scoring prediction model obtained using the supervised learning model is also provided online. For unsupervised learning, a color visualization of the clustering heat map depicts a vivid medical pattern of the patient's EMR data, and a record of each user is also constructed using hierarchical clustering with yellow highlights labeling in the heat map (Figure 6). The user will then be classified as more similar to either a healthy subject (green column on the lower left) or an unhealthy subject (orange column on the upper left). A blue bar depicts abnormal values, while a red bar depicts normal values. In addition, on the web system, users can choose to view it as landscape or portrait. The zoom-in and zoom-out functions and the height of the cluster are also dynamic, with users being able to change the settings online to inspect the medical outcomes in detail.

Figure 5. Outcome page for supervised learning models and the scoring system for disease diagnosis.
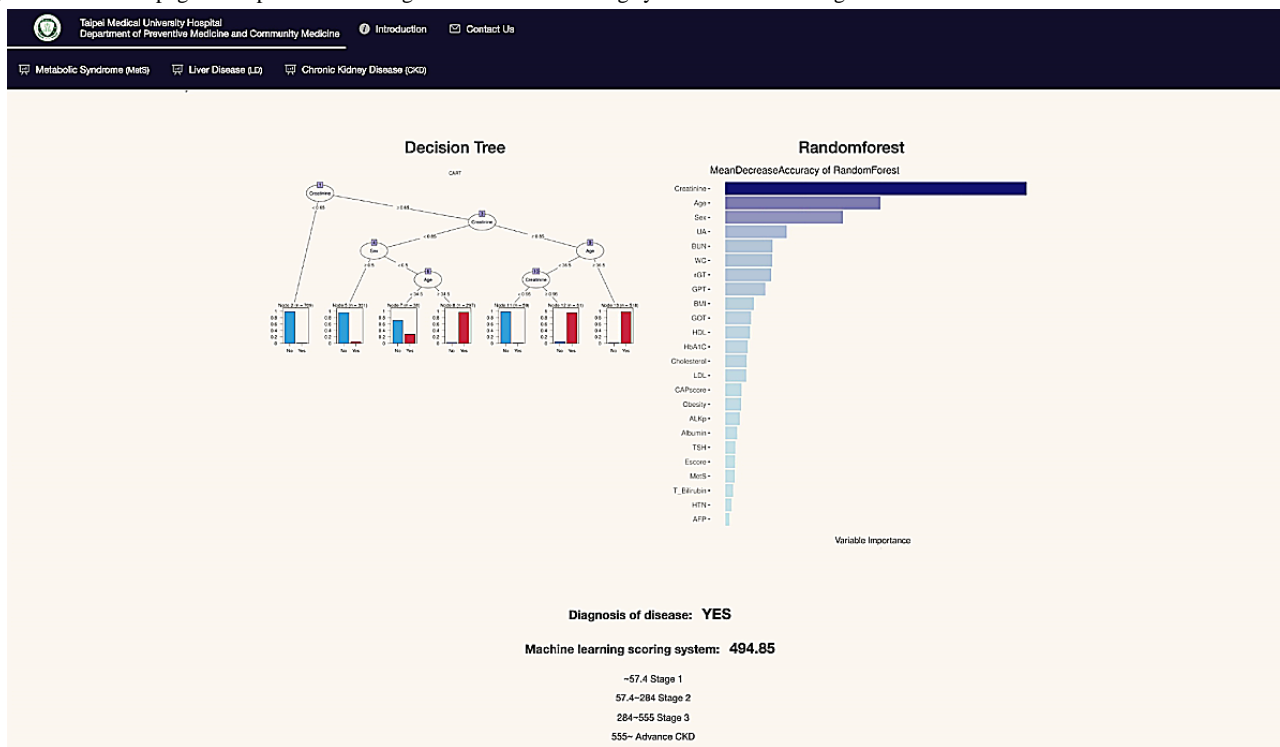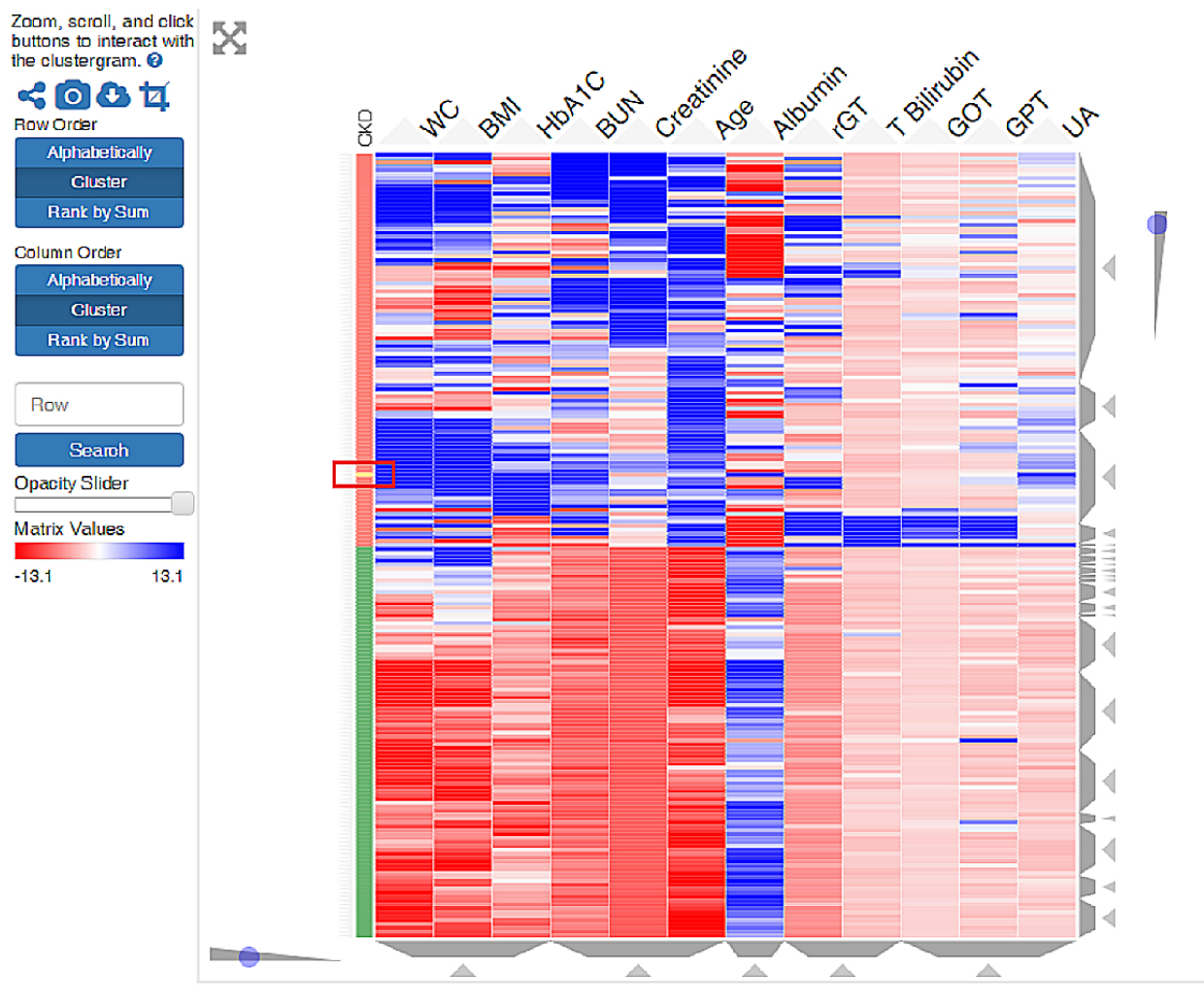
Figure 6. Dynamic interactive heat map obtained using unsupervised clustering. Green: healthy patients; orange: CKD patients; blue: normal values; red: abnormal values. The new patients (yellow bar in red rectangle) are compared and clustered into the system's patient database.

