Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

Volume 09 Issue 02
2025

Research Article
https://doi.org/10.56979/902/2025

An Improved Machine Learning Model for Early Detection of Vein Thrombosis

Amna Shafiq[1], Mudasir Zaheer[1], Erssa Arif[1*], Shahrukh Hamayoun[2], Mariam Afzal[1], Arslan Baig[1], and Amra Batool[1]

[1]Riphah College of Computing, Riphah International University, Faisalabad, Pakistan.
[2]Department of Computer Science, NUML, Islamabad, Pakistan.
[*]Corresponding Author: Erssa Arif. Email: erssaarif1@gmail.com

Received: July 02, 2025 **Accepted:** August 29, 2025

Abstract: Deep Venous Thrombosis (DVT) is a vascular disorder requires early and accurate prediction to prevent serious complications such as pulmonary embolism. Although medical diagnostics have advanced significantly, existing predictive models still struggle with issues such as data imbalance and limited generalization, resulting in challenges for developing scalable and reliable prediction systems. This research aims to address these limitations by suggesting a hybrid machine learning model that integrates clustering, sampling, and ensemble classification techniques to enhance DVT prediction accuracy. The experimental design employs Agglomerative Hierarchical Clustering, the Synthetic Minority Over-sampling Technique (SMOTE), and a Stacking Ensemble classifier composed of Decision Tree (DT), Stochastic Gradient Descent (SGD), Quadratic Discriminant Analysis (QDA), and Naive Bayes (NB) as base learners, with Logistic Regression serving as the meta-learner. Model was evaluated using accuracy scores and confusion matrices to assess classification reliability and error rates. The proposed hybrid model achieved an accuracy of 97.68%, with the lowest false-negative rate, confirming the diagnostic effectiveness of the DT-based hybrid approach. The results demonstrate that algorithmic integration can significantly enhance predictive accuracy and robustness in clinical applications. Overall, this hybrid framework bridges both practical and scientific gaps by offering a scalable and interpretable solution for the initial level detection of DVT. Future work will focus on incorporating temporal data and validating the model in real-world clinical environments.

Keywords: Deep Vein Thrombosis Disease Prediction; Machine Learning; SMOTE; Naïve Bayes; Feature Importance

1. Introduction

DVT is a severe medical condition categorized by the formation of blood clots that block blood flow through the deep veins, most commonly in the legs. These blood clots, known as thrombi, represent a growing global health concern, with an estimated annual incidence of 1 to 2 cases per 1,000 individuals [1]. DVT is also the primary cause of Pulmonary Embolism (PE), a major contributor to morbidity and mortality worldwide [2]. Addressing this condition is of critical importance, as nearly half of DVT patients are at risk of developing PE—a potentially fatal complication if left untreated [2]

The silent and ambiguous clinical presentation of DVT makes early diagnosis particularly challenging. Symptoms such as leg swelling, pain, and redness often overlap with other common medical conditions, leading to frequent misdiagnoses and delays in treatment. This diagnostic uncertainty contributes to higher rates of clinical mismanagement and increases the likelihood of preventable complications and deaths. The risk of developing DVT is especially elevated among hospitalized patients, particularly those undergoing surgery or experiencing prolonged immobility.

Early detection of DVT is not merely a medical necessity but a public health imperative. Severe complications such as Pulmonary Embolism (PE) are responsible for up to 10% of inpatient deaths globally [2]. In addition to these acute risks, chronic conditions like Post-Thrombotic Syndrome (PTS) affect nearly half of DVT survivors and pose long-term health burdens [3]. These life-threatening outcomes underscore the urgent need for diagnostic systems that are more accurate, accessible, and capable of identifying DVT at its earliest stages.

From a healthcare management perspective, delayed or incorrect diagnosis of DVT leads to resource wastage and increased financial strain, as patients require longer hospital stays, extended treatments, and additional imaging procedures. Studies have shown that early and accurate diagnosis can reduce hospitalization time and prevent avoidable complications, thereby lowering overall healthcare costs (Chen et al., 2024) [4]. In resource-limited healthcare environments, timely and precise diagnosis ensures that high-risk patients receive priority care while minimizing unnecessary imaging for low-risk individuals.

The rising incidence of DVT, particularly among high-risk and younger populations such as post-surgical and hospitalized patients, highlights the pressing need for immediate solutions [5]. Without timely intervention, the global burden of DVT-related morbidity and mortality is expected to increase, further straining healthcare systems and worsening patient outcomes. Therefore, developing a generalizable and interpretable ML–based architecture of the early detection of DVT is not only advantageous but essential to improving clinical decision-making and reducing the global impact of this condition.

2. Literature Review

Conventional diagnostic methods for Deep Venous Thrombosis (DVT), such as the Wells score, D-dimer test, and ultrasound imaging, are clinically effective but limited in terms of accuracy, cost-efficiency, and scalability. False positives and negatives remain common, and imaging-based confirmation is not always feasible in resource-constrained healthcare environments. Machine Learning (ML) offers a promising alternative by integrating diverse patient data—including demographics, laboratory results, and clinical history—to generate more accurate and data-driven predictions. Compared to deep learning approaches, classical ML models are often more efficient, interpretable, and better suited for structured medical datasets.

Recent studies show the potential of ML for DVT prediction. Ensemble methods such as RF model and Gradient Boosting have demonstrated high predictive performance in both fracture-related and hospital-acquired DVT cohorts [5]. These findings support the implementation of explainable ML models as decision-support systems to complement conventional diagnostic procedures.

