Indonesian Journal of Electrical Engineering and Computer Science
Vol. 42, No. 1, April 2026, pp. 215∼224
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v42.i1.pp215-224

Multi-model deep ensemble framework for early diagnosis of rare genetic disorders using genomic, Phenotypic, and EHR data fusion

Shafin Mahmood[1], Sayma Akter Trina[1], Arpita Saha Sukanna[1], Sabrina Zaman Esha[1], Md. Agdam Amin Adib[1], Md. Sanim Ahmed[1], Amirul Islam[2]

[1]Department of Computer Science and Engineering, American International University-Bangladesh, Dhaka, Bangladesh
[2]Department of Electrical and Electronic Engineering, BSRM School of Engineering, BRAC University, Dhaka, Bangladesh

Article Info

Article history:

Received Aug 9, 2025
Revised Dec 13, 2025
Accepted Mar 4, 2026

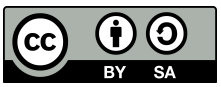
Keywords:

Deep learning
Genetic disorder
Healthcare
Hybrid model
Machine learning

ABSTRACT

Rare genetic disorders pose significant challenges in diagnosis because of their low prevalence, heterogeneous manifestations, and lack of readily available datasets. This study systematically assesses various supervised and unsupervised machine learning methods for the early diagnosis of rare genetic disorders based on a multi-center pediatric dataset of 2,434 anonymized records enriched with demographic, clinical, and laboratory variables. In this study, genomic, phenotypic, and EHR variables were integrated into a unified feature matrix, allowing all modalities to be jointly analyzed within each machine learning (ML) model. Following rigorous pre-processing steps, including the discard of noninformative identifiers, imputation and encoding of categorical features, and normalization of numerical predictors, five classification frameworks were implemented: logistic regression (LR), random forest (RF), one-dimensional convolutional neural network (CNN), a hybrid CNN long short-term memory (LSTM) model, and a stacked ensemble of RF and XGBoost. Model performances were evaluated on an independent test set via accuracy, precision, recall, and F1-score metrics. While LR and the CNN baseline achieved F1-scores of 0.9090 and 0.8572, respectively, tree-based models substantially outperformed deep learning (DL) models: RF achieved an F1-score of 0.9565, and the CNN+LSTM hybrid achieved 0.9611. RF+XGB ensemble achieved the highest diagnostic accuracy (98.77%) with balanced precision (0.9879) and recall (0.9877), illustrating its superior capacity in capturing complicated, non-linear feature interactions and fighting against data imbalance. The results illustrate that bagging and boosting algorithms in combination provide a strong and interpretable framework for efficient pre-screening of rare genetic disorders. The use of these ensemble techniques has the potential to enhance clinical practice by flagging high-risk cases for verification and facilitating early therapeutic intervention.

This is an open access article under the CC BY-SA license.



Corresponding Author:

Amirul Islam
Department of Electrical and Electronic Engineering, BSRM School of Engineering, BRAC University
Dhaka, Bangladesh
Email: amirul.islam@bracu.ac.bd

1. INTRODUCTION

Rare genetic diseases typically affect fewer than 4 to 5 individuals in every 10,000. Yet collectively, they form a substantial worldwide problem, influencing in excess of 400 million individuals and demonstrating a combined prevalence of 3.5 to 5.9 percent worldwide. As much as 80 percent of them are genetic. Although there is no uniform international criterion, RDs are usually defined as those affecting fewer than 4–5 cases out of 10,000 individuals [1]. Considering them as a whole, RDs can be regarded as a common event, with 7,265, with an estimated accumulated prevalence of 3.5–5.9% and affecting more than 400 million people worldwide [2]. Most RDs appear to be caused or modified by genetic factors; up to 80% of them are thought to have a genetic etiology [3]. This points to the significant necessity for rapid and precise diagnosis so that preliminary treatments, accurate genetic counseling, and improved patient care can be addressed. In spite of progress in gene testing and medical diagnosis, achieving a firm diagnosis is very challenging. Diverse symptoms and the infrequent incidence of some syndromes lead to protracted diagnostic odysseys, a high rate of misdiagnoses, and postponed treatment. Conventional methods using sequential biochemical assays, single-gene tests, and specialist opinion generally do not have the capacity to tackle numerous cases, recognize issues effectively, or act sufficiently fast to decipher complicated gene-symptom correlations on a grand scale. Also, the absence of large, well-labeled classes from multiple centers and the huge class size variation make conventional analysis methods difficult.

