An Optimized Quadratic Support Vector Machine for EEG Based Brain Computer Interface

Original Scientific Paper

Omar N. Maher
Mansoura University,
Computers and Control Systems Engineering Dept.,
Faculty of Engineering,
Mansora, Dakahlia, Egypt
o.elbklawi@gmail.com

Amira Y. Haikal
Mansoura University,
Computers and Control Systems Engineering Dept.,
Faculty of Engineering,
Mansora, Dakahlia, Egypt
amirayh@mans.edu.eg

Mostafa A. Elhosseini
Mansoura University,
Computers and Control Systems Engineering Dept.,
Faculty of Engineering, Mansora, Dakahlia, Egypt
Taibah University,
College of Computer Science and Engineering in Yanbu, Yanbu, Madinah, Saudi Arabia
melhosseini@mans.edu.eg

Mahmoud Saafan
Mansoura University,
Computers and Control Systems Engineering Dept.,
Faculty of Engineering, Mansora, Dakahlia, Egypt
Saafan2007@mans.edu.eg

Abstract *– The Brain Computer Interface (BCI) has a great impact on mankind. Many researchers have been trying to employ different classifiers to figure out the human brain's thoughts accurately. In order to overcome the poor performance of a single classifier, some researchers used a combined classifier. Others delete redundant information in some channels before applying the classifier as they thought it might reduce the accuracy of the classifier. BCI helps clinicians to learn more about brain problems and disabilities such as stroke to use in recovery. The main objective of this paper is to propose an optimized High-Performance Support Vector Machines (SVM) based classifier (HPSVM-BCI) using the SelectKBest (SKB). In the proposed HPSVM-BCI, the SKB algorithm is used to select the features of the BCI competition III Dataset IVa subjects. Then, to classify the prepared data from the previous phase, SVM with Quadratic kernel (QSVM) were used in the second phase. As well as enhancing the mean accuracy of the dataset, HPSVM-BCI reduces the computational cost and computational time. A major objective of this research is to improve the classification of the BCI dataset. Furthermore, decreased feature count translates to fewer electrodes, a factor that reduces the risk to the human brain. Comparative studies have been conducted with recent models using the same dataset. The results obtained from the study show that HPSVM-BCI has the highest average accuracy, with 99.24% for each subject with 40 channels only.*

Keywords*: brain computer interface, classification, quadratic support vector machine, feature selection, SelectKBest.*

1. INTRODUCTION

Brain Computer Interface (BCI) is a computerized system that can cooperate between the signals created by the human brain's thoughts and the computer [1]. The incorporating signals developed into actions. BCI collects and transmits electrical signals used in controlling electrical wheel cheers for disabled people; it also helps clinicians learn more about brain problems and illnesses such as stroke to use in recovery [2]. BCI comes in three types; Invasive, which injects electrodes into the grey matter; partially Invasive, in which electrodes are implanted in the brain surface. Non-Invasive one comes in a wearable device full of external sensors and electrodes and eases to communicate with computers. Many Competitions have appeared in this field, and all aim to find out the human brain's thoughts with high accuracy [3]. BCI competition III dataset IVa is one of the most common datasets subjected to extensive study by researchers recently.

Many recent studies have been applied in the BCI field. Amin Hekmatmanesh et al.[4] present a literature review that discusses brain-controlled vehicles. The study shows that electroencephalogram (EEG) signals are used to detect brain signals from the motor cortex area. Different Artificial Intelligence (AI) optimization algorithms are applied then to classify EEG signals. The biomedical signals are then used to control vehicles. R. Agarwal et al.[5] present a literature review that discusses human emotions and how to classify them using EEG signals and different datasets. The study discusses different classification accuracies according to different brain regions. S. Sodagudi et al.[6] proposed a new hybrid method to classify EEG signal data. The hybrid method consists of two stages. First, the Kernel Extreme Learning-based Multi-Layer Perceptron (KEL MLP) was used to extract brain activity features. Then Bayesian Quadratic Discriminant Transfer Neural Network (BQDTNN) was used as a classification technique. W. Al-Salman et al.[7] Constructing a new method to classify the six sleep stages using EEG signals. The method consists of using Discrete Wavelet Transform (DWT) to analyze EEG signals and extract brain wave features. Then, the extracted features were applied to Least Square Support Vector Machine LS-SVM to classify sleep stages. C. Wang et al.[8] Applied a new method to enhance classification accuracy by using Shannon Complex Wavelets (SCW) with Convolutional Neural Networks (CNN). The method consists of three stages. First, EEG signals have been recorded. Then SCW is used to calculate the time-frequency matrix. Finally, CNN was used to classify the BCI data.

