

Jñānābha, Vol. 56(1P) (2026), 21-32

EARLY DETECTION OF HUNTINGTONS DISEASE (HD) BY MACHINE LEARNING USING GENETIC DATA

Arunesh Dutt^{1*}, Komal Singh^{1,2}, Pradeep Kumar Tiwari^{1,2}, Abhishek Pathak³,
Ashish Khare¹, Vijaya Nath Mishra³ and Narendra Kumar Shukla¹

¹Department of Electronics and Communication Engineering, University of Allahabad, Prayagraj, Uttar Pradesh, India - 211002

²Centre of Computer Education and Training, Institute of Professional Studies, University of Allahabad, Prayagraj, Uttar Pradesh, India-211002

³Department of Neurology, Institute of Medical Sciences (IMS), Banaras Hindu University, Varanasi, Uttar Pradesh, India - 221005

*Corresponding Author

Email: aruneshdutt19@gmail.com, s.komalcs@gmail.com, pradeeptiwari@allduniv.ac.in,
abhishekipathakims@gmail.com, khare@allduniv.ac.in, vnishraneuro@bhu.ac.in, nkshukla@allduniv.ac.in,

(Received : August 01, 2026; Revised : February 17, 2026; Accepted : February 20, 2026)

DOI: <https://doi.org/10.58250/jnanabha.2026.561P-3>

Abstract

A genetic condition known as Huntington's disease (*HD*) impairs voluntary movement by causing the death of brain cells called neurons in various brain regions. It is challenging to identify early-stage *HD* in clinical practice since there aren't many effective, simple-to-administer cognitive tests that are sensitive to organized data. Life expectancy has risen steadily as modern treatment has advanced. The prevalence of *HD* has also increased and is anticipated to increase over the next several decades as a result of escalating healthcare expenses. Early *HD* detection is therefore essential. Through this work, we used a clinical and genetic dataset of 10 *HD* patients. The present approach introduces a novel support vector machine (*SVM*)-based machine Learning (*ML*) algorithm for the detection of *HD* patients. The results of a classification study comparing various techniques indicate that *SVM*-based classification is more accurate at identifying subjects, even across varying stages of development. The detection capabilities of the suggested method were thoroughly examined using clinical and genetic data. With two-fold cross-validation, the maximum accuracies for early *HD* achieved by Gaussian *SVM*, Cubic *SVM*, Quadratic *SVM*, Naive Bayes, and Linear *SVM* are 70%, 100%, 100%, 80%, and 60%, respectively. These findings show that a useful and promising strategy for *HD* identification is the use of supervised classification techniques to genetic characteristics. The primary conclusions of this research are: 1) To detect substantial *HD*, a genetic test will be necessary. 2) The quadratic and cubic *SVMs* are the most effective predictors for *HD*. 3) According to experimental findings, quadratic and cubic *SVMs* outperform other classification algorithms for *HD* diagnosis as well. Even with a small dataset, the non-linear *SVM* approach improves diagnosis accuracy.

2020 Mathematical Sciences Classification: 00B99.

Keywords and Phrases: Huntington's disease (*HD*), symptoms, classification issues, and machine learning (*ML*) approaches.

1 Introduction

A neurological disorder that significantly affects people is *HD* [1]. It has a lasting impact on people's families in addition to themselves [20]. An uncontrollable progression of mental symptoms, cognitive decline, and physical dysfunction characterizes this complicated clinical picture, which ultimately leads to severe disability and a tragically reduced lifetime [12]. The seriousness of *HD* has led the medical and scientific communities worldwide to become more concerned, prompting a variety of initiatives not only to understand its etiology and pathophysiology but also to advance early detection and treatment [7]. The *HD* pathological

progression is still unclear, attracting interest from a variety of scientific fields. In order to better understand the motor engagement in *HD*, the work [17] investigates motor speech patterns. Paper [6] explores speech biomarkers at different phases of *HD*, highlighting the range from early manifestation to pre-symptomatic. The paper [19] provides information on Huntingtons complex role, which extends from neurodevelopment to neurodegeneration. In [21] and another study published in Nature Medicine [8] provide light on synaptic loss by highlighting the early roles of innate immune systems, complement activation, and microglia in the reduction of corticostriatal synapses. Views on the harmful consequences of mutant Huntington’s disease are provided by the study [8]. It is impossible to overestimate the significance of early *HD* diagnosis since it enables prompt treatment that could reduce the progression of this debilitating neurological condition. By combining existing knowledge and offering perceptive analysis, it seeks to encourage more study and innovation in this field in order to improve our capacity to identify and treat *HD* in its early stages and possibly enhance the lives of people impacted and their families [8]. Diseases can be diagnosed and tracked using biomarkers, which are measurable indicators of a certain biological process or state [22, 23]. The *HD* is a degenerative neurological condition that runs in families, is brought on by a mutation in the cytosine-adenine guanine (*CAG*) repeat of the *HTT* gene [4, 11]. Biomarkers can be utilized in the context of *HD* to offer more details regarding the patient’s condition, prognosis, and reaction to therapy. It provides a thorough analysis of *ML*-powered therapies for this illness and evaluates their efficacy. By showing how beneficial these technologies are in this situation, this study aims to close the gap between *ML* and *HD*. The important contributions of this study are:

- This study investigates the potential for machine learning (*ML*) to enhance *HD* diagnosis.
- This study presents a variety of *SVM*-based machine learning techniques that have been used for early diagnosis of *HD*.
- In this work, the cubic *SVM* outperformed other existing *SVM*-based algorithms.

This study presents a novel *SVM*-based machine learning method for *HD* classification. The suggested approach uses categorization techniques to accomplish this goal. In this investigation, data from ten *HD* patients were used. With the most important *SVM*-based method (Quadratic *SVM*), the highest accuracy is 100%, and the sensitivity and specificity values are almost 100%, respectively.

