Journal of Theoretical and Applied Information Technology
15th April 2023. Vol.101. No 7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A NEW WEIGHTED FEATURE SELECTION AND ENSEMBLE LEARNING WITH HYBRID HEURISTIC STRATEGY FOR DEEP VEIN THROMBOSIS PREDICTION

KAPARTHI UDAY[1,2], Dr.MUDIGONDA MALINI[3]

[1]Sr.AssistantProfessor, CVR College of Engineering, Department of EIE, Hyderabad, India

[2]Research Scholar, OU College of Engineering, Department of BME, Hyderabad, India

[3]Professor, OU College of Engineering, Department of BME, Hyderabad, India

E-mail: [1]k.uday@cvr.ac.in, [2]kaparthiuday@gmail.com,[3]malini.m@uceou.edu

ABSTRACT

Deep Vein Thrombosis (DVT) is caused due to blood clots in deep veins. DVT symptoms differ from patient to patient. The prediction of DVT is essential in the primary stage. Otherwise, it may lead to Pulmonary Embolism(PE). In this research, a new DVT prediction model is designed with the data augmentation concept, a technique commonly used in machine learning to enhance the dataset used in the learning process. It produces new cases from the original dataset without varying the pattern of the data. The proposed model consists of data augmentation, weighted feature selection, and classification phases. Initially, the data augmentation is performed after gathering the data related to DVT. Secondly, data cleaning and outlier removal are done on the collected data to improve the data quality. Furthermore, weighted feature extraction is used with Hybrid Lion-Sea Lion Optimization (HL-SLnO) to extract high-quality, non-redundant features from data. Finally, the obtained features are sent to the classification phase, where they are processed using Modified Ensemble Classifiers (MEC) with Long Short-Term Memory (LSTM), Deep Neural Network (DNN), and XGBoost with parameter optimization derived from the same HL-SLnO. MEC combines the decisions from the classifiers mentioned above to improve the overall performance to obtain the best-predicted result. Finally, the experimental results show that the proposed model has effectively enhanced the prediction of DVT risk and provided decision support for other nursing and medical intervention.
Keywords: *Deep Vein Thrombosis, Weighted Feature Selection, Hybrid Lion-Sea Lion Optimization, Deep Neural Network, Long short-term memory.*

1. INTRODUCTION

One of the most severe issues in peripheral vein diseases is intraluminal clotting. DVT is one of the peripheral vascular disorders. Thrombosis is the formation of a blood clot inside a blood vessel. DVT is the name for this condition when it occurs in the deep veins. Pulmonary embolism (PE) is the most dangerous complication of DVT. PE is a blockage in one of the lungs' pulmonary arteries. PE occurs when blood clots migrate from the deep veins of the legs to the lungs. As a result, DVT is a leading cause of illness and death worldwide.

Primary care physicians in many countries lack the technology to perform imaging tests such as ultrasound or venography, requiring them to rely on probabilistic models. The finest clinical probability model, such as the wells criteria, has been developed for this purpose. It is usually rule-based and simple to understand. It can be calculated based on symptoms and risk factors shown in Table.1. Age and gender should be considered for prediction. Female subjects have more chances of getting DVT compared to male subjects. DVT is more likely to occur in older subjects [11,12]. All of these should be considered features of the proposed model. Based on the Wells score, the DVT is classified into three categories such as:

- The risk is minimum: Wells score = 1
- The risk is modest: Wells score = 2 or 3
- The risk is considerable: Wells score $\geq$ 3

This helps the doctor to decide whether DVT is present or not and also to proceed with further diagnostic testing like venography and CT Scan [13-15]. The usage of machine learning and Artificial Intelligence (AI) in biomedical and medical engineering has improved the early detection, prediction, diagnosis, and classification of diseases [16-22]. Machine learning is more helpful to analyze the data, which is more applicable to aiding the Information Technology (IT) goals and design tools. It also can synthesize and utilize the data for meeting those goals and serving users. Hence, the exploration of ensemble learning suggested in this research can be further adopted into various IT domains like image recognition, text generation and analysis, speech recognition, data analytics, etc..

Table 1: Clinical Criteria For Detecting Deep Vein Thrombosis Developed By Wells [13,14,15]

Symptoms and risk factors	Points
Cancer that is current or has just been treated	1
Leg paralysis	1
Bedridden for more than three days recently or underwent significant surgery within the past four weeks	1
Pain in the area of a deep vein	1
Leg swollen	1
A calf that is swollen and has a diameter that is more than 3 cm greater than the other calf's.	1
Veins in the legs that are larger but not varicose	1
Dvt was previously diagnosed	1
Pitting edema confined to the symptomatic leg	1
More likely other diagnoses	1

The research questions considered for promoting this research work with the incorporation of ML techniques to predict DVT are listed here.