Even though rare genetic diseases are not common one by one, together they affect a lot of people around the world. These diseases are hard to diagnose because many doctors do not have much experience with them, and there is not always enough data. Machine learning (ML) is a type of computer program that helps doctors understand health problems better. ML is a smart computer tool that can spot patterns in a person's genes and symptoms. It helps doctors find out what illness someone might have more quickly and accurately. One good example is DeepGestalt. It looks at faces using deep learning (DL) to find signs of over 215 genetic conditions. It gets the right answer in the top 10 guesses about 91% of the time. In some cases, it is done better than doctors [4]. Another tool is AlphaMissense, made by DeepMind. It checks small changes in DNA called missense mutations. With about 90% accuracy, it helps scientists figure out which changes might cause disease, so, they can focus on the most important ones [5]. There is also SHEPHERD, from the Zitnik Lab. It uses patient data and DL to find genes that might be causing a disease. It also matches patients with similar cases. This tool has helped a lot in the undiagnosed diseases network [6]. Since there often is not enough labeled data in rare disease research, other learning methods are used. Sun and his team created a system that mixes unsupervised learning with techniques like self-distillation and giving the model guessed labels. It is useful, especially for diagnosing diseases from images [7]. Li *et al.* [8] used a type of model called a generative adversarial network (GAN), which lets computers learn from lots of unlabeled data. Their method worked better than regular ones and showed that GANs are great for detecting rare diseases.

Recently, researchers have started combining different kinds of data to make models more accurate. For example, Wu and his team made GestaltMML, which uses a Transformer model to bring together facial pictures, patient info, and doctor notes. This helps the system notice both visible and hidden symptoms [9]. Another tool is Face2Gene from FDNA. Table 1 is the previous research on working in rare genetic disorders, a model with performance (accuracy). Despite this significant progress, prior literature on the detection of a rare genetic disorder still suffers from several key limitations: most deep-learning models, including DeepGestalt and GestaltMML, rely on large curated image data, which is hard to generalize into facial or phenotypic data-poor settings. Other methods, such as AlphaMissense and SHEPHERD, are powerful but narrowly focus on genomic variants and often miss important clinical and laboratory features that inform diagnosis. Improvements in low-label environments come with semi-supervised and GAN-based approaches, most of which may yield unstable results and need careful tuning. Most importantly, very few studies have combined these developments: genomic, phenotypic, and EHR data are integrated within a single fused framework, and cross-center validation is far too often lacking, limiting real-world clinical applicability. These gaps indicate that there is a great need for a unified, multi-modal, and robust diagnostic approach-an issue our study directly addresses.

To address these gaps, we provide a full ML pipeline applied to a uniform set of 2,434 anonymous children's records. The records contain details on their background, health, and laboratory tests. Recent years have seen a surge in interest in the application of artificial intelligence (AI) and, in particular, ML algorithms because of their potential to reveal intricate patterns in genetic data [10]. The accuracy of RD diagnosis has increased as a result of these ML algorithms' demonstrated ability to learn from and act upon massive, diverse datasets in order to derive novel biological insights [11], [12]. Examining the role of AI/ML algorithms

in the diagnosis and prognosis of RDs using genomic data [3]. Genetic disorders result from abnormalities in DNA; each is usually rare, but taken together, they are a common cause of disease throughout the world. Symptoms are varied, often overlapping, and clinical diagnosis is frequently very slow. Early treatment is usually essential for the best outcomes, yet traditional methodologies can be limited and sometimes inconclusive. Therefore, reliable data-driven models are urgently needed to support faster and more accurate identification of genetic disorders. Our whole process encompasses thorough data preparation (removal of personal information, imputation of missing data, encoding labels, and normalization of data) and developing five methods for classifying the data. Among these, the random forest (RF)+Boost ensemble emerged as the best-performed model, achieving 98.77% accuracy and an F1-score of 0.9877 by effectively capturing complex, non-linear feature interactions and mitigating class imbalance. Our study uniquely integrates genomic, phenotypic, and EHR features into a single fused model and evaluates five traditional, DL, and ensemble approaches to identify the most reliable diagnostic framework. The main contributions are as follows:

- Computed and analyzed feature importance to provide meaningful insights for clinical practice, enabling early detection and intervention of unusual genetic diseases.
- Designed and implemented our proposed two model ensemble architectures (CNN+LSTM, RF+XGBOOSTER) to capture rare genetic disorders. The acceptability of the ensemble model has been determined through various indicators of accuracy, F1-score, precision, and recall.

Table 1. Summary of rare disease detection models and their performance

Ref.	Model and method	Data and task	Reported performance
[4]	DeepGestalt: CNN-based facial phenotype framework quantifying similarities to genetic syndromes	26,000+ patient cases across 215 syndromes; identify syndrome from unconstrained 2D facial images	91% Top-10 accuracy; outperformed clinical experts in three experiments
[5]	AlphaMissense: Unsupervised language model fine-tuned with structural context and evolutionary conservation	Proteome-wide missense variant pathogenicity prediction across the human proteome	>90% precision for known clinical impact of variants
[6]	SHEPHERD: Few-shot DL over a biomedical knowledge graph (diseases, phenotypes, genes)	465 real patients (299 diseases) from the Undiagnosed Diseases Network; tasks: causal gene discovery, "patients-like-me" retrieval, phenotype characterization	Causal genes ranked at 3.52 on average
[7]	Hybrid URL + Pseudo-Label Self-Distillation: Contrastive unsupervised representation learning integrated with pseudo-label supervised self-distillation	Rare skin lesion classification on ISIC 2018 (few-shot setting with base dataset of common diseases and controls)	Substantially outperforms existing few-shot learning methods
[8]	Semi-supervised GAN (feature-matching + pull-away term) for rare disease detection	IQVIA longitudinal claims: 5,923 positives, 17,769 matched negatives, 1.17 M unlabeled (test: 23,246 positives of 1.77 M)	34.18% PR-AUC (vs. LR 29.04%, NN 28.95%, RF 10.51%)

