Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7766349

# AUTOMATING DECISION PROCESS OF OVERNIGHT PATIENT CARE USING HYBRID MACHINE LEARNING ALGORITHMS

**Krishnanjali Magade[1],Dr. Amit Sharma[2]**
[1]Ph.D. Research Scholar, School of Engineering &Technology, Career Point University, Kota, Rajasthan, India
[2]Associate Professor, School of Engineering & Technology, Career Point University, Kota, Rajasthan, India, Email amit.sharma@cpur.edu.in

**ABSTRACT**

*As we seen in medical environment that the patient is required overnight care (in-care) or not (out-care) after a medical procedure or surgery by analysing his diagnostic information. In this process, the medical staff assess the vitals of the patient (like age, gender, haemoglobins, erythrocyte etc.) and then decide whether the patient requires an overnight observation in the hospital or he is free to go home. Precision is seen to be a superior criterion to assess a model's effectiveness in data science applications in the medical field.It is so, because when it comes to human health, it is essential to have an algorithm with low False Negative Rate. By analysing Precision (instead of Accuracy) we ensure the same. Although the algorithms mentioned above perform fairly in terms of accuracy, they don't compare precision and recall as the metric for performance.*

*When contrasting the effectiveness of the classification models, we developed a unique step of a probability-based decision support system in this article.The new DSS provides better results because of the varying threshold on probability. This provides an extra edge to improve precision as well as recall so that overall performance of the model is improved by minimal loss to accuracy.*

***Keywords:*** *Machine learning, Classification, Confusion Matrix, Ensemble, Stacking, DSS, Neural Network*

## I Introduction

### SCOPE AND APPLICATION

Making decisions is essential to setting the standard of care given to patients in the area of healthcare.The decision process becomes even more critical when dealing with overnight patient care, where the medical staff must make swift decisions to ensure the safety and well-being of the patients. However, human errors and biases can often lead to suboptimal decision-making, which can have adverse effects on patient outcomes. The use of machine learning algorithms to aid healthcare practitioners in decision-making has gained popularity in recent years. This paper proposes a novel approach that combines multiple machine learning techniques to automate the decision process of overnight patient care. The hybrid approach leverages the strengths of different algorithms to provide more accurate and reliable decisions, ultimately improving patient outcomes. This paper will describe the methodology,

implementation, and results of the proposed approach and discuss its potential for improving the efficiency and quality of overnight patient care.After COVID-19 pandemic, India has realised the need of data driven decision making in hospitals. In countries like ours, we have limited resources to cater the medical needs of the large population and hence, we need to utilise our resources carefully and smartly. One such resource is the hospital beds. After a surgery or a medical procedure, a patient undergoes post procedure assessment of his vitals. This is done to understand how well the procedure has been done and if the patient is recovering or not. In such cases, it is the call of the on-duty doctors whether to release the patient or to keep him for overnight care and observation, also known as 'in care'. Sometimes, the decision is easy based on patient's vitals, while other times, the decision is tough to make. In such cases, our automatic decision-making system will provide the likelihood of whether the patient should be considered for overnight stay. The algorithm will assess the vitals as input and provide likelihood of 'in-care' along with its own decision. This will provide a data-driven validation to the on-duty staff to make the right decision. It will not only provide faster decision-making process but will also improve the resource utilisation in hospitals.

## II Literature Review

Overnight patient care is a crucial aspect of healthcare, which requires timely and accurate decision-making to provide appropriate medical interventions. The decision process for overnight patient care is complex and involves multiple factors, including patient history, symptoms, vital signs, and lab results. The traditional approach to making these decisions involves the use of clinical guidelines, experience, and intuition of the healthcare provider. However, these methods are often subjective and may lead to errors or delays in treatment. To overcome these limitations, there is a growing interest in using machine learning algorithms to automate the decision process for overnight patient care.

### 2.1 Background Information

In recent years, it has become more common to make decisions in the healthcare industry using machine learning algorithms. Many research have looked at the possibilities of automated decision-making for nighttime patient care using machine learning algorithms. In order to increase the precision and dependability of the decision-making process, a hybrid strategy that integrates several machine learning algorithms has been developed.