1. Is there a general culture of screening conducted for DVT in COVID-19 patients?
2. Is the data collection in predicting DVT in COVID-19 patients adequate?
3. Are the recently developed ML techniques performed better to predict DVT in COVID-19 patients?
4. Are heuristic methods helpful for diagnosing DVT symptoms?

This paper is organized as follows: Section 2 explains the literature survey, Section 3 describes the architectural view of the DVT detection model, Section 4 discusses optimal weighted feature selection by hybrid HL-SLnO for DVT detection, Section 5 describes Enhanced Deep venous thrombosis detection using MEC, section 6 presented the results and discussion, and finally, conclusions are drawn in section 7.

2. LITERATURE SURVEY

Various machine learning and deep learning models have been developed to detect and analyze DVT and PE. Numerous DVT prediction models are reviewed in Table 2. The research gaps identified in these reviews motivate the development of a new prediction model in this field.

3. ARCHITECTURAL VIEW OF DVT DETECTION MODEL

A novel MEC-based approach is introduced in this work to detect DVT accurately. The architectural design of the proposed DVT detection model is shown in figure 1.

The proposed DVT prediction model consists of data collection, data augmentation, weighted feature selection, and classification phases. The dataset is collected based on the reference [1] and consists of 10000 cases used to train and validate ensemble classifiers. The total risk type of the dataset is 6. They are low risk with a positive and negative diagnosis, medium risk with a positive and negative diagnosis, and high risk with a positive and negative diagnosis.

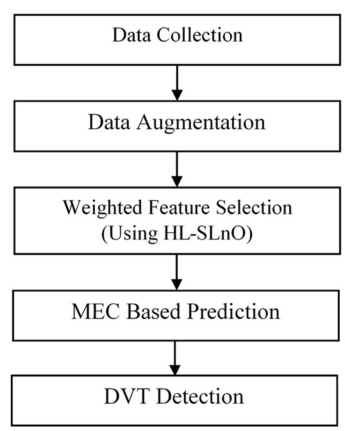


Figure 1: DVT detection model

Then, the data augmentation enhances the data set by producing new cases from the original data set without varying the data pattern and results in identifying the risk factors. The dataset is trained in row types, with each row including gender, age, and ten symptoms (features). At last, the values related to each patient are obtained, the positive cases are denoted as 1, and the negative cases are indicated as 0. The collected input data is determined as Da_r, where, r =1,2,3….R, R is represented as the total

Journal of Theoretical and Applied Information Technology
15th April 2023. Vol.101. No 7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

data count. The augmented data are expressed as Da_r^{ag}.

Later, the feature is selected, and the weighted feature selection is performed using the suggested HL-SLnO method by optimizing the weight to minimize the correlation and maximize classification accuracy. Finally, the chosen weight features are subjected to the classification phase. The classification task is performed with MEC, including LSTM, DNN, and XGboost. The highest ranking from the classifier output is considered the best-predicted result.

4. OPTIMAL WEIGHTED FEATURE SELECTION BY HYBRID LION-SEA LION OPTIMIZATION FOR DVT DETECTION

The proposed HL-SLnO algorithm is used for weighted feature selection to minimize the correlation between the features such as age, gender, and Well's mentioned symptoms and enhance the classification phase's performance by optimizing the parameters like hidden neuron count in LSTM, learning rate in XGBoost, and hidden neuron count in DNN. The hybridization of the Lion Optimization Algorithm (LOA) [26] and Sea Lion Optimization (SLnO) [27] algorithms have several features; chosen in this paper to maximize the performance of DVT detection.

Table 2: Features of DVT prediction models

Author [citation]	Methodology	Features
Berenice *et al.* [1]	Artificial Neural Network(ANN) model	•It improves accuracy, recall, and sensitivity. •It performs very well in unbalanced data. •In this, age and gender are considered, along with Wells mentioned ten features. •Diagnostic accuracy is 91% only.
Luo *et al.* [2]	Generalized Linear Model(GLM)and Support Vector Machine(SVM)	•It has reduced the usage and occupation of medical costs. •It performs very well in real-time datasets. •It is not performed well while handling insufficient data.
Ryan *et al.* [3]	Gradient boosted algorithm	•It has maintained a lower false positive rate. •It prevents unnecessary invasive testing. •It does not seem suitable for predicting the long-term risk of the patient population.
Willan*et al.* [4]	ANN model	•It offers efficient decision-making. •It is more suitable for performing complex data. •The model was based on the clinical score of D-dimer blood test patients. •This model is not applicable for evaluating high-risk patients.
Emilia *et al.* [5]	Multi-Objective Optimization(MOO)	•This model has been evaluated in terms of Pareto fronts with constraints like (1-FOR)-Precision, Accuracy-AUC Specificity-Sensitivity, and Sensitivity-Precision. •It efficiently deals with data imbalance. •It has higher false positives
Li *et al.* [6]	Multivariable logistic regression	•The receiver operating characteristic curve ensures superior prediction performance. •It is more helpful for helping high-risk patients and recognizing patients with ischemic stroke. •This designed model does not apply to many patients and diversities.
Ma *et al.* [7]	The hierarchical machine learning model	•It has attained higher predictive efficiency. •This model can be applied to other domains also. •However, it does not perform well in the absence of additional features or a large sample size
Martins *et al.* [8]	ANN	•It is more significant for processing a diverse set of inputs and solves the recurrence of VTE in several situations. •This model was based on clinical, inherited & acquired molecular factors. •It is only applicable to processing a lower number of patients