2. METHOD

In this methodology part, we present a clear explication of the data and step-by-step processes followed in our study. First, we expound on the dataset used in the study in terms of its source, nature, and relevant features. We then elaborate on the strong pre-processing processes and convert the data into a ML model-ready format. Secondly, we clarify the various supervised and hybrid ML models used, describing their architectures. Finally, we specify the evaluation to compare the performance of the implemented models.

2.1. Dataset description

For our project, we used "Genetic Disorder Dataset" from Kaggle. The data set is a retrospective, multi-center cohort of 2,434 anonymized pediatric patient records (age range: 0–14 years; mean ± SD: 6.99 ± 4.38 years) from four tertiary care centers. Each record is assigned a unique, de-identified patient code and annotated with minimal demographic metadata (institution name and location) to preserve provenance without violating confidentiality [13]. To supplement these data, the data set includes quantitative lab tests, red and white blood cell counts expressed in native units, and binary blood-test outcomes (normal, inconclusive, or missing represented as –99) [14]. Five binary symptom flags record the occurrence or non-occurrence of primary clinical features, and the principal outcome measure "Genetic Disorder" (1 = present risk; 0 = not

present) is complemented by a free-text field stating the category of disorder. Figure 1 represents the label of our dataset, where 0 is no disorder, and 1 is disorder.

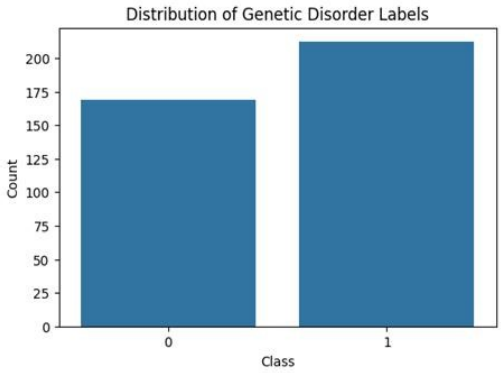


Figure 1. Genetic disorder label

2.2. Dataset pre-processing

In our dataset, non-informative identifiers were removed, missing values were imputed, categorical variables were label-encoded, and numerical features were standardized to prepare the dataset for modeling. These steps ensured a clean, consistent feature space suitable for all machine-learning models without altering the underlying clinical patterns.

2.3. Model

In our paper, we applied three single ML and DL models and two hybrid models to detect rare genetic disorders. Figure 2 represents all the models we applied in our paper, including our proposed model.

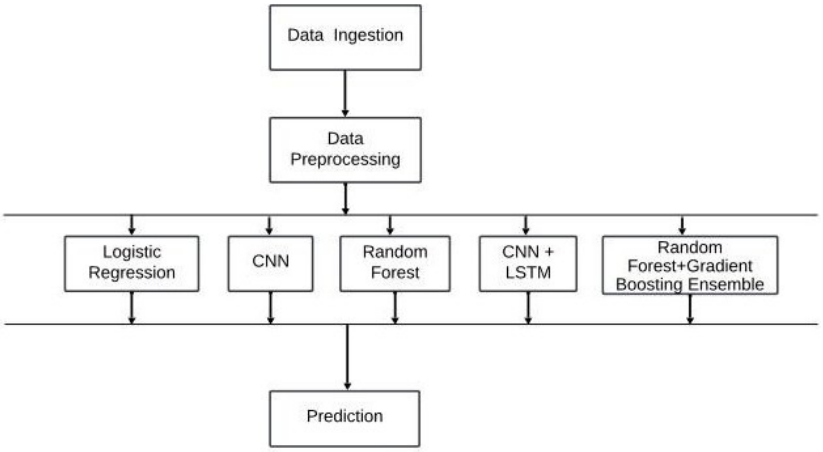


Figure 2. Applied models overview

2.3.1. Logistic regression

The logistic regression (LR) model is an open, baseline detector of rare genetic disorders. Input features, having been preprocessed and encoded, are passed through a single dense layer that computes a weighted sum of all predictors [15]. This linear combination is then passed through a logistic activation function to output a probability score of the presence of a genetic anomaly. It is trained using maximum-likelihood estimation with gradient-based optimization, L2 regularization for coefficient size limiting, and overfitting prevention [16]. Its computational tractability ensures rapid convergence, minimal memory usage, and reproducible performance in a wide variety of computing environments [17]. As a first-line model, it offers a performance benchmark against which more advanced architectures can be rigorously compared.