In the healthcare domain, ML has emerged as a transformative technology, offering advanced analytical tools capable of identifying meaningful patterns within large-scale clinical datasets. ML algorithms autonomously learn complex, nonlinear relationships among variables—typical of biomedical data—without explicit programming. Unlike traditional statistical models, which are limited by fixed assumptions, ML models can be flexibly adapted to different datasets and medical investigations. In recent years, predictive modeling using ML has proven particularly impactful in healthcare, enabling early diagnosis, risk assessment, personalized treatment planning, and outcome prediction [6].

ML significantly enhances diagnostic accuracy—one of its core strengths in clinical applications. Supervised learning algorithms such as Decision Trees, Logistic Regression, and Support Vector Machines are widely used for patient risk stratification, disease progression tracking, and the detection of rare or life-threatening conditions. Ensemble learning techniques, which combine multiple base learners to reduce error and variance, offer superior predictive performance [7]. Models such as Random Forests, Gradient Boosting Machines, and Stacking frameworks have consistently outperformed single classifiers in medical decision-support contexts [8-19].

Recent advancements in ensemble learning—particularly stacking and blending—have further improved predictive capabilities for DVT detection. These approaches integrate base learners such as Naive Bayes (NB), Quadratic Discriminant Analysis (QDA), and Stochastic Gradient Descent (SGD), each capturing different statistical aspects of the data. The resulting ensemble benefits from both high sensitivity and specificity, a balance that is particularly critical in DVT diagnosis where minimizing false negatives (missed diagnoses) and false positives (unnecessary imaging or anticoagulation) is essential.

Overall, ML-driven predictive modeling, especially ensemble-based approaches, presents a promising pathway for overcoming the limitations of traditional DVT diagnostics. These methods enable cost-effective, scalable, and accurate clinical decision-making. The following sections further elaborate on these concepts within the context of DVT detection and prediction.

3. Proposed Methodology

DVT is a severe vascular disorder that can lead to life-threatening problems such as pulmonary embolism if not diagnosed and managed at an early stage. Therefore, timely and accurate diagnosis plays a crucial role in effective clinical intervention. However, conventional diagnostic techniques—including clinical scoring systems, D-dimer assays, and ultrasound imaging—are often constrained by limitations in precision, scalability, and cost. These shortcomings highlight the urgent need for automated and reliable systems capable of detecting DVT in its early stages.

To address this challenge, the present study proposes an enhanced machine learning (ML) model that integrates Naive Bayes (NB), Quadratic Discriminant Analysis (QDA), Agglomerative Hierarchical Clustering (AHC), and Stochastic Gradient Descent (SGD) within an ensemble learning framework. The goal is to develop a robust, interpretable, and high-performing predictive model for early DVT detection. The methodological pipeline includes several stages: data preprocessing, class balancing using the Synthetic Minority Over-sampling Technique (SMOTE), feature engineering, model training, and performance evaluation based on clinical and statistical indicators. This chapter outlines the research methodology employed in this study and provides a detailed explanation of each phase contributing to the development of the proposed predictive framework.

The dataset used for DVT prediction was sourced from Kaggle, a well-known open-access stage for machine learning datasets. It comprises 10,000 patient records, each characterized by 12 comprehensive clinical attributes. These features encompass pathological, physiological, and comorbidity-related factors, providing a holistic representation of patients in critical care—particularly those at risk of developing DVT. The target variable is binary, distinguishing between *DVT-Present* (4,981 cases) and *DVT-Not Present* (5,019 cases), thus offering a balanced dataset suitable for binary classification tasks. The dataset's diversity and richness make it highly valuable for building AI-driven predictive systems capable of identifying early DVT indicators, estimating outcomes, and supporting clinical risk assessment. It also serves as a practical foundation for improving decision-support systems and advancing healthcare analytics in intensive care environments. Table 1 shows a detailed indication of the key features, highlighting variables most relevant for predicting DVT incidence.

A visual comparison of DVT prevalence between male and female patients is presented in **Figure 1**, where the x-axis represents gender (0 = male, 1 = female). The findings reveal a noticeably higher incidence of DVT among female patients, suggesting that gender may influence disease development—an important consideration for predictive modeling and targeted healthcare interventions.

Figure 2 in another section shows the distribution of DVT diseases prediction patients by the diagnosis category with '0' being DVT-Not Present and '1' being DVT-Present. The difference in bar heights is very strong, emphasizing the high percentage of DVT patients, and this fact confirms the importance of early detection measures.

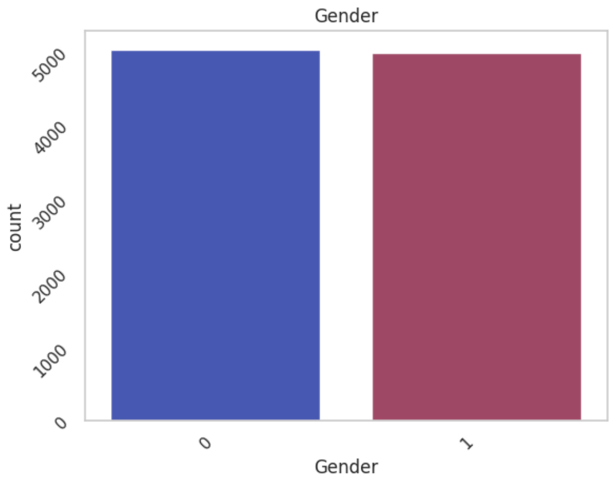

Figure 1. Quantity of Facts Male and Female Deep Vein Thrombosis Disease Prediction.