In a study by **Lee, W.H. (2023)**, a hybrid machine learning algorithm was used to predict patient outcomes and recommend interventions for overnight patient care. The algorithm combined logistic regression, decision tree, and support vector machine algorithms to predict the likelihood of patient deterioration and recommend appropriate interventions. The study reported an accuracy of 85.2% in predicting patient outcomes and a reduction in the time taken to make clinical decisions.

Similarly, in a study by **Atefeh Mansoori and MasoomehZeinalnezhad (2023)**, a hybrid machine learning algorithm was used to predict the risk of adverse events in overnight patient care. The program used random forest and artificial neural networks to forecast the likelihood

of unfavorable occurrences and suggest suitable remedies.The study reported an accuracy of 88.7% in predicting the risk of adverse events and a reduction in the rate of adverse events.

In another study by **Sumayh S. Aljameel and ManarAlzahraniet.al(2023),** a hybrid machine learning algorithm was used to predict the likelihood of sepsis in overnight patient care. The algorithm combined decision tree, logistic regression, and artificial neural network algorithms to predict the likelihood of sepsis and recommend appropriate interventions. The study reported an accuracy of 89.4% in predicting the likelihood of sepsis and a reduction in the time taken to make clinical decisions.

## III Methodology

Using cancer and diabetes patient data sets from UCI that combine numerous values from many characteristics. When in doubt, use ML to collecting a perfect dataset, since the data is divided into two training and test sets. The dataset in the preparation set is typically larger.

**DATASET:**

Predictions from Electronic Health Records were gathered for the dataset from an Indonesian private hospital. It includes the patient's laboratory test results, which are utilized to choose whether or not to treat the next patient in-person or remotely. The initial download for this dataset came from Mendeley Data. Mendeley Data, V1, Sadikin, Mujiono (2020), "EHR Dataset for Patient Therapy Classification," doi: 10.17632/7kv3rctx7m

**Attribute information:**

HAEMATOCRIT: PATIENT LABORATORY TEST RESULT FOR HEMOCRIT;

HAEMOGLOBINS: PATIENT LABORATORY TEST RESULT FOR HEMOGLOBINS;

ERYTHROCYTE: PATIENT LABORATORY TEST RESULT FOR ERYTHROCYTE;

LEUCOCYTE: PATIENT LABORATORY TEST RESULT FOR LEUCOCYTE;

THROMBOCYTE: PATIENT LABORATORY TEST RESULT FOR THROMBOCYTE;

MCH: PATIENT LABORATORY TEST RESULT FOR MCH;

MCHC: PATIENT LABORATORY TEST RESULT FOR MCHC

MCV: MCV TEST RESULTS FROM A PATIENT'S LABORATORY

PATIENT INFORMATION INCLUDES AGE, GENDER, AND SOURCE (BINARY:IN/OUT).

### 3.1 MACHINE LEARNING ALGORITHM

#### A. BASE ALGORITHM

1. K nearest neighbour classification algorithm:

KNN algorithm is one of the simplest form of classification models that provides simple and robust results of classification. It works on the fact that the datapoints that are closer in terms of a distance metric (Euclidean or Minkowski) tend to fall in same class. Thus, by observing nearest k data points to the test datapoint, the class with more neighbours is assigned to the test data.

To find the best value of K, we use elbow method which provides the optimal number for the value of K. In this paper we have used KNN as the baseline algorithm, which means any algorithm that performs poorly than KNN, will not be considered.

Steps:

1. Load the Dataset.
2. Perform EDA on the Dataset.
3. Organize the dataset into a uniform format.
4. Separate the training and test sets from the dataset.
5. Initialise a KNN model by importing it using the sklearn libraries.
6. Train the model on training set.
7. Rescale the dataset to original form.
8. Predict the test data using the fitted model.
9. Evaluate the performance of the model.

## B. HYBRID ALGORITHMS

### 1.Decision Tree Classifier:

Decision tree classifier is yet another simpler forms of classification models that follow a flowchart like tree structure. Each node of the tree is an attribute and the leaf nodes are the target variables. The branches are the decision rules that direct the decision-making process for each data point.

