



e-ISSN: 2278-8875  
p-ISSN: 2320-3765

# International Journal of Advanced Research

in Electrical, Electronics and Instrumentation Engineering

Volume 11, Issue 5, May 2022

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 8.18**

☎ 9940 572 462

☑ 6381 907 438

✉ [ijareeie@gmail.com](mailto:ijareeie@gmail.com)

@ [www.ijareeie.com](http://www.ijareeie.com)



# Substantial Influence of Machine Learning Techniques in Healthcare Industries

Manju Payal<sup>1</sup>, Dr. T. Ananth Kumar<sup>2</sup>, P. Praveen kumar<sup>3</sup>

Software Developer, Academic Hub Ajmer, India<sup>1</sup>

Department of Computer Science and Engineering, IFET College of Engineering, Tamilnadu, India<sup>2</sup>

Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, Pondicherry, India<sup>3</sup>

**ABSTRACT:** In general, the term "health care" refers to a system that involves the improvement of medical services in order to meet the needs of individuals in terms of their physical and mental well-being. The process of backing up and restoring patient health records includes participation from a wide variety of parties, including patients, clinicians, suppliers, healthcare corporations, and information technology companies. The authors of this paper conducted a literature review on research publications pertaining to machine learning algorithms utilised in healthcare applications. Most academics who study the ability to predict outcomes in healthcare use algorithms such as Decision Tree, Support Vector Machine, and others to conduct their investigations. These techniques have the highest degree of dependability.

**KEYWORDS:** Healthcare; Machine learning; SVM classifier; Industry 4.0

## I. INTRODUCTION

Machine learning allows you to create models that quickly assess data and provide outcomes[1], using both historical and real-time data. It is possible for medical professionals to improve their diagnostic and treatment decisions for individual patients by using machine learning [2]. This would lead to an overall improvement in the quality of healthcare services.

Machine learning is now being used in a number of sectors of the healthcare industry, including the development of better medical processes, the administration of patient records and data, and the treatment of chronic diseases[3]. Machine learning is also being used in the development of better medical processes, and the administration of patient records and data. It is possible that machine learning will assist businesses in the healthcare industry in meeting increased demand, reducing costs, and improving operations [4].

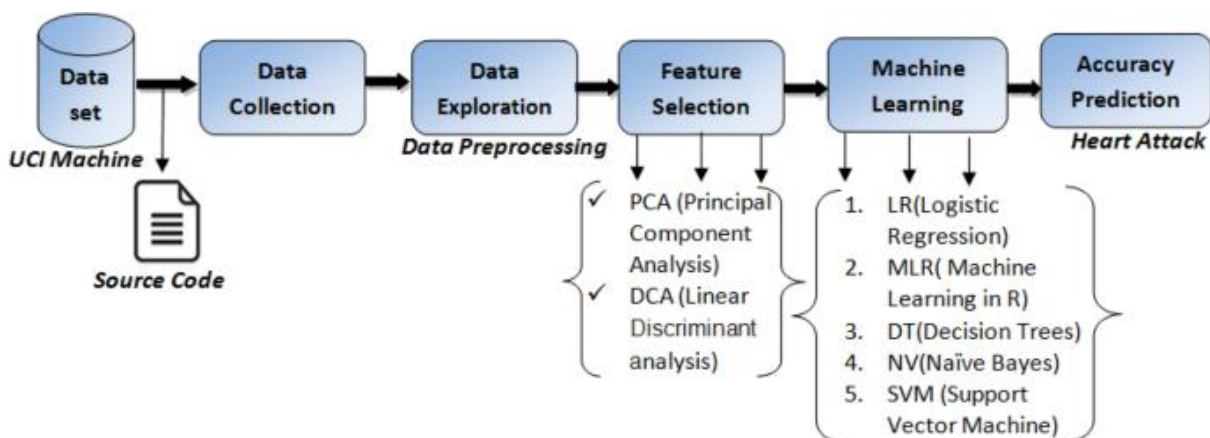


Fig. 1. Proposed architecture

### 1.1. Feature Extraction (FE)

Because irrelevant features have the potential to have a negative impact on the effectiveness of the classification performed by the machine learning classifier[5-8], it is essential to extract the features that are optimal. In



order to extract important characteristics from the dataset, the linear discriminant analysis (LDA) [9] and the principal component analysis (PCA) [10-12] are used.