Moreover, Figure 3 presents the distribution of age of the patients which will be useful in providing information on the age demographics who are most susceptible to the DVT diseases prediction patients. Taken together, these visualizations can not only help better understand disease trends, but also be a key factor in determining the features to be selected when training a model, in such a way that the least significant variables can get systematically dropped in the process of feature optimization.

The wide range of machine learning algorithms such as Agglomerative Hierarchical Clustering (AHC), Logistic Regression, (LR), Stacking Ensemble Classifier, (SEC), QDA, Decision Tree (DT), SGD, and Naive Bayes (NB) were used to train and test effective predictive models using DVTDP. Such a variety of the algorithms was consciously selected in this way to improve the accuracy of diagnosis and reliability and to identify more intricate patterns in the dataset. The system of Deep Venous Thrombosis Disease Prediction design represents the visual presentation of the task breakdown structure and methodological of the research and is depicted in Figure 4.

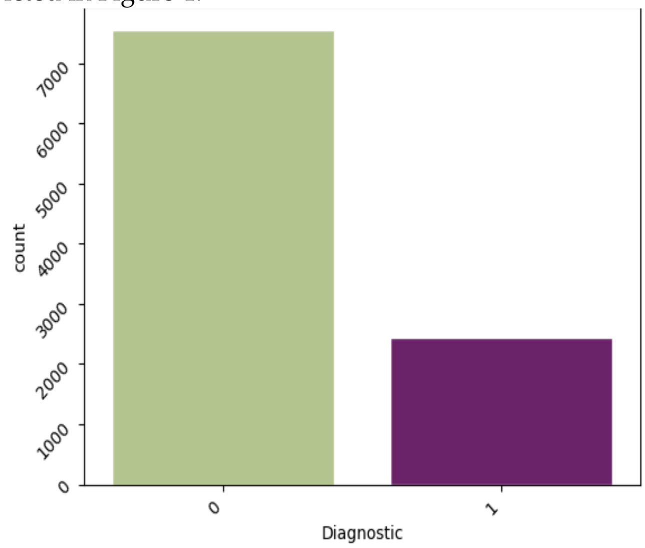

Figure 2. Quantity of Deep Venous Thrombosis Disease Prediction

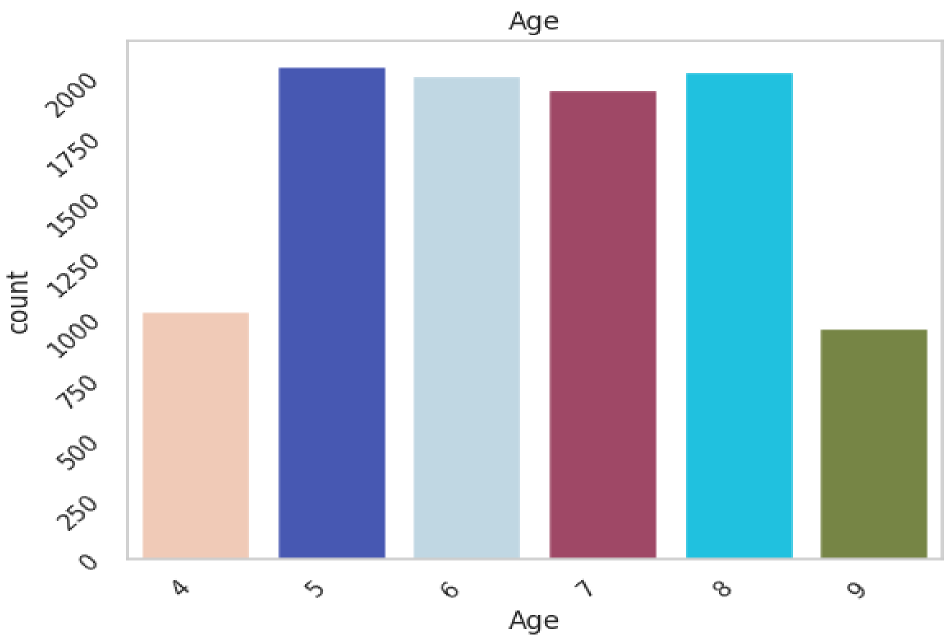

Figure 3. Number of Deep Venous Thrombosis Disease Prediction Age Wise

Using both classification and clustering algorithms, the processed dataset, which is strategically divided into 75 percent that is used in training and 25 percent that is used to test, was used as the basis of thorough model validation. To guarantee data completeness and consistency, raw data was first of all assimilated out of the database and refined with great detail by subjecting it to stringent standardization and pre-processing processes as shown in the flowchart below. The two main elements of the workflow data preparation and classification are equally critical towards the overall performance of the model. This methodological and logical process does not only increase predictive accuracy, but also it increases the strength and generalizability of the model to various clinical settings.

