Cognitive Computation (2024) 16:2047–2062
https://doi.org/10.1007/s12559-023-10115-w

EEG Signal Classification Using a Novel Universum-Based Twin Parametric-Margin Support Vector Machine

Barenya Bikash Hazarika[1] · Deepak Gupta[2] · Bikram Kumar[2]

Received: 21 April 2022 / Accepted: 14 January 2023 / Published online: 30 January 2023

Abstract
The Universum data, which indicates a sample that does not belong to any of the classes, has been proved to be useful in supervised learning. The researchers have explored the support vector machine (SVM) and its twin variants by embedding them with Universum data for classifying the electroencephalogram (EEG) signal. To improve generalization performance even further, this paper presents a novel twin parametric margin SVM based on Universum data (UTPMSVM) for classifying EEG signals. The proposed UTPMSVM forms a pair of non-parallel parametric hyperplanes that solves two small SVM-type problems. The addition of prior information, i.e., the Universum data, boosts the performance of the model. The dimensionality of the EEG datasets is reduced using principal component analysis (PCA), independent component analysis (ICA), and wavelet analysis. Experimental simulations have been carried out on 14 EEG datasets as well as 30 real-world datasets. The classification performance of the proposed model is compared with Universum-based SVM (USVM), Universum non-parallel hyperplane-based SVM (UNHSVM), TPMSVM, and angle-based Universum least squares twin SVM (AULSTSVM) models. Further two different statistical tests are performed to evaluate the performance of the proposed model. For EEG datasets, the UTPMSVM showed the highest accuracy of 78% and the highest F1-score of 0.78658. Moreover, for the real-world datasets, the proposed UTPMSVM showed the highest accuracy of 100%. In addition to that, it is observed that the mean accuracy and F1-score of UTPMSVM are comparatively better than USVM, UNHSVM TPMSVM, and AULSTSVM. The results demonstrate the applicability of UTPMSVM for EEG signal classification problems as well as real-world data classification problems.

Keywords EEG signal · Classification · Universum data · Parametric margin · Twin support vector machine

Introduction

Electroencephalogram (EEG) signals are generally used in the medical field to detect brain-related problems. Based on the results given by EEG signals, necessary actions are taken to resolve those brain-related problems. Our brain cells interconnect with one another via electrical signals, and these cells always remain active. EEG signals record the electrical changes in brain activity which is very useful in the diagnosis of brain-related problems. In EEG, a small-sized metal device called electrodes is placed over the scalp of the subject which records the electrical activity of the brain and passes it to the computer to store those records. The placement of the electrodes over the scalp plays an important role in fruitful diagnoses. Specialists like brain surgeons, psychiatrists, and neurologists realize that EEG is a beneficial diagnostic and can be helpful to predict certain clinical issues. EEG was fundamentally developed for the diagnosis of epilepsy. Epilepsy is one of the baleful neurological disorders in which the brain activity of a human becomes abnormal which causes seizures, sensation, and loss of cognizance as well. Caton disclosed the electrical activity of the brain for monkeys and rabbits in 1875 [1]. Then, Beck [2] studied the electrical signals of the mind

✉ Deepak Gupta
deepakjnu85@gmail.com; deepak@nitap.ac.in

Barenya Bikash Hazarika
barenya1431@gmail.com

Bikram Kumar
biikramkumar@gmail.com

[1] Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

[2] Department of Computer Science & Engineering, National Institute of Technology, Arunachal Pradesh, India

for dogs and rabbits in 1890. Finally, Berger, the German psychologist and scientist recorded the EEG of a human for the very first time in 1924 [3, 4]. One of the most significant advantages of EEG is the ability to observe brain activity in real-time, at the millisecond level, which is not probable with other high-resolution imaging techniques [5]. EEG measures both amplitude and frequency [6]. Due to this advantage, EEG is very popular in the researcher's community to deal with brain-related problems and ameliorate the diagnoses of different mental issues. But EEG data contains an enormous amount of noise whose effect should be avoided during modeling. After the removal of noisy data, many different machine learning approaches could be implemented on the processed data to detect brain-related problems. To select the most important features of EEG signals, several feature extraction techniques, namely, principal component analysis (PCA) [7], independent component analysis (ICA) [8], wavelet-transform, and others, are be performed.

EEG signals are extremely complex for a non-professional observer to interpret and draw conclusions about various brain-related issues. For this reason, there is a need for automatic EEG signal interpretation that can help in the early prediction of brain-related problems. In these days, the selection of a proper ML model is one of the most difficult tasks as there are many models available. Support vector machine (SVM) [9] is one such legendary algorithm which is based on the principle of structural risk minimization (SRM). SVM can be used for performing various tasks related to both classification and regression. SVM constructs a hyperplane that separates the data points based on the maximal margin hyperplane. One of the most important advantages of SVM is that it does not suffer from the problem of local minima unlike artificial neural networks (ANN). SVM can also work with very high dimensional data efficiently and does not suffer from the curse of dimensionality. Hence, SVM is used by researchers to solve an extensive range of real-world problems. Yeo et al. [10] applied SVM to detect the car driver's drowsiness while driving a car using EEG data. Subasi and Gursoy [11] used ICA, PCA, and linear discriminant analysis (LDA) to extract the important features from the EEG data and then applied SVM to detect epileptic seizures. Recently, Afifi et al. [12], performed melanoma detection, using the SVM model. Although SVM has several advantages, the key demerit of SVM is that it takes high computational time for large-scale size datasets as it solves a large quadratic programming problem (QPP) [13].

To solve the problem of SVM and to improve the generalization performance, several variations of SVM have been suggested in the literature. Twin SVM (TWSVM) [14] is one such work which is influenced by generalized eigenvalue proximal SVM (GEPSVM) [15]. TWSVM searches for two non-parallel hyperplanes where each of the hyperplanes is close to one class and as far as possible from the other class. Whereas SVM solves a large QPP, TWSVM solves a pair of small QPPs. As a result, it lowers the computational cost and makes TWSVM nearly four times faster than conventional SVM. Several extensions of TWSVM are proposed by researchers such as least square TWSVM (LSTWSVM) [16], improved TWSVM [17], robust TWSVM [18], robust twin bounded SVM [19], and density-weighted TWSVM [20]. Peng developed a twin parametric margin SVM (TPMSVM) [21] in which two non-parallel parametric-margin (PM) hyperplanes are generated, which are solved by two smaller size SVM-type problems. TPMSVM constructs two PM hyperplanes so that each one decides the positive or negative PM, whereas this is not the case in TWSVM as discussed above, and the QPPs for these two methods are completely different. Peng et al. [22] proposed another variant of TPMSVM called structural TPMSVM (STPMSVM) where the structural information of data was taken into consideration. Furthermore, Peng et al. [23] found out that the decision function of traditional TPMSVM losses the sparsity, and hence, they developed another method, i.e., centroid-based TPMSVM (CTPMSVM) where the decision hyperplane becomes sparse as it optimizes the projection values of the centroid points of the target classes. Shao et al. [24] suggested another variant of TPMSVM termed least squares TPMSVM (LSTPMSVM). Recently, Gupta et al. [25] suggested a novel classifier which is based on TPMSVM and FSVM called fuzzy-based Lagrangian TPMSVM to analyze biomedical data.

Several studies in the literature prove that incorporating prior information about the data distribution to the classifier drastically improves the performance of the same. Universum data, along with the SVM classifier, serves as prior information about the data distribution in USVM [26]. It is believed that the Universum data should not belong to any of the concerned classes and must fall in between the target classes. The concept of Universum data is used to solve many real-world problems due to its higher generalization performance [27, 28]. But it cannot be concluded that the Universum data will always lead to a high generalization performance. Motivated by TWSVM and USVM, Qi et al. [29] developed a new methodology utilizing the benefits of both TWSVM and USVM called Universum TWSVM (UTWSVM). Richhariya and Gupta [30] used iterative UTWSVM to classify the facial expressions automatically. Recently, Zhao et al. [31] proposed an efficient non-parallel hyperplane-based USVM (UNHSVM) for classification. Furthermore, a fuzzy USVM has been proposed to enhance prior information by assigning weights to Universum points based on information entropy [32]. A reduced Universum TWSVM is implemented to address the class imbalance problems [33]. Recently Kumar and Gupta [34] proposed a novel Universum-based Lagrangian twin bounded SVM for EEG signal classification. Moosaei et al. [35] suggested an Universum parametric margin v-SVM for classification. Richhariya et al. [36] diagnosed disease using a new USVM based on recursive feature elimination (USVM-RFE). For solving the same problem, Richhariya and Tanveer [37] suggested a fuzzy Universum least

squares TWSVM (FULSTSVM). Moosaei and Hladik [38] suggested a Lagrangian-based method for Universum twin bounded SVM. Riccharіya and Tanveer [39] suggested a novel angle-based Universum LSTSVM (AULSTSVM) for classification. Ganaie et al. [40] proposed a k-nearest neighbor weighted reduced UTWSVM for imbalanced data classification problems (KWRUTSVM-CIL). To take advantage of the interclass information, weight vectors are used in the KWRUTSVM-CIL's corresponding constraints of the objective functions.

Inspired by the previous works of Qi et al. [29], Moosaei et al. [35], Richhariya et al. [36], and Peng [21] and to utilize the benefit of Universum data, a new variant of TPMSVM has been proposed in this paper called Universum-based TPMSVM (UTPMSVM). In UTPMSVM, the slack variables are taken in 2-norm rather than 1-norm which formulates a strongly convex problem; hence, it always leads to unique solutions. The proposed method contains regularization terms which prevent it from the problem of overfitting. The proposed UTPMSVM intends to generate two nonparallel hyperplanes, each of which decides whether the separating hyperplane has a positive or negative parametric margin. Like the UTWSVM, the UTPMSVM also solves two smaller-sized QPPs for this purpose rather than solving a larger one, unlike traditional SVM or USVM. In this paper, seizure EEG signals and healthy EEG signals are considered for the classification. Interictal data falls between the seizure and healthy signals. So, interictal data has been utilized as Universum data. As stated above EEG data contains lots of noise and outliers, to get rid of those problems, many feature extraction methods have been applied, namely, PCA, ICA, and wavelet transform. To prove the acceptability of the proposed classifier, it is applied to several well-known real-world datasets. The results of the proposed UTPMSVM are compared with USVM, UNHSVM, TPMSVM, and AULSTSVM. The Universum points used in our UTPMSVM method come directly from the EEG dataset. As the Universum, we use the interictal or seizure-free signals from the EEG dataset. This more effectively provides the necessary prior information to the TPMSVM classifier since the variation of the seizure-free state signal occurs between the variation of the healthy and epileptic EEG signals. Since our Universum data is not derived from training data, there are no outliers in the Universum data, and no noise from training data is present [26]. The main contributions of the work are as follows:

- A novel UTPMSVM is proposed to classify seizures and healthy EEG signals.
- UTPMSVM incorporates prior knowledge regarding the data distribution from interictal EEG signals.
- Three different feature extractors have been applied to extract the most important features.
- Statistical analysis is performed to reveal the superiority of the proposed method over other related models.