The feature selection algorithm filters out unnecessary data or redundant features and chooses a subset of specific features or variables that lead to better performance in classification accuracy and training time. Feature selection methods are divided into three main categories: filter method, wrapper method, and embedded method [9], [10]. The filter methods are known to be the quickest in execution but imprecise. The wrapper method uses a computational model that rates subsets based on the miss classification rate. Embedded methods figure out which features contribute best to the model during the construction process. SelectKBest (SKB) feature selection algorithm is univariate. It uses different univariate statistical tests such as (Analysis Of Variance (ANOVA) F-value, Chi-square, and mutual information methods) to select the best features from the dataset [11].

Classification algorithms are accustomed to categorizing data into a class or category. Classification comes into three types: binary classification, multiclass classification, and multilabel classification. Binary classification algorithms are used to classify datasets that have only two classes, the normal state usually called "class 0" and the abnormal state usually called "class 1". Multiclass classification algorithms are used to classify datasets that have more than two classes. Many algorithms used for binary classification can be used for multiclass classification. Multilabel classification algorithms are used to classify datasets that have two or more classes, where each input might have one or more class labels predicted. SVM is one of the most common binary classifiers and it was proposed first by Vapnik [12]. It aims to find the best separable line that can divide the data into two groups. SVM work very well with linear data [13]. However, if the data is non-linear; a kernel function must be added to SVM such as a Quadratic kernel, Cubic kernel, and Gaussian kernel.

The main contribution of this paper is summarized in applying a modern feature selection method (SKB) on the dataset to reduce the unnecessary channels, Then, training selected features with a Quadratic SVM (QSVM) classifier. The proposed approach decreases the computational cost and time needed to train BCI datasets and predict the class. This paper has been organized as follows: Section 2 covers the literature studies related to this work. Section 3 declares the used dataset, and illustrates the presented feature selection, classification algorithms, the structure of the proposed approach, and the utilized performance metrics. Section 4 contains simulation results for the experiment as well as conducting a comparative analysis. Finally, Section 5 introduces the conclusions.

2. RELATED WORK

Before presenting the proposed approach, a group of previous studies that use different algorithms to classify BCI competition III dataset IVa has been summarized.

Sahar Selim et al. [14] proposed a method consisting of Common Spatial Pattern as feature extraction, Attractor Metagene (AM) with Bat optimization Algorithm (BA) as feature selection, and SVM has been used as a classifier (CSP\AM-BA-SVM). Finally, This hybrid algorithm obtains average classification accuracy of 85% with few EEG channels, but that requires large computational time.

Amardeep Singh et al. [15] proposed a Symmetric Positive Definite (SPD) as matrices based on the motor imagery classification method. SPD performed very well with a small sample set. Their method was applied to BCI Competition III dataset IVa and obtained an average accuracy of 87.21% of the subjects. However, this method obtained better ac-curacy just in a small sample set.

Yongkoo Park et al. [16] proposed a method that extracts features using Filter Bank CSP (FBCSP) and then selects the optimal channels which include the best features. Finally, the selected features were classified by LS-SVM. Their method was applied to BCI competition III dataset IVa and the average accuracy of the 5-subjects was 86.73%. However, one of the limitations of this method was badly performing with multiclass data.

Kais Belwafi et al. [17] proposed an algorithm Dynamic Self-Adaptive Algorithm (DSAA), which depends on the LS method. The study applied to BCI competition III dataset IVa with an average of 81.95%. The filtering method performed well only with online systems.

Amin Hekmatmanesh et al. [18] proposed an improved CSP algorithm to recognize and classify BCI Competition III dataset IVa data by aggregating four algorithms. The algorithms are Discriminative FBCSP with the Discriminative Sensitive Learning Vector Quantization (DFBCSP-DSLVQ), the Soft margin SVM (SSVM) classifier, and the Generalized Radial Bases Functions (GRBF) to create a method called DFBCSP DSLVQ SSVM GRBF with an average accuracy of 92.70%. However, for multi-classes, the error ratio rises when using this method.

Considering the feasibility of classifying datasets through efficient evaluation, we can see several serious limitations of the results. These problems can be summarized as; high computational time for the algorithm to be executed, and choosing bad channels when it contains a large amount of common noise. Moreover, feature selection algorithms may not perform well with multiclass motor imagery tasks. However, the hybrid algorithm successfully overcomes two of these challenges by classifying the dataset with high performance in an adequate training time. (Table 1) summarizes previously discussed algorithms focusing on various pros and limitations.