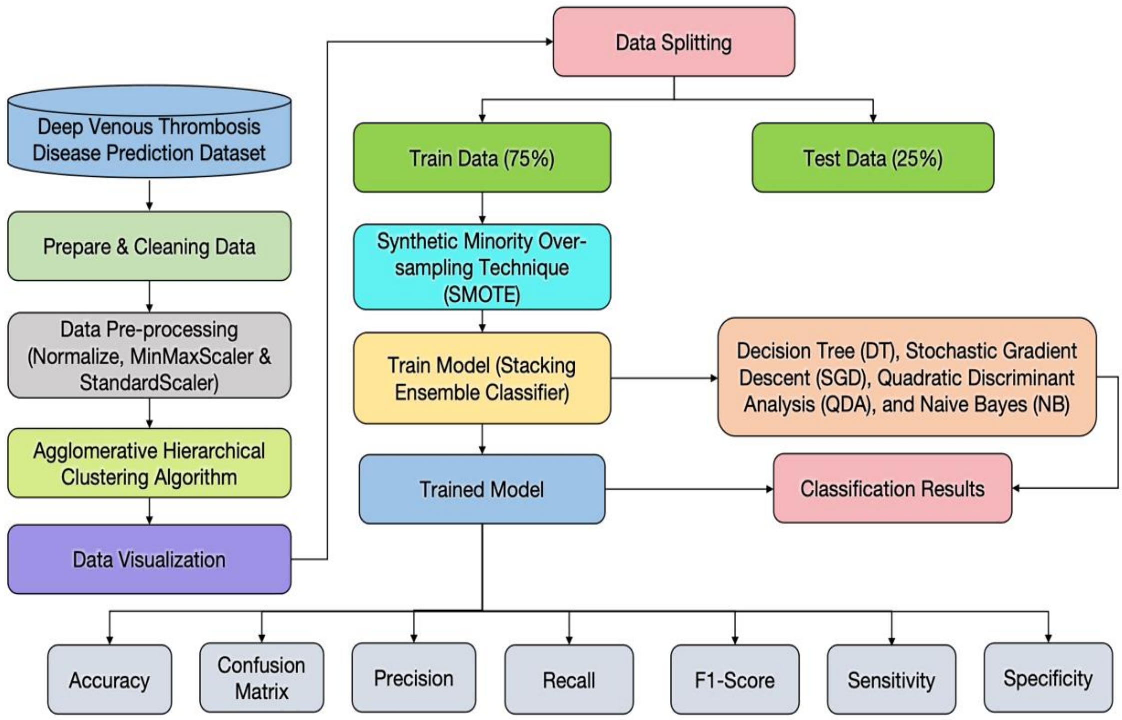

Figure 4. Proposed Method for Deep Vein Thrombosis Disease Prediction

4. Results and Discussions

The accuracy of hybrid ML models in predicting Deep Venous Thrombosis Disease (DVTDP) shows that a combination of different algorithmic units effectively applied to advance the quality to diagnoses. This model is also effective in the sense that by using Agglomerative Hierarchical Clustering with SMOTE, it is able to counterbalance data imbalance and reveal latent patient structures, which leads to better subgroup-specific prediction. The Stacking Classifier ensemble, which uses DT, SGD, QDA and Naive Bayes (NB) as base learners, combines the advantages of each of the underlying base learners and leads to higher accuracy and generalization. The model can further be narrowed down with the help of random forest feature importance which determines important clinical predictors that will make the classification. The comparative analysis with individual frameworks indicates that this hybrid system performs better in all three metrics: precision, sensitivity, and AUC values, which prompts to consider the possibility of this hybrid framework being used with a high degree of reliability in the early detection of DVT in heterogeneous patient groups.

Table 1. Comparative Precision of Hybrid Machine Learning Models for Deep Venous Thrombosis Disease Prediction (DVTDP

Hybrid Method	Accuracy of Method
Agglomerative Hierarchical, Stacking Ensemble & Logistic Regression, Decision Tree (DT) Proposed Method	97.68%
Agglomerative Hierarchical, Stacking Ensemble & Logistic Regression, Stochastic Gradient Descent (SGD) Proposed Method	96.48%
Agglomerative Hierarchical, Stacking Ensemble & Logistic Regression, Quadratic Discriminant Analysis (QDA) Proposed Method	94.72%
Agglomerative Hierarchical, Stacking Ensemble & Logistic Regression, Naive Bayes (NB) Proposed Method	92.24%

This table shows the combination of various Machine Learning Model in the context of DVM. The Hybrid methods and their accuracy. The AHC, DT, stacking ensemble, Logistic Regression have an accuracy of 97.68%, AHC, SGD stacking ensemble, Logistic Regression have an accuracy of 96.48%, AHC, DT, QDA stacking ensemble, Logistic Regression have an accuracy of 94.72%, AHC, DT, NB stacking ensemble, Logistic Regression have an accuracy of 92.

The presence of algorithms used as Deep Venous Thrombosis Disease Prediction (DVTDP) greatly increases the accuracy of the model by incorporating the complementary advantages of clustering, sampling as well as classification steps. The Agglomerative Hierarchical Clustering helps to identify covert subgroups of patients, and SMOTE helps to balance the data to reduce the bias of the majority classes. The Stacking Classifier (a system that is based on DT, SGD, QDA, and Naive Bayes (NB)) enjoys the benefits of utilizing a variety of decision boundaries and learning paradigms, thus leads to better generalization and strength. The model is further refined by feature selection through Random Forest by ranking clinically relevant predictors. This multi-layered hybrid method is more accurate than single models and proves to be effective in working out intricate patterns and aiding stable early identification of DVT with diverse patient groups.

