Annals of R.S.C.B., ISSN: 1583-6258, Vol.25, Issue 4, 2021, Pages. 19194-19204
Received 05 March 2021; Accepted 01 April 2021

Diseases Identification Method Using Machine Learning Classification in E-Healthcare

Muppala Nithendra Varma[1], Kothakapa Balaji Vivekvardhan[2], Karna Girinath Reddy[3], Prabu Sevugan[4*]

[1]School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India.
[2]School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India.
[3]School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India.
[4*]School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India.
E-mail: sprabu@vit.ac.in

ABSTRACT

Heart infection and diabetes may be the main cause of death on the planet today. The prognosis of COVID-19 is the main test for studying clinical information. Options and forecasts based on the vast amount of information provided by the healthcare industry. We also see AA continuous improvement practices used in various Internet of Things (IoT). The different tests only outline the COVID-19 prediction using machine learning strategies to obtain higher accuracy, 4 algorithms are analysed, namely Support Vector Machine (SVM), Decision Tree (DT), Nearest Neighbour Algorithm (KNN), and Random Forest classifier. Discussed and compared the performance and accuracy of the algorithms used. A comparison of the various machine learning methods used in this research shows which algorithm best predicts COVID-19.

KEYWORDS

COVID-19, Support Vector Machine, Decision Tree, Nearest Neighbour Algorithm, Random Forest classifier.

Introduction

1. Overview

Heart is quite possibly the most indispensable organs for the appropriate working of our body. Yet, as per the report by WHO, 31% of the overall passings consistently happens because of cardiovascular diseases (CVDs). Additionally, over 75% of these passings happen in low and center pay nations including India [1].

Its major goal is to anticipate precisely whether cardiovascular diseases are in the human body [2-6]. Older approaches do not forecast heart disease effectively very successfully. Many medical technologies on the market that forecast cardiac illness are highly costly and not enough to predict cardiac illness properly. Some of the main risk factors are obesity, poor diet, low physical activity, drinking and smoking, age, sex, hypertension, cholesterol, blood glucose, and diabetes [7-11]. Disease of the heart. Most dangers in your lifestyle can be controlled. In recent decades, technology innovations have been used extensively to enhance the quality of treatment. Advances in these technologies have made it possible to diagnose and predict correct diseases. Machine study might be a great alternative for predicting cardiovascular diseases when you analyse massive quantities of data and uncover patterns and trends [12-17]. Fast and dependable outcomes can also be achieved via machine training. Various soft computing approaches can be used to forecast heart health, including neural artificial networks (ANNs) [18-21].

Diabetes is a disease in which the body cannot regulate glucose-insulin levels after a few meals [22-26]. Due to unbalanced diets and unsuccessful lifestyles, the number of diabetic patients has increased dramatically [27-32]. People all over the world can benefit from smart medical innovation, thereby increasing their satisfaction. Diabetes can cause heart disease, kidney damage, blurred vision, and nerve damage. Especially the severe respiratory disease Covid 2 (SARS-CoV-2) mainly infects diabetic patients [33-37]. Facing the past known as the Middle East (MERS) and Severe Acute Respiratory Syndrome (SARS) (a variant of Covid) and the extremely severe 2009 H1N1 flu epidemic, people with diabetes also find themselves more powerless [38-45]. SARS-CoV-2 mainly affects the elderly and people with health problems. Various background analyses show that diabetes is the most important past comorbidity for COVID-19 patients. The difficulty in controlling blood sugar levels in diabetic patients after infection is related to the following factors:

1. The variance of glucose influences the insusceptibility of individuals that uncover him against COVID-19 and uneven glycemic profile may prompt a longer season of recovery for the patient.
2. The high blood glucose permits the infection to taint the human body without any problem.

- **Objective**

The major aim is to create a web application for cardiac illness prediction using machine learning [46-54]. This online application with the best accuracy is applied for prediction purposes after analysing and comparing the various ML algorithms [55-58].

- **Problem Definition**

Cardiovascular and diabetic diseases are considered to be one of the leading causes of death worldwide. Predicting them is very difficult for doctors because predicting them is a complex task that requires knowledge and experience. Nevertheless, medical diagnosis will improve the effectiveness of treatment and help reduce costs. We will develop a system that can effectively identify rules for predicting heart disease and diabetes based on patient health data. The goal is to find hidden patterns. Using sufficiently fast and reliable machine learning algorithms can predict Covid-19 by detecting heart and diabetes diseases of users and patients.

