International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Hypokinetic Rigid Syndrome Prognosis using Random Forest Classifiers and Support Vector Machines

Vani Hiremani[1*], Raghavendra M Devadas[2], Harshal Patil[3], Smita Patil[4], Soni Sweta[5], Vijeeta Patil[6]

Submitted: 07/12/2023 **Revised**: 17/01/2024 **Accepted**: 27/01/2024

Abstract: Many individuals worldwide experience Hypokinetic Rigid Syndrome (HRS), a condition more prevalent among those aged 50 and above. Despite numerous technological advancements and breakthroughs, early disease diagnosis is still a formidable challenge. This underscores the need for the development of automatic machine learning techniques to aid healthcare professionals in precisely identifying this condition during its initial stages. The primary aim of this research paper is to perform a comprehensive analysis and comparison of contemporary machine learning methods like Support Vector Machine (SVM) and Random Forest (RFM) used for detecting HRS. To assess and determine the most effective and accurate classifier for HRS categorization, this study concentrates on evaluating SVM and RFM on UCI Machine Learning Repository's Parkinson's Data Set. The results indicate that the support vector machine achieved an 84.3% accuracy and a Kappa score of 0.824, while the random forest exhibited an 87.2% accuracy with a Kappa score of 0.82.

Keywords: *Hypokinetic rigid syndrome, Random Forest Classifiers, Support Vector Machines, Prognosis.*

1. Introduction

Hypokinetic Rigid Syndrome (HRS), also known as Parkinson's Disease (PD) is a condition that affects body movements, including speech, and typically deteriorates with time. It is the second most prevalent neurodegenerative disorder, trailing only Alzheimer's disease. Parkinson's is a degenerative nervous system disorder affecting both the brain and the nervous system, making it one of the most prevalent degenerative illnesses globally, alongside Alzheimer's, brain cancer, and epilepsy [1]. This condition was initially recognized in 1817 by Dr. James Parkinson, who referred to it as "shaking palsy" [2].

Patients with HRS exhibit symptoms such as tremors at rigidity, rest, postural instability and akinesia (or bradykinesia). Additionally, they may display a flexed posture and motor blocks, which are considered classic characteristics of parkinsonism [3]. The tremors associated with HRS typically occur during periods of rest but subside during purposeful movement, thereby usually not significantly affecting daily life.

Rigidity describes the heightened stiffness of a patient's limbs when subjected to passive movement. Bradykinesia, hypokinesia, and akinesia are among the symptoms that may manifest, including reduced movement amplitude, slow movement, and reduced arm swing while walking (a lack of typical unconscious movements) [4].

While advanced stages of HRS can be reliably diagnosed, effective treatment remains challenging to achieve. In addition, treatment during the advanced stages of HRS may have a reduced likelihood of halting the progression of the condition. Diagnosis is often based on clinical observation of these motor symptoms, along with medical history and, in some cases, the use of rating scales like the Unified Parkinson's Disease Rating Scale (UPDRS) to assess symptom severity [5]. In recent times, machine learning (ML) has gained widespread use in disease diagnosis due to its user-friendliness and high accuracy [6]. ML has also found application in managing HRS. The structure of this paper is organized as follows: Section 2 provides a literature review of various machine learning techniques utilized for HRS detection, Section 3 outlines the machine learning algorithms employed, the research methodology utilized in this study is elucidated in Section 4, Section 5 presents the observations and results, and Section 6 offers a summary of work.