The parameters W1 and W2 are incorporated to improve the proposed algorithm.

$$MS_{leader} = \left|\left(W_1(1 + W_2)\right)/W_2\right| \quad (1)$$

Here, the two parameters W1 and W2 were chosen based on the best fitness of all individuals of initialization. If the random variable, MS_{leader}<0.25, the position is updated based on the SLnO algorithm; otherwise, the status is updated based on LOA. SLnO is a nature-based algorithm inspired by the hunting behavior of sea lions. This algorithm assumes the target prey is the current best or close optimal solution.

The hunting behavior of lions inspires the LOA algorithm. Cubs are solutions derived from previous solutions, whereas lions are solutions that have yet to be identified. Territorial defense is the process of comparing the new solution (nomadic lion) to the existing solution (territorial lion), replacing the old solution with the new one if the new solution is superior, and eradicating the old solution's derived solutions. The exploration and exploitation phase updates the solutions in all possible directions. The HL-SLnO improves detection performance by identifying DVT with extensive data. Algorithm 1 provides the pseudo-code for the proposed HL-SLnO.

Algorithm 1: Proposed HL-SLnO

Step 1:Initialize the population
Step 2: Determine the best position for an individual based on fitness value
Step 3:Determine the parameters W1 and W2
Step 4: If MS_{leader}< 0.25
Update the solution using SLnO
else
Update the solution using LOA
Step 5: Find the best solution
Step 6**:** End.

The augmented data Da_r^{ag} is subjected to optimal feature selection and weight optimization in the proposed DVT prediction model. Thus, the optimal features and optimized weight of DVT prediction are further multiplied to get the weighted features. The optimal selected features are denoted as Fe_n, and the optimized weightW_n is given in Eq. (2).

$$Fe_n^{weg} = W_n \times Fe_n \quad (2)$$

Here, the term Fe_n^{weg} is indicated as the optimized weighted features, where, n=1,2,3…….N, the term N is represented as the total count of features. The primary purpose of the optimized weighted features is to optimize the weight and the features by minimizing the correlation of weighted features.

$$fit_1 = \underset{\{W_n, Fe_n\}}{\arg\max}\left(\frac{1}{corr}\right) \quad (3)$$

Eq. (3), the term corr is denoted as the correlation coefficient. It measures how strongly one feature depends on another.

5. DVT DETECTION USING MODIFIED ENSEMBLE CLASSIFIERS

MEC is used for the detection of DVT in the developed model. The detection phase is performed by utilizing efficient techniques like LSTM, DNN, and XGBoost, described as follows. The advanced version of Recurrent Neural Network(RNN) is LSTM [31]. The LSTM input is taken (Fe_n^{weg}) at a given time(t). The essential formation time s is accumulated in the cell memoryMs. The three gates in LSTM are the forget gate $\alpha(s)$, the input gate $\beta(s)$, and the output gate $\gamma(s)$. The hidden layer in LSTM gets the information captured by the prior hidden layer, and the information is stored in the cell memory and computed with the input gate. The cell memory is upgraded when it gets new memory. This computation of the total information collected by the cell memory is performed to initiate the hidden layers. Finally, the LSTM-based detection is obtained accurately.

The DNN [33] has three main components: the hidden layer, the output layer, and the input layer. The data mapping is done with two hidden levels, which is challenging because of the rapid increase in the count of training data. The hidden layer enhances the classification accuracy and the training speed. Finally, the classified outcomes are attained using DNN. XGBoost[32] is a tree-based decision model. The observation 'r' and the predictor 'h' in a dataset are taken, and the tree model has additive functions 'x' to create a new model. These functions compute the relationship between the predicted output and the target output. The proposed MEC-based detection method maximizes DVT detection accuracy. The bounding limit of hidden neuron count ranges from 5-255, the