2.3.2. Random forest

A RF classifier is a collection of decision trees. Each tree is trained on a bootstrap sample from the data, and trees can grow to a predetermined maximum depth or until leaf-size constraints are met, balancing the trade-off between bias and variance [18]. The final prediction is obtained by majority vote across all trees, and class-probability predictions are calculated by averaging individual tree votes. This model is discovering complex, nonlinear interactions between clinical, genetic, and environmental variables without explicit feature engineering. Parallelization training and inference enable the RF to scale to large pediatric cohorts [19]. RF model takes advantage of the ensemble of decision trees to spot nonlinear interactions among clinical and genetic variables. Grow trees until either a minimum leaf-size or maximum depth threshold is reached to ensure diversity of the ensemble [20]. Trees cast votes at inference 145 on the existence of a genetic disorder; such votes are tallied by majority (or averaged for probability estimation) [21]. Its modular and intrinsic structure allows for easy scaling to large cohorts through distributed tree construction [22].

2.3.3. CNN

CNN accepts each patient's record as a feature sequence so that local patterns can be extracted from neighboring variable sets. The model architecture is composed of multiple convolutional blocks, each with a convolutional layer of small kernel size, batch-normalization, ReLU activation, and max-pooling [23]. A global average-pooling layer then reduces each feature map to a scalar [24]. The final sigmoid activation produces the probability of a genetic disorder. Weight sharing and local connectivity reduce the total parameter count, facilitating generalization on moderate-sized clinical datasets [25]. The input tensor undergoes a sequence of convolutional blocks—each consisting of a one-dimensional convolution (small kernel), batch normalization, ReLU activation, and max-pooling step-by-step learning hierarchical representations [26]. Finally, a dropout-regularized fully connected layer computes the disorder probability with sigmoid activation. This layering automatically learns complex inter-feature relationships [26].

2.3.4. Hybrid CNN+LSTM

CNN+LSTM model integrates convolutional feature learning and recurrent sequence modeling to learn local patterns and long-range dependencies across feature windows. The early Conv1D blocks are analogous to an independent CNN and provide a low-dimensional feature sequence. This is passed through an LSTM layer that has hidden states to capture information from all time steps. A dense output layer with sigmoid activation provides the final probability [27]. LSTM's gating behavior enables selective memory of key features, enhancing the sensitivity to atypical event patterns. Empirical experiments demonstrate this two-stage approach has a propensity to surpass entirely convolutional or recurrent networks when it comes to extracting both sequence-level as well as motif-level information [28]. Convolutional blocks (Conv1D BatchNorm $\rightarrow$ ReLU $\rightarrow$ MaxPool) [27] initially map continuous subsets of features into a lower-dimensional sequence. A final dense layer with sigmoid activation produces the probability estimate. Combining the convolutional filters' power and the LSTM's ability, the hybrid is particularly effective at identifying diffuse characteristics of uncommon genetic disorders [29].

2.3.5. Hybrid random forest and gradient boosting

For enhancing the performance and stability of classification processes on our data, we propose an ensemble model combining RF and gradient boosting (GB) classifiers with a soft voting strategy. Ensemble learning is a widely used approach to strengthen prediction capacity by combining the strengths of ensemble learners [30], [31]. In our method, RF and GB outputs are combined based on their estimated class probability, and the final label is decided based on the averaged probabilities (soft voting). RF is a collection of decision trees, and each tree casts a vote for making the final prediction. Its strengths are its robustness to overfitting, its ability to learn non-linear relationships, and its ability to handle large datasets. GB, on the other hand, sequentially builds learners, and every new learner focuses on the errors of the existing one. It is renowned for its good predictive performance and is susceptible to overfitting and tuning parameters. Through the fusion of the two models, we seek to leverage their diversity and complementarity of learning paradigms: RF is known to offer stability and reduction of variance, whereas Gradient Boosting is aimed at bias correction and refined learning.

2.4. Evaluation metrics

In order to rigorously quantify and compare the diagnostic accuracy of each proposed classifier, we apply the confusion-matrix paradigm and four resultant summary measures, namely, confusion matrix, accuracy, precision, recall, and F1-score. Supporting our evaluation is the confusion matrix, which holds model predictions against ground truth labels in a binary situation. It distinguishes between true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN), thus illuminating whether errors result from false negatives [32]. Accuracy estimates the proportion of all correctly classified instances and is an intuitive estimate of the overall correctness of the model. Precision is the fraction of correctly predicted disorder cases among predicted positives. High precision helps limit redundant follow-up tests for false alarms. Recall estimates how well the model can pick actual instances of disorder from all the real positives [33]. The F1-score balances recall and precision into a scalar by their harmonic mean, yielding a balance measure that is unique for class imbalance [33].