2. Literature Survey

This section elaborates various methods and evident for various state-of-the-art work using neural networks have

[1] *Department of Computer Science and engineering, Symbiosis Institute of Technology, Symbiosis International University (Deemed University) Pune, India vani.hiremani@sitpune.edu.in*
[2] *Department of Computer Science and Engineering, Gitam School of Technology, GITAM Bengaluru, Karnataka, India raghudevdas@gmail.com*
.[3] *Department of Computer Science and engineering, Symbiosis Institute of Technology, Symbiosis International University (Deemed University) Pune, India harshal.patil@sitpune.edu.in*
[4] *School of Engineering, Computer Science Department, Presidency University Bangalore, Karnataka, India smitapatil@presidencyuniversity.in*
[5] *Department of Computer Engineering Mukesh Patel School of Technology Management & Engineering, SVKM's Narsee Monjee Institute of Management Studies (NMIMS) Deem-to be- University, soni.sweta16@gmail.com*
[6] *Department of computer science and engineering*
K.L.E. institute of technology, Hubballli, vijeetapatil@gmail.com

yielded the best performance for the issue at hand [7]. To diagnose Parkinson's Disease (PD), fuzzy C-means clustering and k-NN were employed by the researchers. They assessed different values of k for the k-NN algorithm and selected the most effective one [8]. In a separate study [9], voice characteristics were utilized, and feature augmentation was employed to expand the initial 44 features in the dataset to a total of 177. Relief was utilized to filter out the most relevant features, leaving 66 to classify PD within the dataset. For PD prediction and the creation of feature subsets from the full feature set, the authors utilized a fuzzy k-NN technique in conjunction with PCA. They concluded that their proposed approach outperformed other methods in the existing literature [10]. In the context of employing Feature Selection (FS) for machine learning in brain surgery, researchers conducted a thorough analysis of relevant publications. In the case of PD brain surgery, an ML-based approach was employed to determine the precise area of the brain requiring intervention [11]. SVM was utilized for data classification, and, unlike previous studies, an unsupervised strategy was applied to address PD [12]. Another study aimed to predict PD by analyzing upper limb motion data from patients with PD and healthy volunteers. They conducted various performance tests with a device attached to participants' upper limbs [13]. For PD diagnosis, the authors leveraged Extreme Learning Machines (ELMs). They enhanced the handling of imbalanced data through a weighted approach and non-linear kernel function mapping. They also utilized the ABC method for Functional Self-Support (FS) and parameter optimization. Successful PD diagnosis was achieved using multiple methods, including PCA for dimension reduction, FDR for Fisher Discriminant ratio, and SVM for classification [15]. Furthermore, machine learning found applications in ranking software requirements [16-17]. Three machine learning algorithms, specifically SVM, KNN, and Naïve Bayes, were employed to predict heart arrhythmia disease, and the model's performance was evaluated using accuracy and kappa scores [18].

3. Machine Learning Algorithms

A. Support Vector Machine (SVM)

A Support Vector Machine (SVM) is a robust machine learning technique employed for both classification and regression purposes. Its primary function involves identifying the most effective hyperplane for segregating data points into different groups [19]. It is used in combination with several related supervised learning methods for classifying and regressing data. SVMs are part of a family of generalized linear classifiers (GLCs). The fundamental principle behind support vector machines (SVM) is to minimize the distance between elements that belong to different classes. If the elements belong to different classes at the beginning, then the problem is called classification. Many other important fields, such as image processing and pattern recognition, as well as medical diagnosis technology. SVM's accuracy in a handwriting recognition test is comparable to other popular modelling techniques such as neural networks (NN) [20] with long-range features, which makes it crucial for pixel maps when used as source [21]. Owing to the many intricate traits besides superior practical outcomes, Vapnik's SVMs have been demonstrating promising results [22]. SVM essentially operates on the principle of Structural Risk Minimization (SRM) rather conventional Empirical Risk Minimization approach which used by most of the NN's [23].

B. Random Forest

Random Forest is a classification procedure grounded on decision trees which combines multiple tree predictors. Each tree relies on the same vector values across the entire forest, and they operate independently. As the number of trees increases, the generalization error also increases. The strength of distinct trees in the forest and how they interact with each other play a big role in determining how accurate the model is. Remarkably, Random Forest maintains its performance even when there is additional noise in the training dataset. The basic processes of this strategy yield superior internal estimates that monitor inaccuracy, correlation, and strength. Next, it is shown what happens when you utilize more features while separating data. [24-25].