Journal of Theoretical and Applied Information Technology
15th April 2023. Vol.101. No 7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

limit of the learning rate of XGBoost ranges from 0.01-0.99, and the bounding limit of DNN ranges from 5-255. The main aim of the proposed MEC-based DVT-detected model is given in Eq. (4). Here, the term 'HNL' refers to the optimized hidden neuron count of LSTM, 'HND' refers to the optimized hidden neuron count of DNN, and the term 'LRX' refers to the optimized learning rate of XGBoost. The architectural diagram for the MEC-based DVT detection model is shown in Figure 3.

$$\mathrm{fit}_2\{\mathrm{HNL}, \mathrm{HND}, \mathrm{LRX}\} = \arg\max(\mathrm{Acy}) \quad (4)$$

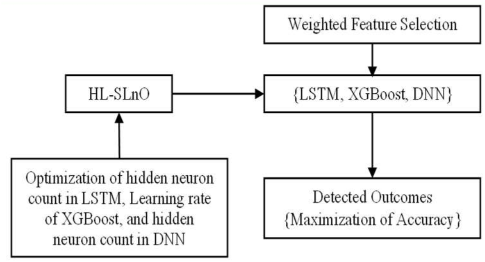
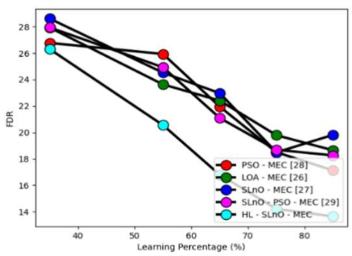
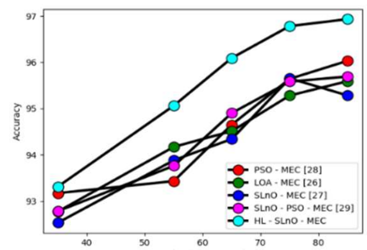


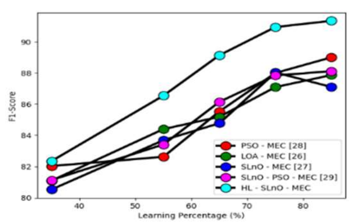
Figure 3: MEC-based DVT detection model



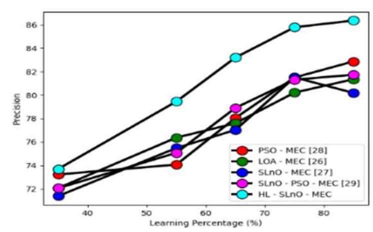
(a)



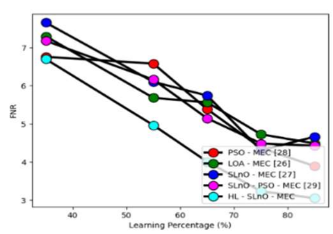
(b)



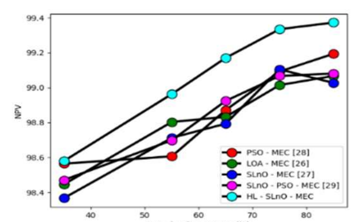
(c)



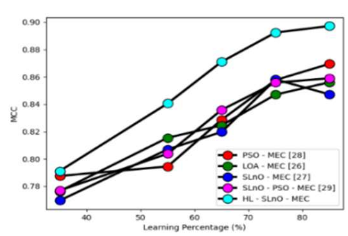
(d)



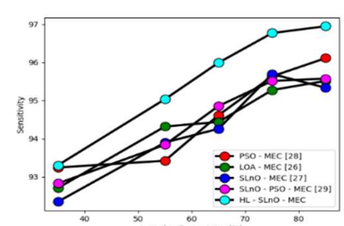
(e)



(f)

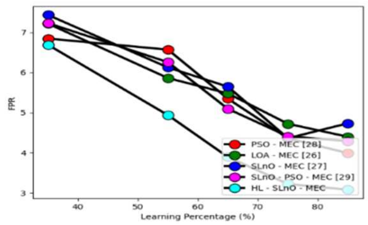


(g)

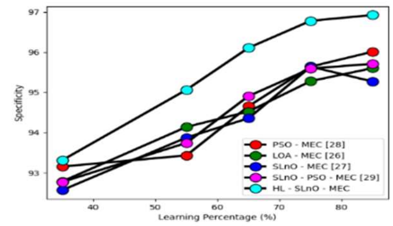


(h)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

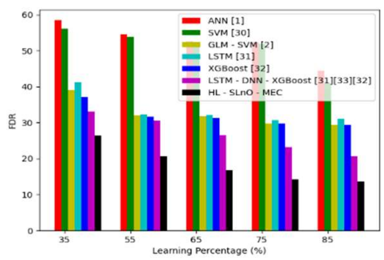


(i)