## II. RELATED WORKS

### 2. 1. Principal component Analysis (PCA)

The principal component analysis (PCA) method is one of the unsupervised machine learning approaches that is utilised in a wide variety of different applications. In addition to that, it is one of the strategies that is utilised the vast majority of the time. Examples of applications that could benefit from this include exploratory data analysis, dimensionality reduction, information compression, data de-noising, and a great many others [13]. The principal component analysis method is currently one of the most widely used approaches to machine learning (PCA). The field of healthcare data makes use of it in order to cut down on the number of dimensions that must be considered. In the field of healthcare, principal component analysis (PCA) is utilised in a variety of sub-industries, one of which is patient insurance data, where there are a large number of associated variables and numerous data sources. Sources such as pharmacies, hospitals, and other types of facilities [14] are examples of sources to consider.

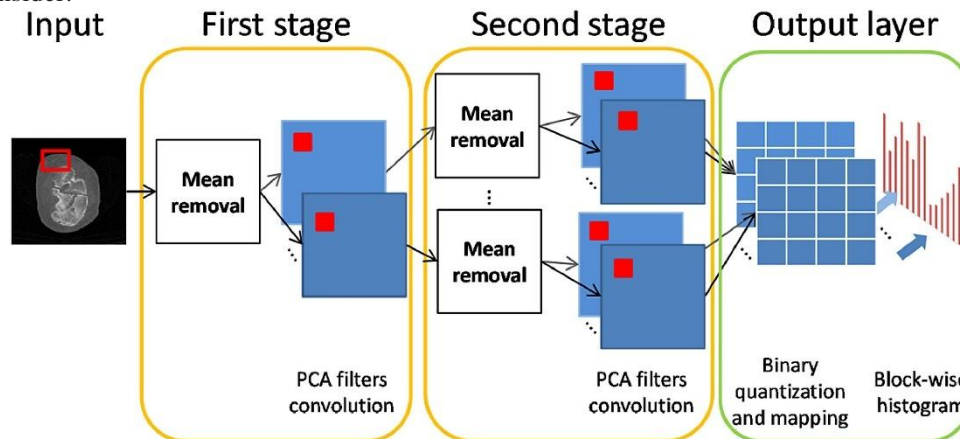


Fig. 2. Conventional layers of proposed architecture

The principal component analysis (PCA) is applied to research various methods of pattern recognition, and it is also utilised for the purpose of reducing the dimensionality of techniques through the utilisation of measurement vectors for distance classification. In addition to extracting features, it is also useful for removing redundancy and for extracting features[15].

- Determine the covariance matrix
- From the covariance matrix, calculate the Eigen vector and Eigen values
- Enter all of the patient data into the matrix
- Estimate the mean and remove it from the dataset
- Determine the covariance matrix
- Produce a new dataset by employing a feature vector in its creation.

### 2.2. Functional Model of Proposed method

1. Image fusion, also known as IF, is a technique that combines information from multiple images that is complementary to one another as well as information that is quite redundant in order to produce a fused image that is more complete and accurate [16]. Image fusion is also referred to by its acronym, "Image fusion."

2. Registration of medical images: When it comes to image-based medical diagnosis, image registration is one of the most crucial methods that can be used.

[10] Using this tool, the information contained in two or more medical images of the same patient can be combined into a single view. Before all of the photos have been correctly superimposed on one another, the physician will not be able to make a reliable diagnosis or carry out precise therapeutic control [17].

3. In the field of medical imaging, the process of segmenting an image into its constituent parts is referred to as image segmentation, and it is utilised in the process. Image segments are typically utilised in the field of medical imaging to differentiate between various tissue types, diseases, organs, or other physiologically relevant entities [18].





The process of removing extraneous and unnecessary image data in order to store or transmit data in a more efficient manner is referred to as picture compression. In the context of medical image compression, this process is carried out. Image compression can be used to reduce the size of an image without having any impact on the image's quality or appearance in any way. Image compression can be used. PCA is utilised in this method of image compression, which results in very little loss of quality [19].