Characteristics of participants in the training and validation data set for MetS can be found in Table 2 and the characteristics of participants of the training set and the validation sets for CKD can be found in Table 3. In general, there are minimal differences in patient characteristics between the training data set and the validation data set for the Taiwanese population of MetS and CKD. However, when comparing the characteristics of the Taiwanese population with other populations (US, Italy, and Japan) for CKD ML performance validation, it was found that there are substantial differences in age and presence of hypertension (Table 3).

Table 2. Characteristics of participants in the training and validation data set for metabolic syndrome.

Characteristics	Training data set (n=904)	Validation data set (n=225)
Sex		
Female, n (%)	411 (45.5)	108 (48.0)
Male, n (%)	493 (54.5)	117 (52.0)
Age, years, median (IQR)	44 (37-50.25)	43 (38-50)
Body mass index, kg/m^2, median (IQR)	23.6 (21.3-25.9)	22.9 (21.2-26)
Waist circumference, cm, median (IQR)	81.75 (74.5-88)	80.5 (74-87)
Albumin, g/dL, median (IQR)	4.6 (4.4-4.8)	4.6 (4.4-4.8)
Alkaline phosphatase, IU/L, median (IQR)	58 (49-69)	58 (48-69)
Glutamic-oxaloacetic transaminase, IU/L, median (IQR)	20 (17-24.25)	20 (17-25)
Glutamate pyruvate transaminase, IU/L, median (IQR)	20 (14-31)	19 (14-28)
Total bilirubin, mg/dL, median (IQR)	0.6 (0.5-0.9)	0.6 (0.4-0.8)
γ-Glutamyl transpeptidase, U/L, median (IQR)	18 (12-27)	17 (11.55-25)
Controlled attenuation parameter (CAP) score, dB/m, median (IQR)	247 (211-284)	241 (216-282)
Elastic modulus (E) score, kPa, median (IQR)	4.2 (3.4-4.9)	4 (3.3-4.8)
Blood urea nitrogen, mg/dL, median (IQR)	12 (10-15)	12 (10-14)
Creatinine, mg/dL, median (IQR)	0.8 (0.6-0.9)	0.8 (0.6-0.9)
Estimated glomerular filtration rate (eGFR) using the Modification of Diet in Renal Disease (MDRD) equation, median (IQR)	90.23 (80.49-104.77)	91.28 (82.72-107.26)
Uric acid, mg/dL, median (IQR)	5.5 (4.5-6.6)	5.4 (4.3-6.6)
Systolic blood pressure, mm Hg, median (IQR)	114 (105-125)	114 (105-126)
Diastolic blood pressure, mm Hg, median (IQR)	73 (67-80)	72 (66-81)
Cholesterol, mg/dL, median (IQR)	189 (165-209)	185 (168-209)
Triglycerides, mg/dL, median (IQR)	90 (65-135.2)	86 (63-126)
High-density lipoprotein, mg/dL, median (IQR)	55(45-67)	54 (47-66)
Low-density lipoprotein, mg/dL, median (IQR)	123 (102-145)	123 (102-145)
Hemoglobin A_{1c}, %, median (IQR)	5.4 (5.2-5.6)	5.4 (5.2-5.5)
Glucose AC, mg/dL, median (IQR)	91 (86-96)	90 (85-95)

XSL•FO
RenderX

Table 3. Characteristics of participants in the training and validation data sets for chronic kidney disease.

Characteristics	Training data set (n=1830)	Validation data set, Taiwan (n=457)	Validation data set, United States (n=4434)	Validation data set, Italy (n=655)	Validation data set, Japan (n=996)
Sex, male, n (%)	902 (49.29)	209 (45.73)	2165 (48.83)	384 (58.63)	696 (69.88)
Chronic kidney disease, n (%)	164 (8.96)	38 (8.32)	410 (9.25)	523 (79.85)	919 (92.27)
Hypertension, n (%)	522 (28.52)	140 (30.63)	1730 (39.02)	599 (91.45)	908 (91.16)
Age, years, median (IQR)	46 (38-55)	45 (37-55)	53 (36-65)	67 (56-74.5)	70 (61-77)
Body mass index, kg/m^2, median (IQR)	23.8 (21.4-26.4)	23.4 (21.3-26.2)	28.6 (24.8-33.5)	28.4 (25.8-31.6)	23.25 (21-25.8)
Waist circumference, cm, median (IQR)	82.5 (75.5-89.5)	81 (75-89)	99.5 (89-111.3)	—[a]	—
Glutamic-oxaloacetic transaminase, IU/L, median (IQR)	21 (17-26)	20 (17-25)	19 (16-24)	—	—
Glutamate pyruvate transaminase, IU/L, median (IQR)	20 (14-30)	19 (13-28)	18 (13-26)	—	—
γ-Glutamyl transpeptidase, U/L, median (IQR)	19 (13-30)	18 (12-33)	21 (15-33)	—	—
Total bilirubin, mg/dL, median (IQR)	0.6 (0.4-0.8)	0.6 (0.4-0.8)	0.4 (0.3-0.6)	—	—
Alkaline phosphatase, IU/L, median (IQR)	62 (51-76)	63 (50-78)	75 (62-91)	—	—
Blood urea nitrogen, mg/dL, median (IQR)	13 (11-16)	13 (10-15)	14 (11-18)	28 (21.2-37.3)	—
Creatinine, mg/dL, median (IQR)	0.8 (0.6-1.0)	0.7 (0.6-0.9)	0.85 (0.71-1.01)	1.49 (1.2-1.9)	1.8 (1.2-2.75)
Uric acid, mg/dL, median (IQR)	5.5 (4.5-6.7)	5.4 (4.5-6.5)	5.3 (4.4-6.4)	6.3 (5.2-7.6)	—
Albumin, g/dL, median (IQR)	4.6 (4.4-4.8)	4.6 (4.4-4.8)	4.1 (3.9-4.3)	4 (3.7-4.3)	4 (3.5-4.3)
Cholesterol, mg/dL, median (IQR)	186 (164-210)	185 (160-209)	185 (160-214)	189 (162.5-218)	—
High-density lipoprotein, mg/dL, median (IQR)	52 (44-64)	53 (43-64)	51 (42-61)	—	—
Low-density lipoprotein, mg/dL, median (IQR)	121 (100-145)	120 (100-142)	—	—	—
Hemoglobin A$_{1c}$, %, median (IQR)	5.4 (5.2-5.6)	5.4 (5.2-5.7)	5.6 (5.3-6)	—	—

[a]Not available.