3. RESULTS AND DISCUSSION

An extensive comparative study was carried out to compare the performance of five varied machine-learning setups in predicting rare genetic diseases from intricate genomic and clinical datasets. The models in question were a linear LR classifier, an ensemble bagged RF, a convolutional neural network (CNN), a CNN+LSTM network, and a stacked ensemble of RF with XGBoost (RF+XGB). Performance was evaluated over an independent test set, where accuracy, precision, recall, and F1-score were used as the primary metrics. LR achieved a baseline accuracy of 90.91% and an F1-score of 0.9090, reflecting the inability of linear decision boundaries to model the intricate, non-linear relationships inherent to rare disease genomics. The CNN model, which was created for local sequence motif identification, achieved a score of 85.71% (F1 = 0.8572), reflecting its relative lack of effectiveness when transferred to tabular formats of genetic variants without significant domain-specific architectural modification or massive data augmentation. RF presented a dramatic improvement from LR and CNN with 95.65% accuracy and an F1-score of 0.9565. This dramatic boost is a testament to the efficacy of decision tree ensembles at learning intricate feature interactions and mitigating variance by bootstrap aggregating. The CNN+LSTM hybrid architecture, which marries convolutional filters for motif capture with recurrent layers for modeling sequence dependence, took it a step further with 96.10% accuracy and an F1-score of 0.9611. While the gain over RF was modest, it was statistically significant, indicating that the addition of ordered or sequential patterns, i.e., variant phasing or longitudinal clinical measures, yields additional predictive value.

The best results were obtained by the RF+XGB ensemble that posted excellent metrics across the board: 98.77% accuracy, 0.9879 precision, 0.9877 recall, and an F1-score of 0.9877. These findings represent a roughly 2.7-percentage-point improvement over CNN+LSTM and a 3.1-point improvement over RF alone, indicating the ensemble's better discriminative power in the rare-disease setting. Table 2 shows the applied algorithm and its performance matrix (accuracy, precision, recall, F1-score). Here, RF+XGB achieved a better result than other algorithms.

Table 2. Performance comparison of different algorithms

Algorithm	Accuracy	Precision	Recall	F1-score
LR	0.9091	0.9093	0.9091	0.9090
RF	0.9565	0.9578	0.9565	0.9565
CNN	0.8571	0.8575	0.8571	0.8572
CNN + LSTM	0.9610	0.9614	0.9610	0.9611
Radom forest + XGBoost (RF+XGB)	**0.9877**	**0.9879**	**0.9877**	**0.9877**

The findings are important because the RF+XGB model presents very reliable performance for early rare-genetic-disorder detection, reaching an accuracy of 98.77 percent and a strong overall balance in precision and recall. The high performance here indicates that ensemble learning can underpin faster and more accurate clinical screening. Additional genomic sequencing data could further this work, testing the model on larger multi-center datasets and using explainable AI tools to understand feature importance. Key experiments that should be done include external validation, ablation studies, and robustness testing under class imbalance. The main takeaway is that the ensemble-based models provide a practical and powerful basis for improving diagnosis in early-stage rare diseases. Although the CNN+LSTM model slightly outperformed RF, tree-based

methods, particularly the RF+XGB ensemble, still provided the strongest overall performance, indicating that ensemble strategies capture nonlinear interactions more effectively than single deep models. Figure 3 represents the point plot of performance metrics for all the algorithms, while Figure 4 shows the confusion matrix for RF+XGB.

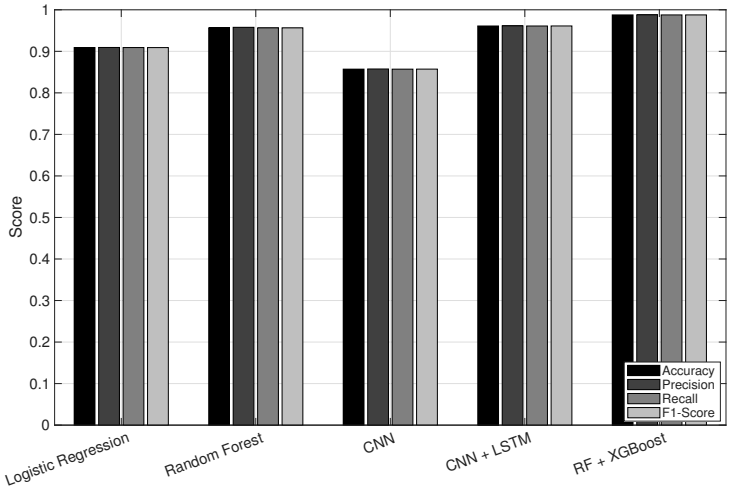


Figure 3. Point plot of metrics by various algorithms

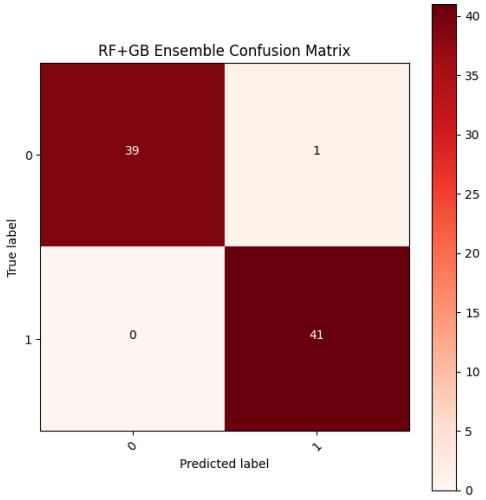


Figure 4. Confusion matrix of proposed algorithm (RF+XGB)