Literature Survey

Literature search is an objective and important review of printed analytical literature related to the research topic. Its purpose is to be familiar with current thinking and analysis of selected topics. It should ensure that future analysis is carried out in a space that has not been mentioned before or is rarely explored. Is the most important part of the report, because it provides direction in the field of analysis. This will help outline the purpose of the analysis and point out gaps. Review the project's literature, the research conducted by different analysts, their methods (essentially a summary of it), and the conclusions they need to find. In addition, he explained how this analysis affects the paper. Himanshu Sharma and Rizvi [1] explain a wealth of information available in the healthcare field and uses certain techniques to manage this information. Intelligent analysis of information is one of the most common methods. Coronary artery disease is the leading cause of death worldwide.

New perspectives on cardiovascular disease The results of this framework reflect the possibility of the frequency of coronary artery disease. The data set used is organized according to clinical boundaries. The framework uses mining pool strategies to estimate these limits. The two main machine learning algorithms, especially the decision tree algorithm and the naive Bayes algorithm, show the best calculation between the two in coronary artery disease accuracy. Dhar, S *et.al.* [3] in the last few decades, heart disease has been the largest cause of death worldwide, this article often explains. A full cardiovascular examination should be carried out to prevent cardiovascular disease or coronary artery disease and determine indications in time. Various smart technologies enhance the detection of cardiovascular diseases by healthcare professionals. In detecting and treating cardiovascular diseases, customized data mining strategies can ensure reasonable accuracy and reliability. Extracting usage data can reduce the number of tests. A fast and effective detection technique is required to ensure higher accuracy and precision to reduce cardiovascular deaths. The aim is to propose an effective method to predict cardiovascular disease. Use methods of machine learning. Therefore, we suggest a hybrid way to forecast heart disease using a random forest classification and a simple K-means machine. Two additional machine learning algorithms, namely the J48 classifier tree and the naive Bayes classifier, evaluate the data set and compare the results.

The confusion matrix shows the method's strength. G. Shanmugasundaram *et.al.*[2], Coronary artery disease refers to a cardiovascular disease. Chest pain is not an indication in all patients with coronary artery disease. Various factors such as R- blood pressure, S-cholesterol, f-glycemia, R-EKG and Ex-Ang can lead to an increase in coronary artery disease and the number of blocked large vessels, thallium scans and other factors. It is anticipated that coronary artery disease will save lives. To predict cardiac diseases based on constraints / factors, use naive bays, decision tree, k-nearest neighbour and other info extraction techniques. This study aims to study various variables and the importance they have to distinguish between coronary artery diseases. Modern waiting methods and models.

Ahmed [4] predicting and detecting cardiac disease has been a key issue for a long time. Early detection of cardiac disease is a major health issue (HCS). More and more health systems are offering very costly treatments and procedures to patients. Heart disease has recently become a common chronic disease, but in the United States it has received increasing attention. Tobacco use, poor lifestyles, sedentary lifestyles and drinking are the main causes of these diseases. An architecture in the cloud is needed to efficiently predict and monitor health information. Machine learning methods for clinical problems and medical diagnosis have been developed recently. This study proposes a four-layer, cloud-based architecture that can improve patient information prediction and monitoring substantially. This is why we use five common technical learning machines for the early detection of cardiac disease. This study mainly aims to evaluate the efficacy of the classification method selected.

Furthermore we use leaders in evaluating these machine learning methods to determine the best performance. Moreover, the effectiveness of the five classifiers is assessed by 10-fold cross-validation. The results of the analysis demonstrate the highest performance of the artificial neural network (ANN). However, by choosing the machine learning techniques they want to apply to researchers and medical professionals they can get a separate insight into this work.