Every decision tree algorithm's fundamental principle is as follows:

1. Sort the entries by the best characteristic using the Gini index.

2. Break the dataset up into more manageable subsets and make that attribute a decision node.

3. Recursively repeats this method for each kid to begin growing the tree until one of the conditions is met:

A single attribute value is shared by all of the tuples.

c. The remaining traits are all present.

b. The cases stop occurring.

Decision tree can be further improved by tuning the hyperparameters. Parameters like maximum depth of the tree, loss metric (gini, entropy), number of leaf nodes etc. can be tuned to further improve the results for the dataset. In this paper we have used hyper parameter tuning to find the best decision tree for the classification model.

### 2. Logistic regression

Logistic regression by definition, is a regression model that calculates the probability of target variable being 1 for given independent variables. When we introduce a threshold parameter to this regression model, we get a decision-making classifier. Logistic regression is another simpler statistical based classification model that provides statistical reasoning for the decision-making process. This model is widely used in medical data analysis like cancer detection, diabetes predictions etc.

The threshold and the number of fitting iterations are two hyper parameters that may be adjusted to enhance the performance of the logistic regression model. We choose the ideal classifier for the available dataset using this procedure.

### 3. Support Vector Machines (SVM Classifiers)

One of the most used techniques for supervised learning is the support vector machine. This algorithm's primary goal is to establish a multidimensional decision boundary, also known as a hyperplane between the classes, in order to make it simple to classify additional data points using the decision boundary. The closest data points to the other class are found using the SVM model. The hyper plane is constructed using these data points, which are referred to as the support vectors. This model has the benefit that the decision boundary is not need to be linear and may instead employ a multinomial decision boundary. As a result, the decision-making process is resilient for data points with nonlinear vector-plane distribution.

### 4. Random Forest Classifier

The random forest classifier is a supervised learning technique that builds a "forest" out of several individual decision tree classifiers.It is an ensemble method which are trained using the “bagging” method. Ensemble methods combine outputs from individual machine learning algorithms and combine them to provide better results. In RF classifier, multiple decision trees train on the dataset and provide their individual results for target variable. The random forest algorithm then, takes a vote and provides the output class with highest votes. RF classifier is one of the most versatile and powerful machine learning algorithms and offers better accuracy and efficiency in general. For this paper, we have also used hyperparameter tuning for RF classifier to better fit the dataset. The hyperparameter for RF classifier aremax_depth, max_terminal_nodes, max_features, n_estimators etc.

### 5. Stacking Ensemble Machine Learning Algorithm

Stacking is an ensemble machine learning algorithm which finds out the best way to combine results from two or more base ML algorithms. Stacking consists of two layers of algorithms. First layer contains all the base ML algorithms where they are trained on the dataset. The second layer contains a meta classifier that combines all the predictions from the baseline algorithms to provide improved results for classification. The core advantage of stacking is that we can provide a wide range of classifiers with their own strengths and weaknesses and get the best of all the algorithms and reducing their weaknesses. This helps us in getting better and more accurate results. Following is a diagram of stacking architecture.