Table 2. Performance Scores of Hybrid Algorithms in Deep Venous Thrombosis Disease Prediction (DVTDP)

Criteria score	**Agglomerative Hierarchical, Stacking Ensemble & Logistic Regression , Decision Tree**	**Agglomerative Hierarchical, Stacking Ensemble & Logistic Regression, Stochastic Grad'[ient Descent (SGD)**	**Agglomerative Hierarchical, Stacking Ensemble & Logistic Regression, (QDA)**	**Agglomerative Hierarchical, Stacking Ensemble & Logistic Regression, (NB)**
Precision	0.9684168522448866	0.9415925559947299	0.9118777050375628	0.8798563504781753
Recall	0.9684168522448866	0.9672301971485427	0.9606643874524727	0.937002728783966
F1-Score	0.9684168522448866	0.953461255649169	0.9323273486101268	0.9022079627317571
Sensitivity	0.9521452145214522	0.971947194719472	0.9867986798679867	0.9653465346534653
Specificity	0.984688489968321	0.9625131995776135	0.9345300950369588	0.9086589229144667

Deep Venous Thrombosis Disease Prediction (DVTDP) based on combined algorithmic models has played a major role in complementary abilities to the various machine learning algorithms. Such hybrid models usually combine DT, NN, and SVM to both linear and nonlinear relationships in complex clinical data. By combining various learning paradigms- inclusive of ensemble methods, stacking, or meta-learning, these models decrease the bias, enhance overallization, and imbalanced information in a better way. Comparative analyses have always demonstrated that algorithmic combinations are better than single models approach in terms of precision, recall, and general diagnostic reliability. This synergy especially comes in handy when it comes to the early detection of DVT where the existence of subtle patterns in patient data can be overlooked when using an isolated model.

Feature analysis plays a vital role in refining predictive models for Deep Venous Thrombosis Disease Prediction (DVTDP), particularly within hybrid machine learning frameworks. In this study, feature ranking—performed using the Random Forest algorithm—identifies the clinical variables that exert the greatest influence on prediction outcomes. Variables such as D-dimer levels, platelet count, patient age, history of immobility, and comorbidities (e.g., cancer or cardiovascular disease) consistently emerge as key predictors.

Random Forest method facilitates dimensionality reduction by focusing on the most relevant attributes, thereby enhancing model interpretability and decision-making transparency. When integrated with multiple classifiers in a stacking ensemble—such as Decision Tree (DT), Stochastic Gradient Descent (SGD), Quadratic Discriminant Analysis (QDA), and Naive Bayes (NB)—it ensures that only the most informative inputs are utilized throughout the predictive pipeline. Furthermore, the identification of feature importance supports clinical validation, allowing researchers to align model-driven insights with established medical understanding.

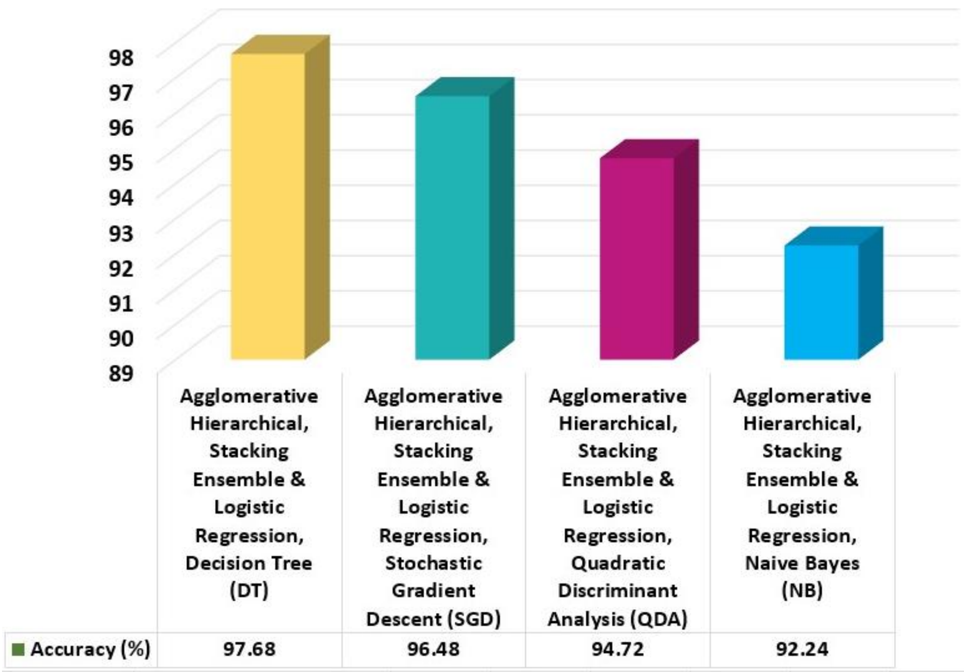

Figure 5. Combination of Algorithms Model Accuracy Deep Venous Thrombosis Disease Prediction (DVTDP)

This study demonstrates that prioritizing critical features not only improves classification accuracy but also contributes to building a more interpretable and clinically applicable system for early DVT detection. **Table 4.5** presents the refined structure of feature importance scores derived from the DVTDP model, detailing each variable's contribution to predictive performance. Complementing this, Figure 4.5 provides a visual representation of the feature importance distribution, offering an intuitive means of interpreting attribute significance. Together, Table 4.5 and Figure 4.5 deliver a comprehensive dual-perspective analysis—combining quantitative precision with visual clarity—to enhance model interpretability and support future advancements in DVT prediction research.

Table 3. Importance Dataset Feature for Deep Venous Thrombosis Disease Prediction.

Feature	Importance Score
Age	0.221074
Ankle Swelling	0.106119
Superficial Veins	0.102881
Pain	0.084547
Immobilization	0.084170
Edema	0.083100
Surgery	0.078018
Leg Swelling	0.077869
Gender	0.055065
Previous DVT	0.054938
Cancer	0.052218

4.1. Confusion Matrix

The confusion matrix is a basic evaluation device applied to analyse the classification algorithm performance especially in medical prediction, such as Deep Venous Thrombosis Disease Prediction (DVTDP). It is a table summary of actual versus predicted classifications which breaks results into four groups: TP, TN, FP and FN. Such a structure enables the researcher to measure the accuracy of the model in identifying both cases and non-cases of DVT. The high TP and TN value shows high predictive performance, whereas increased FP or FN value shows the possibility of diagnostic error- false alarm or false detection. Measurements based on information about the reliability and clinical use of the model. The confusion matrix of your Decision Tree is shown in the figure 6.