Related Work

In this section, we have discussed a few related models. They are USVM and TPMSVM. Moreover, the proposed UTPMSVM is also elaborated.

Universum Support Vector Machine (USVM)

Weston et al. [41] proposed a variant of classical SVM named Universum SVM (USVM) for binary classification problem by incorporating Universum data. Universum data are treated as non-examples that do not belong to any of the concerned classes. The idea of Universum is somewhat similar to the Bayesian idea. However, there is an important applied distinction between these two concepts. If there should arise an occurrence of Bayesian deduction, the earlier information is about the information on choice guidelines, whereas Universum data is the information about the collection of examples. The kernel function used here is $k(z_p, z_a) = \phi(z_p)^t \phi(z_a)$ where ϕ is the mapping function. The QPP of USVM can be expressed as,

$$min\ C_k \sum_{a=1}^{2|k|} \chi_a + \frac{1}{2}||w||^2 + C(\sum_{p=1}^{m} \sigma_p)$$

subject to,

$$\begin{aligned} &y_p\left(\phi\left(z_p\right)^t w + g\right) \geq 1 - \kappa_p, \kappa_p \geq 0, \forall p = 1, \ldots \ldots .., m \\ &y_a\left(\phi\left(z_a\right)^t w + g\right) \geq -\varepsilon - \chi_d, \chi_d \geq 0, \forall a = 1, ..., 2|k| \end{aligned} \tag{1}$$

where m is the total number of data-points; χ, κ are the slack variables; C, C_k are the penalty terms; k is the total number of Universum data-points; and tolerance value of Universum is ε.

Now, the dual primal problem is obtained using Lagrangian multipliers (LMs) and further implementing Karush–Kuhn–Tucker (KKT), is shown as,

$$max \sum_{p=1}^{m+2|k|} \kappa_l \eta_p^* - \frac{1}{2} \sum_{p=1}^{m+2|k|} \sum_{a=1}^{m+2|k|} \eta_l^* \eta_a^* k(z_p, z_a) y_p y_a$$

subject to,

$$\begin{aligned} &0 \leq \eta_p^* \leq C, \text{where } \forall p = 1, ..., m \\ &\kappa_p = 1, \text{where } \forall p = 1, ..., m \\ &0 \leq \eta_p^* \leq C_k, \text{where } \forall p = n+1, ..., m+2|k| \\ &\kappa_p = -\varepsilon, \text{where } \forall p = n+1, ..., m+2|k| \\ &\text{and } \sum_{p=1}^{m+2|k|} \eta_p^* y_p = 0 \end{aligned} \tag{2}$$

Now, let us suppose $z \in R^m$ is a new test instance. The function that will determine the class label of that instance as,

$$f(z) = sign\left(\sum_{p=1}^{m+2|k|} \eta_p^* y_p k(z_p, z_a) + g\right) \tag{3}$$

Twin Parametric-Margin Support Vector Machine (TPMSVM)

TPMSVM determines its non-parallel margin hyperplanes by solving a pair of QPPs. Let us suppose we have a binary classification problem where we have two different classes of data i.e., +1 and -1 respectively. Let us assume that the number of datapoints belonging to +1 class is k_1 and the number of datapoints belonging to −1 class is k_2. Let, D_1 and D_2 be two matrices represents the datapoints belonging to +1 and -1 class.

For the linear case, TPMSVM constructs two hyperplanes:

$$f_1(z) = w_1 z + g_1 = 0 \text{ and } f_2(z) = w_2 z + g_2 = 0 \tag{4}$$

By introducing the positive and negative parametric-margin hyperplanes, data will be separated by TPMSVM if:

$$\begin{aligned} w_1 z_i + c_1 \geq 0, i = 1, 2, ..., k_1 \\ w_2 z_i + c_2 \geq 0, i = 1, 2, ..., k_2 \end{aligned} \tag{5}$$

To find out the margins, we need to solve the following optimization problem:

$$min \frac{1}{2}||w_1||^2 + \frac{\lambda_1}{k_2} e_2^t (D_2 w_1 + e_2 g_1) + \frac{d_1}{k_1} e_1^t \rho$$

subject to,

$$\begin{aligned} D_1 w_1 + e_1 g_1 \geq 0 - \rho \\ \rho \geq 0 e_1 \end{aligned} \tag{6}$$

and

$$min \frac{1}{2}||w_2||^2 - \frac{\lambda_2}{k_1} e_1^t (D_1 w_2 + e_1 g_2) + \frac{d_2}{k_2} e_2^t \kappa$$

subject to,

$$-\left(D_2 w_2 + e_2 g_2\right) \geq 0 - \kappa, \kappa \geq 0 e_2 \tag{7}$$

where ρ and κ are slack variables; $d_1, d_2 \geq 0$, λ_1, λ_2 are the regularization terms; and e_1 and e_2 are the vectors of 1 s of appropriate dimension.

The dual QPPs of (6) and (7) is as follows:

$$max - \frac{1}{2}\theta^t D_1 {D_1}^t \theta + \frac{\lambda_1}{k_2} \theta^t D_1 {D_2}^t e_2$$

subject to,

$$e_1^t \theta = \lambda_1 \text{ and } 0 \leq \theta \leq \frac{d_1}{k_1} e_1 \tag{8}$$

and

$$max - \frac{1}{2}\omega^t D_2 {D_2}^t \omega + \frac{\lambda_2}{k_1} \omega^t D_2 {D_1}^t e_1$$

subject to,

$$e_2^t \omega = \lambda_2 \text{ and } 0 \leq \omega \leq \frac{d_2}{k_2} e_2 \tag{9}$$

where θ and ω are LMs. After solving Eqs. (16) and (17), we will get the vector of LMs, and then, we can compute w_1, w_2, g_1, and g_2. Finally, we use the following function to determine the class label of a new test instance $z \in R^m$:

$$f(z) = sign\left(\frac{w_1 z + g_1}{||w_1||} + \frac{w_2 z + g_2}{||w_2||}\right) \tag{10}$$

Proposed Universum-Based Twin Parametric Margin Support Vector Machine

In this paper, we have proposed an efficient classifier called Universum-based twin parametric margin SVM (UTPMSVM) for classifying EEG signals. In the formulation of the proposed classifier, we have used the L2-norm instead of the L1-norm. Here, $K(x^t, M^t)w_1 + g_1 = 0$ and $K(x^t, M^t)w_2 + g_2 = 0$ are the two non-parallel hyperplanes which are measured by the primal problems of UTPMSVM,

$$min \frac{1}{2}(||w_1||^2 + g_1^2) + c_1 e_2^t (K(B, D^t)w_1 + e_2 g_1) + \frac{c_2}{2}\xi^t \xi + \frac{c_3}{2}\psi_1^t \psi_1$$

subject to,

$$\begin{aligned} K(A, D^t)w_1 + e_1 g_1 \geq 0 - \xi \\ K(U, D^t)w_1 + e_u g_1 + (1 - \varepsilon)e_u \geq \psi_1 \end{aligned} \tag{11}$$

$$\text{and } min \frac{1}{2}(||w_2||^2 + g_2^2) - c_4 e_1^t (K(A, D^t)w_2 + e_1 g_2) + \frac{c_5}{2}\eta^t \eta + \frac{c_6}{2}\psi_2^t \psi_2$$

subject to,

$$\begin{aligned} K(B, D^t)w_2 + e_2 g_2 \leq \eta \\ -(K(U, D^t)w_2 + e_u g_2 + \psi_2 \geq (-1 + \varepsilon)e_u \end{aligned} \tag{12}$$

Assume that,

$$u_1 = \begin{bmatrix} w_1 \\ g_1 \end{bmatrix}, u_2 = \begin{bmatrix} w_2 \\ g_2 \end{bmatrix},$$
$$G = [K(B, D^t e_2), H = [K(A, D^t)e_1],$$
$$V = \left[K\left(U, D^t e_u\right)\right] \text{ and } E = (1 - \varepsilon)e_u$$

After substituting the above assumptions in Eqs. (14) and (15), it become,

$$max\ L_1 = \frac{1}{2}u_1^t u_1 + c_1 e_2^t G\ u_1 + \frac{c_2}{2}\xi^t\xi + \frac{c_3}{2}\psi_1^t\psi_t - \alpha_1^t(Hu_1 + \varepsilon) - \alpha_2^t(Vu_1 + E + \psi_1)$$

subject to,

$$\begin{aligned} Hu_1 &\geq -\xi \\ Vu_1 + E &\geq -\psi_1 \end{aligned} \tag{13}$$

and

$$min\frac{1}{2}u_2^t u_2 - c_4 e_1^t H\ u_2 + \frac{c_5}{2}\eta^t\eta + \frac{c_6}{2}\psi_2^t\psi_2$$

subject to,

$$\begin{aligned} -Gu_2 + \eta &\geq 0 \\ -Vu_2 + \psi_2 + E &\geq 0 \end{aligned} \tag{14}$$

The Lagrangian of Eqs. (13) and (14) is as follows:

$$max\ L_1 = \frac{1}{2}u_1^t u_1 + c_1 e_2^t G\ u_1 + \frac{c_2}{2}\xi^t\xi + \frac{c_3}{2}\psi_1^t\psi_t - \alpha_1^t(Hu_1 + \varepsilon) - \alpha_2^t(Vu_1 + E + \psi_1) \tag{15}$$

and

$$max\ L_2 = \frac{1}{2}u_2^t u_2 - c_4 e_1^t H\ u_2 + \frac{c_5}{2}\eta^t\eta + \frac{c_6}{2}\psi_2^t\psi_2 - \alpha_1^t(-Gu_2 + \eta) - \alpha_2^t(-Vu_2 + E + \psi_2) \tag{16}$$

where α_1 and α_2 are the LMs.

From Eq. (15) we get,

$$\frac{\partial L_1}{\partial u_1} = 0 \Rightarrow u_1 = H^t\alpha_1 + V^t\alpha_2 - c_1 G^t e_2 \tag{17}$$

$$\frac{\partial L_1}{\partial \xi} = 0 \Rightarrow \xi = \frac{\alpha_1}{c_2} \tag{18}$$

$$\frac{\partial L_1}{\partial \psi_1} = 0 \Rightarrow \psi_1 = \frac{\alpha_2}{c_3} \tag{19}$$

After substituting the values of (17), (18), and (19) in Eq. (15), we achieve,

$$max\ L_1 = -\frac{1}{2}\alpha^t \begin{bmatrix} (HH^t + I/c_2)HV^t \\ VH^t(VV^t + I/c_3) \end{bmatrix} \alpha + (c_1 e_2^t G Z_1^t - [0\ E])\alpha \tag{20}$$

where $Z_1 = \begin{bmatrix} H \\ V \end{bmatrix}$.