[Table 4](#) shows the validation performances of supervised learning models in predicting MetS and CKD. In general, it was found that the random forest ML model has higher accuracy than the CART model. Using the random forest ML model, MetS can be predicted with an accuracy of 0.909, and CKD can be predicted up to an accuracy of 0.947. Due to the inconsistent definition of MetS globally, the ML performance of the MetS database has only been validated using the Taiwan data set. However, the ML performances of the CKD database have been validated using data sets from Taiwan, Italy, the United States, and Japan. In general, the CKD database shows good external applicability, and has high AUC for all 4 validation data sets (Taiwan: AUC=0.982; USA: AUC=0.929; Italy: AUC=0.977; Japan: AUC=0.923). However, the validation accuracy and F1 value of CKD prediction differs more substantially, as the unavailable data were excluded from the analysis. When compared to the Taiwanese CKD data set, the respective unavailable data are approximately 6% for the US data set, 50%

for the Italy data set, and 67% for the Japan data set. Therefore, it is observed that the Japanese validation data set has the lowest accuracy (0.743) in predicting CKD, as it also has the highest proportion of unavailable data.

Table 4. The performance of supervised learning models on predicting metabolic syndrome and chronic kidney disease.

Model and disease	Accuracy	Area under the curve (AUC)	F1 score
Classification and regression tree (CART)			
Metabolic syndrome (Taiwan)	0.874	0.887	0.448
Chronic kidney disease (Taiwan)	0.945	0.928	0.965
Random forest			
Metabolic syndrome (Taiwan)	0.909	0.904	0.610
Chronic kidney disease (Taiwan)	0.947	0.982	0.989
Chronic kidney disease (United States)	0.951	0.929	0.679
Chronic kidney disease (Italy)	0.881	0.977	0.920
Chronic kidney disease (Japan)	0.743	0.923	0.838

Discussion

Overview

This ML medical system for three common diseases in family medicine (MetS, CKD, and liver diseases) was constructed from EMR subjects who underwent self-paid health examination. Several ML prediction models are applied to the databases, and the outcomes are summarized and presented visually on the website for users and medical staff. The accuracy of predicting MetS reached 90.9%, and AUC was 0.904 in this system. For chronic kidney disease, the prediction accuracy reached 94.7%, and the AUC was 0.982. In general, users who were invited to test this system rated it with good usability and could easily assess their health online through this web-based ML monitoring system.

CART

Decision trees are an important type of ML algorithm for predictive modeling. They are commonly used in data mining with the objective of creating a model that predicts the dependent variable (the target) based on numerous independent variables [34,37].

A decision tree is a nonparametric ML modeling technique used for regression and classification problems. In classification problems, the target variable is categorical, and the tree is used to identify which group or class a target variable would likely fall into. In regression problems, the target variable is continuous, and the tree is used to predict its value. To find solutions, a decision tree makes a sequential, hierarchical decision about the outcome's variable according to the predictor [55-57].

Hence, CART can provide a visual tree-based diagram for medical practitioners to disseminate health care information to patients. It also helps users to understand the significance of different risk factors for specific diseases. For example, the cut-off controlled attenuation parameter (CAP) score was used to separate patients with MetS and those with other health observations. The CAP score was brought to the attention of users, thereby increasing their awareness of self-health [34].

Random Forest

Random forest, also called random decision forests, is a popular ensemble learning method in ML. Ensemble methods use multiple learning algorithms to improve ML results by combining several decision tree models. This approach allows better predictive performance compared with a single model. Random forest is a parallel ensemble method in which the base learners are generated in parallel. The basic motivation of parallel methods is to exploit independence between the base learners because the error can be reduced dramatically by averaging [58,59]. As random forest provides a bagging technique for feature estimates, it also offers efficient estimates of the test error without incurring the cost of repeated model training associated with cross-validation. Moreover, random forest ranks risk factors in prediction models, which clinicians can use as a reference for diagnosis, and remote users can use to review their risk assessment of related diseases [34,60-62]. For instance, clinicians can refer to significant factors of certain diseases to determine whether those factors exceed the thresholds or not, allowing patients to be more vigilant about their risk of developing such diseases. In addition, sequential ensemble methods such as AdaBoost and XGBoost will be implemented and uploaded to our system in the future.

Clustering

Hierarchical clustering is a widely used unsupervised learning technique that groups data with similar characteristics. Both agglomerative and divisive approaches use dendrograms for the results. A heat map is a color graphical representation of data, which uses a matrix with color gradients to present the similarity of data.

Many studies on genetic bioinformatics and bacterial ecology have used heat maps for the analysis of large and complicated data sets, and some medical studies have used heat maps with clustering to present the relationship between various biomarkers according to their characteristics [34,63-65]. Furthermore, our system provides an interactive clustering heat map for health care. From the perspective of big data, users can evaluate their health status by using ML models and EMR databases. In addition to online health evaluation, in the future, this system

could be implemented into different IoMT to assist medical practitioners in achieving real-time health evaluations and monitoring remote patients or patients in specific wards. For the heat map, the EMR data of users were grouped into clusters of patients with diseases in the database; they would then be classified as clinically high-risk objects requiring close attention in the clinical setting [34]. Therefore, whether it is applied in preventive medicine for health management, in a monitoring system for critical care, or in the telemedicine environment, our system can provide real-time monitoring and help predict patient conditions.