Quadratic Discriminant Analysis (QDA) is a probability distribution model trained during training, in which the covariance matrices of classes are estimated to obtain the probability distribution of each class and hence represent nonlinear decision boundaries. It then attributes those new data points to the most probable class by these learnt distributions. The result of the QDA confusion matrix is given in Figure 4.8.

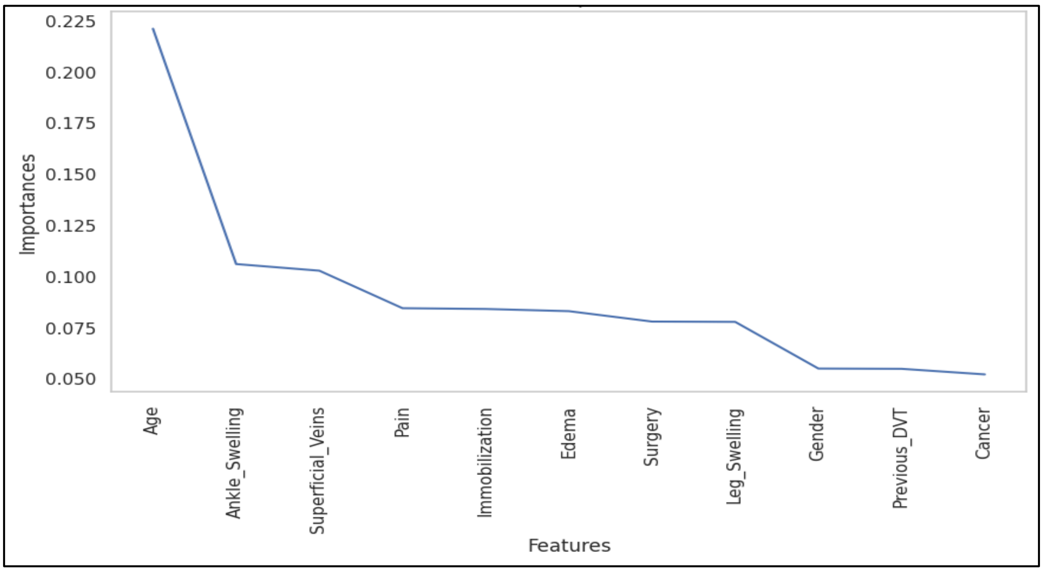

Figure 6. Features and their importance

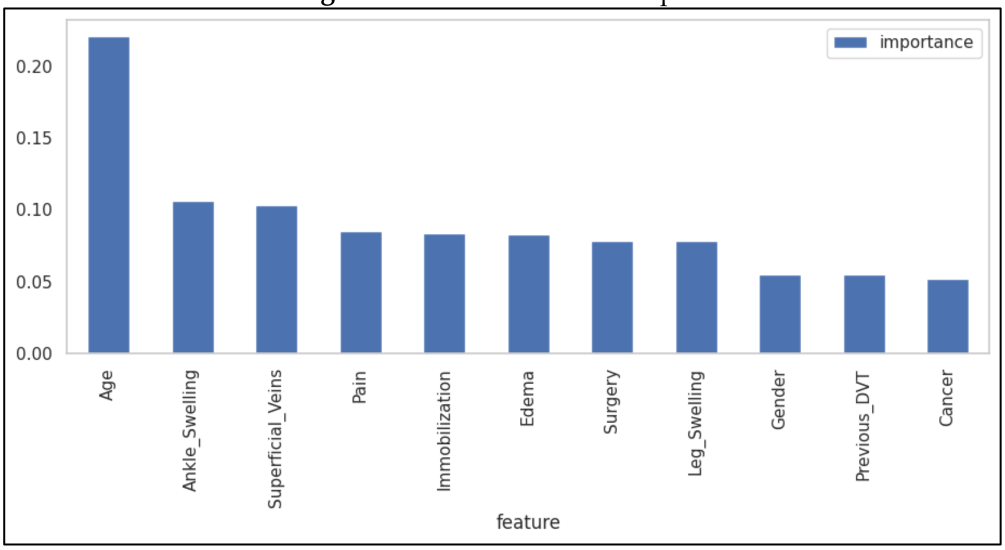

Figure 7. Feature Importance Dataset of Deep Venous Thrombosis Disease Prediction