Pouriyeh *et.al.* [5] the purpose of this Article is to examine and compare the precision of different data mining classification schemes using integrated machine learning techniques for predicting heart disease. This article states. The 303-set Cleveland Disease Data Set was used as the principal database for the 10 times cross validation system developed to increase the originally limited data volume for training and testing. Different classifications: the decision tree (DT), the naive bay (NB), the multi-layer perceptron (MLP), the near neighbour K (K-NN), the simple joint learning element (SCRL) (SVM). Apply classifier prediction, packing, reinforcement, and stacking to the data set. The experimental results show that the SVM method using amplification technology is better than the other methods mentioned above Joshi *et.al.* [6] the document usually states that people with diabetes are at an increased risk of the new disease Covid 2019 (COVID-19), which is spread through the Coronavirus 2 Severe Acute Respiratory Syndrome (SARS-CoV-2). The number of COVID-19 cases has been reduced from 20% to half. Diabetes in different parts of the world. This article discusses the recommendations and associated risks for people with diabetes related to blood glucose changes during the COVID-19 outbreak. Similarly, a context-sensitive cross-national survey on the impact of COVID-19 on diabetic patients is being investigated. This presents a new clinical challenge to prevent COVID-19. Shetty *et.al.* [7] states that data mining is a subset of computer programming; it is a systematic method for discovering patterns in a huge information index of programs that contain the intersection of knowledge, artificial intelligence, expertise, and data set structure. Information retrieval systems think about information from a large amount of information and transform it into meaningful design for future use. Our evaluation focuses on this part of the final medical education plan, based on the information collected about diabetes and creating a brilliant restorative solution, sincere and lasting organization. Help the doctor. The main purpose of this assessment is to build an intelligent diabetes disease prediction system that can use data sets of diabetic patients to investigate diabetic diseases. In this case, we recommend using Bayesian and K-Nemost Neighbor (K-Nemost Neighbor) calculations to apply them to the data set of diabetic patients, and to decompose them according to various characteristics of diabetes to predict diabetic diseases. Zhilbert *et.al.* (2015) explains the rising incidence of diabetes, which has recently affected approximately 346 million people, more than one-third of whom went unnoticed in the early stages, which is an urgent need to support medical decision-making. Focus on using one of the algorithms or comparing the algorithm's performance with a specific set of data, which is usually predefined, static and available on the Internet. Bayesian statistical modeling was performed on the data set obtained from the physical examination of 402 patients to improve the reliability of computer diagnosis. This data set contains some attributes that have not been used in computer estimation before. The realization of the two algorithms greatly improves the overall reliability of the computer's key system output. -Assist in the diagnosis process of diabetes.

Rout and Kaur [9] rapid population growth and health maintenance are extremely important issues worldwide. In recent years, many fatal diseases have posed a serious threat. The introduction of machine learning technology in medical care for early prognosis and diagnosis needs to be more precise based on parameters and framework conditions. This article aims to analyze and test the results of several studies on machine learning methods for diabetes and how these findings can help develop future diabetes prediction models. More variables and mixed disciplines need to be considered to obtain accurate results that can overcome existing limitations. Ladha *et.al.* [10] explains a report from various health organizations that shows the anxiety caused by diabetes worldwide. Many researchers around the world have studied its various parameters and are studying it for early detection. The main purpose is to explore and develop methodological views based on the data provided to predict diabetes. This research helps us pave the way for

identifying research gaps to develop an effective framework for future diabetes management. The research and knowledge of attributes and the realization of the classification structure.

Before, the Doctors only view the report to convey the result to the patients. There are some problems appeared while seeing the laboratory details, they can't predict it properly. There is some difficulty in existing project. They created for some other purpose to test for different disease prediction, but According to covid 19, Diabetes and heart disease result is very important to predict, whether the person will be affected by covid-19.

Proposed System

The project's main goal is to find the most accurate factors affecting public health and obtain good results.

- This chart is used to predict the coronavirus by viewing reports of coronary artery disease and diabetes.
- We use Python and Panda questions to characterize coronary artery disease and diabetes in the Cleveland ICU repository.
- Provide a user-friendly visual representation of data sets, working atmosphere, and predictive test settings. The stages of information preparation and subsequent definition depend on the purity of the information, the order in which performance evaluations are presented, and more accurate results.

The data format plays an important role in this application. If the user is entering medical information, it must be in the correct format and within the specified range, otherwise an error dialog will be displayed. Four algorithms are analyzed below:

- Support Vector Machine (SVM).
- Decision Tree (DT).
- *K*-Nearest Neighbour Algorithm (KNN).
- Random Forest Classifier.

The working of these calculations has been clarified in the areas ahead. The calculations have been prepared utilizing the UCI (University of California, Irvine) Cleveland informational index. 75% of the sections in the informational index have been utilized for preparing and the leftover 25% for testing the precision of the calculation. Besides, a few stages have been taken for streamlining the calculations along these lines improving the exactness. These means incorporate cleaning the dataset just as information pre-handling.

The calculations were judged dependent on their precision and it was seen that the K-Nearest Neighbor Algorithm (KNN) was the most exact out of the four with 87.0% proficiency. Thus, it was chosen for execution of the primary application.

The fundamental application is a web application that acknowledges the client's different boundaries as info and registers the outcome.

Performance Requirements

- To be precise, there are no specific guidelines or standards for the performance of Web applications.
- The system must be reliable.
- If the request cannot be processed, a corresponding error message will be displayed.
- The web page loads in a few seconds.

System Design

An E-R model is a specific sort of information model fit to planning social data sets. The fundamental segment of the model is the Entity-Relationship Diagram. The E-R chart is a straightforward method of addressing the information substances being demonstrated and the connections between these information elements. It is not difficult to change

Annals of R.S.C.B., ISSN: 1583-6258, Vol.25, Issue 4, 2021, Pages. 19194-19204
Received 05 March 2021; Accepted 01 April 2021

E-R charts to the Relational Model (information elements compare to relations and connections relate to the inferred affiliations made by keys and unfamiliar keys of relations).