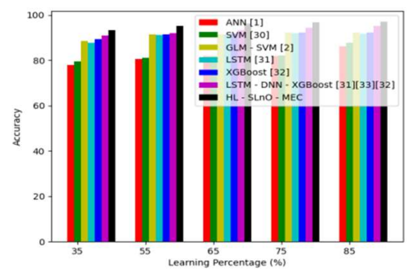


(j)

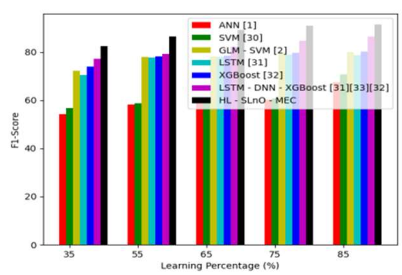
Figure 4: Analysis of the suggested DVT detection method compared to existing optimal algorithms by increasing learning % in terms of (a) FDR, (b) Accuracy, (c) F1-Score, (d) Precision, (e) FNR (f) NPV, (g) MCC, (h) Sensitivity, (i) FPR, (j) Specificity.



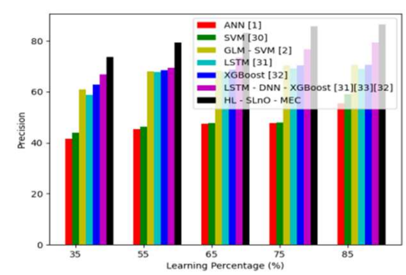
(a)



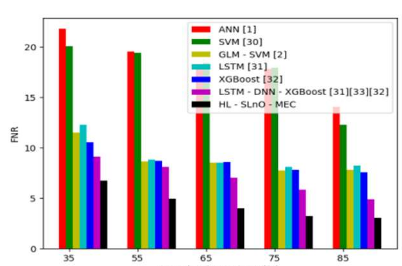
(b)



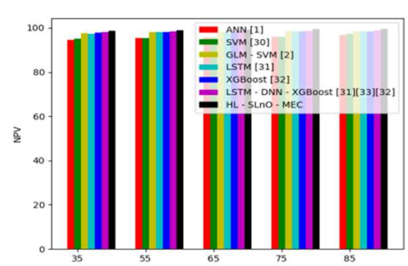
(c)



(d)



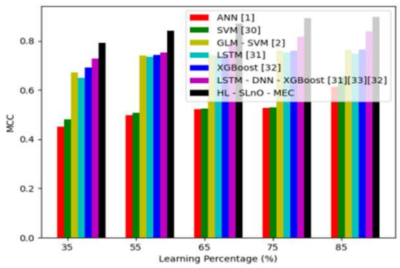
(e)



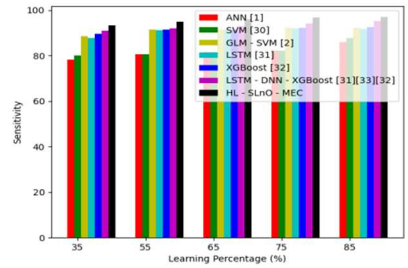
(f)

Journal of Theoretical and Applied Information Technology
15th April 2023. Vol.101. No 7

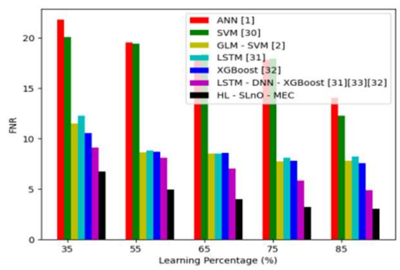
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



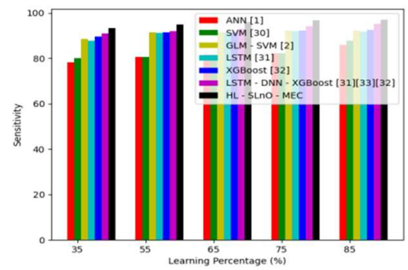
(g)



(h)



(i)



(j)

Figure 5: The suggested DVT detection model was compared to existing machine learning models by adjusting the learning percentage in terms of (a) FDR, (b) Accuracy, (c) F1-Score (d) Precision (e) FNR, (f) NPV, (g) MCC, (h) Sensitivity, (i) FPR, (j) Specificity

	Predicted class	
True class	True Positive(TP)	False Negative(FN)
	False Positive(FP)	True Negative(TN)

Figure 6: Confusion matrix

Table 3: Measures of the suggested DVT prediction's overall effectiveness in comparison to other Meta-heuristic algorithms