2 Problems and Related Work

A fatal autosomal dominant disorder, *HD* often manifests in early adulthood and develops over time to produce a range of motor, behavioral, and cognitive symptoms. Even though *HD* is difficult to forecast in medical science, researchers work hard to create the finest techniques and results. Because machine learning yields practical results, it is becoming more and more popular. In order to classify Huntington & 39;s illness, Miranda, A. *et al.* [13] used the support vector machine (*SVM*) technique on oculomotor variables that were aggregated from a four-task psychophysical experiment. With an accuracy of 73.47%, sensitivity of 74.31%, and specificity of 72.64%, it was possible to automatically separate control participants from presymptomatic *HD* (pre-*HD*) participants; 81.84%, sensitivity of 76.19%, and specificity of 87.48% were achieved when separating control participants from *HD* patients; and 83.54%, sensitivity of 92.62%, and specificity of 74.45% were achieved when separating pre-*HD* participants from *HD* patients. A Money Box Test (*MBT*) was created by Bennasar *et al.* [2] to measure involuntary movement. A novel method for automatically dividing the subjects into two groups, healthy and *HD*, is suggested using time series accelerometer data. After extracting a set of 90-time time-domain features from the accelerometer data, the dataset dimensionality is decreased, and the features; relevance is evaluated using a feature selection technique. Lastly, the participants are split into *HD* and healthy groups using an *SVM* classifier. Data from 15 *HD* patients and seven healthy controls are used in this study. The eight most important criteria yield the maximum accuracy of 86.36%, sensitivity of 87.50%, and specificity of 83.33%. According to Perez *et al.* [16], the proposed method can accurately distinguish between control and *HD* patients using non-linear forecasting techniques and *DNA CpG* methylation data as an input. It should be mentioned that the dataset being studied is not very large. Nonetheless, the findings appear to be very consistent, and if other data sets become available, the study might be conducted using them. The results will be evaluated using the neural network (*NN*) technique. The accuracy, specificity, and sensitivity of the *NN* are 92%, 80%, and 95%, respectively. Using a method of shallow machine learning, Maddury *et al.* [14] provided a prepared, validated dataset that includes biological signal data (*EEG*, *ECG*, AND *FNIRS*) from a clinic with 27 *HD*-positive patients, 36 controls, and 6 unknowns [14]. With a 91.35% accuracy rate and a 0.96 area under the receiver operator characteristic curve, the scaled-out Extremely Randomized Trees technique produced the best results. With

60.86% of the features having $p < 0.05$, the resulting feature analysis showed that the most significant features were those from the raw signal. The results imply that heart and brain signals may be useful in detecting abnormalities in *HD* and monitoring the course of the illness in individuals. Erda *et al.* [6] use convolutional neural networks (*CNNs*) to recognize gait data and convert it into *QR* codes to create a dependable and efficient system for detecting neurodegenerative diseases. By introducing a novel method to gait pattern analysis, this strategy aims to improve the accuracy of identifying neurodegenerative disorders such as *HD*, Parkinsons disease (*PD*), and amyotrophic lateral sclerosis (*ALS*). The dataset consists of recordings from 13 patients with *ALS*, 20 patients with *HD*, 15 patients with *PD*, and 16 healthy controls. Using a 10-fold *CV*, the accuracy rates for *ALS* versus control, *PD* versus *HD* versus *ALS* versus control, *PD* versus control, and *NDD* versus control were 95.81%, 93.86%, and 93.56%, respectively, and *HD* versus control, respectively. These results demonstrate how effectively the proposed method may distinguish between the control group and different neurodegenerative diseases.

Earlier research used data from many sources to solve the problem of categorizing neurodegenerative disorders by utilizing traditional machine learning. *ML* approaches like the Quadratic *SVM* can be used to generate more creative solutions for the classification challenge of neurodegenerative illnesses. In the current work, *HD* genomic data were categorized using the quadratic *SVM* machine learning technique. One of the innovative features of this paper is the reproduction of genetic data in the classification issue using Quadratic *SVM*.

3 Problems and Related Work

Standard protocol approval, ethical clearance, and consent: Following a Memorandum of Understanding between the two universities, protocol clearance was acquired from the institutional ethics committees of the two collaborating institutions (a teaching engineering institute and a tertiary care university hospital) (Ethics permission no. *IERB ID: COM-D-21- 001*). After gaining patients; or caregivers; written informed consent, participants were enlisted.

Study period and site: The study was conducted from August 2022 to September 2024. Data collection and nerve conduction velocity (*NCV*), and genetic tests were performed at a tertiary care university hospital. Data analysis and algorithm development were mainly performed at the Engineering Institute lab.

Study Design and Participants: The study was cross-sectional. For the study, ten *HD* patients were enlisted.

***NCV* and Genetic Features:** In our study, muscles were characterized on the basis of *NCV* parameters and genetic data. All patients were assessed for the diagnosis of *HD* clinically and on *NCV*. The *NCV* parameters were recorded with concentric needle electrodes. A region with many *CAG* repeats (Cytosine, Adenine, and Guanine) is found in the gene that causes *HD*. Although the gene encoding the Huntington protein has these *CAG* repeats in all of us, individuals with *HD* have an abnormally high amount of *CAG* repeats in one of their inherited genes. For the *NCV* of *HD* patients, we perform the Youden Index (*J*), which determines the maximal efficacy of a test and computes the performance of a single diagnostic test. The index's value falls between 0 and 1, where 0 denotes a test with no discriminatory power, and 1 denotes a perfect test.

Validation: In the study, clinical diagnosis (based on clinical and investigations by experts) was considered. Neurology experts (*AP* and *VNM*), from the Department of Neurology, Institute of Medical Sciences, Banaras Hindu University (*BHU*), Varanasi confirmed the clinical diagnosis of the disease. An *ML* Model was run to identify the diagnosis on the patients data.

Data Collection:

- **Clinical data:** The *HD* UHDRS-TMS=Unified Huntington's Disease Rating Scale-Total Motor Score was used to evaluate how *HD* affected day-to-day functioning. Among the demographic characteristics of the *HD* is the mean (*SE*).
- ***NCV* and Genetic data:** *NCV* and genetic data from 10 clinically diagnosed *HD* patients were taken. The demographic and clinical information of the *HD* patients was collected, which is reported in Table 3.1.
- **Data Analysis:** *NCV* and genetic data sets were used for an *ML*-based diagnosis detection system to identify that genetic data is a biomarker for *HD* diagnosis.
- ***HD* Detection with *SVM*-based *ML* Modeling** A separating hyperplane is constructed as a decision boundary using the *SVM* machine learning algorithm, which maximizes the margin of separation between the classes. More specifically, given training data that includes two distinct classes,

the *SVM* method creates a model using data from individuals in one class on one side of the hyperplane and those in the other class on the other side. The optimal separation hyperplane is found by reducing the amount of classification error and maximizing the distance between each class of data points and the decision boundary. Overall, this study has shown that *SVMs* are a useful tool for detecting *HD* and have led the way in applying machine-learning algorithms to genetic data in *HD*. The proposed model used an *SVM*-based *ML* algorithm, which is explained in Figure 3.1.