Random forest classification procedure:

Step 1: Random samples are selected from training set.

Step 2: A decision tree is constructed for every training instance.

Step 3: Voting is performed through the decision tree averaging.

Step 4: Finally, the most voted prediction result is selected as the final prediction result.

4. Methodology

This section details the procedure followed in this work. Fig.1depicts proposed method.

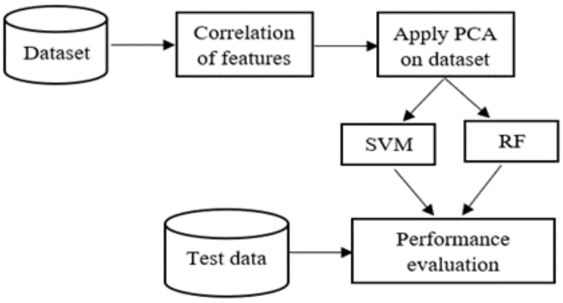


Fig. 1. Research Methodology

A. Dataset Details

This work uses the UCI Machine Learning Source's Parkinson Data Set. Table 1 displays the details of dataset's features used in the underlying work.

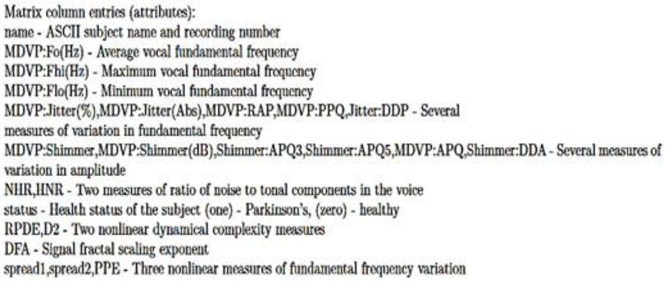
TABLE I. CHARACTERISTICS OF DATASET

Features	Multi-variables
Occurrences	197
Attributes type	Real
Attributes	23
Nature	Classification

The Parkinson’s Disease Dataset represents

- Null values are omitted in the dataset
- All instances in the dataset are unique
- Dataset contains 197 patients, 23 having PD

Few attributes information is depicted below



Matrix column entries (attributes):
name - ASCII subject name and recording number
MDVP:Fo(Hz) - Average vocal fundamental frequency
MDVP:Fhi(Hz) - Maximum vocal fundamental froquency
MDVP:Flo(Hz) - Minimum vocal fundamental frequency
MDVP:Jitter(%),MDVP:Jitter(Abs),MDVP:RAP,MDVP:PPQ,Jitter:DDP - Several measures of variation in fundamental frequency
MDVP:Shimmer,MDVP:Shimmer(dB),Shimmer:APQ3,Shimmer:APQ5,MDVP:APQ,Shimmer:DDA - Several measures of variation in amplitude
NHR,HNR - Two measures of ratio of noise to tonal components in the voice
status - Health status of the subject (one) - Parkinson’s, (zero) - healthy
RPDE,D2 - Two nonlinear dynamical complexity measures
DFA - Signal fractal scaling exponent
spread1,spread2,PPE - Three nonlinear measures of fundamental frequency variation

B. Correlation among features

In statistics, any kind of connection between random variables is called a "dependency" or "correlation". The term correlation is used to refer to the strength of a linear association among 2 variables. The type of correlation that is used here is the Spearman correlation. Pearson's correlation between rank variables (Pearson's correlation coefficient) is calculated by the following formula:

$$p = \frac{n(\Sigma ij)-(\Sigma i)(\Sigma j)}{\sqrt{[n\Sigma i^2-(\Sigma i)^2][n\Sigma j^2-(\Sigma j)^2]}}$$

$'i', 'j'$ are the raw values in the model data and n is the sample size.