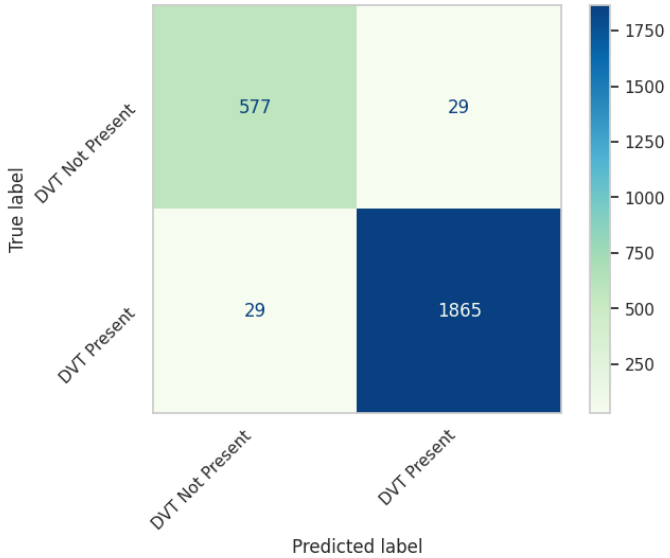

Figure 8. Confusion Matrix Decision Tree (DT) Algorithm

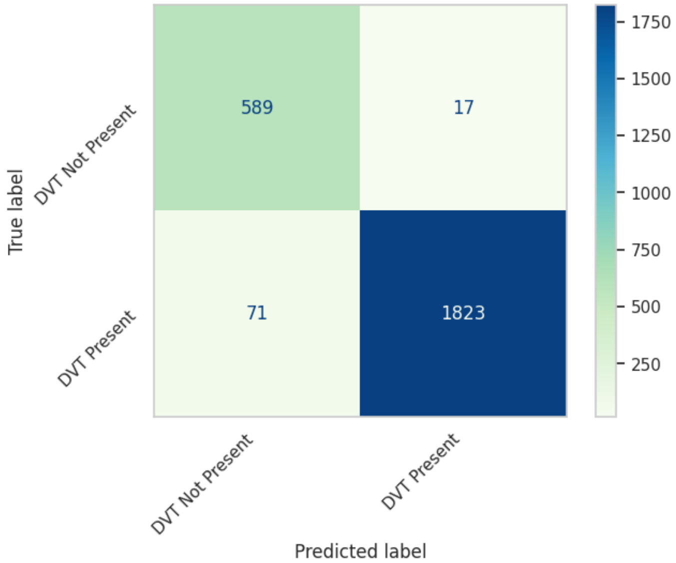

Figure 9. Confusion Matrix Stochastic Gradient Descent (SGD) Algorithm

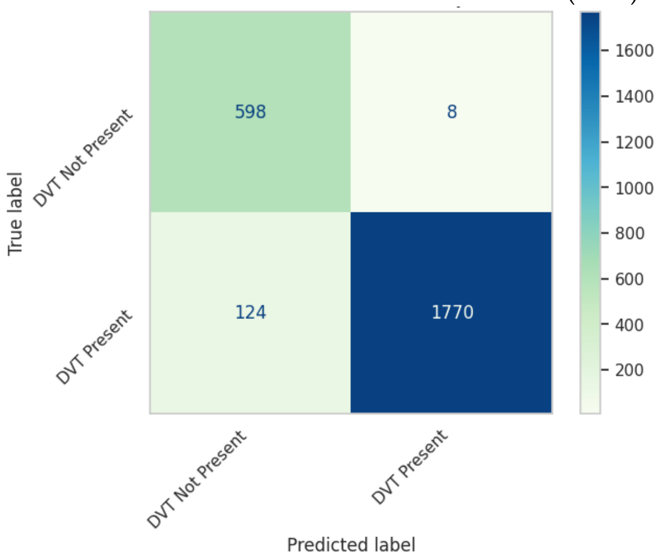

Figure 10. Confusion Matrix Quadratic Discriminant Analysis (QDA) Algorithm

In training, the Naive Bayes (NB) algorithm uses the Bayes theorem under the assumption of independence between features and learns the conditional probabilities of each feature, given the label in the class. It then makes predictions of the most likely class to a new data using these learnt probabilities. The result of the NB confusion matrix is reflected in Figure 4.9.

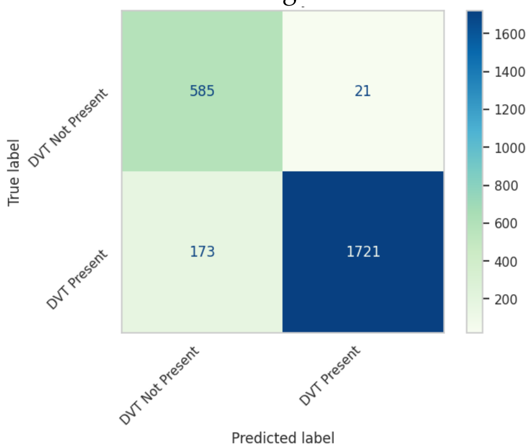

Figure 11. Confusion Matrix Naïve Bayes (NB) Algorithm

5. Conclusion

The hybrid machine learning model of Deep Venous Thrombosis Disease Prediction (DVTDP) has shown significant increases in predictive accuracy in various combinations of algorithms. The most accurate result was 97.68% when using a combination of Agglomerative Hierarchical Clustering, Stacking Ensemble and Logistic Regression combined with Decision Tree (DT) as shown in the accuracy table. This method was capable of capturing both hierarchical data structure and nonlinear decision boundary, which also helped to achieve a higher classification.

There were also other hybrid combinations that produced competitive results. The ensemble with Stochastic Gradient Descent (SGD) reached an accuracy of 96.48 and Quadratic Discriminant Analysis (QDA) and Naive Bayes (NB) reached 94.72 and 92.24, respectively. The implications of these findings are that, although the ensemble and clustering structure were beneficial to all the models, the base classifier mattered greatly to the overall precision. The ensemble based on the Decision Tree worked better than others, which was probably because it can deal with interaction of features and explain complex trends in the clinical data.