Table 3.1: Features of the *HD* patients' demographics.

Case	Gender (male: female)	Age	CAG Repetition	Duration of Disease (years)	Years Before the Start	UHDRS (Motor)
<i>HD</i> Patients	7:3	35.79 (2.63)	41.29 (0.52)	N/A	21.01 (2.92)	1.93 (0.46)

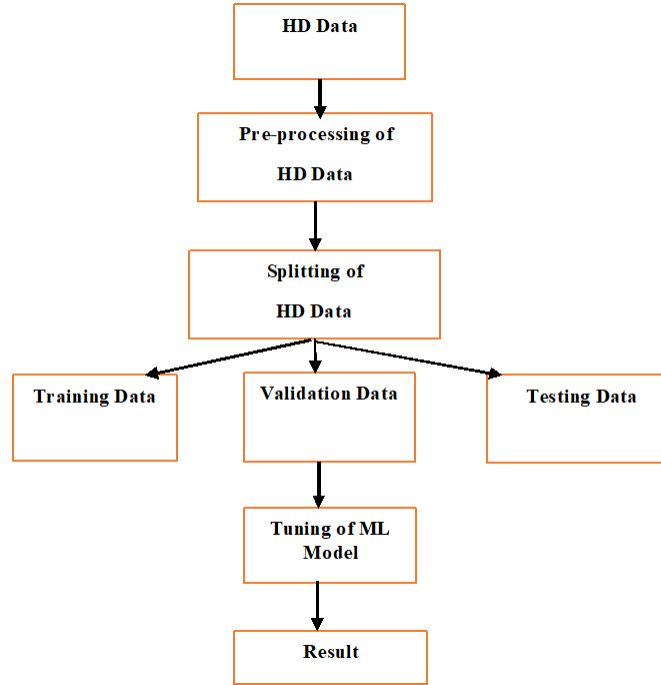


Figure 3.1: Proposed Model for *HD* Detection

4 Experimental Results

We evaluated the model's performance using an *SVM*-based *ML* algorithm using the confusion matrix for sensitivity, specificity, accuracy [3, 24], and Youden Index (J)[18], which is shown in equations (4.1), (4.2), (4.3), and (4.4).

- **Sensitivity :** Sensitivity is the only really positive statistic that considers every situation and can be calculated as follows:

$$Sensitivity = \frac{TP}{TP + FN}, \quad (4.1)$$

- **Specificity:** The number of true negatives that are accurately identified and calculated is ascertained in this manner:

$$Specificity = \frac{TN}{TN + FP}, \quad (4.2)$$

- **Accuracy:** The accuracy is used to determine the proportion of correctly classified values. It shows the frequency of accuracy of our classifier. The following formula can be used to calculate accuracy:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (4.3)$$

- **Youden Index (J):** It is the most notable vertical difference between the receiver operating characteristic (*ROC*) curve and the diagonal line, which indicates how well a random guess performs. The *J* can be computed in this way:

$$YoudenIndex(J) = Sensitivity + Specificity - 100, \quad (4.4)$$

Here, *TN* stands for true negative, *FN* for false negative, *FP* for false positive, and *TP* for true positive.

The quadratic *SVM* model's confusion matrix in terms of *TPR* and *FNR* is displayed in Figure 4.1. The quadratic *SVM* model's confusion matrix is displayed in Figure 4.2 in terms of *PPV* and *FDR*. The *ROC* curve for the *SVM* quadratic model is displayed in Figure 4.3. For the cubic *SVM* model, Figure 4.4 displays the confusion matrix in terms of *TPR* and *FNR*. The confusion matrix for the cubic *SVM* model in terms of *PPV* and *FDR* is shown in Figure 4.5. The cubic *SVM* model's *ROC* curve is displayed in Figure 4.6. The training duration and prediction speed of the proposed model are shown graphically in Figure 4.7. The Youden index, sensitivity, specificity, and accuracy of the suggested models are shown graphically in Figure 4.8. The visual representation for contrasting the performance of the proposed model with earlier research is displayed in Figure 4.9. The total performance graphical depiction and comparison with the suggested model work are displayed in Figure 4.10.

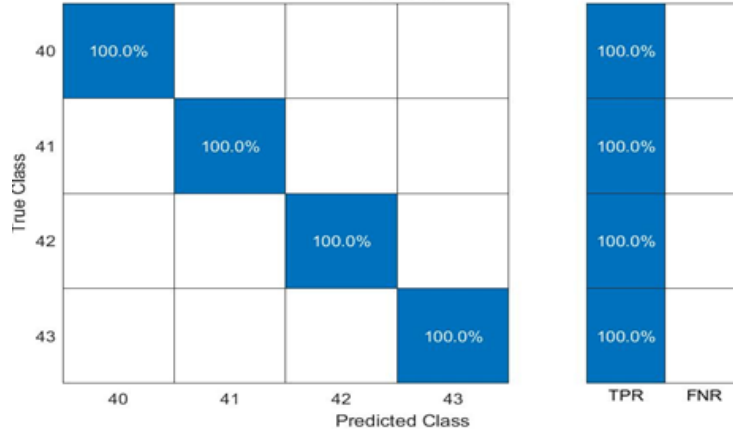


Figure 4.1: Confusion matrix in terms of *TPR* and *FNR* of the quadratic *SVM* model