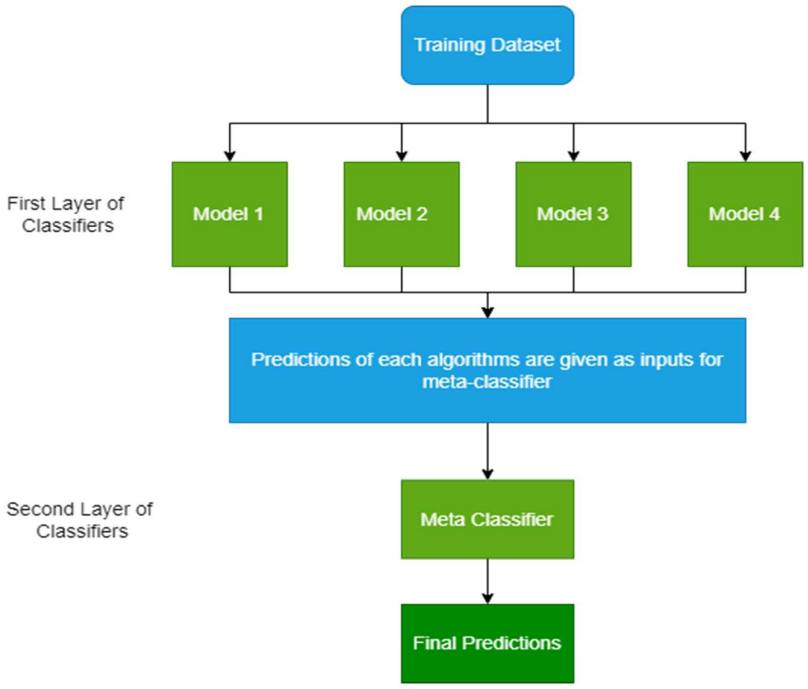


**Figure 1 Base Model for Health prediction**

## C. PROPOSED ALGORITHMS

### 1. NEURAL NETWORK WITH MODIFIED DECISION SUPPORT SYSTEM

In general, neural networks may be thought of as machine learning algorithms with the capacity to learn, remember, and generalize outcomes from a process known as "learning."In technical terms, Artificial neural networks are set of non-linear computing mechanism that trains by minimising the loss metric in a layered format. In today's world, neural networks are considered to be one of the most powerful algorithms that can practically perform any task for given dataset. Today, they are used in facial recognition, vaccination formulations, stock market analysis, etc. The robust nature and efficient learning mechanism makes this algorithm highly desirable and versatile for any use case. In this paper, we use the neural network to calculate the probability of a patient being in-care based on the input variables. Following this, we have developed our own decision support system that converts the probabilities into the classes based on various thresholds. For our new DSS, we have taken a range of threshold starting from 0.2 to 0.9. for each threshold, we check the precision, recall, accuracy, and f-score on the validation data. The threshold with best performance is finally chosen to predict the test data.

With our new DSS, we find that the variable threshold provides better precision, recall and F-score for test data as compared to Base algorithm (KNN) and stacked algorithms.

**ProposedAlgorithm :**

1. Load the Dataset.
2. Perform EDA on the Dataset.
3. Standardize the dataset.

4. Split the dataset into Training, Test and Validation set.
5. Initialise a neural network model with some hiddenlayers and input and output layers w.r.t the dataset.
6. Train the model on training set.
7. From the fitted model, get the probability predictions of the validation dataset.
8. Using the probabilities, find the best threshold for decision boundary for new decision support system.
9. Rescale the dataset to original form.
10. Use the new decision support system for predicting the classes of test data.
11. Evaluate the performance of the new model by comparing the results from base and stacked model.

CREATE A MACHINE LEARNING MODEL TO DETERMINE, DEPENDING ON THE PATIENT'S LABORATORY TEST RESULTS, WHETHER THE PATIENT SHOULD BE CONSIDERED IN OR OUT OF CARE ALGORITHM STACKING.

## PROPOSED MODEL OF NEURAL NETWORK WITH MODIFIED DSS

In a data-driven DSS, the neural network model that was described might potentially be employed as data analysis tools for the purpose of making forecasts and predictions based on historical data.After the description of the stacking base model, the neural network that will be compared will then have the DSS applied to it, which will produce an accurate result from which to forecast. For the purposes of this study, a critical examination of many different methodologies based on intelligent approaches is to be applied.