5. Eliminating as much background noise as possible from medical images For this reason, noise reduction is a critically important topic in medical image processing, as erroneous or noisy information invariably has an impact on a physician's ability to make an accurate medical diagnosis. The messy data can be cleaned up using any one of a variety of strategies, and there are a lot of them to choose from.

Approaches for computing abstractions of picture information and making local decisions about whether or not a particular type of image feature exists at each individual image location are collectively referred to as "feature detection," and the term "feature detection" is used as a noun. [20] The features that are generated will most frequently take the form of individual dots, continuous curves, or linked areas, and they will each represent a subset of the picture domain. [21]

7) The process of classifying medical images involves ranking the training photos according to how closely they resemble one another in terms of image classification..

### 2.3. Linear discriminant analysis (LDA)

The LDA algorithm is a variation of Fisher's linear discriminant; it employs a linear combination of features to classify data into vector format based on a target factor or class variable. This is accomplished through the use of a target factor or class variable. This method is comparable to Analysis of Variance (ANOVA) and regression in that it employs linear combinations to describe the effects that predictors have on the outcome of the study[22]. There are two different ways to approach LDA. The approaches presuppose that the data follow a Gaussian distribution, which means that when the data are plotted, the curve for each attribute will look like a bell, each variable will have the same variance, and the data points for each attribute will fluctuate by the same amount around the average [23].

To put it another way, in order for the method to work, the data and characteristics in question need to have a regular distribution and either a constant variance or a standard variation. An extension of linear discriminant analysis, Fisher's linear discriminant is a technique utilised in the fields of statistics and machine learning to differentiate between two or more classes. The DA method is illustrated by the algorithm. [24]

- Complete the matrix by entering all of the observations (patient data).
- Determine the Prior Probabilities ( $P_i$ ) of a situation.
- Determine the values for the parameters of the conditional probability density functions.
- Fourth, calculate discriminant functions using the formulae provided in the previous step.
- Estimate misclassification probability using cross validation.
- Assign the observations to a category that is not known (patients).

### 2.4. Logistic Regression (LR)

A Logistic Regression classifier is a tool that can be utilised to determine whether a tumour is cancerous or benign[25]. A wide variety of imaging techniques used in medicine are utilised in order to extract distinct aspects of tumours. For instance, the size of the tumour, the area of the body that is affected, and so on and so forth. After that, these characteristics are inputted into a Logistic Regression classifier, which then determines whether or not the tumour is cancerous[26].

#### A. Decision Trees:

One of the most extensively used machine learning algorithms nowadays is the decision tree. It's a problem-classification method based on supervised learning. It works effectively with categorical and continuous dependent variables.[27] We divide the popular into two or more homogeneous sets depending on the most important features in our machine learning approach.

Decision trees are the machine learning algorithm that is utilised the most frequently in the field of healthcare for solving classification and regression problems. These are exceptionally complex algorithms that are able to accommodate a wide variety of data. [28] In the field of medicine, decision trees can be put to use in situations where the available treatment options are unclear, such as when a patient is ill or about to pass away. Although at first glance the idea might seem strange, it actually enables medical professionals to select the treatment that has the greatest potential [29].