Further, from Eq. (16), we get,

$$\frac{\partial L_2}{\partial u_2} = 0 \Rightarrow u_2 = c_4 H^t e_1 - G^t\alpha_1 - V^t\alpha_2 \tag{21}$$

$$\frac{\partial L_2}{\partial \eta} = 0 \Rightarrow \eta = \frac{\alpha_1}{c_5} \tag{22}$$

$$\frac{\partial L_2}{\partial \psi_2} = 0 \Rightarrow \psi_2 = \frac{\alpha_2}{c_6} \tag{23}$$

After putting the values of (21), (22), and (23) in Eq. (16), we get

$$max\ L_2 = -\frac{1}{2}\alpha^t \begin{bmatrix} (GG^t + I/c_5)GV^t \\ VG^t(VV^t + I/c_6) \end{bmatrix} \alpha + (c_4 e_1^t H Z_2^t - [0\ E])\alpha \tag{24}$$

where $Z_2 = \begin{bmatrix} G \\ V \end{bmatrix}$.

After finding the values of the Lagrangian parameters, we can find the values of the parameters w_1, w_2, g_1, g_2 using the following expressions:

$$\begin{bmatrix} w_1 \\ g_1 \end{bmatrix} = H^t\alpha_1 + V^t\alpha_2 - c_1 G^t e_2$$

and

$$\begin{bmatrix} w_2 \\ g_2 \end{bmatrix} = c_4 H^t e_1 - G^t\alpha_1 - V^t\alpha_2$$

Experimental Setup, Results, and Analysis

Experimental simulations have been performed on a 64-bit windows OS based computer with 4 GB RAM and i5 processor. We have non-linear (NL) kernel for the experiments. The Gaussian kernel which may be expressed as $k(x_h, x_j) = -\exp(-\mu||x_h - x_j||^2)$, where x_h, x_j represents the samples, is used as the non-linear kernel. The C and μ parameters of USVM, UNHSVM, TPMSVM, AULST-SVM, and UTPMSVM are selected from $\{10^{-3}, ..., 10^3\}$ and $\{2^{-5}, 2^{-3}, 2^{-2}, .., 2^3, 2^5\}$. Also, the c_7 parameter of AULST-SVM is selected from $\{0.1, 0.2, ..., 1\}$. Moreover, the ε parameter for the proposed UTPMSVM classifier is chosen from $\{0.1, 0.3, 0.5, 0.7, 0.9\}$. The classification performance, as well as the optimal parameters, is computed using a fivefold cross-validation method. The MOSEK optimization toolbox is utilized to solve the QPPs of USVM, UNHVM, and UTPMSVM [42]. The performance of the classifiers is evaluated using accuracy and F_1-score which can be determined as

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

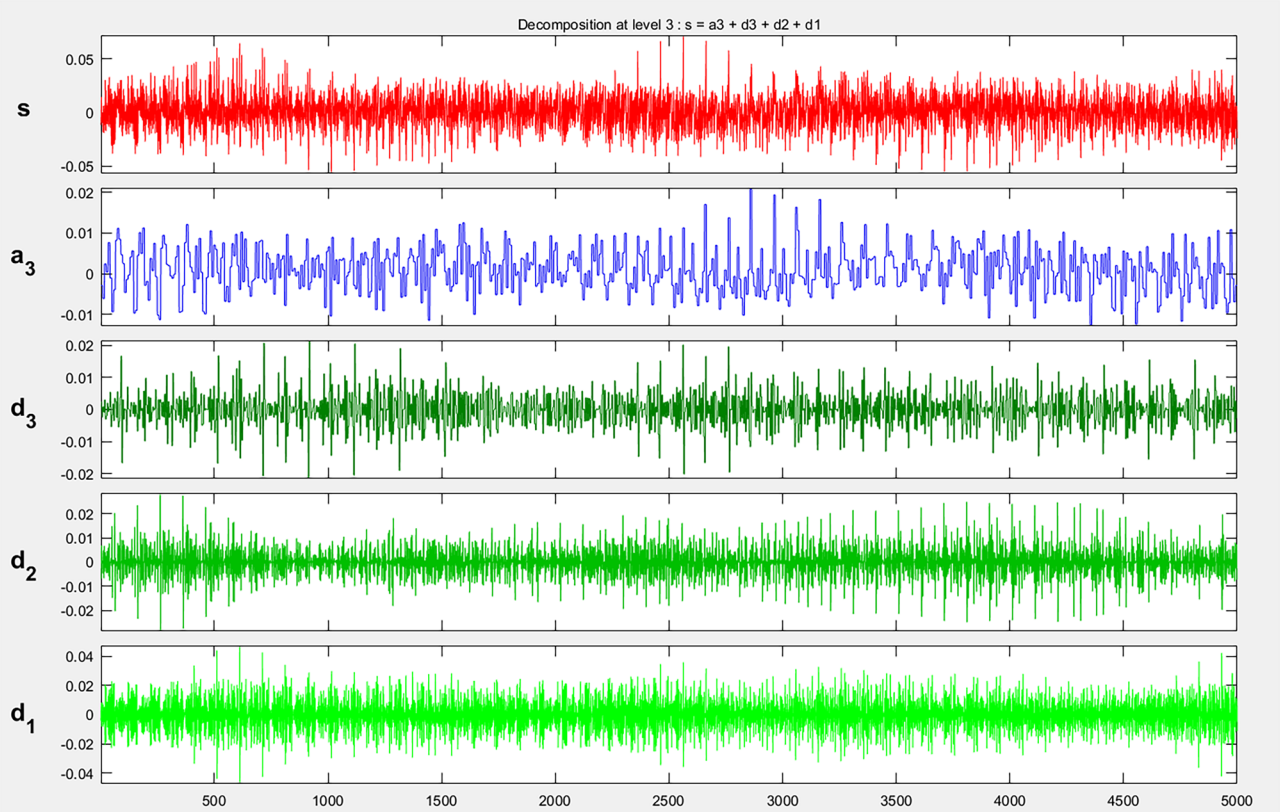

Fig. 1 Decomposition of N signal using db1 at level 3

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$F_1 - score = \frac{2(precison \times recall)}{precision + recall}$$

EEG Signal Classification

We have considered the classification between the healthy and seizure signals in this work. The EEG signal dataset is collected from [43] which consists of Z, N, O, F, and S signals. Each set consists of 100 single-channel EEG signals that were each collected for 23.6 s at a sampling rate of 173.61 Hz. Five healthy participants' surface EEG recordings are shown in sets Z and O, with their eyes open and closed, respectively. In sets N and F, the hippocampal formation of the opposing hemisphere of the brain and the epileptogenic zone, respectively, were the recording sites for five patients in the interictal stage. Seizure recordings from all the recording sites exhibiting ictal activity make up the set S for the ictal state. For N, F, and S, intra-cranial EEG recording is the mode used. N is used as Universum data that appears between seizure and healthy signals. The dataset S was compiled from seizure recording destinations with physiological movement. Each training and testing session contains 100 samples. One of the main reasons for the model's poor performance is its high dimensionality. To solve this issue, we used common feature extraction techniques such as PCA, ICA, and wavelet transform (WT). NL kernels are used to reduce the dimensions in PCA. Few wavelet families are used to implement the discrete wavelet transform (DWT) [44] at various stages of decomposition. The approximation and decomposition coefficients are combined to form the function vector. The degree of decomposition for Daubechies wavelets db1, db2, db4, and db6 is set to level-3. Figures 1, 2, 3, 4 shows the decomposition of db1 and db2 at level 3 for N and S signals. Here, Figs. 1 and 2 indicate decomposition using db1 at level 3 for signals N and S respectively. Moreover, Figs. 3 and 4 indicate decomposition using db2 at level 3 for signals N and S respectively.

PCA is used to reduce dimensionality in the case of ICA and WT. ICA is used in the same manner as in [34, 45]. The class discriminatory ratio (CDR) is used to filter the PCA components and select the most appropriate PCA components. The proposed framework is presented in Fig. 5. The