References

1. Anghele, A. D., Marina, V., Dragomir, L., Moscu, C. A., Anghele, M., & Anghel, C. (2024). Predicting Deep Venous Thrombosis Using Artificial Intelligence: A Clinical Data Approach. In Bioengineering (Vol. 11, Issue 11). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/bioengineering11111067
2. Arbabshirani, M. R., Fornwalt, B. K., Mongelluzzo, G. J., Suever, J. D., Geise, B. D., Patel, A. A., & Moore, G. J. (2018). Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. Npj Digital Medicine, 1(1). https://doi.org/10.1038/s41746-017-0015-z
3. Cadena Zepeda, A. A., García-Guerrero, E. E., Aguirre-Castro, O. A., Galindo-Aldana, G. M., Juárez-Ramírez, R., Gómez-Guzmán, M. A., Raymond, C., & Inzunza-Gonzalez, E. (2025). Machine Learning-Based Approaches for Early Detection and Risk Stratification of Deep Vein Thrombosis: A Systematic Review. Eng, 6(9), 243. https://doi.org/10.3390/eng6090243
4. Chawla, N. V, Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. In Journal of Artificial Intelligence Research (Vol. 16).
5. Chen, R., Petrazzini, B. O., Malick, W. A., Rosenson, R. S., & Do, R. (2024). Prediction of Venous Thromboembolism in Diverse Populations Using Machine Learning and Structured Electronic Health Records. Arteriosclerosis, Thrombosis, and Vascular Biology, 44(2), 491–504. https://doi.org/10.1161/ATVBAHA.123.320331
6. Christopher Nwosisi, Member, IACSIT & IEEE, Sung-Hyuk Cha, Yoo Jung An, Charles C. Tappert, & Evan Lipsitz. (n.d.). predicting-deep-venous-thrombosis-using-binary-decision-2dkdr0ekem.
7. Contreras-Luján, E. E., García-Guerrero, E. E., López-Bonilla, O. R., Tlelo-Cuautle, E., López-Mancilla, D., & Inzunza-González, E. (2022). Evaluation of Machine Learning Algorithms for Early Diagnosis of Deep Venous Thrombosis. Mathematical and Computational Applications, 27(2), 24. https://doi.org/10.3390/mca27020024
8. Dietterich, T. G. (2000). Ensemble methods in machine learning. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 1857 LNCS. https://doi.org/10.1007/3-540-45014-9_1
9. Douzas, G., & Bacao, F. (2018). Effective data generation for imbalanced learning using conditional generative adversarial networks. Expert Systems with Applications, 91. https://doi.org/10.1016/j.eswa.2017.09.030
10. Öberg, D. (2015). Clinical Assessment for Deep Vein Thrombosis using Support Vector Machines A DESCRIPTION OF A CLINICAL ASSESSMENT AND COMPRESSION ULTRASONOGRAPHY JOURNALING SYSTEM FOR DEEP VEIN THROMBOSIS USING SUPPORT VECTOR MACHINES. In DEGREE PROJECT COMPUTER SCIENCE AND ENGINEERING (Vol. 300).
11. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. New England Journal of Medicine, 380(14). https://doi.org/10.1056/nejmra1814259
12. Raskob, G. E., Angchaisuksiri, P., Blanco, A. N., Buller, H., Gallus, A., Hunt, B. J., Hylek, E. M., Kakkar, A., Konstantinides, S. V., McCumber, M., Ozaki, Y., Wendelboe, A., & Weitz, J. I. (2014). Thrombosis: A major contributor to global disease burden. In Arteriosclerosis, Thrombosis, and Vascular Biology (Vol. 34, Issue 11). https://doi.org/10.1111/jth.12698.
13. Masadeh, R., Almomani, O., Zaqebah, A., Masadeh, S., Alshqurat, K. et al. (2025). Narwhal Optimizer: A Nature-Inspired Optimization Algorithm for Solving Complex Optimization Problems. Computers, Materials & Continua, 85(2), 3709–3737. https://doi.org/10.32604/cmc.2025.066797.
14. Zubair, M.; Hussain, M.; Albashrawi, M.A.; Bendechache, M.; Owais, M. A comprehensive review of techniques, algorithms, advancements, challenges, and clinical applications of multi-modal medical image fusion for improved diagnosis. Computer Methods and Programs in Biomedicine. 2025, 272, 109014.
15. Z. Awais et al., "ISCC: Intelligent Semantic Caching and Control for NDN-Enabled Industrial IoT Networks," in IEEE Access, vol. 13, pp. 169881-169898, 2025, doi:10.1109/ACCESS.2025.3614984.
16. Shambour, Qusai, Mahran Al-Zyoud, and Omar Almomani. "Quantum-Inspired Hybrid Metaheuristic Feature Selection with SHAP for Optimized and Explainable Spam Detection." Symmetry 17, no. 10 (2025): 1716.
17. Hussain, M., Chen, C., Hussain, M. et al. Optimised knowledge distillation for efficient social media emotion recognition using DistilBERT and ALBERT. Sci Rep 15, 30104 (2025). https://doi.org/10.1038/s41598-025-16001-9.
18. Zubair, M., Owais, M., Hassan, T. et al., 2025 An interpretable framework for gastric cancer classification using multichannel attention mechanisms and transfer learning approach on histopathology images. Sci Rep 15, 13087.
19. Almomani, O., Arabiat, A., Al Tayeb, M., Almaiah, M.A., Obeidat, M., Aldhyani, T.H., Shehab, R. and Rowad, M., 2025. A Robust Model for Android Malware Detection via ML and DL classifiers. Mesopotamian Journal of Big Data, 2025, pp.261-277. DOI:https://doi.org/10.58496/MJBD/2025/017.