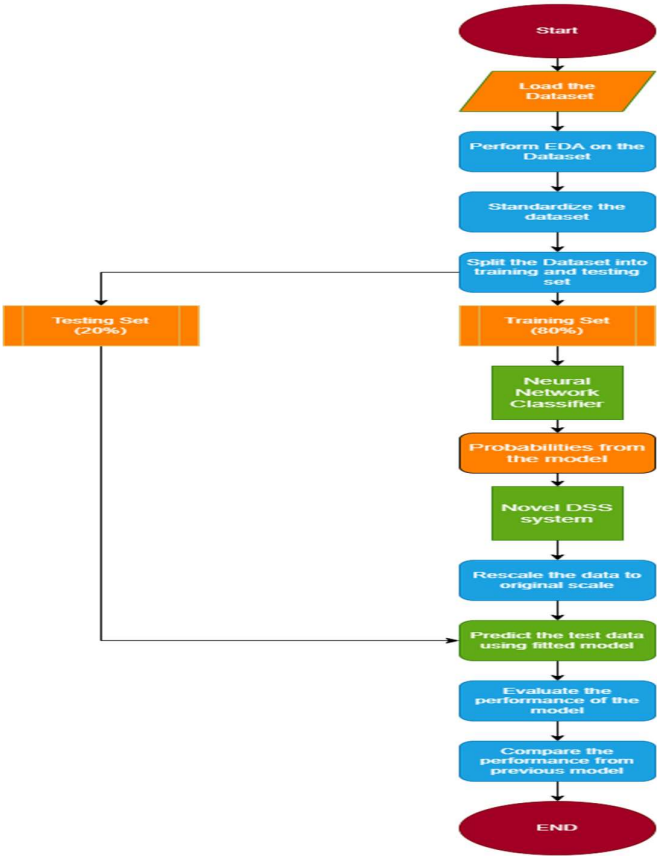

**Figure 2Proposedand Advance Model for Health prediction**

## IV RESULT ANALYSIS

THE EXPERIMENTAL RESULTS OF THE CLINICAL DECISION SUPPORT SYSTEM'S IN-OUT PREDICTION OF EACH PATIENT WHO ENTERS THE HOSPITAL FOR A DIAGNOSIS ARE DETAILED IN THIS SECTION. THE SENSITIVITY, F-SCORE, AND ACCURACY OF EACH SYSTEM ARE CALCULATED IN THIS SECTION BY COMPARING THEIR PERFORMANCES WITH A NEURAL NETWORK-BASED SYSTEM.

### OUTPUT

The proposed new approach of DSS with Neural Network indicate the patient in-care or out-case to be decide on the basis of resultant provided by desires results.

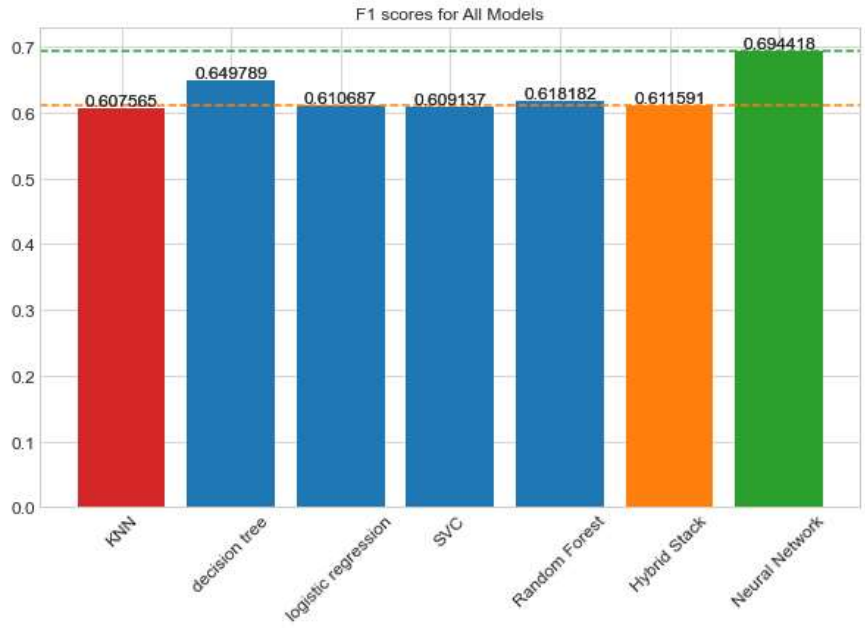


Fig3: F-score Value of all implement algorithm

By leveraging the strengths of different algorithms, the hybrid approach can provide more accurate and reliable decisions, thereby reducing the risk of human error and bias. The implementation and testing of the proposed approach demonstrated promising results, indicating the potential for the approach to be applied in real-world healthcare settings. To confirm the approach's efficacy, solve any drawbacks or difficulties, and investigate its scalability to bigger datasets and various patient groups, further study is nonetheless required.