Elements

Elements are closely resembling relations in the social model. They address the central information objects about which data is to be gathered. Substances address ideas or concrete or theoretical items like individual, place, actual things, occasions. In an E-R graph, a substance is addressed as a named rectangular shape, which may incorporate a rundown of characteristics. For lucidity, regularly just characteristics associated with connections between substances are incorporated, i.e., essential key (PK) and unfamiliar keys (FK). This keeps a cleaned up outline.

System Architecture

A system of architecture is a concept model defining the structure, behavior and other systematic representations. Thus, a formal description and representation of a system is the architectural description, and its construction facilitates study of the structure and behavior of the system.

At first, we getting the data of Diabetes and Heart Disease from UCI Dataset, then we are Pre-Processing of Data. After that we are doing the Feature Selection for classification / Prediction. Then the Performance Evaluation Occurs for the result.

Dataset Description

- The purpose of the data set is to estimate whether a patient has diabetes or not based on the specific analysis used on the data set.
- The data set contains some clinical indicator factors and an objective variable, the result. The indicator factors are the number of pregnancies of the patient, her BMI, insulin level, age, etc.

    - Pregnancies: No of times she is pregnant
    - Glucose: Plasma glucose concentration a 2 hours in an oral glucose tolerance test
    - B.P: Diastolic B.P (mm Hg)
    - Skin Thickness: Triceps skin fold thickness (mm)
    - Insulin: 2-Hour serum insulin (mu U/ml)
    - BMI: weight in kg/(height in m)^2
    - Diabetes Pedigree Function
    - Age: Age in years
    - Outcome: Class variable (0 or 1) 268 of 768 are 1, the others are 0.

Preprocess the cardiovascular and diabetes data after collecting multiple sets of data.

The data set contains a total of 769 patient data sets, 6 of which lack some values. In addition, 763 medical histories were used for pretreatment.

- From the 8 information index credits, one age ascribe is used to acknowledge the patient's specific data.
- The staying seven credits are deemed important since they contain healthcare records which are needed.
- Clinical records are critical for diabetic disease analysis and learning.

Annals of R.S.C.B., ISSN: 1583-6258, Vol.25, Issue 4, 2021, Pages. 19194-19204
Received 05 March 2021; Accepted 01 April 2021

Classification Modeling

The grouping of data sets is based on factors and models in a decision tree (DT). At this stage, a classifier is applied to each grouped record to evaluate its representation. The most effective model is recognized based on past low failure rate results.

- Decision Trees Classifier
- Support Vector Classifier
- Random Forest Classifier
- K Nearest neighbors

Fuzzy-KNN Algorithm

The term fuzzy refers to something uncomfortable or fuzzy. We usually experience a state in which we cannot decide whether it is important or wrong. Confusion and weakness in any situation. Truth 1.0 refers to the fundamental value of truth in the logical structure, 0.0 to the false value. There is no justification for real truth and universal false significance in the fuzzy system, though. However, the fuguing thought that is mainly self-evident and nearly no mistake exists is also highly enticing.

Its architecture consists of four parts:

- **RULE BASE:** Contains a sequence of activities by rules and IF THEN expert conditions to monitor the strong semantic information-based structure. The ongoing advancement of fluid theory offers a wide range of strategic possibilities for designing and coordinating fluid controls. The number of fuzzy rules are reduced most of these enhancements.
- **FUZZIFICATION:** Used to change the post, such as B. The new number in the fuzzy set. The new information source is mainly an individual assessment information source. Through the sensor and transmitted to the control structure for processing similar to temperature, pressure, speed, etc.
- **INITIAL MOTOR:** Choose a schedule with current fuzzy commitment level for each rate, and choose the rules that the data should cover. At this stage, the completed criteria are combined into control exercises.
- **DEFUSION:** Used to change the fuzzy set received from the allocator to a new value. There are several open source defuzzification methods, and the most appropriate method is to use it with a special expert system to avoid confusion.

Enlistment Work

Definition: A chart showing how you want to reverse any point in the information space to a point in the range of 0 to 1. The data space is usually thought of as a universal sentence or slang world (u), which contains all the elements expected to serve in any particular application.
There are generally three types of diffusers:

- Monochromatic diffusers
- Gaussian diffusers
- Trapezoidal or triple diffusers.

Fuzzy Control

- This is a technology that displays human thought in a control system.
- The exact reason is unpredictable, but it should be pleasant.
- It can replicate human deductive reasoning, the collaboration that people use to accumulate degrees from knowledge.
- Any weakness can be easily handled through fuzzy parameters.