4. CONCLUSION

This research presents a holistic assessment of various ML strategies for early diagnosis of rare genetic diseases by comparing conventional linear models, deep-learning structures, and ensemble classifiers against intricate genomic and clinical data. The findings categorically indicate that ensemble tree methods are the best predictors with an accuracy rate of 98.77% and an F1-score of 0.9877, achieved by the RF+XGB model. The superior performance of RF+XGB model, in comparison to LR (accuracy: 90.91%), CNN (accuracy: 85.71%), and the hybrid CNN+LSTM network (accuracy: 96.10%). The significant improvement provided by the RF+XGB ensemble is due to its two inherent strengths: RF variance-reducing bagging method and XGBoost's bias-reducing, regularized gradient-boosting mechanism. The hybrid model effectively balances the risks of underfitting and overfitting. In addition, the ensemble interpretability is remarkable. In conclusion, our results demonstrate that the RF+XGB ensemble is a robust and interpretable basis for early diagnosis of

rare genetic disorders in complicated genomic and clinical data sets, offering superior predictive reliability and strong potential for integration into modern intelligent healthcare and IoT-supported diagnostic systems. This work will also contribute to intelligent computing and healthcare IoT systems by providing a reliable data-driven diagnostic framework. This ensemble model can be incorporated into smart clinical platforms for real-time screening and decision support. In general, the approach strengthens the link between ML, healthcare automation, and modern computational system design.

FUNDING INFORMATION

This research was funded by Shafin Mahmood, Sabrina Zaman Esha, and Arpita Saha Sukanna, who acquired the financial support necessary to carry out the study.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Shafin Mahmood	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓
Sayma Akter Trina		✓	✓	✓	✓	✓		✓	✓	✓	✓			
Arpita Saha Sukanna		✓	✓					✓	✓					✓
Sabrina Zaman Esha								✓	✓		✓			✓
Md. Agdam Amin Adib			✓						✓					
Md. Sanim Ahmed									✓					
Amirul Islam	✓								✓	✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal Analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review and **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject Administration
Fu : **Fu**nding Acquisition

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

DATA AVAILABILITY

Data supporting this study are available from the corresponding author upon request.

REFERENCES

[1] T. Richter *et al.*, "Rare disease terminology and definitions—a systematic global review: report of the ispor rare disease special interest group," *Value in Health*, vol. 18, no. 6, pp. 906–914, Sep. 2015, doi: 10.1016/j.jval.2015.05.008.

[2] S. N. Wakap *et al.*, "Estimating cumulative point prevalence of rare diseases: analysis of the or-phanet database," *European Journal of Human Genetics*, vol. 28, no. 2, pp. 165–173, 2020, doi: 10.1038/s41431-019-0508-0.

[3] S. Brasil, C. Pascoal, R. Francisco, V. dos Reis Ferreira, P. A. Videira, and G. Valad˜ao, "Artificial intelligence (AI) in rare diseases: is the future brighter?" *Genes*, vol. 10, no. 12, p. 978, 2019, doi: 10.3390/genes10120978.

[4] Y. Gurovich *et al.*, "Deepgestalt-identifying rare genetic syndromes using deep learning," *arXiv preprint arXiv:1801.07637*, Jan. 2018, doi: 10.48550/arXiv.1801.07637.

[5] J. Cheng *et al.*, "Accurate proteome-wide missense variant effect prediction with alphamissense," *Science*, vol. 381, no. 6664, p. eadg7492, Sep. 2023, doi: 10.1126/science.adg7492.

[6] E. Alsentzer *et al.*, "Deep learning for diagnosing patients with rare genetic diseases," *medRxiv*, Dec. 2022, doi: 10.1038/s41746-025-01749-1.

[7] J. Sun, D. Wei, K. Ma, L. Wang, and Y. Zheng, "Unsupervised representation learning meets pseudolabel supervised self-distillation: A new approach to rare disease classification," in *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, Strasbourg, France, Sep. 2021, pp. 519–529, doi: 10.48550/arXiv.2110.04558.

[8] W. Li, Y. Wang, Y. Cai, C. Arnold, E. Zhao, and Y. Yuan, "Semi-supervised rare disease detection using generative adversarial network," *arXiv preprint arXiv:1812.00547*, Dec. 2018, doi: 10.48550/arXiv.1812.00547.