Table 1. Related work pros and Limitations summary

Author	Mean accuracy	Pros	Limitations
Sahar Selim et al. [14]	85%	Use only 0.1 of EEG channels with high accuracy	Requires considerable computational time
Amardeep Singh et al. [15]	87.21%	Performs well with small a sample set	Performs badly with large sets
Yongkoo Park et al. [16]	86.73%	Performs well with binary classes	Badly performing with multiclass data
Kais Belwafi et al. [17]	81.95%	The filtering method performs well with online systems.	The filtering method performs badly only with offline systems.
Amin Hekmatmanesh et al. [18]	92.70%	Performs well with binary classes	For multi-classes, the error ratio rises when using this method.

3. MATERIALS AND METHODS

Facing the previous difficulties of classifying datasets led to considering alternatives to achieve more accuracy with acceptable computational time. Therefore, by combining two algorithms, we found out that, after all, creating a union can boost both processes' strengths and overcome both weaknesses. The two main components of this union are QSVM and SKB algorithms.

3.1. BCI COMPETITION III DATASET IVA

BCI Competition III Dataset IVa has been collected from five healthy subjects. Those subjects were sat in a comfortable chair with arms placed on comfortable armrests [19]. The Data set include data from the four initial sessions with no feedback. The subject sat with open eyes opposite a screen that presents a letter for 3.5 seconds, as declared in (Fig. 1).

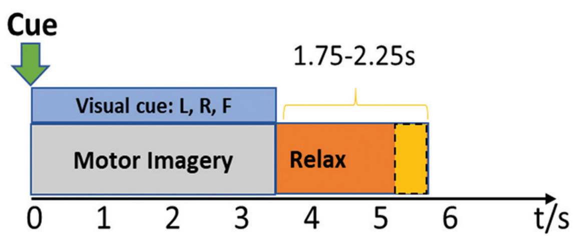


Fig. 1. Dataset timeline

Three letters equal three motor imageries the subject must perform [20]. For example, where, (L, R) left or Right hand and (F) foot. The subject relaxed for (1.75-2.25) seconds randomly between performed tasks. The dataset consists of continuous signals of 118 EEG channels according to the 10/20 system as shown in (Fig. 2) and markers that indicate the time points of 280 cues for each of the 5 subjects (aa, al, av, aw, ay). The data was recorded using Ag/AgCl electrode cap.

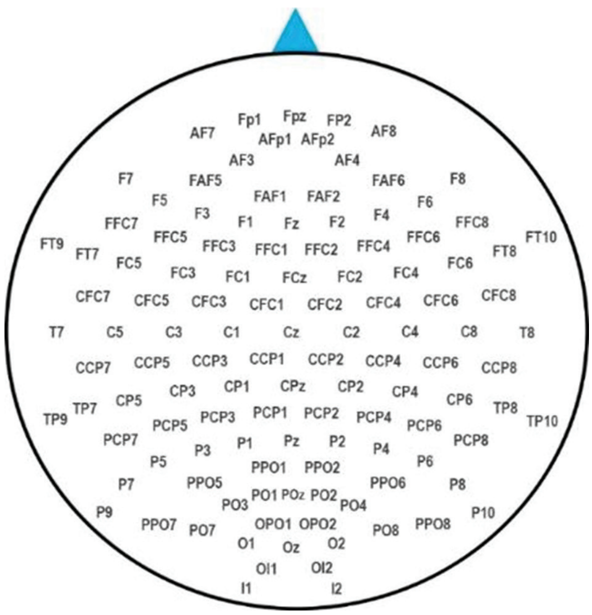


Fig. 2. 118 EEG channels

3.2. ALGORITHMS INVOLVED

This section discussed the paper's algorithms from a theoretical view. The algorithms that have been used in the hybrid approach are SKB as a feature selection algorithm and QSVM as a classifier.

3.2.1. Feature selection algorithm

Ag/AgCl electrode cap covers 118 channels in the human brain as declared in (Figure 2). While using BCI Competition III Dataset IVa, we noticed that some channels contain redundant information and others have only noisy information. SKB algorithm has been used to remove redundant and noisy channels according to the chi-square value. SKB keeps only 40-channel for each subject.

SKB Algorithm is a modern algorithm used in the 20th century. SKB chooses the powerful features by ranking the whole features according to statistical tests such as (ANOVA) F-value, Chi-square,…etc.) [21], [22]. Then select the best features that represent the data. This study used the chi-square test-based method to select the best features. Chi-square is given by:

$$X_c^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \tag{1}$$

Where n is the number of features, c is the freedom degree, O_i is the observed values and E_i is the expected values if there is no association between the two events [23]. The Chi-square test is used to test how much two events depend on each other. From (Equation 1), we can conclude that if there are two independent features, the observed count and expected count is very close values, leading to a small Chi-square value. The greater the correlation of features, the higher the value of Chi-square in promoting the selection of these features. (Algorithm 1) declare SKB steps in brief.