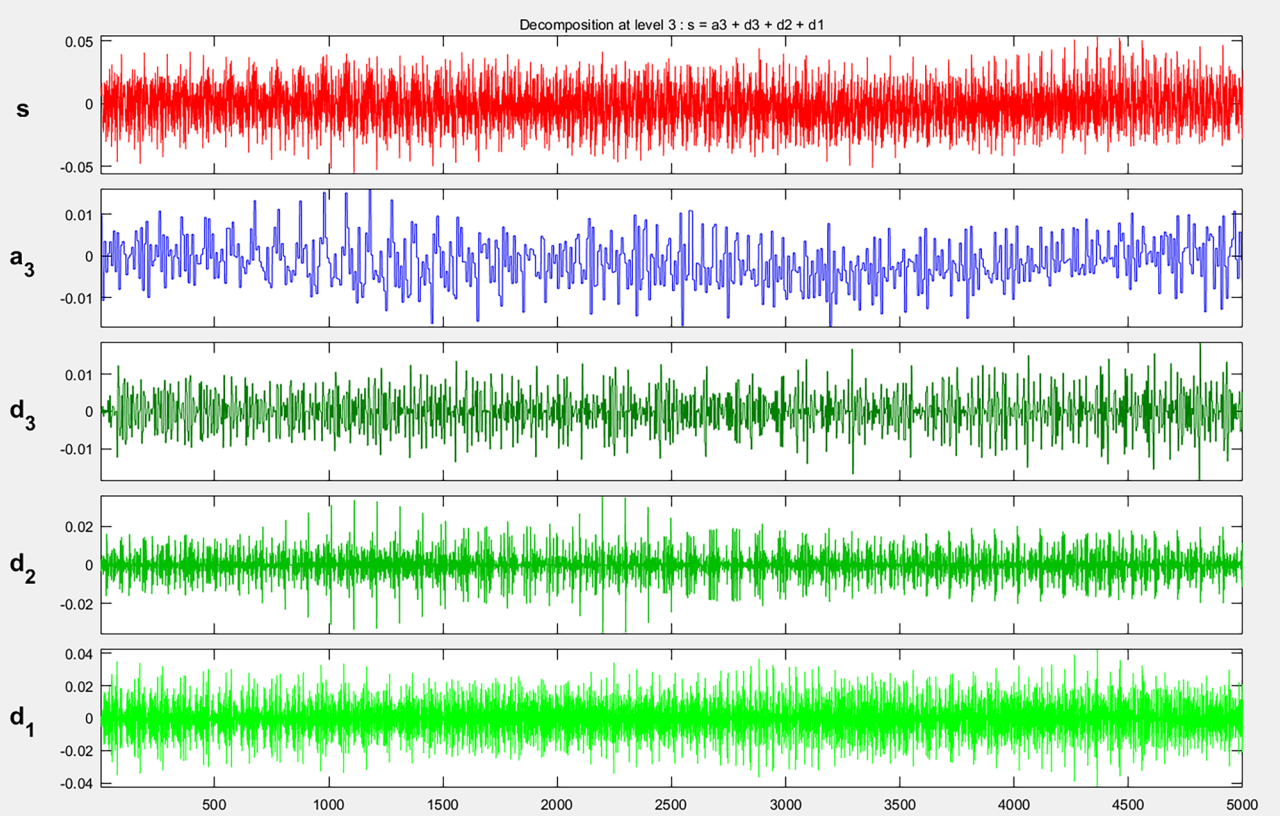

Fig. 2 Decomposition of S signal using db1 at level 3

EEG signal is passed as an input, and the features of the EEG signal are extracted using PCA, ICA, and wavelet transform. The extracted features are provided as input to the proposed UTPMSVM model. The classification performance is measured using accuracy and F1-score. Table 1 shows the classification accuracies of the USVM, UNHSVM, AULSTSVM, and proposed UTPMSVM, for different EEG signals with nonlinear kernels. It is observed that our proposed UTPMSVM showed the best results in 15 cases out of 28 which indicate the supremacy of the UTPMSVM model compared to USVM, UNHSVM, and AULSTSVM to classify seizure and non-seizure signals. One can also notice that the mean accuracy of proposed UTPMSVM (74.2142) is increased by 9.599%, 0.629%, and 0.1929% compared to USVM (67.7142), UNHSVM (73.75), and AULSTSVM (74.07) models. Moreover, the mean F1-score of proposed UTPMSVM (0.7042) is increased by 2.578%, 2.176%, and 2.0156% compared to USVM (0.6865), UNHSVM (0.6892), and AULSTSVM (0.6903) models. Additionally, we have also shown the rank of the reported classifiers on different EEG signals in Table 2. Moreover, the mean ranks of the classifiers are exhibited in Fig. 6. It can be observed that the proposed UTPMSVM showed the lowest mean rank based on both, F1-score and accuracy, which reveals the effectiveness of the proposed model.

Experiment on Real-World Datasets

Experiments have been performed on 30 real-world benchmark datasets to see the efficiency of the proposed UTPMSVM model. These datasets are taken from the UCI ML data repository [46] and KEEL imbalanced data repository [47]. The classification performance of UTPMSVM is compared with USVM, UNHSVM, TPMSVM, and AULSTSVM. The classification accuracies of USVM, UNHSVM, TPMSVM, AULSTSVM, and UTPMSVM with the optimal parameters and training time are shown in Table 3. It can be noticed that the proposed UTPMSVM shows comparable or better classification performance compared to other reported models. The mean accuracies of the models are also shown in the last row. It is observed that the proposed UTPMSVM shows the highest mean accuracy (87.5887%) compared to USVM (83.8189%), UNHSVM (85.2132%), TPMSVM (86.0933%), and AULSTSVM (85.384%). Further, the ranks based on their classification accuracies are shown in Table 4. One can notice that the proposed UTPMSVM shows the best

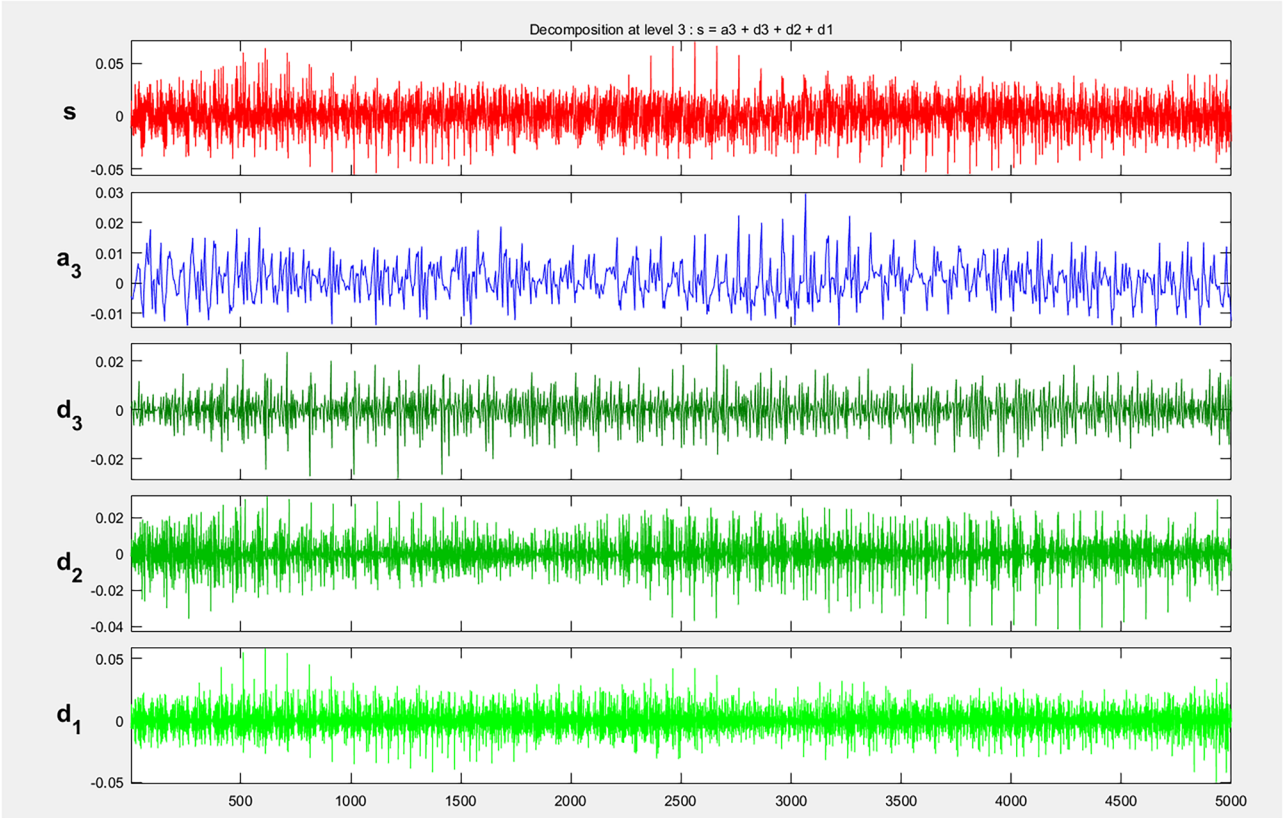

Fig. 3 Decomposition of N signal using db2 at level 3

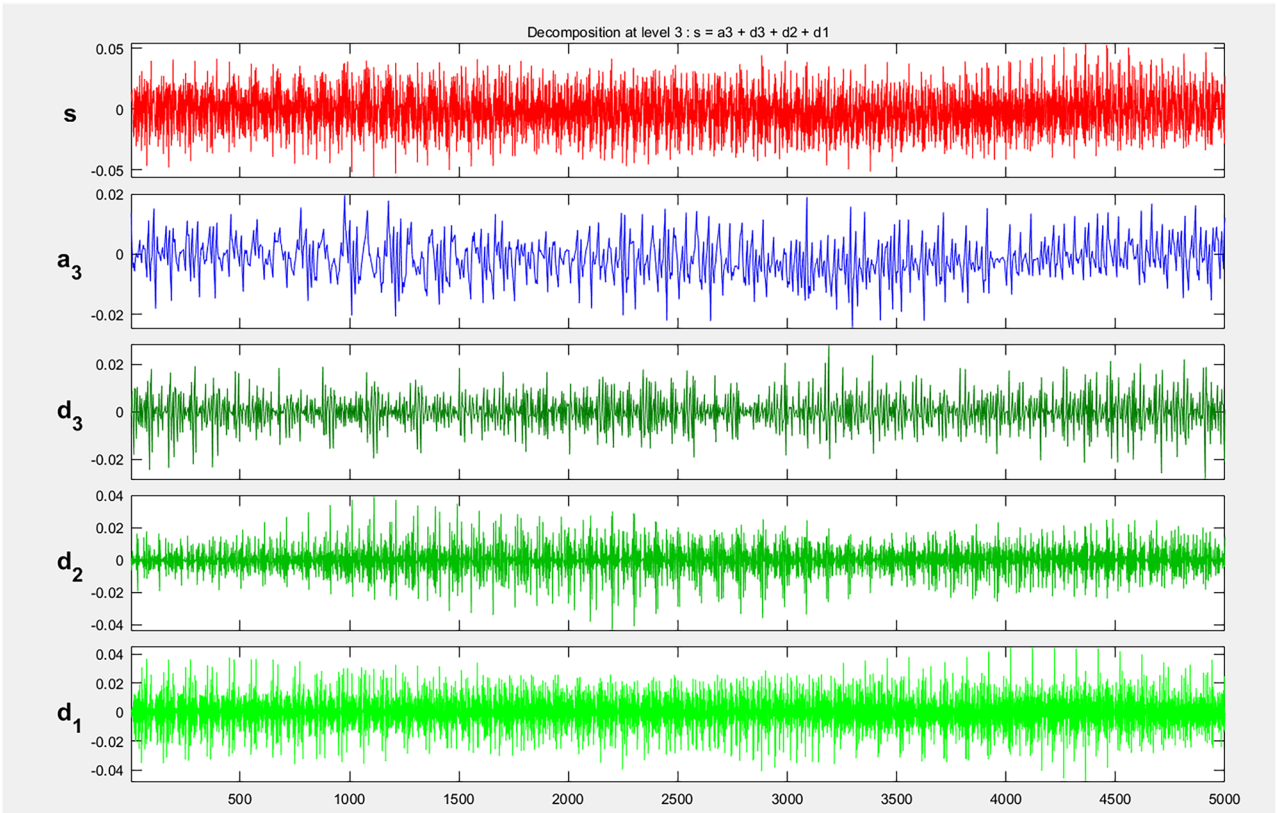

Fig. 4 Decomposition of S signal using db2 at level 3

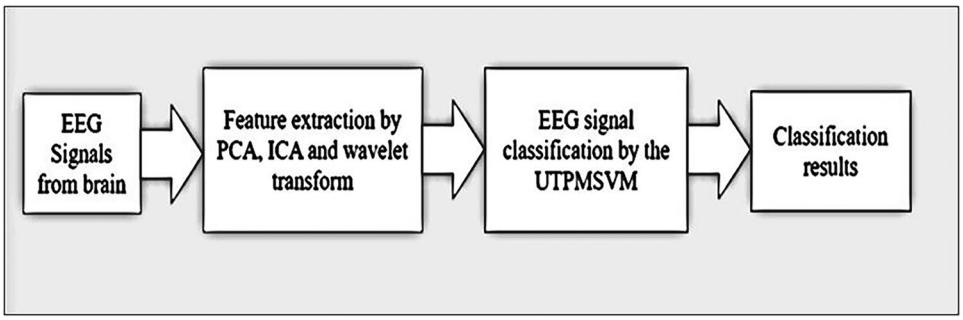

Fig. 5 Block diagram of the proposed framework for classifying EEG signal

Table 1 Accuracy and F1-score of USVM, UNHSVM, AULSTSVM, and proposed UTPMSVM on EEG datasets