Spearman Correlation Coefficient formula is given below:

$$p_s = \rho_{pg_i, pg_j} = \frac{cov(pg_i, pg_j)}{\sigma_{rg_i}, \sigma_{pg_j}}$$

Here, 'ρ' symbolizes Pearson Correlation Analysis coefficient which is used to rank the variables, where pgi and pgj are the standard deviations of the ranking variables, and covariance is represented by cov (pgi, pgj). In Fig. 2, a correlation plot is presented, displaying the relationships between attributes based on the generated correlation matrix. Additionally, in Fig. 3, another correlation plot is displayed, considering p-values and correlation values enhancing the understanding of the relationships between attributes

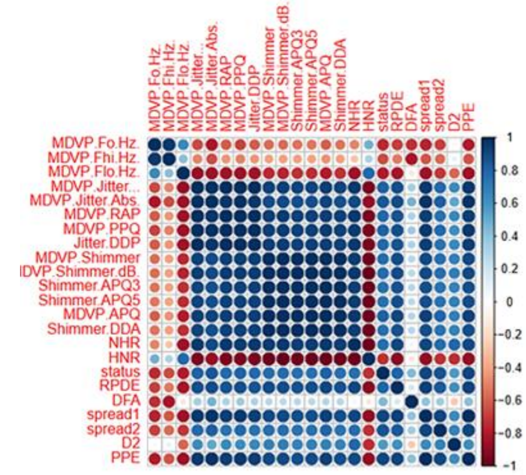


Fig. 2. Correlation between attributes

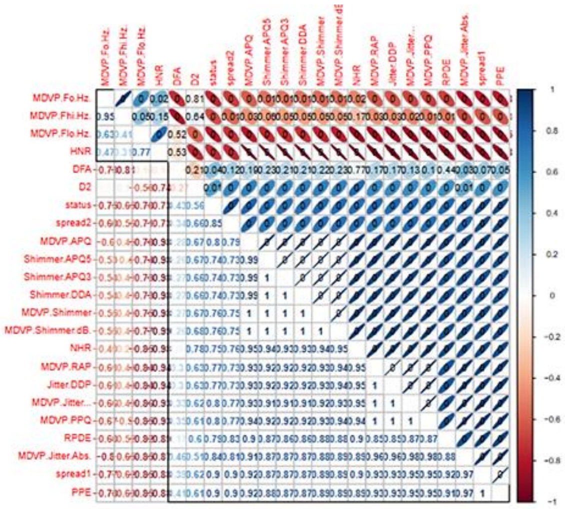
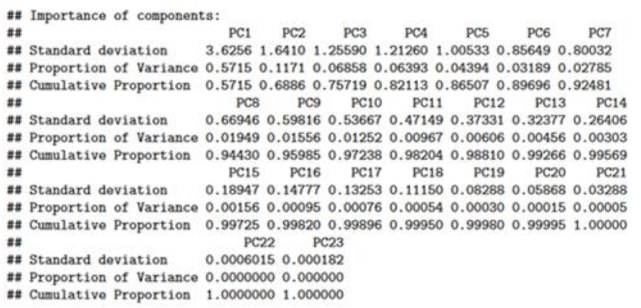


Fig. 3. Correlations with p-values and corr-values

C. Principal Component Analysis (PCA)

PCA is a mathematical approach to numerical analysis that involves the reduction of a large number of variables, which may or may not have relationships with each other, into a smaller set of uncorrelated variables known as Principal Components. This iterative approach involves finding a linear distribution of variables through the largest deviation, removing it, then reiterating the process.