Limitations and Future Work

To the best of our knowledge, this is the first web-based machine learning system based on self-paid health examination subjects that can provide an online self-health evaluation for several common diseases (MetS, CKD, and liver diseases). The version 1.0 web-based system still has several limitations that may be improved in the next update. First, the 1.0 system is not yet ready for embedding into a hospital for real-time assessment. We are currently working on an improved system to accept unstructured data input and multimodal data, which are especially essential for the prediction of eye diseases such as macular degeneration. Second, the 1.0 system did not have a user login or account security function. Retrievable prediction and security will be improved as the system is matured for hospital embedment. Third, the 1.0 system does not have whole dynamic analyses such as an interactive decision tree; whole dynamic analyses will be incorporated in subsequent versions to improve communication between the medical staff and patients.

In the future, more clustering algorithms will be implemented in subsequent versions to make the prediction results more robust and reliable. Although the 1.0 system can currently only evaluate three chronic diseases (MetS, CKD, and liver diseases) frequently encountered in family medicine, more chronic disease prediction models, such as those for coronary artery disease, will be added in the near future.

Conclusion

We constructed an ML health monitoring system to offer an online health assessment service to medical units, telemedicine patients, and all health-conscious users worldwide. Our aim is that this system will be implemented in medical centers as a real-time patient monitoring system and provide regular health evaluations for telemedicine patients. Online users can now access our platform and use ML technology to estimate their health status, increasing self-health awareness.

Acknowledgments

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire information.
[PDF File (Adobe PDF File), 10106 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Heat map clustering example and system loading test results.
[DOCX File , 2487 KB-Multimedia Appendix 2]

References

1. Goodman KW. Ethics, Computing, and Medicine: Informatics and the Transformation of Health Care. Cambridge, England: Cambridge University Press; 1997.

2. Kim J. Big Data, Health Informatics, and the Future of Cardiovascular Medicine. J Am Coll Cardiol 2017 Feb 21;69(7):899-902 [FREE Full text] [doi: 10.1016/j.jacc.2017.01.006] [Medline: 28209228]

3. Chien TW, Wang WC, Huang SY, Lai WP, Chow JC. A web-based computerized adaptive testing (CAT) to assess patient perception in hospitalization. J Med Internet Res 2011 Aug 15;13(3):e61 [FREE Full text] [doi: 10.2196/jmir.1785] [Medline: 21844001]

4. De Beurs DP, de Vries AL, de Groot MH, de Keijser J, Kerkhof AJ. Applying computer adaptive testing to optimize online assessment of suicidal behavior: a simulation study. J Med Internet Res 2014;16(9):e207 [FREE Full text] [doi: 10.2196/jmir.3511] [Medline: 25213259]

5. Pan L, Liu G, Mao X, Li H, Zhang J, Liang H, et al. Development of Prediction Models Using Machine Learning Algorithms for Girls with Suspected Central Precocious Puberty: Retrospective Study. JMIR Med Inform 2019 Feb 12;7(1):e11728 [FREE Full text] [doi: 10.2196/11728] [Medline: 30747712]

6. Zheng L, Wang Y, Hao S, Shin AY, Jin B, Ngo AD, et al. Web-based Real-Time Case Finding for the Population Health Management of Patients With Diabetes Mellitus: A Prospective Validation of the Natural Language Processing-Based

Algorithm With Statewide Electronic Medical Records. JMIR Med Inform 2016 Nov 11;4(4):e37 [FREE Full text] [doi: 10.2196/medinform.6328] [Medline: 27836816]

7. Jones M, Koziel C, Larsen D, Berry P, Kubatka-Willms E. Progress in the Enhanced Use of Electronic Medical Records: Data From the Ontario Experience. JMIR Med Inform 2017 Feb 22;5(1):e5 [FREE Full text] [doi: 10.2196/medinform.6928] [Medline: 28228372]

8. Installé AJ, Van den Bosch T, De Moor B, Timmerman D. Clinical data miner: an electronic case report form system with integrated data preprocessing and machine-learning libraries supporting clinical diagnostic model research. JMIR Med Inform 2014 Oct 20;2(2):e28 [FREE Full text] [doi: 10.2196/medinform.3251] [Medline: 25600863]

9. Fei Y, Hu J, Gao K, Tu J, Li WQ, Wang W. Predicting risk for portal vein thrombosis in acute pancreatitis patients: A comparison of radical basis function artificial neural network and logistic regression models. J Crit Care 2017 Jun;39:115-123. [doi: 10.1016/j.jcrc.2017.02.032] [Medline: 28246056]

10. Santelices LC, Wang Y, Severyn D, Druzdzel MJ, Kormos RL, Antaki JF. Development of a hybrid decision support model for optimal ventricular assist device weaning. Ann Thorac Surg 2010 Sep;90(3):713-720 [FREE Full text] [doi: 10.1016/j.athoracsur.2010.03.073] [Medline: 20732482]