Dataset	Indicator	USVM	UNHSVM	AULSTSVM	UTPMSVM
zs_pcas	Accuracy (%)	72	75.5	72.5	74
	F1-score	0.683	0.71	0.669	0.688
os_pcas	Accuracy (%)	72	78	79.5	78
	F1-score	0.736	0.737	0.779	0.766
zs_ica30	Accuracy (%)	69.5	76	77	73.5
	F1-score	0.745	0.7	0.717	0.677
os_ica30	Accuracy (%)	71	76	77.5	74
	F1-score	0.71	0.723	0.732	0.699
zs_db1	Accuracy (%)	69.5	75	76.5	76.5
	F1-score	0.684	0.708	0.723	0.699
os_db1	Accuracy (%)	65	70.5	68.5	71.5
	F1-score	0.679	0.663	0.643	0.707
zs_db2	Accuracy (%)	64.5	73	75.5	70
	F1-score	0.679	0.686	0.706	0.673
os_db2	Accuracy (%)	67.5	73.5	74.5	74.5
	F1-score	0.703	0.683	0.687	0.694
zs_db4	Accuracy (%)	68	74	75	75.5
	F1-score	0.653	0.683	0.701	0.754
os_db4	Accuracy (%)	70.5	72	71.5	75
	F1-score	0.692	0.689	0.681	0.63
zs_db6	Accuracy (%)	64	73	72.5	74
	F1-score	0.644	0.674	0.669	0.711
os_db6	Accuracy (%)	64.5	71.5	71.5	75.5
	F1-score	0.664	0.662	0.625	0.719
zs_haar	Accuracy (%)	60	70	71	72.5
	F1-score	0.612	0.621	0.638	0.697
os_haar	Accuracy (%)	70	74.5	74	74.5
	F1-score	0.726	0.708	0.692	0.743
Mean	Accuracy (%)	67.714	73.75	74.071	74.214
	F1-score	0.687	0.689	0.69	0.704

performance in 17 cases out of 30 which shows the efficiency of the proposed UTPMSVM model. Moreover, the mean rank is presented in the last row of Table 4. It can be observed that UTPMSVM shows the lowest mean rank compared to USVM, UNHSVM, TPMSVM, and AULSTSVM.

Figure 7 shows the parameter insensitivity performance of UTPMSVM on Autism, Led7digit-02-4-5-6-7-8-9_vs_1, Vowel, and WDBC. It is possible to confirm that our suggested UTPMSVM's performance is not dependent on the values of its parameters ε and μ. Extensive simulations reveal that the user-specified parameter μ does not significantly affect the performance of UTPMSVM.

Friedman Test for Statistical Comparison

The ranks of each classifier for the real-world datasets are shown in Table 4. It can be noted that the proposed method achieves the best average rank. To prove this argument statistically, the Friedman test [48] has been performed. Initially, we assume that all the methods are identical under the null hypothesis. Then, the Friedman statistic is computed using the given formula:

$$\chi_F^2 = \frac{12 \times X}{p \times (p+1)} \left[\sum_{i=1}^{p} R_i^2 - \frac{p(p+1)^2}{4} \right]$$

where X represents the total number of datasets and p is the number of classifiers. Here, $X = 30$ and $p = 5$.

$$\chi_F^2 \frac{12 \times 30}{5 \times (5+1)} [(3.516^2 + 3,216^2 + 3.266^2 + 3^2 + 2^2) - \frac{5(5+1)^2}{4}] \cong 16.62$$

After that, F_F value is computed as given below:

$$F_F = \frac{(X-1)\chi_F^2}{X(p-1) - \chi_F^2} = \frac{(30-1) \times 16.62}{30 \times (5-1) - 16.62} \cong 4.662$$

The critical value (CV) of F (4, 116) is 2.45 for the significance level $\alpha = 0.05$. The value F_F is greater than the CV. So, we reject the null hypothesis. As the null hypothesis

Table 2 Ranks based on Accuracy and F1-score of USVM, UNHSVM, AULSTSVM, and proposed UTPMSVM on EEG datasets

Dataset	Indicator	USVM	UNHSVM	AULSTSVM	UTPMSVM
zs_pcas	Accuracy (%)	4	1	3	2
	F1-score	3	1	4	2
os_pcas	Accuracy (%)	4	2.5	1	2.5
	F1-score	4	3	1	2
zs_ica30	Accuracy (%)	4	2	1	3
	F1-score	1	3	2	4
os_ica30	Accuracy (%)	4	2	1	3
	F1-score	3	2	1	4
zs_db1	Accuracy (%)	4	3	1.5	1.5
	F1-score	4	2	1	3
os_db1	Accuracy (%)	4	2	3	1
	F1-score	2	3	4	1
zs_db2	Accuracy (%)	4	2	1	3
	F1-score	3	2	1	4
os_db2	Accuracy (%)	4	3	1.5	1.5
	F1-score	1	4	3	2
zs_db4	Accuracy (%)	4	3	2	1
	F1-score	4	3	2	1
os_db4	Accuracy (%)	4	2	3	1
	F1-score	1	2	3	4
zs_db6	Accuracy (%)	4	2	3	1
	F1-score	4	2	3	1
os_db6	Accuracy (%)	4	2.5	2.5	1
	F1-score	2	3	4	1
zs_haar	Accuracy (%)	4	3	2	1
	F1-score	4	3	2	1
os_haar	Accuracy (%)	4	1.5	3	1.5
	F1-score	2	3	4	1
Mean rank	Accuracy (%)	4	2.25	2.03571	1.71429
	F1-score	2.71429	2.57143	2.5	2.21429

got rejected, we can proceed with post-hoc test and for that Nemenyi post hoc test is used. Further, the critical difference (CD) is calculated as follows:

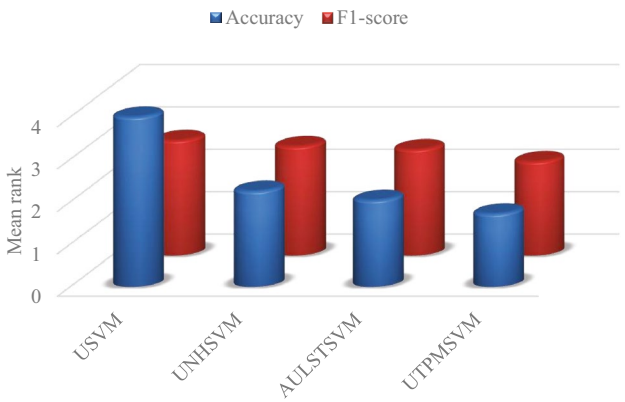

Fig. 6 Mean rank comparison among the reported models based on the experiments on EEG datasets

$$CD = q_{\alpha}\sqrt{\frac{p(p+1)}{6X}} = 2.728\sqrt{\frac{5(5+1)}{6\times 30}} \cong 1.1137$$

The pair-wise difference of USVM, UNHSVM, and TPMSVM with UTPMSVM are 3.516 - 2 = 1.516, 3.2146 - 2 = 1.2146, and 3.2666 - 2 = 1.266 which is higher than the CD. We can conclude that UTPMSVM shows statistically better performance compared to USVM, UNHSVM, and TPMSVM. But, the pair-wise difference of AULSTSVM with UTPMSVM is 1, and therefore, it cannot be concluded that UTPMSVM shows statistically better performance than AULSTSVM. However, it can be observed from Table 4 that UTPMSVM shows low average rank than AULSTSVM.

Figure 8 shows the comparative results of the Nemenyi statistics among all reported classifiers based on mean ranks. The classifiers with higher and lower ranks are plotted on the right and the left, respectively. The classifiers within a horizontal line with a length less than or equal to the CD (1.1137) perform statistically identically.

Table 3 USVM, UNHSVM, TPMSVM, AULSTSVM, and proposed UTPMSVM results for real-world datasets