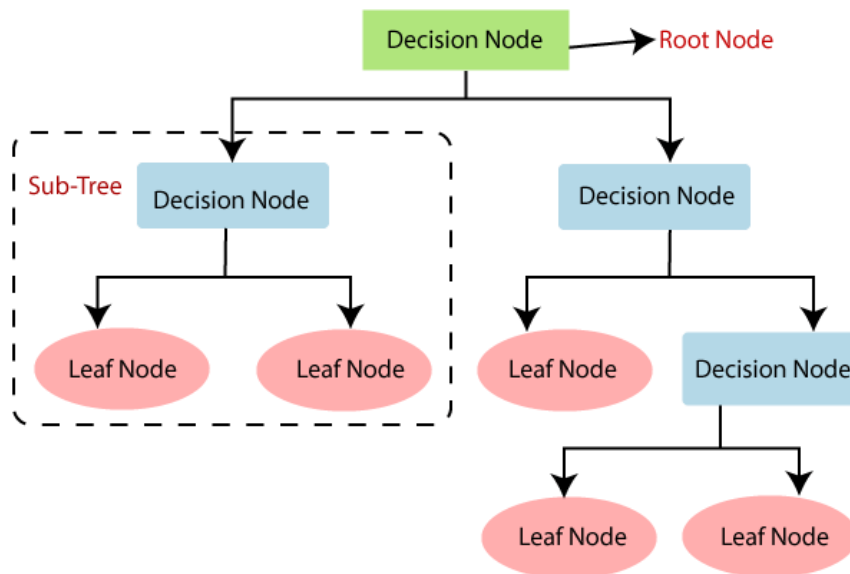


Fig.3. Classification of Decision node

Techniques for classification and forecasting based on decision trees are highly effective and are used extensively. Because they can be represented graphically and broken down into a series of straightforward, logical decision principles that can be comprehended with relative ease, they are widely utilised in the fields of data science and healthcare. "Training the data" refers to the process of determining the best possible set of decision criteria through the use of statistical analysis. A tree is traversed and each node, which represents a category, is examined, with the goal of breaking (splitting) each category into two or more distinct groups [30].

When applied to the field of medicine, the phrase "target variable" can mean a variety of distinct things depending on the circumstances surrounding its use. For instance, it may mean determining whether or not a patient develops Parkinson's disease, or it may mean determining which combination of symptoms would result in the highest payout category for a health insurer [31]. Additionally, it may mean determining whether or not a patient develops Alzheimer's disease. There are additional goals that should be pursued, such as identifying individuals who are at a high risk for particular medical illnesses before major problems arise and providing preventive healthcare interventions before negative consequences occur [32]. These are just two examples of goals that should be pursued.

Companies in the healthcare and life science industries compete in a highly competitive global market. As a result of decreased profit margins, increased regulatory requirements, and heightened patient awareness, medical companies are coming under increasing pressure to implement innovative technologies that make patient care simpler and less expensive. It is possible that Decision Tree will be of assistance to healthcare and life science organisations by improving the way in which physicians make decisions regarding their patients. This will be accomplished by displaying all relevant data on an electronic dashboard. In addition, this will help to reduce costs. We also assist with the hand-off of patients by providing corporations with access to all records, both within and across the various types of healthcare institutions that can be found all over the country. Using behaviour analytics, we can also assist you in improving patient outcomes by altering patients' lifestyles and the decisions they make regarding their health[33].

### B. Nave Bayes Classifier (NBC)

In machine learning, the Nave Bayes Classifier (NBC) is a common algorithm that has been used in a variety of domains, including text categorization, medical diagnosis, and system performance management, to name a few examples. The nave Bayesian classifier, also known as the nave bayes classifier, is one of the most successful and efficient classification methods available (NB). A straightforward probabilistic classifier, it relies on Bayes' theorem and strong (nave) independence assumptions[34] to achieve its classification accuracy.

The purpose of NBNs is to improve illness therapy as well as early disease diagnosis in order to facilitate treatment that is both quicker and more effective. As a consequence of this, the goals of any given NBN are to: Produce a disease forecast that is both quicker and more accurate; Support medical professionals in making sound decisions.

The Naive Bayes (NB) algorithm, which is considered to be one of the most effective and efficient classification algorithms, has been successfully applied to a wide variety of medical situations. The information that is provided by the classifier can be utilised by medical professionals in order to diagnose patients. Naive Bayes is a statistical method



that can be utilised by medical professionals in order to ascertain whether or not a patient is at an increased risk for a variety of illnesses, including cancer, cardiovascular disease, and other conditions [35][40-42].

Support Vector Machine (SVM): The Support Vector Machine (SVM) algorithm is a common supervised learning technique that may help with difficulties with classification and regression. Support Vector Machines, also known as SVMs, are a well-known approach to machine learning that can be utilised for categorization in addition to other tasks. A discriminative classifier, SVM's formal definition includes an ideal hyperplane as one of its components. The optimal hyperplane produces support vectors, which are used to categorise new examples and datasets in accordance with the optimal hyperplane [36]. This hyperplane is a line in the two-dimensional (2D) space that divides into two segments, with each segment lying on each side[37]. Diseases of the cardiovascular system, more commonly referred to as heart conditions, are the leading cause of death in every region of the world with the exception of Africa. According to the World Health Organization, these conditions are a collection of heart and blood vessel disorders that include coronary heart disease, cerebrovascular disease, rheumatic heart disease, and other conditions[38][39].