The snapshot shown below is PCA applied on the dataset.



```
## Importance of components:
##                           PC1    PC2     PC3     PC4     PC5     PC6     PC7
## Standard deviation     3.6256 1.6410 1.25590 1.21260 1.00533 0.85649 0.80032
## Proportion of Variance 0.5715 0.1171 0.06858 0.06393 0.04394 0.03189 0.02785
## Cumulative Proportion  0.5715 0.6886 0.75719 0.82113 0.86507 0.89696 0.92481
##                            PC8     PC9    PC10    PC11    PC12    PC13    PC14
## Standard deviation     0.66946 0.59816 0.53667 0.47149 0.37331 0.32377 0.26406
## Proportion of Variance 0.01949 0.01556 0.01252 0.00967 0.00606 0.00456 0.00303
## Cumulative Proportion  0.94430 0.95985 0.97238 0.98204 0.98810 0.99266 0.99569
##                           PC15    PC16    PC17    PC18    PC19    PC20    PC21
## Standard deviation     0.18947 0.14777 0.13253 0.11150 0.08288 0.05868 0.03288
## Proportion of Variance 0.00156 0.00095 0.00076 0.00054 0.00030 0.00015 0.00005
## Cumulative Proportion  0.99725 0.99820 0.99896 0.99950 0.99980 0.99995 1.00000
##                             PC22     PC23
## Standard deviation     0.0006015 0.000182
## Proportion of Variance 0.0000000 0.000000
## Cumulative Proportion  1.0000000 1.000000
```

Fig. 4. Shows 2D-Plot of PCA upon considering 23 features

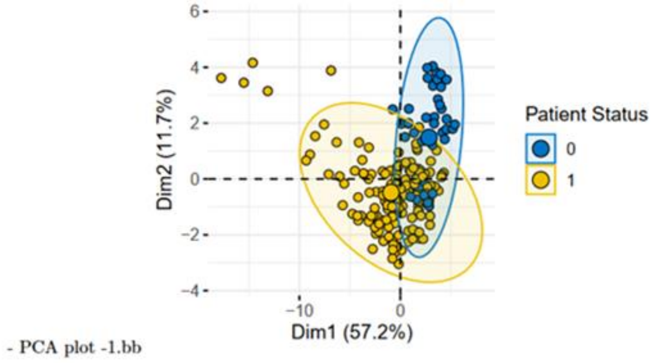
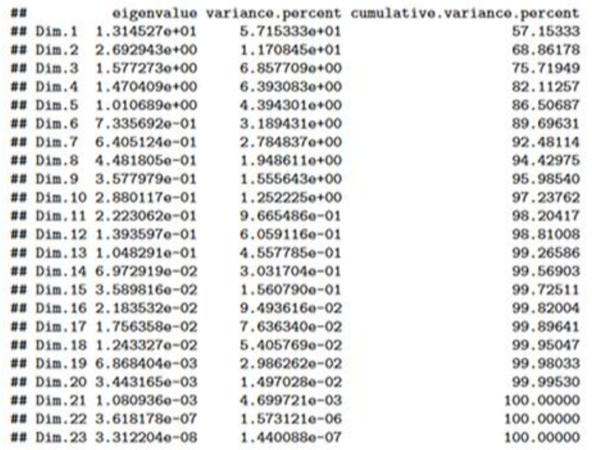


Fig. 4 2D-Plot for PCA

Eigenvalues, variances, and cumulative variances for each dimension or principal component are shown as follows.



```
##        eigenvalue variance.percent cumulative.variance.percent
## Dim.1  1.314527e+01     5.715333e+01                    57.15333
## Dim.2  2.692943e+00     1.170845e+01                    68.86178
## Dim.3  1.577273e+00     6.857709e+00                    75.71949
## Dim.4  1.470409e+00     6.393083e+00                    82.11257
## Dim.5  1.010689e+00     4.394301e+00                    86.50687
## Dim.6  7.335692e-01     3.189431e+00                    89.69631
## Dim.7  6.405124e-01     2.784837e+00                    92.48114
## Dim.8  4.481805e-01     1.948611e+00                    94.42975
## Dim.9  3.577979e-01     1.555643e+00                    95.98540
## Dim.10 2.880117e-01     1.252225e+00                    97.23762
## Dim.11 2.223062e-01     9.665486e-01                    98.20417
## Dim.12 1.393597e-01     6.059116e-01                    98.81008
## Dim.13 1.048291e-01     4.557785e-01                    99.26586
## Dim.14 6.972919e-02     3.031704e-01                    99.56903
## Dim.15 3.589816e-02     1.560790e-01                    99.72511
## Dim.16 2.183532e-02     9.493616e-02                    99.82004
## Dim.17 1.756358e-02     7.636340e-02                    99.89641
## Dim.18 1.243327e-02     5.405769e-02                    99.95047
## Dim.19 6.868404e-03     2.986262e-02                    99.98033
## Dim.20 3.443165e-03     1.497028e-02                    99.99530
## Dim.21 1.080936e-03     4.699721e-03                   100.00000
## Dim.22 3.618178e-07     1.573121e-06                   100.00000
## Dim.23 3.312204e-08     1.440088e-07                   100.00000
```