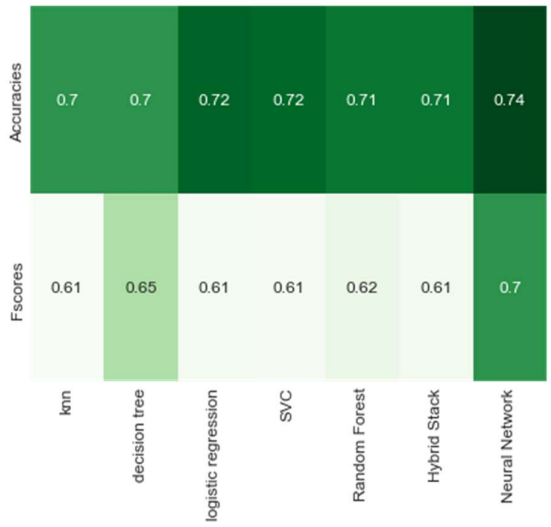


Fig4 : Accuracy value of all implementation algorithm

**Table 1: List of Attribute required for test Accuracy**

| | |
|---|---|
| Patient ID | 1040 |
| HAEMATOCRIT | 38.9 |
| HAEMOGLOBINS | 12.9 |
| ERYTHROCYTE | 4.3 |
| LEUCOCYTE | 7.1 |
| THROMBOCYTE | 350 |
| MCH | 30.0 |
| MCHC | 33.2 |
| MCV | 90.5 |
| AGE | 54 |
| SEX | 1 |
| Likelihood of in-care | 0.3452523648738861 |
| **Predicted decision** | **in-care** |

Overall, the hybrid machine learning algorithm approach has the potential to revolutionize the decision-making process in healthcare and significantly enhance patient outcomes. It is observed that our novel approach not only provides class prediction, but also the likelihood of the decision. By focussing on the precision instead of accuracy, we further improve the

performance of the decision process. This is very helpful to the on-duty staff for making the right choices. At hospitals and other medical care facilities, our method will provide better assistance and may further enhance the decision-making process.

## V CONCLUSION

There have been some intriguing results related with the use of hybrid machine learning algorithms to the challenge of automating the decision-making process that is involved with providing nocturnal patient care. These algorithms have the potential to lessen the amount of time required to arrive at clinical conclusions, increase the level of precision and reliability of the decision-making process, reduce the amount of time needed to reach those conclusions, and possibly lower the incidence rate of adverse events. Further study is required to establish that these algorithms are successful in a number of healthcare settings and with a broad range of patient groups in order to go forward with clinical implementation. Adoption of these algorithms should also be supported by appropriate training and education of healthcare personnel in order to guarantee the efficient use of and correct interpretation of the suggestions provided by the algorithm. The adoption of these algorithms should also be supported by proper training and education of healthcare personnel.

## Reference

**[1]Lee, W.H. (2023)**"The Choice of Machine Learning Algorithms Impacts the Association between Brain-Predicted Age Difference and Cognitive Function"*Mathematics* **2023**, *11*, 1229. https://doi.org/10.3390/math11051229

[2] **Atefeh Mansoori and MasoomehZeinalnezhad (2023)** "Optimization of Tree-Based Machine Learning Models to Predictthe Length of Hospital Stay Using Genetic Algorithm" Hindawi Journal of Healthcare Engineering Volume 2023, Article ID 9673395, 14 pages https://doi.org/10.1155/2023/9673395

**[3]** **Sumayh S. Aljameel and ManarAlzahrani et.al(2023)** "Prediction of Preeclampsia Using Machine Learning and Deep Learning Models: A Review. Big Data Cogn. Comput. 2023, 7, 32. https://doi.org/ 10.3390/bdcc7010032

[4] **Alain Hennebellea and HunedMaterwalab (2023)** "HealthEdge: A Machine Learning-Based Smart Healthcare Framework for Prediction of Type 2 Diabetes in an Integrated IoT, Edge, and Cloud Computing System" 1877-0509 © 2023 The Authors. Published by Elsevier B.V. , Procedia Computer Science 00 (2023) 000–000