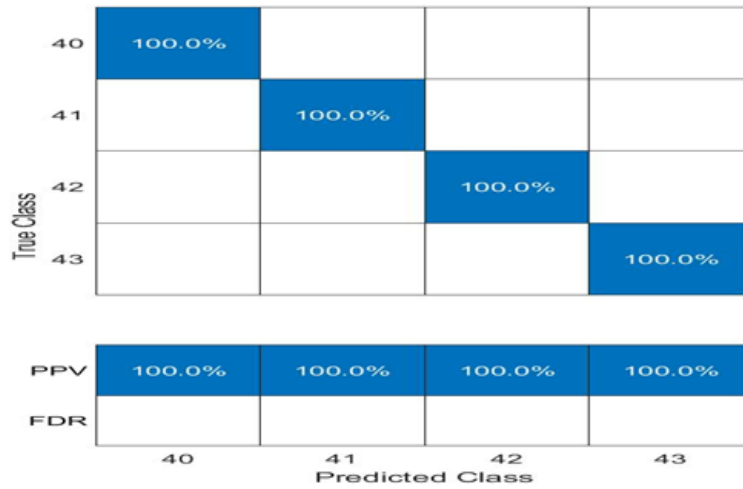


Figure 4.2: Confusion Matrix in terms of *PPV* and *FDR* for the quadratic *SVM* model.