11. Baxt WG. Use of an artificial neural network for the diagnosis of myocardial infarction. Ann Intern Med 1991 Dec 01;115(11):843-848. [doi: 10.7326/0003-4819-115-11-843] [Medline: 1952470]

12. Artis SG, Mark RG, Moody GB. Detection of atrial fibrillation using artificial neural networks. In: 1991 Proceedings of Computers in Cardiology.: IEEE; 1991 Presented at: Computers in Cardiology (CinC); September 23-26; Venice, Italy p. 173-176 URL: https://ieeexplore.ieee.org/document/169073 [doi: 10.1109/CIC.1991.169073]

13. Eftekhar B, Mohammad K, Ardebili HE, Ghodsi M, Ketabchi E. Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data. BMC Med Inform Decis Mak 2005 Feb 15;5:3 [FREE Full text] [doi: 10.1186/1472-6947-5-3] [Medline: 15713231]

14. Foster KR, Koprowski R, Skufca JD. Machine learning, medical diagnosis, and biomedical engineering research - commentary. Biomed Eng Online 2014 Jul 05;13:94 [FREE Full text] [doi: 10.1186/1475-925X-13-94] [Medline: 24998888]

15. Ravi D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, et al. Deep Learning for Health Informatics. IEEE J Biomed Health Inform 2017 Dec;21(1):4-21. [doi: 10.1109/JBHI.2016.2636665] [Medline: 28055930]

16. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. N Engl J Med 2016 Sep 29;375(13):1216-1219 [FREE Full text] [doi: 10.1056/NEJMp1606181] [Medline: 27682033]

17. Zhou L, Pan S, Wang J, Vasilakos AV. Machine learning on big data: Opportunities and challenges. Neurocomputing 2017 May;237:350-361. [doi: 10.1016/j.neucom.2017.01.026]

18. Verbrugge LM, Reoma JM, Gruber-Baldini AL. Short-term dynamics of disability and well-being. J Health Soc Behav 1994 Jun;35(2):97-117. [Medline: 8064125]

19. Chatterji S, Byles J, Cutler D, Seeman T, Verdes E. Health, functioning, and disability in older adults--present status and future implications. Lancet 2015 Feb 07;385(9967):563-575 [FREE Full text] [doi: 10.1016/S0140-6736(14)61462-8] [Medline: 25468158]

20. Leavell HR, Clark EG. Textbook of Preventive Medicine. New York: McGraw-Hill; 1958.

21. Jung P, Lushniak BD. Preventive Medicine's Identity Crisis. Am J Prev Med 2017 Mar;52(3):e85-e89. [doi: 10.1016/j.amepre.2016.10.037] [Medline: 28012813]

22. Margolis SA. Preventive healthcare: A core component of Australian general practice. Aust J Gen Pract 2018 Dec;47(12):821 [FREE Full text] [doi: 10.31128/AJGP-11-18-4751] [Medline: 31212397]

23. Hensrud DD. Clinical preventive medicine in primary care: background and practice: 1. Rationale and current preventive practices. Mayo Clin Proc 2000 Feb;75(2):165-172. [doi: 10.4065/75.2.165] [Medline: 10683656]

24. Kassi E, Pervanidou P, Kaltsas G, Chrousos G. Metabolic syndrome: definitions and controversies. BMC Med 2011 May 05;9:48 [FREE Full text] [doi: 10.1186/1741-7015-9-48] [Medline: 21542944]

25. Ding C, Yang Z, Wang S, Sun F, Zhan S. The associations of metabolic syndrome with incident hypertension, type 2 diabetes mellitus and chronic kidney disease: a cohort study. Endocrine 2018 May;60(2):282-291. [doi: 10.1007/s12020-018-1552-1] [Medline: 29492904]

26. Grundy SM. Metabolic syndrome: a multiplex cardiovascular risk factor. J Clin Endocrinol Metab 2007 Feb;92(2):399-404. [doi: 10.1210/jc.2006-0513] [Medline: 17284640]

27. Lorenzo C, Okoloise M, Williams K, Stern MP, Haffner SM, San Antonio Heart Study. The metabolic syndrome as predictor of type 2 diabetes: the San Antonio heart study. Diabetes Care 2003 Nov;26(11):3153-3159. [doi: 10.2337/diacare.26.11.3153] [Medline: 14578254]

28. Lin YJ, Lin CH, Wang ST, Lin SY, Chang SS. Noninvasive and Convenient Screening of Metabolic Syndrome Using the Controlled Attenuation Parameter Technology: An Evaluation Based on Self-Paid Health Examination Participants. J Clin Med 2019 Oct 24;8(11):1775 [FREE Full text] [doi: 10.3390/jcm8111775] [Medline: 31653028]

29. Levey AS, Coresh J, Balk E, Kausz AT, Levin A, Steffes MW, National Kidney Foundation. National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Ann Intern Med 2003 Jul 15;139(2):137-147. [doi: 10.7326/0003-4819-139-2-200307150-00013] [Medline: 12859163]