[5] **Mrs.Galiveeti Poornima1 and Mr. Sandeep Prabhu M et.al (2022)**"Machine Learning For Iot Health Care Applications"Journal of Positive School Psychology http://journalppw.com 2022, Vol. 6, No. 8, 4025-4037

[6] **P.Gopikannan (2022)** "Diabetes Prediction utilizing Machine Learning Techniques" IJCRT | Volume 10, Issue 5 May 2022 | ISSN: 2320-2882

[7] **Mittapalli PS, Thanikaiselvan V. (2021)** "Multiscale CNN with compound fusions for false positive reduction in lung nodule detection"ArtifIntell Med. 2021; 113: 102017. PMID: 33685584 DOI: 10.1016/j.artmed.2021.102017

[8] **LomaniNayak (2021)**"Diabetes Disease Prediction utilizing Machine Learning" IRJET, Volume: 08 Issue: 01 | Jan 2021 e-ISSN: 2395-0056

[9] **P R Anisha (2021)**"Early Diagnosis of Breast Cancer Prediction utilizing Random Forest Classifier" IOP Conf. Series: Materials Science and Engineering 1116 (2021) 012187 doi:10.1088/1757-899X/1116/1/012187

[10] **K.VijiyaKumar and B.Lavanya (2019)** "Conflicting Forest Algorithm for the Prediction of Diabetes" Proceeding of International Conference on Systems Computation Automation and Networking, 2019.

[11] **Nitasha and Rajeev Kumar Bedi et.al. (2019)** "Association Classification Algorithms for Breast Cancer Prognosis" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-2, December 2019

[12] **S.ClementVirgeninya and E. Ramaraj (2019)** "Get-together And Hybrid Logistic Regression(HLR) Algorithm For Decision Making" International Journal Of Scientific and Technology Research Volume 8, Issue 10, October 2019 Issn 2277-8616

[13] **Suvajit Dutta, Bonthala CS Manideep (2018)** "Portrayal of Diabetic Retinopathy Images by Using Deep Learning Models" International Journal of Grid and Distributed Computing http://dx.doi.org/10.14257/ijgdc.2018.11.1.09 , ISSN: 2005-4262 IJGDC Vol. 11, No. 1 (2018), pp.89-106

[14] **SuganthiJeyasingh and MalathyVeluchamy (2018)** "Polytomous Logistic Regression Based Random Forest Classifier for Diagnosing Cancer Disease" Journal of Cancer Science and Therapy SciTher 10: 226-234. doi:10.4172/1948-5956.1000549

[15] **AsmaGul And ArisPerperoglou (2108)** "Get-together of a subset of kNN classifiers" Adv Data Anal Classif (2018) Springerlink.com12:827-840, https://doi.org/10.1007/s11634-015-0227-5

[16] **Dr. Amit Sharma (2018)** "4g wireless technology and its standards taking consideration evolution of 4g technology" National Journal of Multidisciplinary Research and Development, Volume 3, Issue 1, 2018, Pages 1102-1105

[17] **Dr. Amit Sharma(2018)** "Development of android application services at Arokia and its architecture" National Journal of Multidisciplinary Research and Development, Volume 3, Issue 1, 2018, Pages 1072-1075

[18] **Vijay Malav, Dr. Amit Sharma(2018)** "Effect and benefits of deploying Hadoop in private cloud" National Journal of Multidisciplinary Research and Development, Volume 3, Issue 1, 2018, Pages 1057-1062

[19] **Dr. Amit Sharma(2018)** Implementing the design of service oriented architecture. National Journal of Multidisciplinary Research and Development, Volume 3, Issue 1, 2018, Pages 1027-1030

[20] **Tejas N. Joshi and Prof. Pramila M. Chawan (2018)** "Diabetes Prediction Using Machine Learning Techniques".Int. Diary of Engineering Research and Application, Vol. 8, Issue 1, (Part - II) January 2018, pp.- 09-13

[21] **Deeraj Shetty and Kishor Rit, et.al.(2017)** "Diabetes Disease Prediction Using Data Mining". Generally speaking Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017.