Algorithm 1: SKB

for each Subject
  Select the **Score Function SF // SF = Chi-square**
    Apply **Chi-square** statistical equation
    $X_c^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$
    Rank All Features due to Chi-square value
    Select the number of K
  Select only the Best features according to K-value
end

3.2.2. Classification algorithm

Classification algorithms are accustomed to categorizing data into a class or category. SVM is one of the most common classifiers. SVM Separates the two classes based on the distance between the objects and the hyperplane (Distance Margin). SVM works with a technique called the kernel functions that convert low dimensional input space to a higher dimensional space [12]. Linear SVM can classify linear data only, but if we have non-linear data, we should add a kernel with SVM. The results show that the quadratic kernel is the best one with BCI competition III dataset IVa.

QSVM classifies the data into two groups with hyperplane equation as declared in (Equation 2).

$$f(X) = \frac{1}{2} X^T W X + b^T X + c \tag{2}$$

Where W is a weight vector, X is the input vector, b is bias and T is the transpose. As shown in (Figure 3), QSVM has three decision boundaries [24]; the group of nodes lies on the hyper-plane described in (Equation 3), the group of nodes lies in the positive class described in (Equation 4) and the group of nodes lies in the negative class described with (Equation 5).

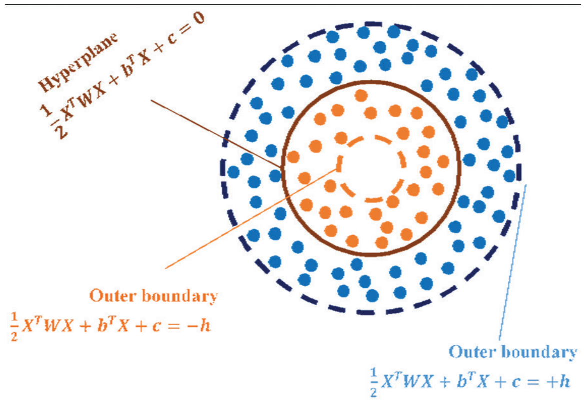


Fig. 3. Quadratic surface taxonomy

$$\frac{1}{2} X^T W X + b^T X + c = 0 \tag{4}$$

$$0 < \frac{1}{2} X^T W X + b^T X + c \leq +h \tag{5}$$

$$0 > \frac{1}{2} X^T W X + b^T X + c \geq -h \tag{6}$$

Where W is a weight vector, X is the input vector, b is the bias, T is a transpose, and (-h, +h) represents the hyper-plane of the inner and outer quadratic surface.

Cross-validation is commonly used to improve model prediction in machine learning. With this technique, we start dividing each subject in the BCI dataset randomly into k parts (k-fold cross-validation) [25]. In this study, we use 5-fold cross-validation. Four parts were used as training sets and the left one was used as a testing set. This process repeats five times with different sets each time.

3.2.3. HPSVM-BCI Approach

This section discusses High-Performance SVM-BCI (HPSVM-BCI) framework and the following method to implement the HPSVM-BCI approach. HPSVM-BCI framework contains the dataset subjects as we described before and its dimensions. The framework contains the classification algorithms that have been applied to the dataset subjects using 5-fold cross-validation such as Linear Discriminant (LD), Quadratic Discriminant (QD), Logistic Regression (LR), Naïve Byes (NB), Linear SVM (LSVM), QSVM, Cubic SVM (CSVM), and Deep Neural Network (DNN) as shown in (Fig. 4). Finally, the framework illustrates the performance metrics that have been used to evaluate classification algorithms and the feature selection algorithm that has been applied to the winner.