### III. CONCLUSION

The field of healthcare is indeed not excluded from the widespread application of machine learning, which has resulted in significant advancements not only in the healthcare industry but also in practically every other industry as well. These advancements have been brought about as a result of the widespread application of machine learning. A previously unimaginable amount of data exists in the field of healthcare that is neither labelled nor consistent; this data can be processed through machine learning in order to obtain valuable insights. The field of medicine is home to one of the most important applications for machine learning. According to the researchers, the technology has the potential to be used for a wide variety of applications, including virtual treatment and surgery.

### REFERENCES

1. Jiang, Z., Dong, Z., Wang, L., & Jiang, W. (2021). Method for Diagnosis of Acute Lymphoblastic Leukemia Based on ViT-CNN Ensemble Model. *Computational Intelligence and Neuroscience*, 2021.
2. Rasheed, Rizwan, Asfra Rizwan, HajraJaved, Faiza Sharif, and Asghar Zaidi. "Socio-economic and environmental impacts of COVID-19 pandemic in Pakistan—an integrated analysis." *Environmental Science and Pollution Research* 28, no. 16 (2021): 19926-19943.
3. Meraj, T., Alosaimi, W., Alouffi, B., Rauf, H. T., Kumar, S. A., Damaševičius, R., & Alyami, H. (2021). A quantization assisted U-Net study with ICA and deep features fusion for breast cancer identification using ultrasonic data. *PeerJ Computer Science*, 7, e805.
4. Kumar, S. A., García-Magariño, I., Nasralla, M. M., & Nazir, S. (2021). Agent-Based Simulators for Empowering Patients in Self-Care Programs Using Mobile Agents with Machine Learning. *Mobile Information Systems*, 2021.
5. Kumar, S. A., Nasralla, M. M., García-Magariño, I., & Kumar, H. (2021). A machine-learning scraping tool for data fusion in the analysis of sentiments about pandemics for supporting business decisions with human-centric AI explanations. *PeerJ Computer Science*, 7, e713.
6. Mostafa, A. M., Kumar, S. A., Meraj, T., Rauf, H. T., Alnuaim, A. A., & Alkhayyal, M. A. (2022). Guava Disease Detection Using Deep Convolutional Neural Networks: A Case Study of Guava Plants. *Applied Sciences*, 12(1), 239.
7. Kumar, S. A., Kumar, H., Dutt, V., & Soni, H. (2021, February). Self-Health Analysis with Two Step Histogram based Procedure using Machine Learning. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 794-799). IEEE.
8. Kumar, S. A., Kumar, A., Dutt, V., & Agrawal, R. (2021, February). Multi Model Implementation on General Medicine Prediction with Quantum Neural Networks..
9. Kumar, S. A., Kumar, H., Swarna, S. R., & Dutt, V. (2020). Early Diagnosis and Prediction of Recurrent Cancer Occurrence in a Patient Using Machine Learning. *European Journal of Molecular & Clinical Medicine*, 7(7), 6785-6794.
10. Jebaleela, R. Sharon, G. Rajakumar, T. Ananth Kumar, and S. Arunmozhiselvi. "An optimized CNN based automated COVID-19 lung infection identification technique from CT images." In *Novel AI and Data Science Advancements for Sustainability in the Era of COVID-19*, pp. 253-276. Academic Press, 2022.
11. Kumar, S. A., Kumar, H., Dutt, V., & Dixit, P. (2020). The Role of Machine Learning in COVID-19 in Medical Domain: A Survey. *Journal on Recent Innovation in Cloud Computing, Virtualization & Web Applications* [ISSN: 2581-544X (online)], 4(1).
12. Kumar, S. A., Kumar, H., Dutt, V., & Swarnkar, H. (2020). COVID-19 Pandemic analysis using SVM Classifier:



Machine Learning in Health Domain. Global Journal on Application of Data Science and Internet of Things [ISSN: 2581-4370 (online)], 4(1).