Performance Measure	Equation	PSO-MEC [28]	LA-MEC [26]	SLNO-MEC [27]	SLNO-PSO-MEC[29]	HL-SLnO-MEC
Accuracy	$\frac{P_T + N_T}{P_T + P_F + N_T + N_F}$	0.9603	0.9559	0.9528	0.9569	0.9693
Sensitivity	$\frac{P_T}{(P_T + N_F)}$	0.961124	0.955144	0.953349	0.955742	0.969498
Specificity	$\frac{N_T}{N_T + P_F}$	0.960134	0.956052	0.95269	0.957133	0.96926
Precision	$\frac{P_T}{(P_T + P_F)}$	0.828778	0.813551	0.801811	0.817391	0.863612
FPR(False Positive Ratio)	$\frac{P_F}{P_F + N_T}$	0.039866	0.043948	0.04731	0.042867	0.03074
FNR(False Negative Ratio)	$\frac{N_F}{N_F + P_T}$	0.038876	0.044856	0.046651	0.044258	0.030502
NPV(Negative Predictive Value)	$\frac{N_T}{N_T + N_F}$	0.991936	0.990668	0.990265	0.990802	0.993722
FDR(False Discovery Rate)	$\frac{P_F}{(P_T + P_F)}$	0.171222	0.186449	0.198189	0.182609	0.136388
F1-Score	$\frac{P_T}{\left(P_T + 0.5(P_F + N_F)\right)}$	0.890058	0.87868	0.871038	0.881169	0.913497
MCC(Matthews Correlation Coefficient)	$\frac{P_T * N_T - P_F * N_F}{\sqrt{(P_T + P_F)(P_T + N_F)(N_T + P_F)(N_T + N_F)}}$	0.869535	0.856038	0.847143	0.85894	0.897123

Table 4: Overall performance measures of proposed DVT prediction over other deep learning models.

TERMS	ANN [1]	SVM[30]	GLM-SVM[2]	LSTM[31]	XGBOOST[32]	LSTM-DNN-XGBOOST[31][33][32]	HL-SLnO-MEC
Accuracy	0.8612	0.878	0.9226	0.9173	0.9232	0.9503	0.9693
Sensitivity	0.85945	0.877392	0.922249	0.918062	0.924641	0.950957	0.969498
Specificity	0.861551	0.878122	0.922671	0.917147	0.922911	0.950168	0.96926
Precision	0.554826	0.591056	0.705398	0.689888	0.706581	0.793017	0.863612
FPR(False Positive Rtio)	0.138449	0.121878	0.077329	0.082853	0.077089	0.049832	0.03074
FNR(False Negative Ratio)	0.14055	0.122608	0.077751	0.081938	0.075359	0.049043	0.030502
NPV(Negative Predictive Value)	0.968286	0.972732	0.983363	0.982379	0.983871	0.989744	0.993722
FDR(False Discovery Rate)	0.445174	0.408944	0.294602	0.310112	0.293419	0.206983	0.136388
F1-Score	0.674331	0.706307	0.799378	0.787785	0.801036	0.864835	0.913497
MCC(Matthews Correlation Coefficient)	0.614137	0.652648	0.762855	0.749322	0.76498	0.83986	0.897123

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

RESULTS AND DISCUSSIONS

Python was used to run the proposed DVT prediction model, and the experimental results were assessed. The total size of the population was ten, and the maximum count of iterations was ten. The confusion matrix shown in Figure 6 is used to determine the performance of the classification model for a set of data. The proposed model performance metrics were calculated and compared over various meta-heuristic algorithms like PSO [28], LOA [26], SLnO [27], SLnO-PSO [29] and further compared over other deep learning models like ANN [1], SVM [30], GLM-SVM [2], LSTM [31], XGBoost [32], DNN [33] by varying the learning percentage shown in Figure 4, and Figure 5, and illustrated in Table 3 and Table 4.

6. CONCLUSION

The suggested model has evaluated the accurate DVT disease prediction with MEC and weighted feature selection using a hybrid meta-heuristic HL-SLnO algorithm. A data collection of 10,000 suspected DVT cases was collected using the data augmentation approach. The features from the dataset were selected, and then the weighted feature selection was performed with a hybrid meta-heuristic HL-SLnO algorithm. Using the HL-SLnO technique, the weights and features were optimized during the weighted feature selection phase to minimize correlation. Different deep learning models were introduced to detect DVT accurately. The performance of the MEC detection method was enhanced, and the hyperparameters like the hidden neuron count of LSTM, the learning rate of XGBoost, and the hidden neuron count of DNN were optimized with HL-SLnO. The accuracy of the suggested HL-SLnO-MEC was 11% better than PSO-MEC, 14% better than LOA-MEC, 16% better than SLnO-MEC, and 19% better than SLnO-PSO-MEC at a learning percentage of 55. Thus, the proposed DVT prediction model has achieved high detection accuracy than other existing works. In future works, algorithm improvement will be conducted to determine the risk prediction of PE to further help clinicians and people to reduce the mortality rate.