Performance Measure

Annals of R.S.C.B., ISSN: 1583-6258, Vol.25, Issue 4, 2021, Pages. 19194-19204
Received 05 March 2021; Accepted 01 April 2021

- To determine the execution effectiveness of this model, several standard metrics, such as accuracy, accuracy and order error have been considered.
- Logistic Regression: 71.42857142857143
- K Nearest neighbors: 78.57142857142857
- Support Vector Classifier: 73.37662337662337
- Naive Bayes: 71.42857142857143
- Decision tree: 68.18181818181817
- Random Forest: 75.97402597402598
- Fuzzy KNN: 95.00

Implementation Strategy

Based on the analysis, K-Nearest Neighbor (KNN) was found to be most accurate and reliable. Therefore, KNN was used for the final implementation of the project. Python 3 was used for modelling and classification. The dataset was split into training and testing data in the ratio of 3:1 i.e., 75% of the dataset was used for training purpose & the remaining 25% was used for testing and validation. Front-end is based on HTML5, CSS and JS. Python's micro web-framework Flask is also used for database connection.

Step1: Import required libraries, Import diabetes dataset.

Step2: Pre-measure information to eliminate missing information.

Step3: Perform rate split of 80% to partition dataset as Preparing set and 20% to Test set.

Step4: Select the AI calculation for example KNearest Neighbor, Support Vector Machine, Decision Tree, Strategic relapse, Random Forest and Gradient boosting calculation.

Step5: Build the classifier model for the referenced AI calculation dependent on preparing set.

Step6: Test the Classifier model for the referenced AI calculation dependent on test set.

Step7: Perform Comparison Evaluation of the trial execution results got for every classifier.

Step8: After dissecting dependent on different measures finish up the best performing calculation.

Conclusion

First, four algorithms were implemented, and all the algorithm data sets were individually trained, and then all the algorithms were tested. According to several criteria, the most effective algorithm was selected, and the ANN algorithm was found to be the most effective. The four algorithms have an accuracy rate of 87.0%. The accuracy rates of decision tree, support vector machine, and random forest classifier are 79.0%, 83.0%, and 84.0%. Therefore, the ANN algorithm is also implemented in a Web application using a better user interface. It uses HTML5, CSS, JS and Flask (a Python micro-web framework) to help Finals users make preliminary predictions on promising technologies such as machine technology, because heart disease and diabetes are the first time India and the world have learned to predict that Covid-19 will affect society Have a profound impact. Inform users if they are at risk and need to see a doctor. This will help reduce Covid's death rate. Therefore, through the above method, the individual heart disease and diabetes were successfully analyzed. The result of predicting the risk of Covid-19 based on the parameters specified by the user was obtained.

References

[1] Sharma, H., & Rizvi, M. A. (2017). Prediction of heart disease using machine learning algorithms: A survey. *International Journal on Recent and Innovation Trends in Computing and Communication,* 5(8), 99-104.

[2] Shanmugasundaram, G., Selvam, V. M., Saravanan, R., & Balaji, S. (2018). An Investigation of Heart Disease Prediction Techniques. *In 2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA)* (pp. 1-6). IEEE.

[3] Dhar, S., Roy, K., Dey, T., Datta, P., & Biswas, A. (2018). A hybrid machine learning approach for prediction of heart diseases. *In 2018 4th International Conference on Computing Communication and Automation (ICCCA)* (pp. 1-6). IEEE.

[4] Ahmed, M. R., Mahmud, S. H., Hossin, M. A., Jahan, H., & Noori, S. R. H. (2018). A cloud based four-tier architecture for early detection of heart disease with machine learning algorithms. *In 2018 IEEE 4th International Conference on Computer and Communications (ICCC)* (pp. 1951-1955). IEEE.

[5] Pouriyeh, S., Vahid, S., Sannino, G., De Pietro, G., Arabnia, H., & Gutierrez, J. (2017). A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease. *In 2017 IEEE symposium on computers and communications (ISCC)* (pp. 204-207). IEEE.

[6] Joshi, A. M., Shukla, U. P., & Mohanty, S. P. (2020). Smart healthcare for diabetes during COVID-19. *IEEE Consumer Electronics Magazine, 10*(1), 66-71.

[7] Shetty, D., Rit, K., Shaikh, S., & Patil, N. (2017). Diabetes disease prediction using data mining. *In 2017 international conference on innovations in information, embedded and communication systems (ICIIECS)* (pp. 1-5). IEEE.

[8] Tafa, Z., Pervetica, N., & Karahoda, B. (2015). An intelligent system for diabetes prediction. *In 2015 4th Mediterranean Conference on Embedded Computing (MECO)* (pp. 378-382). IEEE.