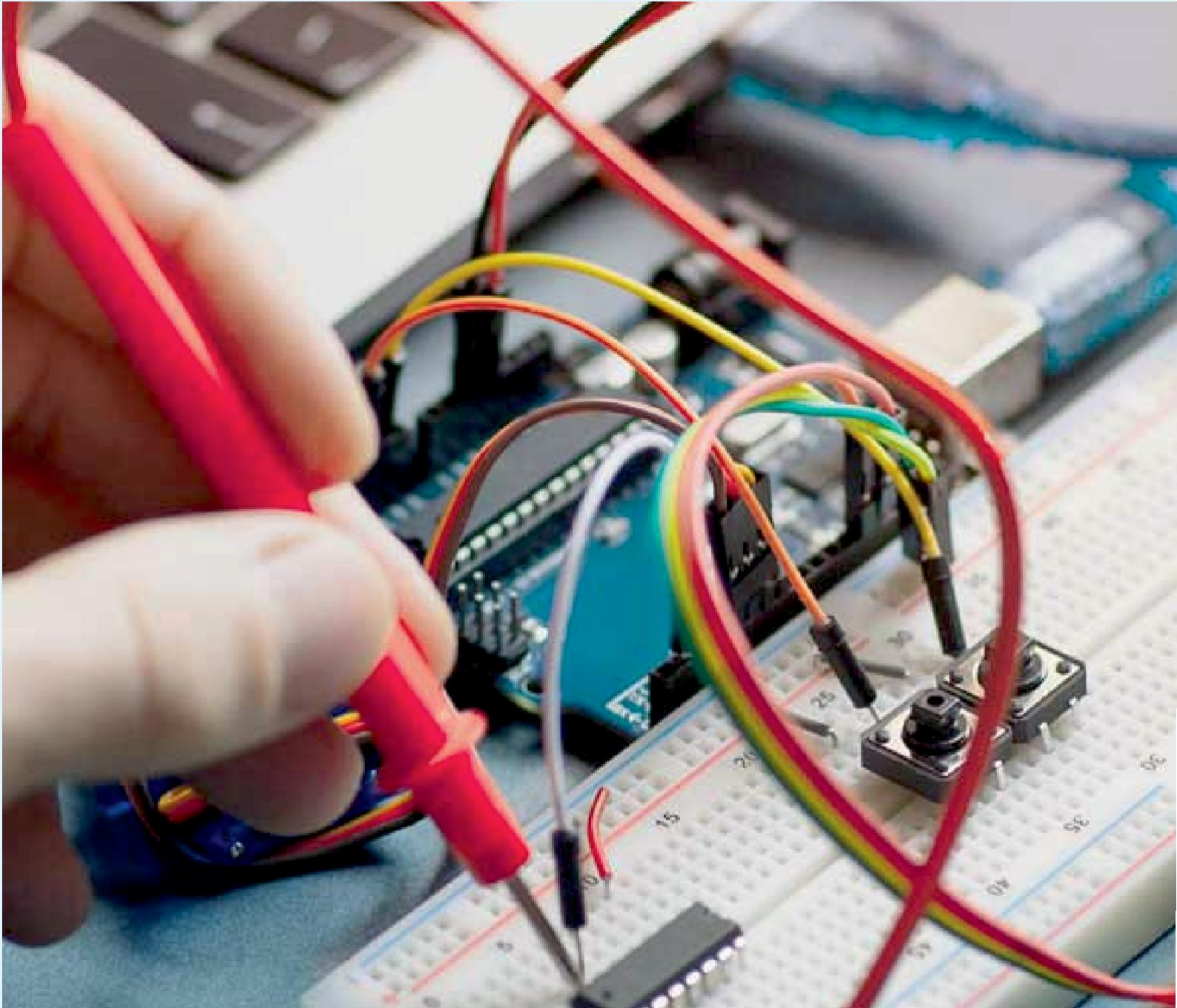
13. Ananth, Christo, Stalin Jacob, Jenifer Darling Rosita, M. S. Muthuraman, and T. Ananth Kumar. "Low Cost Visual Support System for Challenged People." In 2022 International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN), pp. 1-4. IEEE, 2022.
14. Kumar, S. A., Kumar, H., Dutt, V., & Dixit, P. (2020). Deep Analysis of COVID-19 Pandemic using Machine Learning Techniques. Global Journal on Innovation, Opportunities and Challenges in Applied Artificial Intelligence and Machine Learning [ISSN: 2581-5156 (online)], 4(2).
15. Kumar, S. A., Kumar, H., Dutt, V., & Swarnkar, H. (2020). Role of Machine Learning in Pattern Evaluation of COVID-19 Pandemic: A Study for Attribute Explorations and Correlations Discovery among Variables. Global Journal on Application of Data Science and Internet of Things [ISSN: 2581-4370 (online)], 4(2).
16. KUMAR, S. A., KUMAR, H., DUTT, V., & SWARNKAR, H. (2019). CONTRIBUTION OF MACHINE LEARNING TECHNIQUES TO DETECT DISEASE IN-PATIENTS: A COMPREHENSIVE ANALYSIS OF CLASSIFICATION TECHNIQUES. Global Journal on Innovation, Opportunities and Challenges in Applied Artificial Intelligence and Machine Learning [ISSN: 2581-5156 (online)], 3(1).
17. Kumar, S. A., H. Kumar, V. Dutt, and H. Swarnkar. "Contribution Of Machine Learning Techniques To Detect Disease In-Patients: A Comprehensive Analysis Of Classification Techniques." Global Journal on Innovation, Opportunities and Challenges in Applied Artificial Intelligence and Machine Learning [ISSN: 2581-5156 (online)] 3, no. 1 (2019).
18. Kumar, Abhishek, SwarnAvinash Kumar, Vishal Dutt, Ashutosh Kumar Dubey, and Vicente García-Díaz. "IoT-based ECG monitoring for arrhythmia classification using Coyote Grey Wolf optimization-based deep learning CNN classifier." Biomedical Signal Processing and Control 76 (2022): 103638.
19. A.Kumar, S.Kumar, V.Dutt, S.Narang, A.Dubey "A Hybrid Secured Cloud Platform Maintenance based on Improved Attributes. Based Encryption Strategies" published in regular issue in IJIMAI, Indexed by the Science Citation Index Expanded(Web Of Science), Universidad Internacional de La Rioja (UNIR). ISSN 1989-1660.
20. Battineni, Gopi, Nalini Chintalapudi, and Francesco Amenta. "Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM)." Informatics in Medicine Unlocked 16 (2019): 100200.