REFERENCES:

[1] Fong-Mata, María Berenice, et al. "An artificial neural network approach and a data augmentation algorithm to systematize the diagnosis of deep-vein thrombosis by using wells' criteria." Electronics 9.11 (2020): 1810.

[2] Li Luo, Ran Kou, Yuquan Feng, Jie Xiang, and Wei Zhu "Cost-Effective Machine Learning Based Clinical Pre-Test Probability Strategy for DVT Diagnosis in Neurological Intensive Care Unit," Clinical and Applied Thrombosis/Hemostasis, Vol. 27, pp. 1-10, 2021.

[3] Logan Ryan, Samson Mataraso, Anna Siefkas, Emily Pellegrini, Gina Barnes, Abigail Green-Saxena, Jana Hoffman, Jacob Calvert, and Ritankar Das, "A Machine Learning Approach to Predict Deep Venous Thrombosis Among Hospitalized Patients," Clinical and Applied Thrombosis/Hemostasis, Vol. 27, pp. 1-7, 2021.

[4] John Willan, Harley Katz, and David Keeling, "The use of artificial neural network analysis can improve the risk-stratification of patients presenting with suspected deep vein thrombosis," British Journal of Haematology, vol. 185, No. 2, February 2019.

[5] María Emilia Frésard, Rafaela Erices, María Loreto Bravo, Mauricio Cuello, Gareth I. Owen, Carolina Ibáñez, Maria Rodriguez-Fernande, "Multi-Objective Optimization for Personalized Prediction of Venous Thromboembolism in Ovarian Cancer Patients," IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 5, pp. 1500-1508, May 2020.

[6] Sheng-Yun Li, Liang Feng, Mei-Juan Xiao, Si-Yan Chen, Jin-Cai He, and Zhen Wang, "Derivation and Validation of a Clinical Prediction Scale for Isolated Distal Deep Venous Thrombosis in Patients after Acute Ischemic Stroke," Journal of Stroke and Cerebrovascular Diseases, Vol. 26, Issue. 10, pp. 2087-2092, October 2017.

[7] Handong Ma, Wenbo Sheng, Jiyu Li, LengchenHou, Jiafang Yang, JunjieCai, Wenxiang Xu, Shaodian Zhang, "A novel hierarchical machine learning model for hospital-acquired venous thromboembolism risk assessment among multiple-departments," Journal of Biomedical Informatics, Vol. 122, October 2021.

[8] T.D. Martins, J.M. Annichino-Bizzacchi, A.V.C. Romano, R. Maciel Filho, "Artificial neural networks for prediction of recurrent venous thromboembolism," International

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Journal of Medical Informatics, Vol. 141, September 2020.

[9] M. Cushman, "Epidemiology and Risk Factors for Venous Thrombosis," Seminars in Hematology, Vol. 44, Issue. 2, pp. 62-69, April 2007.

[10] J.A. Heit, Epidemiology of venous thromboembolism, Nature Reviews Cardiology, vol. 12, issue. 8, pp. 464–474, 2015.

[11] Frederick A. Anderson, Frederick A. Spencer, "Risk Factors for Venous Thromboembolism," Circulation, vol. 107, pp. I-9, 2003.

[12] T. Baglin, "Venous thromboembolism in hospitalised patients: A public health crisis?" British Journal of Haematology, vol. 141, issue. 6, pp. 764–770, 2008.

[13] S. Barbar, F. Noventa, V. Rossetto, A. Ferrari, B. Brandolin, M. Perlati, E.D. Bon, D. Tormene, A. Pagnan, P. Prandoni, "A risk assessment model for the identification of hospitalized medical patients at risk for venous thromboembolism: The Padua Prediction Score, Journal of Thrombosis and Haemostasis, vol. 8, issue. 11, pp. 2450–2457, 2010.

[14] N. van Es, M. Di Nisio, G. Cesarman, A. Kleinjan, H.-M. Otten, I. Mah´e, I.T. Wilts, D. C. Twint, E. Porreca, O. Arrieta, A. St´epanian, K. Smit, M. De Tursi, S.M. Bleker, P. M. Bossuyt, R. Nieuwland, P.W. Kamphuisen, H.R. Büller, "Comparison of risk prediction scores for venous thromboembolism in cancer patients: A prospective cohort study," Haematologica, vol. 102, issue. 9, pp. 1494–1501, 2017.

[15] N.A.Zakai, L.A. McCLURE, "Racial differences in venous thromboembolism", Journal of Thrombosis and Haemostasis, vol. 9, issue. 10, pp. 1877–1882, 2011.