[9] Rout, M., & Kaur, A. (2020, June). Prediction of Diabetes Risk based on Machine Learning Techniques. *In 2020 International Conference on Intelligent Engineering and Management (ICIEM)* (pp. 246-251). IEEE.

[10] Ladha, G. G., & Pippal, R. K. S. (2018). A computation analysis to predict diabetes based on data mining: A review. *In 3rd International Conference on Communication and Electronics Systems (ICCES)* (pp. 6-10). IEEE.

[11] Rajasekaran Rajkumar, Nallani Chackravatula Sriman Narayana Iyengar: Dynamic Integration of Mobile JXTA with Cloud Computing for Emergency Rural Public Health Care. 10/2013; 4(5):255-264., DOI:10.1016/j.phrp.2013.09.004

[12] Rajasekaran Rajkumar, Vasudev Sharma: Visualization of data mining techniques to predict breast cancer with high accuracy rates. *Journal of Computer Science 01/2019; 15*(1)., DOI:10.3844/jcssp.2019.118.130

[13] D.S. Hooda, Keerti Upadhyay and D.K. Sharma, "On Parametric Generalization of 'Useful' R- norm Information Measure" *British Journal of Mathematics & Computer Science,* Vol. 8(1), pp. 1-15, 2015.

[14] Pandya, S.; Ambient Acoustic Event Assistive Framework for Identification, Detection, and Recognition of Unknown Acoustic Events of a Residence, Advanced Engineering Informatics. *Elsevier*. (http://www.sciencedirect.com/science/article/pii/S147403462030207X)

[15] D.S. Hooda, Keerti Upadhyay and D.K. Sharma, "A Generalized Measure of 'Useful R-norm Information", *International Journal of Engineering Mathematics and Computer Sciences,* Vol 3(5), pp.1-11, 2014.

[16] Ghayvat, H.; Pandya, S.; Awais, M. ReCognizing SUspect and PredictiNg ThE SpRead of Contagion Based on Mobile Phone LoCation DaTa (COUNTERACT): A System of identifying COVID-19 infectious and hazardous sites, detecting disease outbreaks based on internet of things, edge computing and artificial intelligence, Sustainable Cities and Society

[17] D.S. Hooda, Keerti Upadhyay and D.K. Sharma, "Bounds on Cost Measures in terms of 'Useful' R-norm Information Measures" *Direct Research Journal of Engineering and Information Technology,* Vol.2 (2), pp.11-17, 2014.

[18] Pandya S, Wakchaure MA, Shankar R, Annam JR. Analysis of NOMA-OFDM 5G wireless system using deep neural network. *The Journal of Defense Modeling and Simulation.* 2021. doi:10.1177/1548512921999108.

[19] D.S. Hooda and D.K. Sharma, "Lower and Upper Bounds Inequality of a Generalized 'Useful' Mean Code Length" *GAMS Journal of Mathematics and Mathematical Biosciences,* Vol. 4(1), pp.62-69, 2013.

[20] Awais, M.; Ghayvat, H.; Krishnan Pandarathodiyil, A.; Nabillah Ghani, W.M.; Ramanathan, A.; Pandya, S.; Walter, N.; Saad, M.N.; Zain, R.B.; Faye, I. Healthcare Professional in the Loop (HPIL): Classification of Standard and Oral Cancer-Causing Anomalous Regions of Oral Cavity Using Textural Analysis Technique in Autofluorescence Imaging. *Sensors,* 2020, 20, 5780. https://doi.org/10.3390/s20205780

[21] D.S. Hooda, Keerti Upadhyay and D.K. Sharma, 'Useful' R-Norm Information Measure and its Properties" *IOSR Journal of Electronics and Communication Engineering,* Vol. 8, pp. 52-57, 2013.

[22] Patel, C.I.; Labana, D.; Pandya, S.; Modi, K.; Ghayvat, H.; Awais, M. Histogram of Oriented Gradient-Based Fusion of Features for Human Action Recognition in Action Video Sequences. *Sensors* 2020, 20, 7299. https://doi.org/10.3390/s20247299

[23] D.S. Hooda, Sonali Saxena and D.K. Sharma, "A Generalized R-Norm Entropy and Coding Theorem" *International Journal of Mathematical Sciences and Engineering Applications,* Vol.5(2), pp.385-393, 2011.

[24] Ghayvat, H.; Awais, M.; Pandya, S.; Ren, H.; Akbarzadeh, S.; Chandra Mukhopadhyay, S.; Chen, C.; Gope, P.; Chouhan, A.; Chen, W. Smart Aging System: Uncovering the Hidden Wellness Parameter for Well-Being Monitoring and Anomaly Detection. *Sensors* 2019, 19, 766. https://doi.org/10.3390/s19040766.