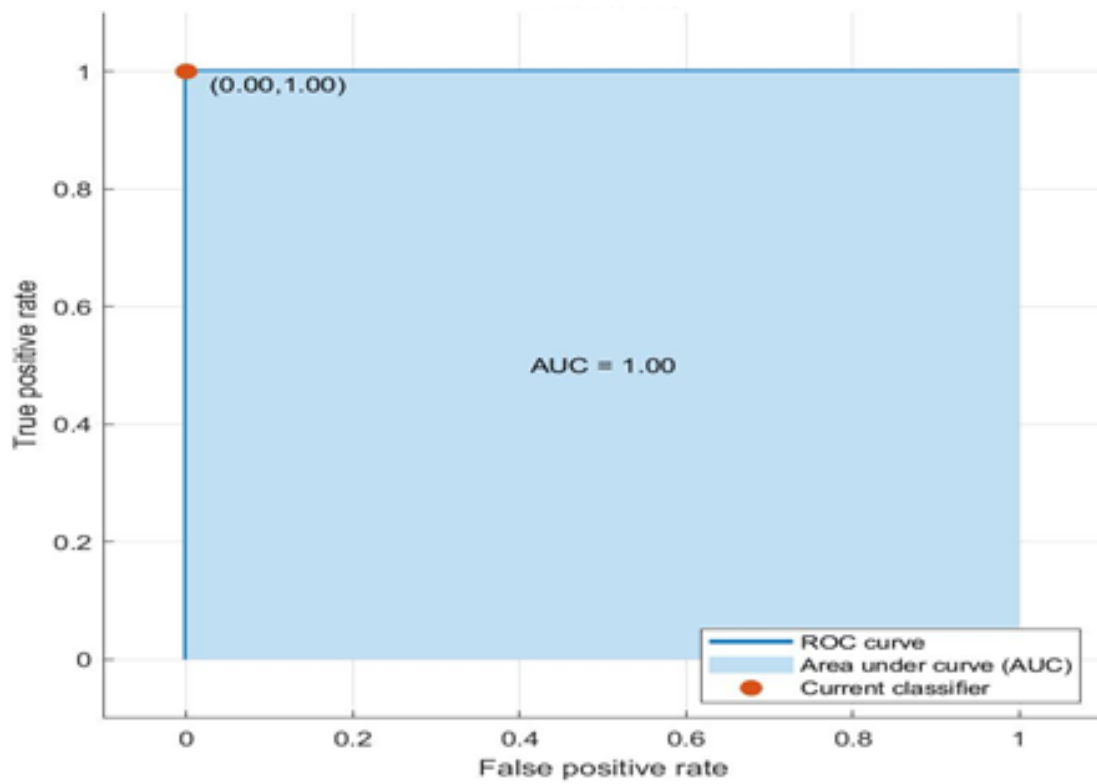


Figure 4.3: ROC curve for the SVM quadratic model

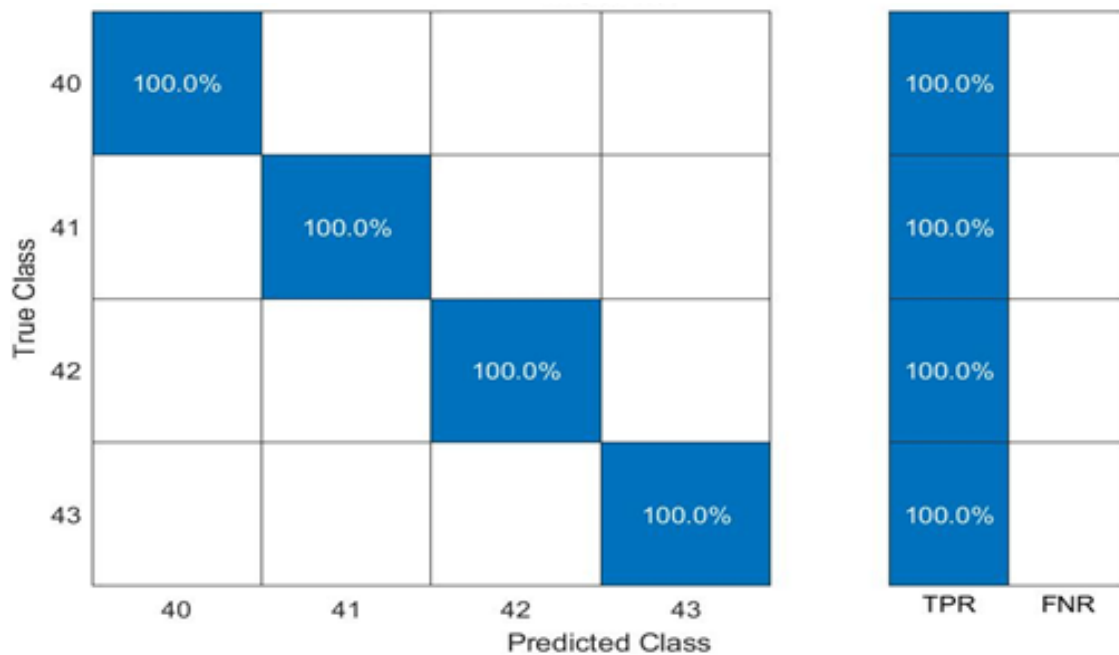


Figure 4.4: Confusion matrix in terms of TPR and FNR for the cubic SVM model

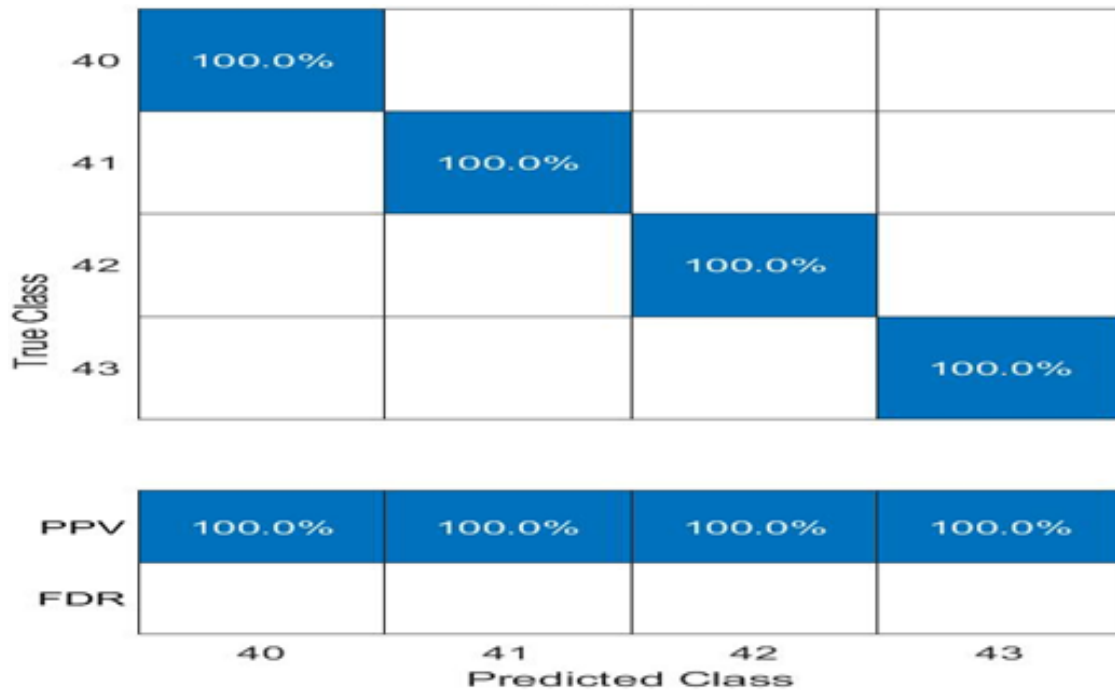


Figure 4.5: Confusion matrix in terms of PPV and FDR for the cubic SVM model

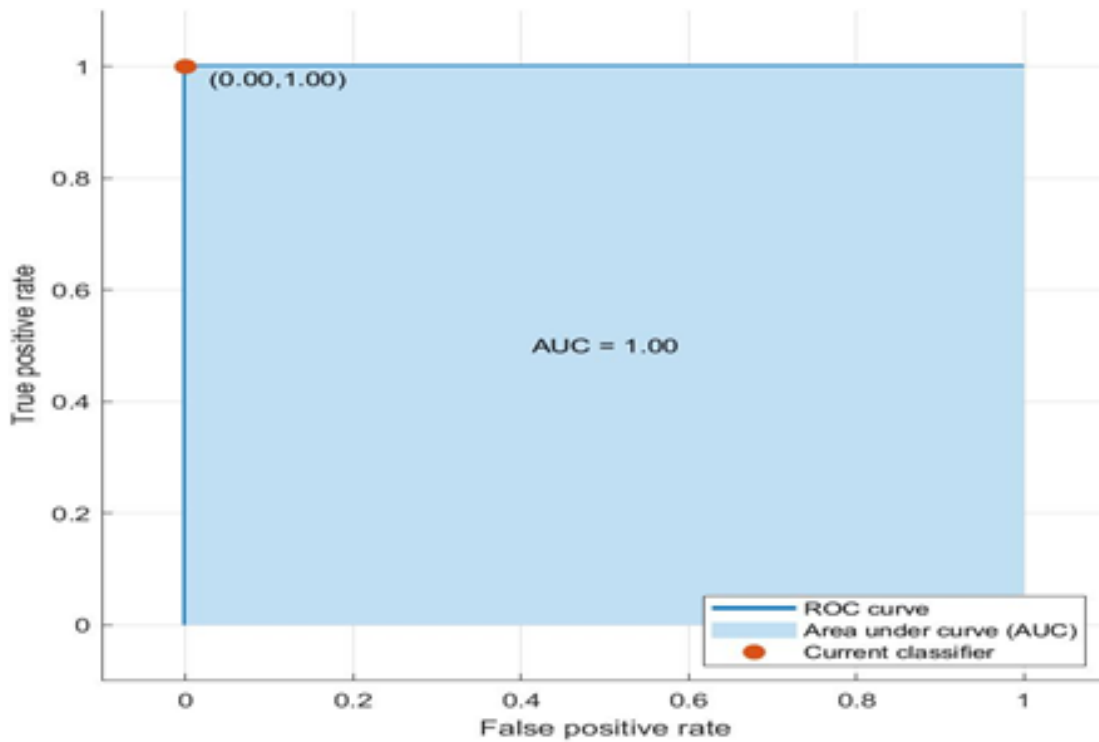


Figure 4.6: ROC curve for cubic SVM model

Table 4.1: The performance of the proposed model in terms of prediction speed and training time.