[16] P. Ferroni, F.M. Zanzotto, N. Scarpato, S. Riondino, F. Guadagni, M. Roselli, "Validation of a Machine Learning Approach for Venous Thromboembolism Risk Prediction in Oncology", Disease Markers, vol. 2017, pp. 1–7, 2017.

[17] J.S. Kim, R.K. Merrill, V. Arvind, D. Kaji, S.D. Pasik, C.C. Nwachukwu, L. Vargas, N. S. Osman, E.K. Oermann, J.M. Caridi, S.K. Cho, "Examining the Ability of Artificial Neural Networks Machine Learning Models to Accurately Predict Complications Following Posterior Lumbar Spine Fusion," Spine, vol. 43, issue. 12, pp. 853–860, 2018.

[18] Y. Yang, X. Wang, Y. Huang, N. Chen, J. Shi, T. Chen, "Ontology-based venous thromboembolism risk assessment model developing from medical records," BMC Medical Informatics and Decision Making, vol. 19, issue. 4, 2019.

[19] S. Eichinger, G. Heinze, and P.A. Kyrle, "d-Dimer Levels Over Time and the Risk of Recurrent Venous Thromboembolism: An Update of the Vienna Prediction Model," Journal of the American Heart Association, vol. 3, 2014.

[20] C. Kearon, S.R. Kahn, G. Agnelli, S. Goldhaber, G.E. Raskob, and A.J. Comerota, "Antithrombotic Therapy for Venous Thromboembolic Disease," CHEST, vol. 133, pp. 454S-545S, 2008.

[21] M. G. Beckman, W. C. Hooper, S. E. Critchley, and T. L. Ortel, "Venous thromboembolism: A public health concern," American journal of preventive medicine, Apr. 2010.

[22] J. R. Zech, M. A. Badgeley, M. Liu, A. B. Costa, J. J. Titano, and E. K. Oermann, "Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study," PLOS Medicine, 2020.

[23] S. Z. Goldhaber and H. Bounameaux, "Pulmonary embolism and deep vein thrombosis," The Lancet Journal (London, England), vol. 379, no. 9828, pp. 1835-1846, May 2012.

[24] Tarek Nafee, C. Michael Gibson, Ryan Travis, Megan K. Yee, Mathieu Kerneis, Gerald Chi, Fahad AlKhalfan, Adrian F. Hernandez, Russell D. Hull, Ander T. Cohen, Robert A. Harrington, Samuel Z. Goldhaber, "Machine learning to predict venous thrombosis in acutely ill medical patients," Research and Practice in Thrombosis and Haemostasis, vol. 4, issue. 4, January 2020.

[25] Shrey Modi, Ryan Deisler, Karen Gozel, Patty Reicks, Eric Irwin, Melissa Brunsvold, KaysieBanton and Greg J. Beilman, "Wells criteria for DVT is a reliable clinical tool to assess the risk of deep venous thrombosis in trauma patients," World Journal of Emergency Surgery, vol. 11, issue. 1, June 2016.

[26] Rajakumar Boothalingam, "Optimization using lion algorithm: a biological inspiration from lion's social behavior", Evolutionary Intelligence, vol.11, no. 1-2, pp.31–52, 2018.

[27] Kamal Kumar Gola, Nishant Chaurasia, Bhumika Gupta, Deepak Singh Niranjan, "Sea lion optimization algorithm based node deployment strategy in underwater acoustic sensor network", vol. 34, Issue 5, 2021.

[28] K. Mistry, L. Zhang, S. C. Neoh, C. P. Lim and B. Fielding, "A Micro-GA Embedded PSO Feature Selection Approach to Intelligent Facial

Emotion Recognition," IEEE Transactions on Cybernetics, vol. 47, no. 6, pp. 1496-1509, June 2017.

[29] Dr. NidhalKamelTaha El-Omari, "Sea Lion Optimization Algorithm for Solving the Maximum Flow Problem", IJCSNS International Journal of Computer Science and Network Security, vol.20, No.8, August 2020.

[30] BohyungHan; Larry S. Davis, "Density-Based Multifeature Background Subtraction with Support Vector Machine", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 5, 2012.

[31] Swapna G, Soman KP, VinayakumarR, "Automated detection of diabetes using CNN and CNN-LSTM network and heart rate signals", Procedia Computer Science, vol. 132, pp. 1253–1262, 2018.

[32] ZhenlongJiang, Ran Ji, Kuo-Chu Chang, "A Machine Learning Integrated Portfolio Rebalance Framework with Risk-Aversion Adjustment", e Machine Learning Applications in Finance, vol. 7, pp. 155, 2020.

[33] PelinAlaboz, OrhanDengiz, Sinan Demir, HüseyinS¸enol, "Digital mapping of soil erodibility factors based on decision tree using geostatistical approaches in terrestrial ecosystem", Catena, vol. 207, August 2021.