[25] D.S. Hooda and D.K. Sharma, "Bounds on Two Generalized Cost Measures" *Journal of Combinatorics, Information & System Sciences,* Vol. 35(3-4), pp. 513-530, 2010.

[26] Barot, V., Kapadia, V., & Pandya, S., QoS Enabled IoT Based Low Cost Air Quality Monitoring System with Power Consumption Optimization, Cybernetics and Information Technologies, 2020, 20(2), 122-140. https://doi.org/10.2478/cait-2020-0021.

[27] D.K. Sharma and D.S. Hooda, "Generalized Measures of 'Useful' Relative Information and Inequalities" *Journal of Engineering, Management & Pharmaceutical Sciences,* Vol.1(1), pp.15-21, 2010.

[28] Sur, A., Sah, R., Pandya, S., Milk storage system for remote areas using solar thermal energy and adsorption cooling, Materials Today, Volume 28, Part 3, 2020, Elsevier, Pages 1764-1770, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2020.05.170.

[29] D.S. Hooda and D.K. Sharma (2010) "Exponential Survival Entropies and Their Properties" *Advances in Mathematical Sciences and Applications,* Vol. 20, pp. 265-279, 2010.

[30] H. Ghayvat, Pandya, S., and A. Patel, "Deep Learning Model for Acoustics Signal Based Preventive Healthcare Monitoring and Activity of Daily Living," *2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India,* 2020, pp. 1-7, doi: 10.1109/IDEA49133.2020.9170666

[31] D.S. Hooda and D.K. Sharma, "Generalized 'Useful' Information Generating Functions" *Journal of Appl. Math. and Informatics,* Vol. 27 (3-4), pp. 591-601, 2009.

[32] Pandya, S., Shah, J., Joshi, N., Ghayvat, H., Mukhopadhyay, S.C. and Yap, M.H., 2016, November. A novel hybrid based recommendation system based on clustering and association mining. *In Sensing Technology (ICST), 2016 10th International Conference on* (pp. 1-6). IEEE.

[33] D.S. Hooda and D.K. Sharma, "Non-additive Generalized Measures of 'Useful' Inaccuracy" *Journal of Rajasthan Academy of Physical Sciences,* Vol. 7(3), pp.359-368, 2008.

[34] U. Naseem, S. K. Khan, M. Farasat, and F. Ali, "Abusive Language Detection: A Comprehensive Review," *Indian Journal of Science Technology,* vol. 12, no. 45, pp. 1-13, 2019.

[35] D.S. Hooda and D.K. Sharma, Generalized R-Norm information Measures. *Journal of Appl. Math, Statistics & informatics (JAMSI),* Vol. 4 No.2, 153-168, 2008.

[36] U. Naseem, I. Razzak, S. K. Khan, and M. Prasad, "A Comprehensive Survey on Word Representation Models: From Classical to State-Of-The-Art Word Representation Language Models," arXiv preprint arXiv:.15036, 2020.

[37] Dilip Kumar Sharma, "Some Generalized Information Measures: Their characterization and Applications", *Lambert Academic Publishing,* Germany, 2010. ISBN: 978-3838386041.

[38] U. Naseem, M. Khushi, S. K. Khan, K. Shaukat, and M. A. Moni, "A Comparative Analysis of Active Learning for Biomedical Text Mining," *Applied System Innovation,* vol. 4, no. 1, p. 23, 2021

[39] S. K. Khan, M. Farasat, U. Naseem, and F. Ali, "Performance evaluation of next-generation wireless (5G) UAV relay," *Wireless Personal Communications,* vol. 113, no. 2, pp. 945-960, 2020.

[40] S. K. Khan et al., "UAV-aided 5G Network in Suburban, Urban, Dense Urban, and High-rise Urban Environments," *2020 IEEE 19th International Symposium on Network Computing and Applications (NCA),* 2020, pp. 1-4: IEEE.

[41] Kumar, S., Kumar, P., Wisetsri, W., Raza, M. & Norabuena-Figueroa, R.P. (2021). Social entrepreneurship education: Insights from the indian higher educational courses. *Academy of Strategic Management Journal,* 20(S1),1-14.

[42] Listiningrum, H. D., Wisetsri, W., & Boussanlegue, T. (2020). Principal's Entrepreneurship Competence in Improving Teacher's Entrepreneurial Skill in High Schools. *Journal of Social Work and Science Education,* 1(1), 87-95.

[43] W. Wisetsri, "The Perception of Brand Personality in the Context of Hotel of Undergraduate Students", vol. 3, no. 1, pp. 1-12, Jun. 2020.