5. Observations and Results

The Parkinson's data set [26] sourced from the UCI Machine Learning Repository is utilized in the underlying experiment. It employs both SVM and RF for classifying the dataset and then compares their prediction performance. The experiment is conducted using the R programming language, taking advantage of various R packages suited for diverse machine learning experiments. Notably, some of the packages utilized in this study include dplyr, corrplot, mlbench, and caret.

A.Comparison of Classifiers

After three replications and ten cross-validations, 80% of training data besides 20% of the test data were combined. Table 2 shows the assessment of both SVM and RF in terms of accuracy and kappa score on Parkinson dataset.

TABLE 2. SVM AND RF MODEL PERFORMANCE

Model	Accuracy (%)	Kappa Score
SVM	84.3	0.824
RF	87.2	0.82

6. CONCLUSION

Based on the outcomes, it is evident that the accuracy of classification by Random Forest can be equated with the accuracy of the SVM. Another benefit of random forest classifiers is that they only need to be configured with two parameters, which is less in comparison to SVM which relies on many users defined parameters. Random Forest classifiers outperforms SVM by being able to manage categorical data, handling missing data values and addressing imbalanced datasets. In near future, innumerable feature selection techniques can be discovered to choose the most relevant attributes, which might potentially enhance the accuracy rate.

References

[1] J. Parkinson, “An Essay on Shaking Palsy”. London: Whitting-ham and Rowland Printing, 1817.

[2] B. Calne, “Is Hypokinetic rigid syndrome the consequence of an event or a process,” Neurology, Vol. 44, no. 15, pp. 5–5, 1994.

[3] Jankovic J, “Parkinson’s disease: clinical features and diagnosis”, Journal of Neurology, Neurosurgery & Psychiatry 2008;79:368-376.

[4] William Dauer, Serge Przedborski, “Parkinson's Disease: Mechanisms and Models”, Neuron, Volume 39, Issue 6, 2003, Pages 889-909, ISSN 0896-6273, https://doi.org/10.1016/S0896-6273(03)00568-3.

[5] G. Goetz, et al., “Movement disorder society-sponsored revision of the unified Parkinson’s disease rating scale (MDS-UPDRS): Scale presentation and clinimetric testing results,” Mov. Disord., Vol. 23, no. 15, pp. 2129–70, 2008.

[6] Rayan Z, Alfonse M, Salem A-BM. “Machine learning approaches in smart health”. Procedia Comput Sci 2019; pp. 154:361–8.

[7] Das R. “A comparison of multiple classification methods for diagnosis of Parkinson disease”. Expert Syst Appl 2010;37(2):1568–72.

[8] Polat K. “Classification of Parkinson’s disease using feature weighting method on the basis of fuzzy C-means clustering”. Int J Syst Sci 2012;43(4):597–609.

[9] Yaman O, Ertam F, Tuncer T. “Automated Parkinson’s disease recognition based on statistical pooling method using acoustic features”. Med Hypotheses 2019:109483.

[10] H.-L. Chen, C-C. Huang, X-G. Yu, X. Xu, X. Sun, G. Wang, and S-J. Wang, “An efficient diagnosis system for detection of Parkinson’s disease using fuzzy k-nearest neighbor approach,” Expert Syst. Appl., Vol. 40, pp. 263–71, 2013.