Dataset (Size)	USVM (c_1, ε, μ) Time (Sec.)	UNHSVM $(c_1, c_2, \varepsilon, \mu)$ Time (Sec.)	TPMSVM (c_1, c_2, μ) Time (Sec.)	AULSTSVM (c_1, c_2, c_7, μ) Time (Sec.)	UTPMSVM $(c_1, c, c_3, \varepsilon, \mu)$ Time (Sec.)
Australian Credit (690X14)	87.2 $(10^1,0.1,2^0)$ 0.71671	86.5 $(10^3,10^1,0.7,2^1)$ 0.39916	86.6 $(10^2,10^{-3},2^2)$ 0.76824	86.5 $(10^3,10^1,0.7,2^2)$ 0.11391	**87.7** $(10^{-3},10^3,10^{-3},0.1,2^5)$ 0.88898
Heart-stat (270X13)	**84.8** $(10^0,0.7,2^1)$ 0.06788	83.7 $(10^3,10^{-1},0.7,2^{-1})$ 0.04059	82.9 $(10^{-1},10^{-1},2^2)$ 0.10334	83.7 $(10^2,10^{-1},1,2^3)$ 0.02074	83.7 $(10^3,10^1,10^3,0.1,2^1)$ 0.16783
Indian Liver Patient Dataset (ILPD) (579X10)	48.9 $(10^{-3},0.1,2^{-2})$ 0.38393	56.5 $(10^2,10^2,0.7,2^1)$ 0.28451	**67.9** $(10^0,10^0,2^3)$ 0.52477	56.5 $(10^{-2},10^3,0.3,2^5)$ 0.10886	61.7 $(10^0,10^1,10^0,0.7,2^0)$ 0.81664
Iris (150X4)	**100** $(10^{-2},0.7,2^{-2})$ 0.02714	**100** $(10^{-2},10^{-3},0.1,2^{-2})$ 0.02108	**100** $(10^{-3},10^{-3},2^{-2})$ 0.04055	**100** $(10^{-3},10^{-3},0.1,2^{-2})$ 0.0095	**100** $(10^{-3},10^{-3},10^{-3},0.1,2^{-2})$ 0.06099
Lymphography (148X18)	**89.1** $(10^0,0.5,2^1)$ 0.02194	87.1 $(10^2,10^1,0.1,2^1)$ 0.01903	86.4 $(10^{-2},10^{-2},2^2)$ 0.09893	87.1 $(10^{-3},10^0,0.1,2^3)$ 0.0099	88.5 $(10^{-3},10^0,10^{-3},0.1,2^0)$ 0.04998
Seeds (210X7)	95.7 $(10^1,0.7,2^{-1})$ 0.04137	**96.7** $(10^3,10^2,0.7,2^{-1})$ 0.03725	91.9 $(10^{-2},10^{-2},2^{-1})$ 0.07996	**96.7** $(10^{-2},10^{-1},0.9,2^1)$ 0.01413	93.8 $(10^3,10^2,10^3,0.5,2^0)$ 0.10340
Transfusion (748X4)	58.3 $(10^{-2},0.9,2^{-3})$ 0.80047	62.8 $(10^3,10^2,0.9,2^{-3})$ 0.53957	61.5 $(10^0,10^0,2^2)$ 1.23414	62.8 $(10^{-2},10^2,0.3,2^2)$ 0.2724	**75.7** $(10^{-1},10^1,10^{-1},0.9,2^{-3})$ 1.48682
Ecoli3 (336X7)	87.5 $(10^0,0.9,2^{-2})$ 0.28505	87.5 $(10^3,10^3,0.7,2^{-1})$ 0.09843	**88.7** $(10^{-3},10^{-3},2^{-2})$ 0.24835	87.5 $(10^3,10^1,1,2^{-2})$ 0.09843	87.8 $(10^{-3},10^{-2},10^{-3},0.1,2^{-2})$ 0.22788
Ecoli0137vs26 (311X7)	93.2 $(10^0,0.7,2^{-2})$ 0.14478	94.9 $(10^2,10^2,0.3,2^{-1})$ 0.06901	93.9 $(10^{-3},10^{-3},2^{-1})$ 0.17731	94.9 $(10^1,10^0,0.3,2^{-1})$ 0.06901	**96.1** $(10^{-1},10^{-1},10^{-1},0.9,2^{-1})$ 0.17433
Glass-0-1-4-6_vs_2 (205X9)	75.1 $(10^1,0.5,2^{-3})$ 0.04802	**89.8** $(10^3,10^0,0.9,2^{-1})$ 0.02726	84.9 $(10^{-1},10^{-1},2^{-2})$ 0.10341	**68.3** $(10^3,10^{-1},0.9,2^{-2})$ 0.02726	85.4 $(10^3,10^3,10^3,0.3,2^0)$ 0.09486
Glass-0-1-5_vs_2 (172X9)	68.1 $(10^0,0.9,2^{-2})$ 0.04527	78.6 $(10^3,10^3,0.9,2^{-2})$ 0.03074	86.1 $(10^{-2},10^{-2},2^{-2})$ 0.06882	64.7 $(10^1,10^{-1},0.7,2^{-1})$ 0.03074	**88.4** $(10^3,10^3,10^3,0.9,2^{-1})$ 0.07893
Led7digit-0-2-4-5-6-7-8-9_vs_1 (443X7)	93.7 $(10^{-1},0.9,2^{-1})$ 0.29490	93 $(10^0,10^{-2},0.9,2^{-1})$ 0.14014	89.8 $(10^{-3},10^{-3},2^0)$ 0.39518	95.3 $(10^0,10^{-1},0.5,2^3)$ 0.14014	**95.3** $(10^{-1},10^0,10^{-1},0.9,2^0)$ 0.36798
New-thyroid1 (215X5)	99.1 $(10^{-3},0.1,2^{-1})$ 0.06812	**99.6** $(10^2,10^{-3},0.1,2^0)$ 0.02519	99.1 $(10^{-2},10^{-2},2^{-1})$ 0.08421	**98.6** $(10^1,10^{-1},1,2^{-1})$ 0.02519	**99.6** $(10^3,10^3,10^3,0.7,2^0)$ 0.08161
Vowel (988X10)	92.3 $(10^1,0.9,2^{-2})$ 1.69117	94.8 $(10^3,10^2,0.7,2^{-1})$ 0.86685	92 $(10^{-3},10^{-3},2^{-1})$ 2.13776	90.4 $(10^0,10^{-1},0.9,2^{-1})$ 0.33203	**96.3** $(10^3,10^2,10^3,0.3,2^2)$ 4.16783
WDBC (569X30)	97.9 $(10^0,0.7,2^1)$ 0.37899	97.9 $(10^3,10^2,0.5,2^0)$ 0.32452	96.3 $(10^0,10^0,2^5)$ 0.62298	98.2 $(10^{-1},10^{-1},0.7,2^0)$ 0.0955	**97.9** $(10^1,10^1,10^1,0.9,2^1)$ 0.94994
Wine quality (178X13)	**98.9** $(10^0,0.7,2^1)$ 0.02599	**98.9** $(10^2,10^0,0.5,2^{-2})$ 0.02288	**98.9** $(10^{-3},10^{-3},2^{-2})$ 0.07527	**98.9** $(10^{-1},10^{-1},0.5,2^{-1})$ 0.0088	**98.9** $(10^3,10^1,10^3,0.1,2^0)$ 0.05610
WPBC (194X33)	55.8 $(10^2,0.7,2^{-1})$ 0.04496	66.1 $(10^1,10^0,0.7,2^0)$ 0.03412	**71.2** $(10^{-2},10^{-2},2^0)$ 0.09322	77.3 $(10^0,10^{-1},0.1,2^2)$ 0.01412	69.1 $(10^3,10^3,10^3,0.5,2^2)$ 0.07929
Ecoli-0-1_vs_5 (240X6)	95.8 $(10^1,0.7,2^{-3})$ 0.04733	97.1 $(10^2,10^1,0.7,2^{-2})$ 0.03960	97.5 $(10^{-1},10^{-1},2^{-1})$ 0.14137	95.4 $(10^0,10^{-1},0.9,2^{-1})$ 0.0173	**97.9** $(10^0,10^{-1},10^0,0.9,2^{-2})$ 0.09020

Table 3 (continued)

Dataset (Size)	USVM (c_1, ε, μ) Time (Sec.)	UNHSVM $(c_1, c_2, \varepsilon, \mu)$ Time (Sec.)	TPMSVM (c_1, c_2, μ) Time (Sec.)	AULSTSVM (c_1, c_2, c_7, μ) Time (Sec.)	UTPMSVM $(c_1, c, c_3, \varepsilon, \mu)$ Time (Sec.)
Haberman (306X3)	**74.2** $(10^0,0.9,2^{-2})$ 0.06862	54.6 $(10^{-3},10^1,0.7,2^0)$ 0.08278	66.4 $(10^{-1},10^{-1},2^0)$ 0.20192	73.2 $(10^1,10^{-1},1,2^3)$ 0.0261	70.6 $(10^3,10^1,10^3,0.7,2^{-1})$ 0.17099
Dermatology (358X34)	**100** $(10^{-2},0.9,2^0)$ 0.13792	**100** $(10^{-2},10^{-3},0.1,2^0)$ 0.11118	**100** $(10^{-3},10^{-3},2^0)$ 0.28500	**100** $(10^{-3},10^{-1},0.3,2^0)$ 0.0357	**100** $(10^{-3},10^{-1},10^{-3},0.1,2^1)$ 0.25872
Cleveland (297X13)	**84.2** $(10^2,0.9,2^5)$ 0.07983	82.4 $(10^3,10^{-1},0.1,2^{-1})$ 0.06232	82.8 $(10^{-1},10^{-1},2^5)$ 0.18098	84.5 $(10^2,10^{-1},0.7,2^1)$ 0.0193	83.5 $(10^2,10^1,10^2,0.1,2^2)$ 0.18088
Breast tissue (106X9)	98.1 $(10^2,0.7,2^0)$ 0.00983	**99** $(10^3,10^1,0.5,2^2)$ 0.01033	97.1 $(10^{-2},10^{-2},2^0)$ 0.03842	**97.1** $(10^{-3},10^{-1},0.3,2^0)$ 0.0029	98.1 $(10^0,10^0,10^0,0.5,2^{-1})$ 0.02587
Bupa or liver-disorders (345X6)	67.2 $(10^1,0.9,2^{-2})$ 0.08222	65.5 $(10^3,10^2,0.1,2^{-1})$ 0.07199	**72.2** $(10^{-1},10^{-1},2^{-1})$ 0.24196	73.6 $(10^0,10^{-1},0.1,2^0)$ 0.0022	71.6 $(10^{-3},10^1,10^{-3},0.5,2^{-1})$ 0.31955
Breast-cancer-wisconsin (683X9)	97.4 $(10^{-3},0.1,2^{-2})$ 0.40374	97.7 $(10^{-3},10^{-3},0.7,2^0)$ 0.38303	97.1 $(10^{-2},10^{-2},2^0)$ 0.90881	97.5 $(10^2,10^{-1},0.9,2^5)$ 0.1045	**97.8** $(10^3,10^1,10^3,0.1,2^3)$ 1.17417
Breast_cancer_coimbra (116X9)	75.1 $(10^0,0.9,2^{-1})$ 0.01487	76.8 $(10^2,10^1,0.1,2^0)$ 0.01204	76.1 $(10^{-2},10^{-2},2^{-1})$ 0.04255	81.2 $(10^1,10^{-1},0.7,2^0)$ 0.0054	**79.5** $(10^3,10^2,10^3,0.7,2^0)$ 0.04745
Autism-Adolescent-Data (98X20)	95.8 $(10^{-3},0.1,2^0)$ 0.01201	96.9 $(10^0,10^{-2},0.9,2^0)$ 0.01036	98 $(10^{-2},10^{-2},2^1)$ 0.03643	96.9 $(10^2,10^{-1},0.5,2^5)$ 0.0031	**99** $(10^{-3},10^1,10^{-3},0.1,2^1)$ 0.04198
03subcl5-600-5-50-BI (98X20)	74.2 $(10^{-1},0.1,2^{-3})$ 1.4717	74 $(10^3,10^1,0.9,2^{-2})$ 0.3733	76.8 $(10^{-2},10^{-3},2^{-1})$ 0.0601	76.5 $(10^3,10^3,0.3,2^{-2})$ 0.1063	**77** $(10^3, 10^3,10^2,0.9,2^{-2})$ 0.2312
03subcl5-600-5-60-BI (98X20)	72.5 $(10^0,0.7,2^{-3})$ 0.3504	74.8 $(10^3,10^2,0.9,2^{-2})$ 0.3573	76 $(10^{-2},10^{-3},2^{-3})$ 0.0603	76 $(10^2,10^2,0.1,2^{-5})$ 0.1104	**76.5** $(10^3, 10^3,10^2,0.9,2^{-1})$ 0.2392
04clover5z-600-5-30-BI (98X20)	78.5 $(10^0,0.9,2^{-3})$ 0.3397	82.3 $(10^3,10^3,0.1,2^{-3})$ 0.4217	83.7 $(10^{-2},10^{-3},2^{-3})$ 0.0584	82 $(10^0,10^0,0.3,2^{-5})$ 0.1118	**86** $(10^2,10^3,10^0,0.3,2^{-3})$ 0.1974
04clover5z-600-5-50-BI (98X20)	76 $(10^{-1},0.9,2^{-3})$ 0.3692	80.8 $(10^1,10^0,0.1,2^{-3})$ 0.3226	81 $(10^{-2},10^{-3},2^{-3})$ 0.0594	79.5 $(10^2,10^2,0.1,2^{-5})$ 0.1307	**84.3** $(10^3,10^3,10^0,0.3,2^{-2})$ 0.1826
Mean Accuracy	83.8	85.2	86.1	85.4	**87.6**

It is noticeable that UTPMSVM, AULSTSVM, and UNHSVM are on the right side of the graph which indicates the efficiency of these models. Moreover, while comparing the UTPMSVM and UNHSVM, it is observed that the horizontal line does not connect these two models. This indicates that the proposed UTPMSVM shows significantly better performance than UNHSVM. It can be further noted that UTPMSVM shows significantly better performance than USVM, TPMSVM, and UNHSVM. Moreover, the solid line connects UTPMSVM and AULSTSVM; hence, they are not significantly different.