21. SwarnAvinash Kumar, Harsh Kumar, Vishal Dutt, Himanshu Swarnkar, "Contribution Of Machine Learning Techniques To Detect Disease In Patients : A Comprehensive Analysis Of Classification Techniques" Vol 3 No 1 (2019): Global Journal on Innovation, Opportunities and Challenges in AAI and Machine Learning. ISSN 2581-5156.
22. SwarnAvinash Kumar, Kapil Chauhan, Aastha Parihar, "Functionality of Classification and Regression tree in Bioinformatics" Vol 5 No 2 (2021): Global Journal on Innovation, Opportunities and Challenges in Applied Artificial Intelligence and Machine Learning. ISSN 2581-5156..
23. KalaiPriya, R., S. Devadharshini, R. Rajmohan, M. Pavithra, and T. Ananthkumar. "Certain investigations on leveraging blockchain technology for developing electronic health records." In 2020 International conference on system, computation, automation and networking (ICSCAN), pp. 1-5. IEEE, 2020.
24. Usharani, S., P. Manju Bala, R. Rajmohan, T. Ananth Kumar, and M. Pavithra. "Blockchain Technology Use Cases in Healthcare Management: State-of-the-Art Framework and Performance Evaluation." In Blockchain, Artificial Intelligence, and the Internet of Things, pp. 117-140. Springer, Cham, 2022.
25. Tanoli, Ziaurrehman, Markus Vähä-Koskela, and Tero Aittokallio. "Artificial intelligence, machine learning, and drug repurposing in cancer." Expert opinion on drug discovery 16, no. 9 (2021): 977-989.
26. Kumar, S.A. (2021), "Corona Recognition Method Based On Visible Light Color Using Artificial Intelligence". AusPat Application No. AU 2021103067(A4).
27. Wright, Caroline F., David R. FitzPatrick, and Helen V. Firth. "Paediatric genomics: diagnosing rare disease in children." Nature Reviews Genetics 19, no. 5 (2018): 253-268.
28. Kumar, S.A. (2021), "An Artificial Intelligence And IoT Based Method For Prevention Of Security Attack On Cloud Medical Data". AusPat Application No. AU 2021102115(A4).
29. Kumar, S.A. (2021), "IOT Based Generic Framework For Computer Security Using Artificial Immune System". AusPat Application No. AU 2021102104(A4).
30. Lenders, Jacques WM, Michiel N. Kerstens, A. M. A. R. Laurence, Aleksander Prejbisz, Mercedes Robledo, David Taieb, Karel Pacak et al. "Genetics, diagnosis, management and future directions of research of pheochromocytoma and paraganglioma: a position statement and consensus of the Working Group on Endocrine Hypertension of the European Society of Hypertension." Journal of hypertension 38, no. 8 (2020): 1443.
31. Kumar, S.A. (2021), "IOT Enabled Wall Climbing Robot For Security". AusPat Application No. AU