[44] Vijai C.& Wisetsri, W. (2021). Rise of Artificial Intelligence in Healthcare Startups in India. *Advances In Management.* 14 (1) March (2021):48-52.

[45] Wisetsri, W. (2020). The Perception of Brand Personality in the Context of Hotel of Undergraduate Students. *Journal of Multidisciplinary in Humanities and Social Sciences,* 3(1): 1-12.

[46] Ishaq, A., Sadiq, S., Umer, M., Ullah, S., Mirjalili, S., Rupapara, V., & Nappi, M. (2021). Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques. *IEEE Access,* 9, 39707–39716. https://doi.org/10.1109/access.2021.3064084

[47] Rustam, F., Khalid, M., Aslam, W., Rupapara, V., Mehmood, A., & Choi, G. S. (2021). A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis. *PLOS ONE,* 16(2), e0245909. https://doi.org/10.1371/journal.pone.0245909

[48] Yousaf, A., Umer, M., Sadiq, S., Ullah, S., Mirjalili, S., Rupapara, V., & Nappi, M. (2021b). Emotion Recognition by Textual Tweets Classification Using Voting Classifier (LR-SGD). IEEE Access, 9, 6286–6295. https://doi.org/10.1109/access.2020.3047831

[49] Sadiq, S., Umer, M., Ullah, S., Mirjalili, S., Rupapara, V., & NAPPI, M. (2021). Discrepancy detection between actual user reviews and numeric ratings of Google App store using deep learning. *Expert Systems with Applications,* 115111. https://doi.org/10.1016/j.eswa.2021.115111

[50] Rao, A. N., Vijayapriya, P., Kowsalya, M., & Rajest, S. S. (2020). Computer Tools for Energy Systems. *In International Conference on Communication, Computing and Electronics Systems* (pp. 475-484). Springer, Singapore.

[51] Manne, R., & Kantheti, S. C. (2021). Application of Artificial Intelligence in Healthcare: Chances and Challenges. *Current Journal of Applied Science and Technology,* 40(6), 78-89.

https://doi.org/10.9734/cjast/2021/v40i631320

[52] Gupta J., Singla M.K., Nijhawan P., Ganguli S., Rajest S.S. (2020) *An IoT-Based Controller Realization for PV System Monitoring and Control.* In: Haldorai A., Ramu A., Khan S. (eds) Business Intelligence for Enterprise Internet of Things. EAI/Springer Innovations in Communication and Computing. Springer, Cham.

[53] Sharma M., Singla M.K., Nijhawan P., Ganguli S., Rajest S.S. (2020) *An Application of IoT to Develop Concept of Smart Remote Monitoring System.* In: Haldorai A., Ramu A., Khan S. (eds) Business Intelligence for Enterprise Internet of Things. EAI/Springer Innovations in Communication and Computing. Springer, Cham.

[54] U. Zulfiqar, S. Mohy-Ul-Din, A. Abu-Rumman, A. E. M. Al-Shraah, And I. Ahmed, "Insurance-Growth Nexus: Aggregation and Disaggregation," *The Journal of Asian Finance, Economics and Business*, vol. 7, no. 12, pp. 665–675, Dec. 2020. https://doi.org/10.13106/jafeb.2020.vol7.no12.665

[55] Al-Shqairat, Z. I., Al Shraah, A. E. M., Abu-Rumman, A., "The role of critical success factors of knowledge stations in the development of local communities in Jordan: A managerial perspective," *Journal of management Information and Decision Sciences,* vol. 23, no.5, pp. 510-526, Dec. 2020. 1532-5806-23-5-218

[56] Ganguli S., Kaur G., Sarkar P., Rajest S.S. (2020) *An Algorithmic Approach to System Identification in the Delta Domain Using FAdFPA Algorithm.* In: Haldorai A., Ramu A., Khan S. (eds) Business Intelligence for Enterprise Internet of Things. EAI/Springer Innovations in Communication and Computing. Springer, Cham.

[57] Singla M.K., Gupta J., Nijhawan P., Ganguli S., Rajest S.S. (2020) *Development of an Efficient, Cheap, and Flexible IoT-Based Wind Turbine Emulator.* In: Haldorai A., Ramu A., Khan S. (eds) Business Intelligence for Enterprise Internet of Things. EAI/Springer Innovations in Communication and Computing. Springer, Cham.

[58] Rajasekaran R., Rasool F., Srivastava S., Masih J., Rajest S.S. (2020) *Heat Maps for Human Group Activity in Academic Blocks.* In: Haldorai A., Ramu A., Khan S. (eds) Business Intelligence for Enterprise Internet of Things. EAI/Springer Innovations in Communication and Computing. Springer.