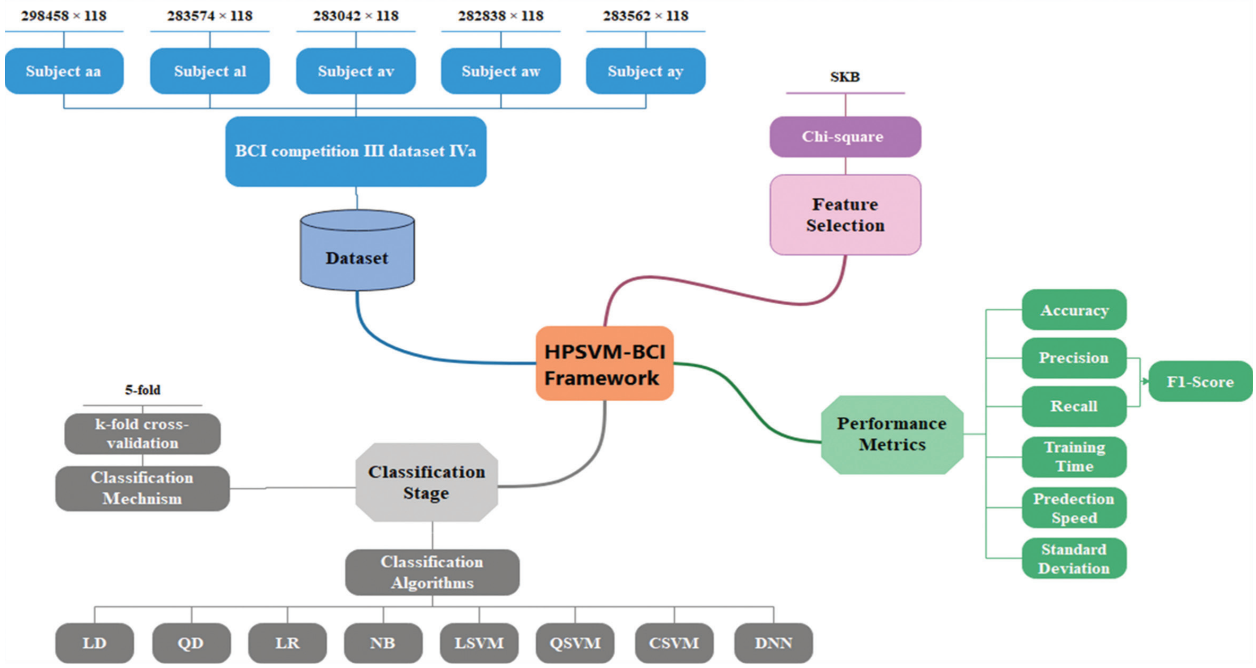


Fig. 4. HPSVM-BCI mechanism.

The proposed approach is a combination of both the feature selection algorithm (SKB) and the winner classification algorithm (QSVM). The HPSVM-BCI Process flow diagram is declared in (Fig. 5). SKB has been applied to the original dataset to evaluate the importance of each feature according to the Chi-square equation. The best features are selected then and a prepared dataset has been created. The prepared dataset has been subject to the classification stage. QSVM separates the prepared dataset into 5-fold cross-validation. Four parts randomly have been used as input to QSVM as training data. The last part has been used as testing data to evaluate the classifier. The classification stage has been repeated five times each with random training and testing data and prepares the data for the classification stage. The whole operation repeated for each subject on BCI competition III dataset IVa.