2021101471(A4).

32. Du-Harpur, X., F. M. Watt, N. M. Luscombe, and M. D. Lynch. "What is AI? Applications of artificial intelligence to dermatology." *British Journal of Dermatology* 183, no. 3 (2020): 423-430.
33. Young, Albert T., MulinXiong, Jacob Pfau, Michael J. Keiser, and Maria L. Wei. "Artificial intelligence in dermatology: a primer." *Journal of Investigative Dermatology* 140, no. 8 (2020): 1504-1512.
34. Kumar, S.A. (2021), "Corona Recognition Method Based On Visible Light Color Using Artificial Intelligence". AusPat Application No. AU 2021103067(A4).
35. Kumar, S.A. (2021), "An Artificial Intelligence And IoT Based Method For Prevention Of Security Attack On Cloud Medical Data". AusPat Application No. AU 2021102115(A4).
36. Kumar, S.A. (2021), "IOT Based Generic Framework For Computer Security Using Artificial Immune System". AusPat Application No. AU 2021102104(A4).
37. Kumar, S.A. (2021), "IOT Enabled Wall Climbing Robot For Security". AusPat Application No. AU 2021101471(A4).
38. Danasegaran, Sathish Kumar, Elizabeth Caroline Britto, and Susan Christina Xavier. "Exploration of trigonal patch antenna characteristics with the impact of 2D photonic crystal of various air hole shapes." *Journal of Electronic Materials* 50, no. 9 (2021): 5365-5374.
39. Rajmohan, R., T. Ananth Kumar, M. Pavithra, S. G. Sandhya, E. G. Julie, J. J. V. Nayahi, and N. Z. Jhanjhi. "Blockchain: Next-generation technology for industry 4.0." *Blockchain Technology* (2020): 177-198.
40. Velammal, M. Navaneetha, Tamilarasan Ananth Kumar, M. Steffi Anto, and A. Andrew Roobert. "Design of High-Speed Nanoscale Adder Logic Circuit for Low Power Consumption." In *2021 IEEE Pune Section International Conference (PuneCon)*, pp. 1-6. IEEE, 2021.
41. Arumugam, Devi, Kavya Govindaraju, and Ananth Kumar Tamilarasan. "AIIoT-Based Smart Framework for Screening Specific Learning Disabilities." In *Machine Learning for Critical Internet of Medical Things*, pp. 103-124. Springer, Cham, 2022.
42. Padmapriya, N., K. Tamilarasi, P. Kanimozhi, T. Ananth Kumar, R. Rajmohan, and Ajagbe Sunday Adeola. "A Secure Trading System using High level Virtual Machine (HLVM) Algorithm." In *2022 International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN)*, pp. 1-4. IEEE, 2022.





INNO  SPACE  
SJIF Scientific Journal Impact Factor

Impact Factor: 8.18

 **doi**<sup>®</sup>  
**cross** **ref**

 **INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA**



# International Journal of Advanced Research

in Electrical, Electronics and Instrumentation Engineering

 9940 572 462  6381 907 438  [ijareeie@gmail.com](mailto:ijareeie@gmail.com)



[www.ijareeie.com](http://www.ijareeie.com)

Scan to save the contact details