SVM-based Algorithms	Prediction speed (obs/sec)	Training time (sec.)
Gaussian <i>SVM</i>	43000	0.86
Cubic <i>SVM</i>	29000	0.91
Quadratic <i>SVM</i>	21000	0.73
Naive Bayes	46000	3.74
Linear <i>SVM</i>	41000	2

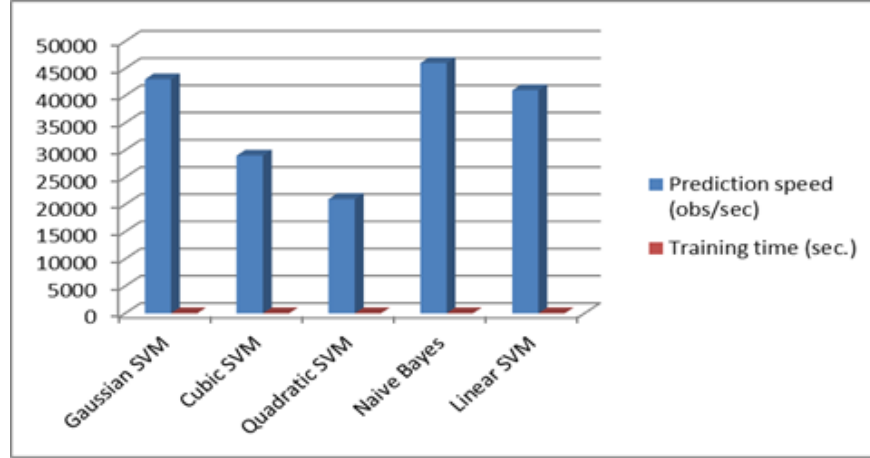


Figure 4.7: A visual representation of the suggested model's training time and prediction speed.

Table 4.2: Performance Matrices of the Proposed Model

Models	Sensitivity (%)	Specificity (%)	Accuracy (%)	J	AUC
Gaussian <i>SVM</i>	50	50	70	0	0.56
Cubic <i>SVM</i>	100	100	100	100	1
Quadratic <i>SVM</i>	100	100	100	100	1
Naive Bayes	85.42	99.99	80	85.41	1
Linear <i>SVM</i>	41.67	87.5	60	29.17	0

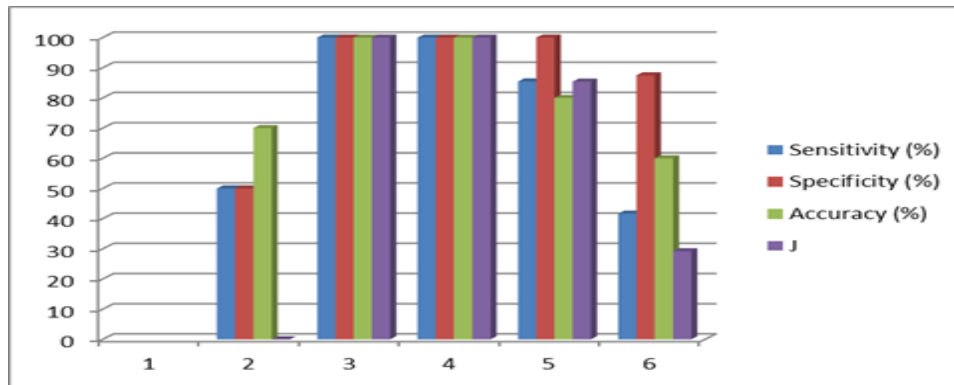


Figure 4.8: Visual illustration of the Youden index, sensitivity, specificity, and accuracy.

Table 4.3: Performance comparison with existing works and the proposed model.

Models	Sensitivity (%)	Specificity (%)	Accuracy (%)	J
Bennasar, M., <i>et al.</i> [2]	87.50	83.33	86.36	-
Erda, . B., <i>et al.</i> , [6]	93.86	93.56	95.81	
Miranda, A., <i>et al.</i> [13]	76.19	87.48	81.84	-
Maddury, S. <i>et al.</i> , [14]	-	-	91.35	-
Alfonso Perez, G., <i>et al.</i> [16]	95	80	92	-
Proposed Model	100	100	100	100

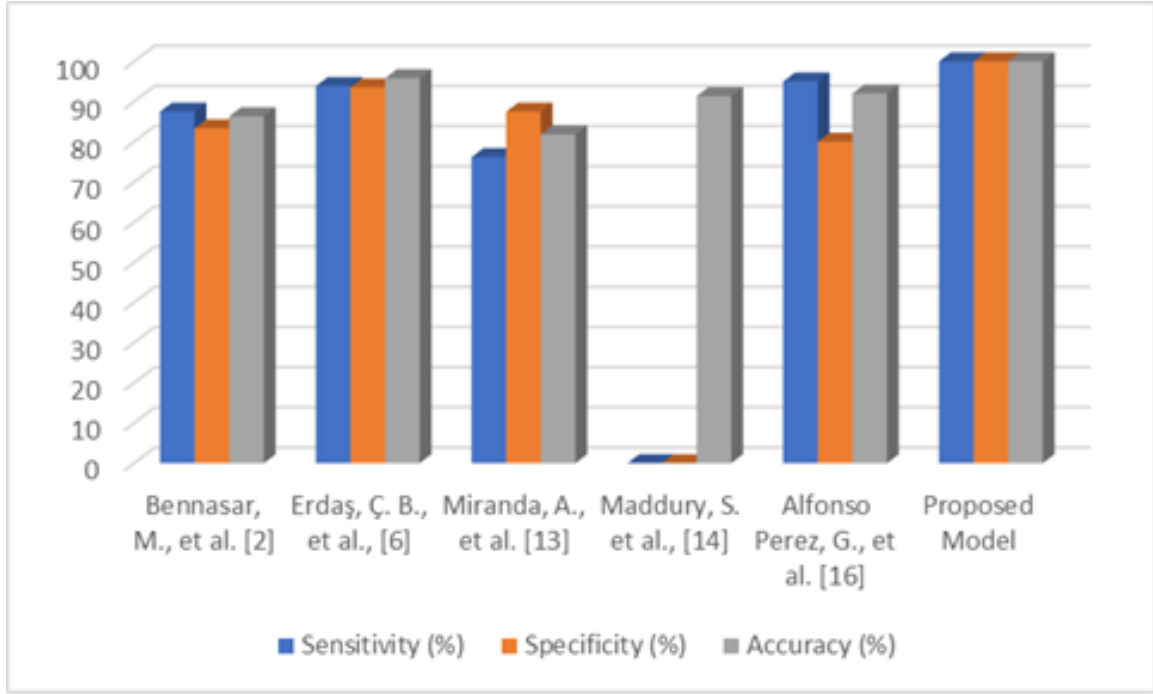


Figure 4.9: Visual representation for comparing the suggested model's performance to that of previous works

Table 4.4: Overall performance and comparison with the proposed model.