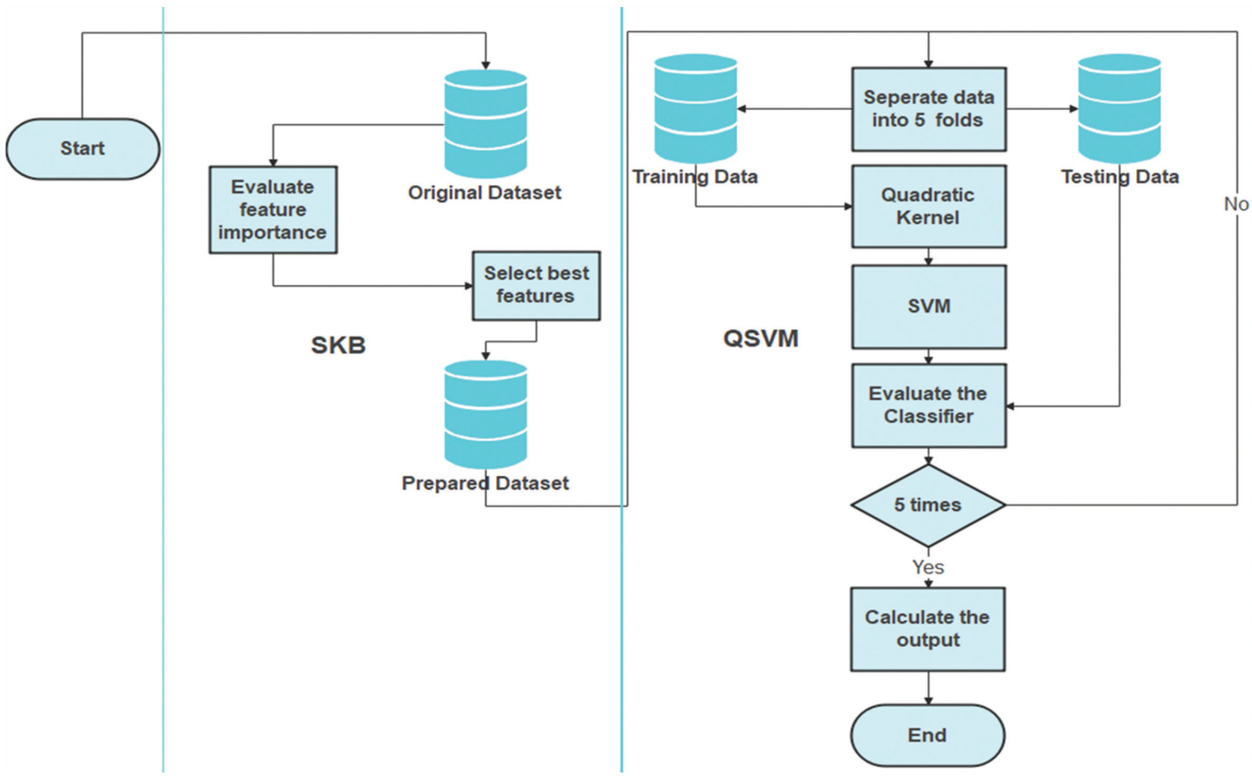


Fig. 5. HPSVM-BCI mechanism

3.3. PERFORMANCE METRICS

The computer results for this research have been evaluated according to different metrics; confusion matrix, F1 score, training time, and prediction speed.

3.3.1. Confusion matrix

The confusion matrix describes the effects of a forecast over a classification problem. Confusion metrics are very important metrics in evaluating classifier performance [26]. The accuracy equation is described in (Equation 6):

$$Accuracy = \frac{T_P+T_N}{T_P+T_N+F_P+F_N} \quad (6)$$

Where T_P is a True Positive, T_N is a True Negative, F_P is a False Positive and F_N is a False Negative.

3.3.2. Precision and Recall

Precision is defined as the ratio between the TP and all the Positives. It also helps to measure the relevant data points. The recall is defined as the fraction of retrieved instances among all relevant instances.

$$\boldsymbol{precision} = \frac{T_P}{T_P+F_P} \quad (7)$$

$$\boldsymbol{recall} = \frac{T_P}{T_P+F_N} \quad (8)$$

3.3.3. F1 Score

F1 Score aims for a balance between Precision and Recall [27], and there are many negative classified cases.

The Precision and Recall equations were described in (Equations 7, and 8). F1 Score equation:

$$F1 = \frac{2}{(1/Precision + 1/Recall)} \quad (9)$$

3.3.4. Training Time

It is the whole time that the model needs to be trained.

3.3.5 Prediction Speed

It is the number of observations that the AI model can deliver every second.

4. COMPUTER SIMULATIONS AND RESULTS

This section consists of three parts; first, displays the results of applying several algorithms on the dataset and obtains the winner. Second, the winner algorithm has been compared with the suggested approach HPSVM-BCI. Finally, a comparison between the proposed approach against the related work.

This experiment discusses several algorithms that have been executed with 5k-fold cross-validation on BCI Competition III Dataset Iva as illustrated in (Table 2). Performance metrics have been calculated 50 times and the average value is calculated for each metric. Standard deviation (SD) was also calculated for classification accuracy to show the algorithm's stability.

QSVM proved its ability in dealing with high-complexity data, such as Electroencephalography datasets, but it takes a huge training time. Accordingly, we suggest adding SKB to select the most relevant features of datasets. (Table 3) shows that HPSVM-BCI achieved higher average accuracy and average F1-Score than QSVM except

with subject "al" and reduces the average training time from (127,341 to 49,642) sec as shown in (Figure 6. a) this means that training time decreases by 250%. The mean prediction speed increases as well from (8,536 to 16,024.6) obs/sec as shown in (Fig. 6. b).

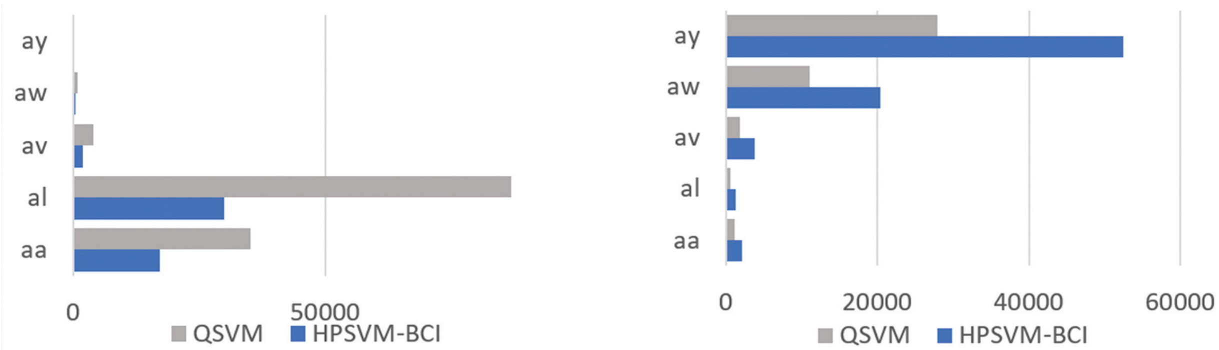
Table 2. The average values of performance metrics for several algorithms