30. Drawz P, Rahman M. Chronic kidney disease. Ann Intern Med 2015 Jun 02;162(11):ITC1-IT16. [doi: 10.7326/AITC201506020] [Medline: 26030647]

31. Webster AC, Nagler EV, Morton RL, Masson P. Chronic Kidney Disease. Lancet 2017 Mar 25;389(10075):1238-1252. [doi: 10.1016/S0140-6736(16)32064-5] [Medline: 27887750]

32. Trettel A, Eissing L, Augustin M. Telemedicine in dermatology: findings and experiences worldwide - a systematic literature review. J Eur Acad Dermatol Venereol 2018 Feb;32(2):215-224. [doi: 10.1111/jdv.14341] [Medline: 28516492]

33. van der Vaart R, Drossaert C. Development of the Digital Health Literacy Instrument: Measuring a Broad Spectrum of Health 1.0 and Health 2.0 Skills. J Med Internet Res 2017 Jan 24;19(1):e27 [FREE Full text] [doi: 10.2196/jmir.6709] [Medline: 28119275]

34. Yu CS, Lin CH, Lin YJ, Lin SY, Wang ST, Wu JL, et al. Clustering Heatmap for Visualizing and Exploring Complex and High-dimensional Data Related to Chronic Kidney Disease. J Clin Med 2020 Feb 02;9(2):403 [FREE Full text] [doi: 10.3390/jcm9020403] [Medline: 32024311]

35. Breiman L, Friedman JH, Stone CJ, Olshen RA. Classification and Regression Trees. Boca Raton, Florida: CRC Press; Jan 1, 1984.

36. Yu CS, Lin YJ, Lin CH, Wang ST, Lin SY, Lin SH, et al. Predicting Metabolic Syndrome With Machine Learning Models Using a Decision Tree Algorithm: Retrospective Cohort Study. JMIR Med Inform 2020 Mar 23;8(3):e17110 [FREE Full text] [doi: 10.2196/17110] [Medline: 32202504]

37. Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data-NHANES 2017-2018. Hyattsville, MD: US Department of Health and Human Services, Centers for Disease Control and Prevention URL: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017 [accessed 2020-03-27]

38. De Nicola L, Provenzano M, Chiodini P, Borrelli S, Garofalo C, Pacilio M, et al. Independent Role of Underlying Kidney Disease on Renal Prognosis of Patients with Chronic Kidney Disease under Nephrology Care. PLoS One 2015;10(5):e0127071 [FREE Full text] [doi: 10.1371/journal.pone.0127071] [Medline: 25992629]

39. Iimori S, Naito S, Noda Y, Sato H, Nomura N, Sohara E, et al. Prognosis of chronic kidney disease with normal-range proteinuria: The CKD-ROUTE study. PLoS One 2018 Jan 17;13(1):e0190493 [FREE Full text] [doi: 10.1371/journal.pone.0190493] [Medline: 29342207]

40. Peute L, Scheeve T, Jaspers M. Classification and Regression Tree and Computer Adaptive Testing in Cardiac Rehabilitation: Instrument Validation Study. J Med Internet Res 2020 Jan 30;22(1):e12509 [FREE Full text] [doi: 10.2196/12509] [Medline: 32012065]

41. Caruana R, Niculescu-Mizil A. An empirical comparison of supervised learning algorithms. In: Proceedings of the 23rd International Conference on Machine Learning. 2006 Presented at: International Conference on Machine Learning; June 25-29; Pittsburgh, Pennsylvania p. 161-168 URL: https://dl.acm.org/doi/10.1145/1143844.1143865 [doi: 10.1145/1143844.1143865]

42. Rokach L, Maimon O. Clustering Methods. In: Data Mining and Knowledge Discovery Handbook. Boston, MA: Springer; 2005:321-352.

43. Ward JH. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 1963 Mar;58(301):236-244. [doi: 10.1080/01621459.1963.10500845]

44. Cormack RM. A Review of Classification. Journal of the Royal Statistical Society. Series A (General) 1971;134(3):321. [doi: 10.2307/2344237]

45. Wilkinson L, Friendly M. The History of the Cluster Heat Map. The American Statistician 2009 May;63(2):179-184. [doi: 10.1198/tas.2009.0033]

46. Perrot A, Bourqui R, Hanusse N, Lalanne F, Auber D. Large interactive visualization of density functions on big data infrastructure. In: 2015 IEEE 5th Symposium on Large Data Analysis and Visualization (LDAV). 2015 Presented at: LDAV 2015 - Big Data Analysis and Visualization; October 25-26; Chicago, IL, USA URL: https://ieeexplore.ieee.org/document/7348077 [doi: 10.1109/ldav.2015.7348077]

47. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, et al. gplots: Various R Programming Tools for Plotting Data. 2020. URL: https://cran.r-project.org/web/packages/gplots/index.html [accessed 2020-02-26]