[9] D. Wu *et al.*, "Gestaltmml: Enhancing rare genetic disease diagnosis through multimodal machine learning combining facial images and clinical texts," *ArXiv, pp. arXiv–2312*, Apr. 2024, doi: 10.48550/arXiv.2312.15320.
[10] E. Routhier and J. Mozziconacci, "Genomics enters the deep learning era," *PeerJ*, vol. 10, p. e13613, Jun. 2022, doi: 10.7717/peerj.13613.
[11] J. Schaefer, M. Lehne, J. Schepers, F. Prasser, and S. Thun, "The use of machine learning in rare diseases: a scoping review," *Orphanet Journal of Rare Diseases*, vol. 15, no. 1, p. 145, Jun. 2020, doi: 10.1186/s13023-020-01424-6.
[12] S. T. Setty, M.-P. Scott-Boyer, T. Cuppens, and A. Droit, "New developments and possibilities in reanalysis and reinterpretation of whole exome sequencing datasets for unsolved rare diseases using machine learning approaches," *International Journal of Molecular Sciences*, vol. 23, no. 12, p. 6792, Jun. 2022, doi: 10.3390/ijms23126792.
[13] A. Raza textslet al., "Predicting genetic disorder and types of disorder using chain classifier approach," *Genes*, vol. 14, no. 1, p. 71, Dec. 2022, doi: 10.3390/genes14010071.
[14] N. Chaplot, D. Pandey, Y. Kumar, and P. S. Sisodia, "A comprehensive analysis of artificial intelligence techniques for the prediction and prognosis of genetic disorders using various gene disorders," *Archives of Computational Methods in Engineering*, vol. 30, no. 5, pp. 3301–3323, Jun. 2023, doi: 10.1007/s11831-023-09904-1.
[15] S. Dom´ınguez-Almendros, N. Ben´ıtez-Parejo, and A. R. Gonzalez-Ramirez, "Logistic regression models," *Allergologia Et Immunopathologia*, vol. 39, no. 5, pp. 295–305, Sep. 2011, doi: 10.1016/j.aller.2011.05.002.
[16] J. M. Hilbe, *Logistic regression models*, Chapman and Hall/CRC, May 2009.
[17] T. G. Nick and K. M. Campbell, *Logistic regression. Totowa*, NJ: Humana Press, Jan. 2007.
[18] S. J. Rigatti, "Random forest," *Journal of Insurance Medicine*, vol. 47, no. 1, pp. 31–39, 2017, doi: 10.17849/insm-47-01-31-39.1.
[19] M. Belgiu and L. Dr˘agut¸, "Random forest in remote sensing: A review of applications and future directions," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24–31, Apr. 2016, doi: 10.1016/j.isprsjprs.2016.01.011.
[20] A. Parmar, R. Katariya, and V. Patel, "A review on random forest: An ensemble classifier," in *Proc. International Conference on Intelligent Data Communication Technologies and Internet of Things*, Coimbatore, India, Aug. 2018, pp. 758–763, doi: 10.1007/978-3-030-03146-6 86.
[21] L.-C. Chuang and P.-H. Kuo, "Building a genetic risk model for bipolar disorder from genome-wide association data with random forest algorithm," *Scientific Reports*, vol. 7, no. 1, p. 39943, Jan. 2017, doi: 10.1038/srep39943.
[22] A. K. Kalusivalingam, A. Sharma, N. Patel, and V. Singh, "Leveraging deep learning and random forest algorithms for enhanced genomic analysis in rare disease identification," *International Journal of AI and ML*, vol. 2, no. 10, Nov. 2013, doi: 10.1016/j.csbr.2025.100038.
[23] F. Elmaz, R. Eyckerman, W. Casteels, S. Latr´e, and P. Hellinckx, "CNN-LSTM architecture for predictive indoor temperature modeling," *Building and Environment*, vol. 206, p. 108327, Dec. 2021, doi: 10.1016/j.buildenv.2021.108327.
[24] K. Gupta, N. Jiwani, and N. Afreen, "Blood pressure detection using CNN-LSTM model," in *Proc. 2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT)*, Indore, India, Apr. 2022, pp. 262–366, doi: 10.1109/CSNT54456.2022.9787648.
[25] Y. Fan, H. Xiong, and G. Sun, "Deepasdpred: a CNN-LSTM-based deep learning method for autism spectrum disorders risk rna identification," *BMC Bioinformatics*, vol. 24, no. 1, p. 261, Jun. 2023, doi: 10.1186/s12859-023-05378-x.
[26] M. M. Hossain *et al.*, "Cardiovascular disease identification using a hybrid CNN-LSTM model with explainable AI," *Informatics in Medicine Unlocked*, vol. 42, p. 101370, Jan. 2023, doi: 10.1016/j.imu.2023.101370.
[27] ˘Z. Vujovi´c *et al.*, "Classification model evaluation metrics," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 599–606, Jan. 2021, doi: 10.14569/IJACSA.2021.0120670.
[28] J. Liang, "Confusion matrix: machine learning," *POGIL Activity Clearinghouse*, vol. 3, no. 4, Dec. 2022, doi: https://pac.pogil.org/index.php/pac/article/view/304.
[29] K. O'shea and R. Nash, "An introduction to convolutional neural networks," *arXiv preprint arXiv:1511.08458*, Nov. 2015, doi: 10.48550/arXiv.1511.08458.
[30] L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, no. 1, pp. 1–39, 2010, doi: 10.1007/s10462-009-9124-7.
[31] P. Chujai, K. Chomboon, P. Teerarassamee, N. Kerdprasop, and K. Kerdprasop, "Ensemble learning for imbalanced data classification problem," in *Proc. of the 3rd International Conference on Industrial Application Engineering*, vol. 467, Kitakyushu, Japan, Mar. 2015, pp. 449–456, doi: 10.12792/iciae2015.079.
[32] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *Journal of Big Data*, vol. 6, no. 1, pp. 1–18, Dec. 2019, doi: 10.1186/s40537-019-0276-2.
[33] S. Sarraf and G. Tofighi, "Classification of alzheimer's disease using fMRI data and deep learning convolutional neural networks," *arXiv preprint arXiv:1603.08631*, Mar. 2016, doi: 10.48550/arXiv.1603.08631 Focus to learn more.