Method	Performance Metrics	Subjects					Mean
		aa	al	av	aw	ay	
LD	Accuracy (%)	78.82	76.21	83.87	87.20	89	83
	F1 Score (%)	74.63	76	76.52	87.56	90.50	81.99
	Training Time (sec)	54	65	12	17	4	30.40
	Prediction Speed (obs/sec)	75000	84000	130000	62000	110000	92200
	SD (±%)	0.16	0.18	0.15	0.19	0.15	0.17
QD	Accuracy (%)	87.77	88.92	87.91	91	90.88	89.29
	F1 Score (%)	87.14	88.66	85.52	91.13	91.25	88.74
	Training Time (sec)	**53**	**64**	**9**	**16**	**4**	**29.20**
	Prediction Speed (obs/sec)	75000	81000	140000	62000	110000	93600
	SD (±%)	0.12	0.12	0.11	0.11	0.11	0.12
LR	Accuracy (%)	79.12	76.11	87.83	92.65	97.09	86.56
	F1 Score (%)	74.86	76.12	82.41	92.91	97.90	84.84
	Training Time (sec)	278	324	177	75	16	174
	Prediction Speed (obs/sec)	110000	100000	99000	100000	170000	115800
	SD (±%)	0.23	0.22	0.21	0.23	0.20	0.22

NB	Accuracy (%)	50.92	51.83	45.72	55.55	48.08	50.42
	F1 Score (%)	60.31	36.62	55.80	63.80	47.23	52.75
	Training Time (sec)	78	92	33	20	3	45.20
	Prediction Speed (obs/sec)	110000	**130000**	**150000**	110000	**180000**	**136000**
	SD (±%)	0.14	0.15	0.14	0.15	0.14	0.14
LSVM	Accuracy (%)	80	77	87.42	92.51	96.37	86.66
	F1 Score (%)	76.1	76.5	81.2	92.8	97.4	84.8
	Training Time (sec)	35510	67507	8322	925	83	22469
	Prediction Speed (obs/sec)	210	96	450	1600	9500	2371
	SD (±%)	0.078	0.09	0.087	0.087	0.127	0.094
QSVM	Accuracy (%)	**99.12**	98.91	**99.30**	99. 41	99.41	**99.20**
	F1 Score (%)	**98.88**	99	**98.91**	99.44	99.68	**99.18**
	Training Time (sec)	35291	86898	4086	991	75	25468
	Prediction Speed (obs/sec)	1200	580	1900	11000	28000	8536
	SD (±%)	0.06	**0.06**	0.09	0.11	0.11	0.09
CSVM	Accuracy (%)	51.32	**99.61**	59.27	**99.52**	**99.72**	81.88
	F1 Score (%)	41.21	**99.61**	25.83	**99.46**	**99.82**	73.18
	Training Time (sec)	16722	85728	7936	1362	93	22368
	Prediction Speed (obs/sec)	**180000**	1700	120000	18000	34000	70740
	SD (±%)	0.12	0.08	0.12	0.24	0.19	0.15
DNN	Accuracy (%)	76.24	92.36	64.57	87.68	90.83	82.34
	F1 Score (%)	78.24	92.42	67.57	89.62	91.83	83.94
	Training Time (sec)	1199	990	115	244	212	552
	Prediction Speed (obs/sec)	**115400**	125200	139300	**118500**	145100	128700
	SD (±%)	**0.06**	0.06	**0.08**	**0.07**	**0.07**	**0.07**

Table 3. The average values of performance metrics for QSVM vs HPSVM-BCI

Subject	Precision (%)	Recall (%)	QSVM F1-Score (%)	Accuracy (%)	Training Time (sec)	Prediction Speed (obs/sec)
aa	**98.90**	98.90	98.9	99.12	35291	1200
al	**99.58**	98.33	**99.01**	**98.91**	86898	580
av	99.11	98.70	98.90	99.30	4086	1900
aw	**99.52**	99.31	99.34	99.41	991	11000
ay	99.52	**99.71**	99.60	99.41	75	28000
Subject	**Precision (%)**	**Recall (%)**	**HPSVM-BCI F1-Score (%)**	**Accuracy (%)**	**Training Time (sec)**	**Prediction Speed (obs/sec)**
aa	98.83	**99.12**	**90.00**	**99.20**	**17104**	**2190**
al	98.31	**99.19**	98.71	98.70	**30017**	**1302**
av	**99.30**	**98.81**	**99**	**99.41**	**2042**	**3781**
aw	99.40	**99.60**	**99.50**	**99.39**	**439**	**20350**
ay	**99.77**	99.52	**99.71**	**99.50**	**40**	**52500**