Model / Year	Modality	Algorithm Used	No. of features used	No. of <i>HD</i> Patients	Accuracy (%)	Research focus
Bennasar, M., <i>et al.</i> 2016 [2]	Accelerometer data	<i>SVM</i>	90	15	86.36	Classification
Cheng, J., <i>et al.</i> , 2020 [5]	Genetic data	RF	157	-	97.46	Identification
Erda, . B., <i>et al.</i> , 2024 [6]	Gait data	CNN	13	20	95.81	Classification
Miranda, A., <i>et al.</i> 2016 [13]	Oculomotor data	<i>SVM</i>	-	14	81.84	Classification
Maddury, S. 2023 [14]	Biological signal data (EEG, ECG, FNIRS)	Tree	-	27	91.35	Progression rate
Ouwerkerk, J., <i>et al.</i> , 2023 [15]	Genetic data	RNN	21000	486	85.2	Prediction
Alfonso Perez, G., <i>et al.</i> 2022 [16]	DNA CpG methylation data	NN	237	33	92	Classification
Proposed Model 2025	NCV and Genetic data	<i>SVM</i>	28	10	100	Detection

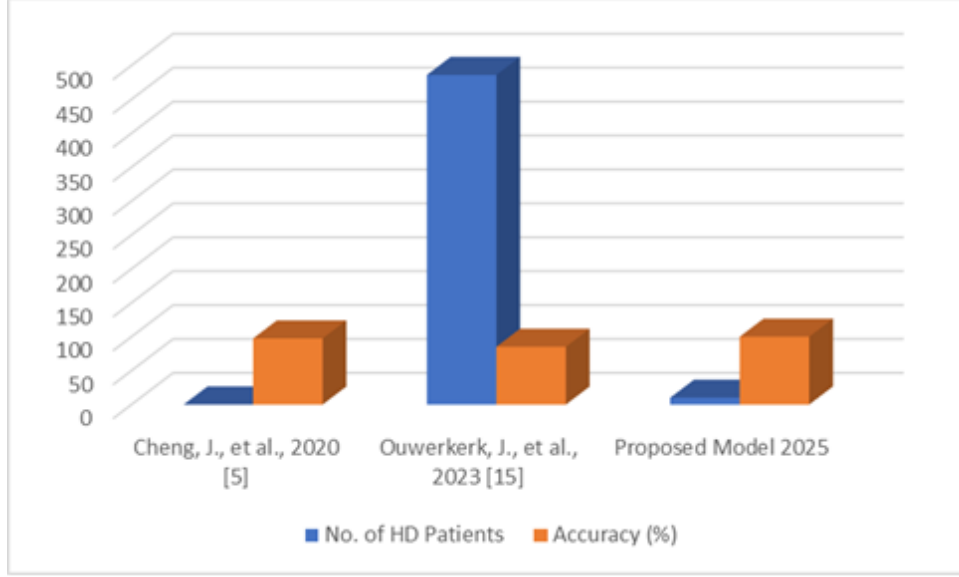


Figure 4.10: Graphical representation of the overall performance and comparison with the proposed model.

5 Discussion

Our findings demonstrate how genetic performance may be used to categorize people with Huntington’s disease. We showed that genetic information directly affects the classifiers’ performance. To achieve the highest accuracy, the *HD* selects which features to feed into the *SVM* algorithm. The best results were obtained for all classifiers. This study developed a novel representation strategy, using genetic data from *HD* patients. To aid early diagnosis, these data were then examined using a quadratic *SVM*-based machine learning technique. Notable categorization performance was achieved in this investigation. In particular, the classification accuracies of five distinct *SVM*-based models were 70%, 100%, 100%, 80%, and 60%, respectively, as shown in Table 4.2. The difficulty of differentiating between several neurodegenerative illnesses at once is the reason for the three *SVM*-based algorithms’ comparatively poorer classification results when compared to other binary classification tasks. On the other hand, a more straightforward assessment based on the existence or lack of a disease is advantageous for binary classification sub-problems. This study’s lack of comparable datasets makes it difficult to compare it to other studies. This distinction draws attention to an important contribution made to this research. Furthermore, the suggested method performs better than the best findings from previous research for additional sub-problems, as indicated in Table 4.3. Table 4.4 displays the total performance and comparison with the suggested model. These results imply that the suggested approach is not only generally effective but also has potential for use in treating additional neurological conditions. According to Table 4.2, the best predictor for *HD* in terms of training time and prediction speed is quadratic *SVM*. According to experimental findings, quadratic and cubic *SVM*s outperform other classification algorithms for *HD* diagnosis as well, because their underlying kernel function was better suited to capture the complex, non-linear patterns present in Huntingtons disease (*HD*). The cubic kernel implicitly maps the data into a higher-dimensional feature space, allowing for more sophisticated decision boundaries than simpler kernels like the linear one. The study’s key conclusions are that a genetic test will be necessary to identify substantial *HD*.

6 Conclusion and future work

In this work, we showed that a genetic test can be a biomarker for the early detection of *HD* with reasonable accuracy and feasibility. Specifically, the genetic test tasks were the most accurate in detecting the *HD*. For the analysis of small real *HD* data, machine learning is a viable method. ML models maintained a lower prediction speed and training time while achieving a greater performance with various *SVM*-based ML algorithms and other previous research. Lastly, ML was used to develop a model that would help clinicians give advice and a future trajectory regarding *HD* patients’ driving abilities. The findings also demonstrate that the quadratic *SVM*-based machine learning algorithm yields good sensitivity, specificity, and accuracy values. More *HD* participants and healthy controls that are available to patients and clinicians

will be improved in future studies to enable participation and adherence to the final clinical tool.

Authors' Contributions: Prof. Vijaya Nath Mishra participated in the study's design, data collection, and manuscript drafting. Data analysis, interpretation, and *NCV* were done by Prof. Abhishek Pathak. Mr. Arunesh Dutt, Ms. Komal Singh, and Dr. Pradeep Kumar Tiwari participated in the discussion of the findings and the interpretation of the data gathering. Prof. Ashish Khare participated in the study's design and result analysis, and Prof. Narendra Kumar Shukla oversaw the manuscript's composition.