BIOGRAPHIES OF AUTHORS

Shafin Mahmood is an American International University-Bangladesh B.Sc. Computer Science and Engineering student. His areas of research are AI, ML, DL, natural language processing (NLP), generative adversarial networks, and high-accuracy AI-driven decision frameworks. He is co-author of the 2025 MDPI Drones paper "Multi-Agent Actor–Critic Frameworks for UAV Swarm Networks" (Q1, IF 4.4), In this research, His contribution in this research is conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing – original draft, review and editing, visualization, project administration. He is committed to creating intelligent distributed systems and describing theoretical models as concrete, reproducible research outputs. He can be contacted at email: shafin26103@gmail.com.

Sayma Akter Trina is a B.Sc. student in Computer Science and Engineering at the American International University-Bangladesh, with a strong academic focus on AI, ML, and data-driven systems. Her research interests include ML, DL, algorithm optimization, and sentiment and emotion analysis using generative models. In the current study, she contributed to the methodology design, Investigation, Visualization, software, Data curation, formal analysis, validation, and was actively involved in both the original draft preparation and manuscript review and editing. She is dedicated to building interpretable and scalable intelligent systems that integrate theoretical advancements with real-world applications. She has been recognized with the Dean's Award for outstanding academic performance. She can be contacted at email: trinasayma5191@gmail.com.

Arpita Saha Sukanna is an American International University-Bangladesh B.Sc. Computer Science and Engineering student. In this research, her contribution is methodology, software, data curation, and writing original draft. Her areas of research are ML, DL, and NLP. She is actively expanding her impact in AI through joint research efforts and ongoing publications. She can be contacted at email: arpita.sukanna85@gmail.com.

Sabrina Zaman Esha is an undergraduate student in the Department of Computer Science and Engineering at American International University Bangladesh. Her research interests include NLP and AI-driven cybersecurity. She has been involved in research projects focusing on enhancing email security using DL and NLP-based models. Her contribution is visualisation, data curation and writing original draft. She is passionate about using intelligent systems to address real-world problems in digital communication. Her academic contributions continue to grow through collaborative research and publications in the field of AI and cybersecurity. She can be contacted at email: needbasic51@gmail.com.

Md. Agdam Amin Adib is currently studying Computer Science and Engineering (CSE) at American International University-Bangladesh. He is an undergraduate student. His research interest is in applying ML and DL to healthcare to solve real-world medical challenges by utilizing Python, Pandas, scikit-learn, and basic genomic data processing tools. His research expertise includes ML, DL, and biomedical data analysis with a focus on healthcare solutions and genomic data applications. His contribution to this research is software and writing the original draft. He aims to contribute practical, research-based solutions that can make a meaningful impact in healthcare to contribute to the betterment of humanity. He can be contacted at email: agdam.adib@gmail.com.

Md. Sanim Ahmed is a B.Sc. student in Computer Science and Engineering at the American International University-Bangladesh with research interests in ML, DL. He has contributed to this research in the following capacities as a co-author of the 2025 MDPI Drones paper "Multi-Agent Actor–Critic Frameworks for UAV Swarm Networks" (Q1, IF 4.4). In the current study, contributions include writing the original draft. His work focuses on creating interpretable and scalable intelligent systems that connect theory to real-world implementation. He can be contacted at email: ahmedsanim1234@gmail.com.

Amirul Islam received his P.hD. degree in Computing and Electronic Systems from the University of Essex, UK, in 2022. He completed his M.Sc. from Kookmin University, South Korea. He currently serves as an Assistant Professor in the Department of Electrical and Electronic Engineering at the BRAC University, Bangladesh. Prior to this, he held the position of a Post-Doctoral Researcher at the Visual AI Laboratory, Oxford Brookes University, UK. His research interests include ML for communication, optical camera communication, deep reinforcement learning, automotive vehicular communications, and optimization strategies. He can be reached at email: amirul.islam@bracu.ac.bd.