Wilcoxon Test for Statistical Comparison

Additionally, to show the substantial difference between UTPMSVM and the other implemented classifiers, we further performed the two-tailed Wilcoxon signed-rank test (WST) [49]. The test outcomes are shown in Table 5. The second column of Table 5 shows the difference between UTPMSVM and USVM based on WST, where "x" denotes that the accuracy (ACC) of UTPMSVM is more than that of USVM, "y" indicates that the ACC of UTPMSVM is less than that of USVM,

Table 4 Specific ranks and average ranks of USVM, UNHSVM, TPMSVM, AULSTSVM, and UTPMSVM for real-world datasets

Dataset	USVM	UNHSVM	TPMSVM	AULSTSVM	UTPMSVM
Australian Credit	2	5	4	3	1
Heart-stat	1	3.5	5	2	3.5
Indian Liver Patient Dataset (ILPD)	5	4	1	3	2
Iris	3	3	3	3	3
Lymphography	1.5	4	5	1.5	3
Seeds	2.5	1	5	2.5	4
Transfusion	4	2	3	5	1
Ecoli3	3.5	3.5	1	5	2
Ecoli0137vs26	5	3	4	2	1
Glass-0-1-4-6_vs_2	4	1	3	5	2
Glass-0-1-5_vs_2	4	3	2	5	1
Led7digit-0-2-4–5-6-7-8-9_vs_1	3	4	5	1.5	1.5
New-thyroid1	3.5	1.5	3.5	5	1.5
Vowel	3	2	4	5	1
WDBC	3	3	5	1	3
Wine quality	3	3	3	3	3
WPBC	5	4	2	1	3
Ecoli-0-1_vs_5	4	3	2	5	1
Haberman	1	5	4	2	3
Dermatology	3	3	3	3	3
Cleveland	2	5	4	1	3
Breast tissue	2.5	1	5	4	2.5
Bupa or liver-disorders	4	5	2	1	3
Breast-cancer-wisconsin	4	2	5	3	1
Breast_cancer_coimbra	5	3	4	1	2
Autism-Adolescent-Data	5	4	2	3	1
03subcl5-600-5-50-BI	4	5	2	3	1
03subcl5-600-5-60-BI	5	4	2.5	2.5	1
04clover5z-600-5-30-BI	5	3	2	4	1
04clover5z-600-5-50-BI	5	3	2	4	1
Mean rank	3.517	3.217	1.267	3	2

and "z" indicates that the ACC of UTPMSVM is similar that of USVM. The level of significance is taken as 0.05. The mean differences between UTPMSVM and USVM are also shown. It can also be noticed that the "p" value is less than the level of significance, which indicates the dominance of UTPMSVM over USVM. Similar conclusions can be derived from the third and fourth columns. As a result, it can be concluded that UTPMSVM's accuracy distribution drastically differs from USVM, UNHSVM, and TPMSVM, demonstrating that UTPMSVM is significantly different from USVM, UNHSVM, and TPMSVM. However, it cannot be concluded that the accuracy distribution of UTPMSVM differs from AULSTSVM which is due to the "p" value being greater than the level of significance.

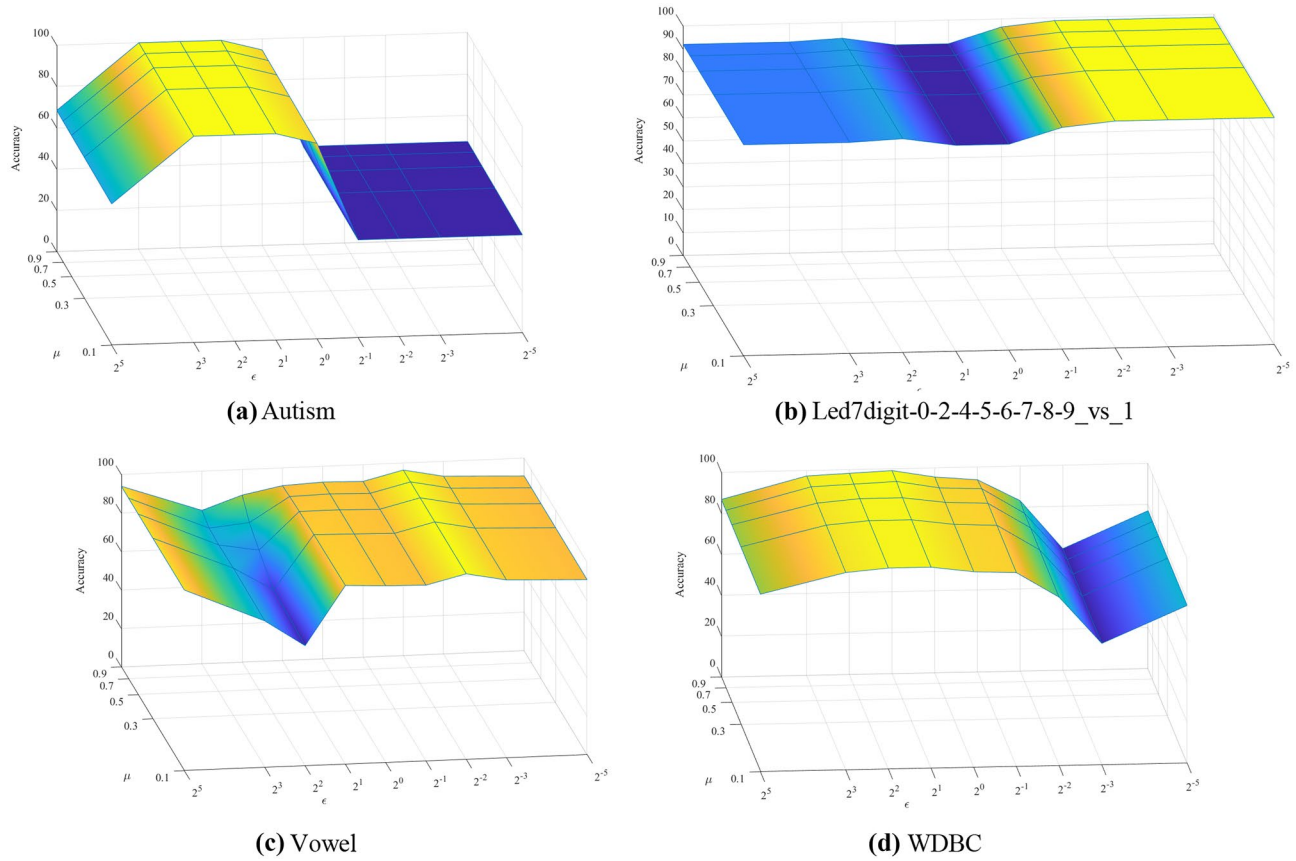

Fig. 7 ε and μ parameter insensitivity of UTPMSVM on **a** autism, **b** Led7digit-0-2-4-5-6-7-8-9_vs_1, **c** vowel, and **d** WDBC

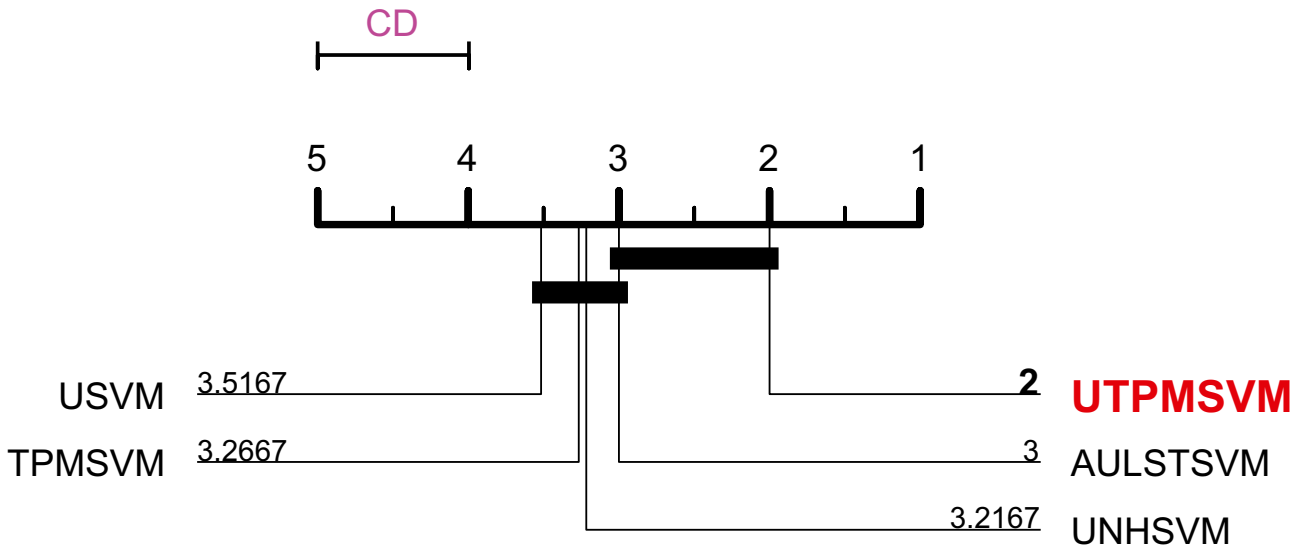

Fig. 8 Graphical visualization of the Nemenyi test. The CD is 1.1137

Table 5 Wilcoxon sign-rank test comparison

		UTPMSVM vs USVM	UTPMSVM vs UNHSVM	UTPMSVM vs TPMSVM	UTPMSVM vs AULSTSVM
N	Positive rank (PR)	20^{x}	21^{x}	23^{x}	17^{x}
	Negative rank (NR)	5^{y}	3^{y}	4^{y}	9^{y}
	Tie (T)	5^{z}	6^{z}	3^{z}	4^{z}
Mean differences		3.35	4.86	5.21	4.06
Sum of PRs		284	263	319.5	241
Sum of NRs		41	37	58.5	110
z value		−3.2692	−3.2286	−3.1353	−1.6636
p value		$0.0011 < 0.05$	$0.0012 < 0.05$	$0.0039 < 0.05$	$0.097 > 0.05$