[11] Wan KR, Maszczyk T, See AAQ, Dauwels J, King NKK. “A review on microelectrode recording selection of features for machine learning in deep brain stimulation surgery for Parkinson’s disease”. Clin Neurophysiol 2019;130(1):145–54.

[12] Nilashi M, Ibrahim O, Ahmadi H, Shahmoradi L, Farahmand M. “A hybrid intelligent system for the prediction of Parkinson’s Disease progression using machine learning techniques”. Biocybern Biomed Eng 2018;38(1):1–15.

[13] Cavallo F, Moschetti A, Esposito D, Maremmani C, Rovini E. “Upper limb motor preclinical assessment in Parkinson’s disease using machine learning”. Park Relat Disord 2019; 63:111–6.

[14] Wang Y, Wang AN, Ai Q, Sun HJ. “An adaptive kernel-based weighted extreme learning machine approach for effective detection of Parkinson’s disease”. Biomed Signal Process Control 2017;38:400–10.

[15] Singh G, Vadera M, Samavedham L, Lim ECH. “Machine learning-based framework for multi-class diagnosis of neurodegenerative diseases: a study on Parkinson’s Disease”. IFAC-Papers OnLine 2016;49(7):990–5.

[16] Raghavendra M Devdas, Nagaraj Cholli. “Interdependency aware QUBIT and BROWNBOOST rank for large scale requirement prioritization”. International Journal of Computing and Digital Systems. Vol 11, Issue 1, pp. 625-634.

[17] Devadas R, Cholli NG. “PUGH Decision Trapezoidal Fuzzy and Gradient Reinforce Deep Learning for Large Scale Requirement Prioritization”. Indian Journal of Science and Technology. 15(12): 542-553.

[18] R. M Devadas, "Cardiac arrhythmia classification using SVM, KNN and Naive Bayes algorithms", International Research Journal of Engineering and Technology (IRJET), pp. 3937-3941, 2021, [online] Available: https://www.irjet.net/archives/V8/i5/IRJET-V8I5721.pdf.

[19] Yao, Z.; Lei, L.; Yin, J., “R-C4.5 Decision tree model and its applications to health care dataset”. Proceedings of International Conference on Services Systems and Services Management 2005, pp. 1099-1103.

[20] Hiremani, V.A., Senapati, K.K. (2021). Quantifying apt of RNN and CNN in Image Classification. In: Nath, V., Mandal, J.K. (eds) Proceeding of Fifth International Conference on Microelectronics, Computing and Communication Systems. Lecture Notes in Electrical Engineering, vol 748. Springer, Singapore. https://doi.org/10.1007/978-981-16-0275-7_59.

[21] Burke HB, Goodman PH, Rosen DB, et al. 1997. “Artificial neural networks improve the accuracy of cancer survival prediction”. Cancer, 79:857-62.

[22] Leenhouts HP, “Radon-induced lung cancer in smokers and nonsmokers: risk implications using a two-mutation carcinogenesis model”. Radiat Environ Biophys, 1999 38:57-71.

[23] Boser, I. Guyon, and V. Vapnik. “A training algorithm for optimal margin classifiers”. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, pages 144{152. ACM Press, 1992.

[24] Michie, D. J. Spiegelhalter, C. C. Taylor, and J. Campbell, editors. “Machine learning, neural and statistical classi_cation”. Ellis Horwood, Upper Saddle River, NJ, USA, 1994. ISBN 0-13-106360-X. Data available at http://archive.ics. uci.edu/ml/machine-learning-databases/statlog/.

[25] V. A. Hiremani and K. K. Senapati, "Significance of Conventional and Nonconventional Features in Classification of Face through Human Intelligence," 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2019, pp. 931-937, doi: 10.1109/ICSSIT46314.2019.8987756.

[26] Little MA, McSharry PE, Roberts SJ, Costello DAE, Moroz IM. “Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection”, BioMedical Engineering OnLine 2007, 6:23 (26 June 2007).