Acknowledgement. No conflicts of interest have been disclosed by any of the authors. Every procedure used in research involving human subjects complied with the Institutional Ethics Committee's ethical guidelines (IERB ID: COM-D-21-001) under the *MoU* of the University of Allahabad and Banaras Hindu University, Varanasi (Uttar Pradesh), India.

References

- [1] S. Bhachawat, E. Shriram, K. Srinivasan, and Yuh-Chung Hu., Leveraging computational intelligence techniques for diagnosing degenerative nerve diseases: a comprehensive review, *Diagnostics*, **13**(2) (2023), 288.
- [2] M. Bennasar, Y. Hicks, S. Clinch, P. Jones, A. Rosser, M. Busse, and C. Holt. Huntington's disease assessment using tri axis accelerometers, *Procedia Computer Science*, **96** (2016), 1193-1201.
- [3] B. Bossaer, J., Danielle Eskens, and Austin Gardner, Sensitivity and specificity of drug interaction databases to detect interactions with recently approved oral antineoplastics, *Journal of Oncology Pharmacy Practice*, **28**(1) (2022), 82-86.
- [4] J. Collett, P. Esser, H. Khalil, M. Busse, L. Quinn, K. DeBono, A. Rosser, A. H. Nemeth, and H. Dawes, Insights into gait disorders: walking variability using phase plot analysis, *Huntington's disease, Gait & posture*, **40**(4) (2014), 694-700.
- [5] J. Cheng, Hsin-Ping Liu, Wei-Yong Lin, and Fuu-Jen, Tsai. "Identification of contributing genes of Huntingtons disease by machine learning, *BMC Medical Genomics*, **13**(1) (2020), 176.
- [6] B. Erda, . and E. Sümer, CNNBased Neurodegenerative Disease Classification Using QRRepresented Gait Data, *Brain and Behavior*, **14**(10) (2024), e70100.
- [7] R. Ghosh, and T. Sarah J., Clinical features of Huntingtons disease, *Polyglutamine disorders*, (2018), 1-28.
- [8] S. Ganesh, T. Chithambaram, N. R. Krishnan, D. R. Vincent, J. Kaliappan, and K. Srinivasan, Exploring Huntingtons disease diagnosis via artificial intelligence models: a comprehensive review, *Diagnostics*, **13**(23) (2023), 3592.
- [9] T. Kouba, W. Frank, T. Tykalova, A. Mühlbeck, J. Klemp, K. S. Lindenberg, G. B. Landwehrmeyer, and J. Rusz, Speech biomarkers in Huntington's disease: A crosssectional study in presymptomatic, prodromal and early manifest stages, *European journal of neurology*, **30**(5) (2023), 1262-1271.
- [10] Kiani, Lisa, Insights into the toxic effects of mutant huntingtin, *Nature Reviews Neurology*, **19** (10) (2023), 576-576.
- [11] A. Klein, Lori-Ann R. Sacrey, S. B. Dunnett, I. Q. Wishaw, and G. Nikkhah., Proximal movements compensate for distal forelimb movement impairments in a reach-to-eat task in Huntington's disease: new insights into motor impairments in a real-world skill, *Neurobiology of disease*, **41**(2) (2011), 560-569.
- [12] McColgan, Peter, and S. J. Tabrizi, Huntington's disease: a clinical review, *European journal of neurology*, **25** (1) (2018), 24-34.
- [13] A. Miranda, R. Lavrador, F. Jlio, C. Janurio, M. Castelo-Branco, and G. Caetano, Classification of Huntingtons disease stage with support vector machines: A study on oculomotor performance, *Behavior research methods*, **48**(4) (2016), 1667-1677.
- [14] S. Maddury, Automated Huntingtons Disease Prognosis via Biomedical Signals and Shallow Machine Learning, *ArXiv*, (2023), arXiv-2302.
- [15] J. Ouwerkerk, S. Feleus, K. F. van der Zwaan, Y. Li, M. Roos, W. MC van Roon-Mom, S. T. de Bot, K. J. Wolstencroft, and E. Mina, Machine learning in Huntingtons disease: exploring the Enroll-*HD* dataset for prognosis and driving capability prediction, *Orphanet Journal of Rare Diseases*, **18**(1) (2023), 218.
- [16] A. Perez, Gerardo, and J. C. Villarraso, Neural network aided detection of Huntington disease, *Journal of Clinical Medicine*, **11**(8) (2022), 2110.
- [17] J. Rusz, and T. Tykalova, Reader response: Motor speech patterns in Huntington disease, *Neurology*, **95**(13) (2020), 607-608.

- [18] D. M. Ruopp, N. J. Perkins, B. W. Whitcomb, and E. F. Schisterman, Youden Index and optimal cutpoint estimated from observations affected by a lower limit of detection, *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, **50**(3) (2008), 419-430.
- [19] L. J. Schultz, M. Neema, and P. C. Nopoulos, Unravelling the role of huntingtin: from neurodevelopment to neurodegeneration, *Brain*, **146**(11) (2023), 4408-4410.
- [20] O. F. Walker, Huntington's disease, *The Lancet*, **369** (9557) (2007), 218-228.
- [21] K. D. Wilton, M. K., M. D. Heller, F. W. Gergits, C. R. Willing, J. B. Fahey, A. Frouin, Microglia and complement mediate early corticostriatal synapse loss and cognitive dysfunction in Huntingtons disease, *Nature medicine*, **29**(11) (2023), 2866-2884.
- [22] W. D. Weir, A. Sturrock, and B. R. Leavitt, Development of biomarkers for Huntington's disease, *The Lancet Neurology*, **10**(6) (2011), 573-590.
- [23] P. Zeun, R. I. Scahill, S. J. Tabrizi, and E. J. Wild, Fluid and imaging biomarkers for Huntington's disease, *Molecular and Cellular Neuroscience*, **97** (2019), 67-80.
- [24] W. Zhu, N. Zeng, and N. Wang, Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations, In *NESUG Proc.: Health Care and Life Sciences*, **19** (2010), 67.