Conclusion

We suggest a new Universum-based twin parametric margin SVM (UTPMSVM) for EEG signal classification problems. The suggested model shows an improvement in generalization performance over the already existing TPMSVM model for real-world as well as EEG signal classification problems. It is well known that the Universum samples work as prior information about the distribution of data. Hence, Universum-based models are suggested for classifying the EEG signals. In addition, to diminish the influence of noise from the EEG signals, we have used several feature reduction algorithms as a pre-processing step. To validate the efficiency of the proposed UTPMSVM, and its classification performance with USVM, UNHSVM, and TPMSVM. The results portray the efficacy of the proposed models for both EEG and other real-world datasets. Further, statistical analyses confirm the dominance of the proposed UTPMSVM over other models. The basic drawback of the UTPMSVM is that due to the incorporation of 2-norm, its sparsity is lost. Improving the sparsity of the UTPMSVM could be an interesting aspect of future work. Moreover, one can remodel the UTPMSVM for solving the multiclass classification problem in future. In addition to that, taking inspiration from the recent works of Tanveer et al. [50] and Ganaie et al. [51], we can improve the model to deal with large-scale as well as noisy datasets. Also, deep learning-based strategies can be embedded with our proposed model for the multiclass classification of EEG signals.

Data Availability The data source has been mentioned in the manuscript.

Declarations

Research Involving Human Participants and/or Animals Not applicable.

Consent to Participate Not applicable.

Conflict of Interest The authors declare no competing interests.

References

1. Caton R. Electrical currents of the brain. J Nerv Ment Dis. 1875;2(4):610.
2. Beck A. Die Bestimmung der Localisation der Gehirn-und Ruckenmarksfunctionen vermittelst der elektrischen Erscheinungen. Centralblatt fur Physiologie. 1890;4:473–6.
3. Berger H. Über das elektroenkephalogramm des menschen. Arch Psychiatr Nervenkr. 1929;87(1):527–70.
4. Tudor M, Tudor L, Tudor KI. Hans Berger (1873–1941)–the history of electroencephalography. Acta medica Croatica: casopis Hravatske akademije medicinskih znanosti. 2005;59(4):307–13.
5. Vecchiato G, Astolfi L, Tabarrini A, Salinari S, Mattia D, Cincotti F, Babiloni F. EEG analysis of the brain activity during the observation of commercial, political, or public service announcements. Comput Intell Neurosci. 2010;2010.
6. Light GA, Williams LE, Minow F, Sprock J, Rissling A, Sharp R, Swerdlow NR, Braff DL. Electroencephalography (EEG) and event-related potentials (ERPs) with human participants. Curr Protoc Neurosci. 2010;52(1):6-25.
7. Pearson K. Principal components analysis. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1901;6(2):559.
8. Jutten C, Herault J. Space or time adaptive signal processing by neural models. In Proceeding AIP Conference on Neural Networks for Computing 1986 (p. 206211).
9. Cortes C, Vapnik V. Support-vector networks Machine learning. 1995;20(3):273–97.
10. Yeo MV, Li X, Shen K, Wilder-Smith EP. Can SVM be used for automatic EEG detection of drowsiness during car driving? Saf Sci. 2009;47(1):115–24.
11. Subasi A, Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst Appl. 2010;37(12):8659–66.
12. Afifi S, GholamHosseini H, Sinha R. A system on chip for melanoma detection using FPGA-based SVM classifier. Microprocess Microsyst. 2019;65:57–68.
13. Gupta D, Borah P, Prasad M. A fuzzy based Lagrangian twin parametric-margin support vector machine (FLTPMSVM). In 2017 IEEE symposium series on computational intelligence (SSCI) 2017 (pp. 1-7). IEEE.
14. Jayadeva K, R., & Chandra, S. Twin support vector machines for pattern classification. IEEE Trans Pattern Anal Mach Intell. 2007;29(5):905–10.
15. Mangasarian OL, Wild EW. Multisurface proximal support vector machine classification via generalized eigenvalues. IEEE Trans Pattern Anal Mach Intell. 2005;28(1):69–74.
16. Kumar MA, Gopal M. Least squares twin support vector machines for pattern classification. Expert Syst Appl. 2009;36(4):7535–43.

17. Shao YH, Zhang CH, Wang XB, Deng NY. Improvements on twin support vector machines. IEEE Trans Neural Networks. 2011;22(6):962–8.
18. Qi Z, Tian Y, Shi Y. Robust twin support vector machine for pattern classification. Pattern Recogn. 2013;46(1):305–16.
19. Borah P, Gupta D. Robust twin bounded support vector machines for outliers and imbalanced data. Appl Intell. 2021;51(8):5314–43.
20. Hazarika BB, Gupta D. Density weighted twin support vector machines for binary class imbalance learning. Neural Process Lett. 2022;54(2):1091–130.
21. Peng X. TPMSVM: a novel twin parametric-margin support vector machine for pattern recognition. Pattern Recogn. 2011;44(10–11):2678–92.
22. Peng X, Wang Y, Xu D. Structural twin parametric-margin support vector machine for binary classification. Knowl-Based Syst. 2013;49:63–72.
23. Peng X, Kong L, Chen D. Improvements on twin parametric-margin support vector machine. Neurocomputing. 2015;151:857–63.
24. Shao YH, Wang Z, Chen WJ, Deng NY. Least squares twin parametric-margin support vector machine for classification. Appl Intell. 2013;39(3):451–64.
25. Gupta D, Borah P, Sharma UM, Prasad M. Data-driven mechanism based on fuzzy Lagrangian twin parametric-margin support vector machine for biomedical data analysis. Neural Comput Appl. 2022;1–11.
26. Richhariya B, Tanveer M. EEG signal classification using Universum support vector machine. Expert Syst Appl. 2018;106:169–82.
27. Long W, Tang YR, Tian YJ. Investor sentiment identification based on the Universum SVM. Neural Comput Appl. 2018;30(2):661–70.
28. Richhariya B, Tanveer M, Rashid AH, Alzheimer's Disease Neuroimaging Initiative. Diagnosis of Alzheimer's disease using Universum support vector machine based recursive feature elimination (USVM-RFE). Biomed Signal Process Control. 2020;59:101903.
29. Qi Z, Tian Y, Shi Y. Twin support vector machine with Universum data. Neural Netw. 2012;36:112–9.
30. Richhariya B, Gupta D. Facial expression recognition using iterative Universum twin support vector machine. Appl Soft Comput. 2019;76:53–67.
31. Zhao J, Xu Y, Fujita H. An improved non-parallel Universum support vector machine and its safe sample screening rule. Knowl-Based Syst. 2019;170:79–88.
32. Richhariya B, Tanveer M. A fuzzy universum support vector machine based on information entropy. In Machine Intelligence and Signal Analysis 2019 (pp. 569-582). Springer, Singapore.
33. Richhariya B, Tanveer M. A reduced Universum twin support vector machine for class imbalance learning. Pattern Recogn. 2020;102: 107150.
34. Kumar B, Gupta D. Universum based Lagrangian twin bounded support vector machine to classify EEG signals. Comput Methods Programs Biomed. 2021;208: 106244.
35. Moosaei H, Bazikar F, Ketabchi S, Hladík M. Universum parametric-margin ν-support vector machine for classification using the difference of convex functions algorithm. Appl Intell. 2021;1–21.
36. Richhariya B, Tanveer M, Rashid AH, Alzheimer's Disease Neuroimaging Initiative. Diagnosis of Alzheimer's disease using Universum support vector machine based recursive feature elimination (USVM-RFE). Biomed Signal Process Control. 2020;59:101903.
37. Richhariya B, Tanveer M. A fuzzy universum least squares twin support vector machine (FULSTSVM). Neural Comput Appl. 2021;1–2.
38. Moosaei H, Hladík M. A lagrangian-based approach for universum twin bounded support vector machine with its applications. Ann Math Artif Intell. 2022;1-23.
39. Richhariya B, Tanveer M, Alzheimer's Disease Neuroimaging Initiative Discipline of Mathematics, Indian Institute of Technology Indore, Simrol, Indore, India Program. An efficient angle-based Universum least squares twin support vector machine for classification. ACM Transactions on Internet Technology (TOIT). 2021;21(3):1–24.
40. Ganaie MA, Tanveer M, Alzheimer's Disease Neuroimaging Initiative. KNN weighted reduced Universum twin SVM for class imbalance learning. Knowl-Based Syst. 2022;245:108578.
41. Weston J, Collobert R, Sinz F, Bottou L, Vapnik V. Inference with the universum. In Proceedings of the 23rd international conference on Machine learning 2006 (pp. 1009-1016).
42. Mosek APS. The MOSEK optimization software. Online at http://www.mosek.com. 2010;54(2–1), p.5.
43. Andrzejak RG, Lehnertz K, Mormann F, Rieke C, David P, Elger CE. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Phys Rev E. 2001;64(6):061907.
44. Hazarika BB, Gupta D. Modelling and forecasting of COVID-19 spread using wavelet-coupled random vector functional link networks. Appl Soft Comput. 2020;96:106626.
45. Bartlett MS, Movellan JR, Sejnowski TJ. Face recognition by independent component analysis. IEEE Trans Neural Networks. 2002;13(6):1450–64.
46. Dua D, Graff C. UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science. 2019. [Accessed 25 April, 2019]
47. Alcalá-Fdez J, Fernández A, Luengo J, Derrac J, Garcia S, Sanchez L, Herrera F. (2011). Keel data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. J. Mult. Valued Logic Soft Comput. 2015;17.
48. Demšar J. Statistical comparisons of classifiers over multiple data sets. J Mach Learn Res. 2006;7:1–30.
49. Woolson RF. Wilcoxon signed-rank test. Wiley encyclopedia of clinical trials. 2007;1-3.
50. Tanveer M, Ganaie MA, Bhattacharjee A, Lin CT. Intuitionistic Fuzzy Weighted Least Squares Twin SVMs. IEEE Trans Cybern. 2022.
51. Ganaie MA, Tanveer M, Lin CT. Large-Scale Fuzzy Least Squares Twin SVMs for Class Imbalance Learning. IEEE Trans Fuzzy Syst. 2022.