48. Joyia GJ, Liaqat RM, Farooq A, Rehman S. Internet of Medical Things (IOMT): Applications, Benefits and Future Challenges in Healthcare Domain. JCM 2017;12(4):240-247 [FREE Full text] [doi: 10.12720/jcm.12.4.240-247]

49. Istepanian RSH, Sungoor A, Faisal A, Philip N. Internet of m-health Things (m-IoT). In: IET Seminar on Assisted Living 2011. 2011 Presented at: IET Seminar on Assisted Living; April 6; London, UK p. 1-40 URL: https://ieeexplore.ieee.org/document/6183148

50. Boutros-Saikali N, Saikali K, Naoum RA. An IoMT platform to simplify the development of healthcare monitoring applications. In: 2018 Third International Conference on Electrical and Biomedical Engineering, Clean Energy and Green Computing (EBECEGC). 2018 Presented at: EBECEGC 2018: The Third International Conference on Electrical and Biomedical Engineering, Clean Energy and Green Computing (EBECEGC2018); April 25-27; Beirut, Lebanon p. 25-27. [doi: 10.1109/ebecegc.2018.8357124]

51. Lewis JR, Sauro J. The Factor Structure of the System Usability Scale. In: HCD 2009: Human Centered Design. 2009 Presented at: International Conference on Human Centered Design; July 19-24; San Diego, CA, USA p. 94-103 URL: https://link.springer.com/content/pdf/10.1007%2F978-3-642-02806-9.pdf

52. Flavián C, Guinalíu M, Gurrea R. The role played by perceived usability, satisfaction and consumer trust on website loyalty. Information & Management 2006 Jan;43(1):1-14. [doi: 10.1016/j.im.2005.01.002]

53. Yu CS, Chang SS. ML-Doctor. URL: https://www.mldoctor.com.tw [accessed 2020-02-15]

54. T W. Measuring and Interpreting System Usability Scale (SUS). URL: https://uiuxtrend.com/measuring-system-usability-scale-sus/ [accessed 2020-04-01]

55. Friedl MA, Brodley CE. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment 1997 Sep;61(3):399-409. [doi: 10.1016/s0034-4257(97)00049-7]

56. Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. IEEE Transactions on Systems, Man, and Cybernetics 1991;21(3):660-674. [doi: 10.1109/21.97458]

57. Fan CY, Chang PC, Lin JJ, Hsieh JC. A hybrid model combining case-based reasoning and fuzzy decision tree for medical data classification. Applied Soft Computing 2011 Jan;11(1):632-644. [doi: 10.1016/j.asoc.2009.12.023]

58. Ho TK. The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Machine Intell 1998;20(8):832-844. [doi: 10.1109/34.709601]

59. Haste T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York, NY: Springer; 2009.

60. Liaw A, Wiener M. Classification and regression by randomForest. R News 2002;2/3:18-22 [FREE Full text]

61. Archer KJ, Kimes RV. Empirical characterization of random forest variable importance measures. Computational Statistics & Data Analysis 2008 Jan;52(4):2249-2260. [doi: 10.1016/j.csda.2007.08.015]

62. Yang F, Wang HZ, Mi H, Lin CD, Cai WW. Using random forest for reliable classification and cost-sensitive learning for medical diagnosis. BMC Bioinformatics 2009 Jan 30;10 Suppl 1:S22 [FREE Full text] [doi: 10.1186/1471-2105-10-S1-S22] [Medline: 19208122]

63. Benton PH, Ivanisevic J, Rinehart D, Epstein A, Kurczy ME, Boska MD, et al. An Interactive Cluster Heat Map to Visualize and Explore Multidimensional Metabolomic Data. Metabolomics 2015 Aug 01;11(4):1029-1034 [FREE Full text] [doi: 10.1007/s11306-014-0759-2] [Medline: 26195918]

64. Nithya NS, Duraiswamy K, Gomathy P. A Survey on Clustering Techniques in Medical Diagnosis. International Journal of Computer Science Trends and Technology 2013;1(2):17-22 [FREE Full text]

65. Ziemba YC, Lomsadze L, Jacobs Y, Chang TY, Haghi N. Using Heatmaps to Identify Opportunities for Optimization of Test Utilization and Care Delivery. J Pathol Inform 2018;9:31 [FREE Full text] [doi: 10.4103/jpi.jpi_7_18] [Medline: 30294500]

Abbreviations

AI: artificial intelligence

AUC: area under the curve

CAP: controlled attenuation parameter

CART: classification and regression tree

CKD: chronic kidney disease

eGFR: estimated glomerular filtration rate

EMRs: electronic medical records

HMC: Health Management Center

IoMT: Internet of Medical Things

IRB: institutional review board

MDRD: Modification of Diet in Renal Disease

MetS: metabolic syndrome

ML: machine learning

ROC: receiver operating characteristic

TMUH: Taipei Medical University Hospital

XSL•FO
RenderX