(a) (b)

Fig. 6. QSVM vs HPSVM-BCI: (a) Training time (b) prediction speed

The suggested method HPSVM-BCI has been compared with literature studies that execute different types of classifiers at BCI Competition III dataset IVa. (Table 4) displays that HPSVM-BCI overwhelms the entire literature studies for aa, av, aw, and ay subjects by 5.60%, 17.43%, 5.73%, and 3.43, respectively. However, in the 'al' subject, SPD and CSP\AM-BA-SVM overcome the proposed method by only 1.30%. Accordingly, the mean accuracy for HPSVM-BCI is the best accuracy with 99.18%.

Table 4. The average values of performance metrics for QSVM vs HPSVM-BCI

Author	Method	Subjects' average accuracy					Mean
		aa	al	av	aw	ay	
Amin Hekmatmanesh et al. [18]	DFBCSP DSLVQ SSVM GRBF	93.51%	98.59%	81.82%	93.63%	96.14%	92.72%
Amardeep Singh et al. [15]	SPD	81.31%	**100%**	76.46%	87.13%	91.29%	87.22%
Yongkoo Park et al. [16]	FBCSP + LS-SVM	92.92%	89.27%	71.39%	83%	94.14%	86.71%
Kais Belwafi et al. [17]	DSAA	69 .55%	96.38%	60.52%	70.53%	78.60%	82%
Sahar Selim et al. [14]	CSP\AM-BA-SVM	86.63%	**100%**	66.78%	90.60%	81%	85%
	Proposed Method	**99.20%**	98.70%	**99.41%**	**99. 39%**	99.50%	**99.24%**

5. CONCLUSIONS

The goal of BCI is to integrate machine intelligence with the brain via electrodes. The field is now flooded with competitions that aim to uncover the human brain's thinking with high accuracy. One of the most widely used datasets for BCI competition III Dataset IVa has been extensively investigated by researchers. We aim to improve the classifications of the BCI dataset in this study. This can be achieved by developing a new approach HPSVM-BCI, which features two steps; selecting the best features and classifying the data. In SKB, in the first step, the features are sorted by Chi-square value, and then the best features are selected for classification by QSVM. After that, the quadratic function is used to determine the best surface for splitting into two classes. This improves the mean accuracy of data and reduces computational time, training time, and prediction time for HPSVM-BCI. As a result, the number of electrodes that reduce the risk of human brain injury is also decreasing.

6. REFERENCES:

[1] K. Venkatachalam, A. Devipriya, J. Maniraj, M. Sivaram, A. Ambikapathy, S. A. Iraj, "A novel method of motor imagery classification using eeg signal", Artificial Intelligence in Medicine, Vol. 103, 2020, p. 101787.

[2] M. Miao, W. Zhang, W. Hu, R. Wang, "An adaptive multi-domain feature joint optimization framework based on composite kernels and ant colony optimization for motor imagery EEG classification", Biomedical Signal Processing and Control, Vol. 61, 2020, p. 101994.

[3] H. Göksu, "BCI oriented EEG analysis using log energy entropy of wavelet packets", Biomedical Signal Processing and Control, Vol. 44, 2018, pp. 101-109.

[4] A. Hekmatmanesh, P. H. J. Nardelli, H. Handroos, "Review of the state-of-the-art of brain-controlled vehicles", IEEE Access, Vol. 9, 2021, pp. 110173-110193.

[5] R. Agarwal, M. Andujar, S. Canavan, "Classification of emotions using EEG activity associated with different areas of the brain", Pattern Recognition Letters, Vol. 162, 2022, pp. 71-80.

[6] S. Sodagudi, S. Manda, B. Smitha, N. Chaitanya, M. A. Ahmed, N. Deb, "EEG signal processing by feature extraction and classification based on biomedical deep learning architecture with wireless communication", Optik, Vol. 270, 2022, p. 170037.

[7] W. Al-Salman, Y. Li, A. Y. Oudah, S. Almaged, "Sleep stage classification in EEG signals using the clustering approach based probability distribution features coupled with classification algorithms", Neuroscience Research, 2022. (in press)

[8] C. Wang, Y. Wu, C. Wang, Y. Zhu, C. Wang, Y. Niu, Z. Shao, X. Gao, Z. Zhao, Y. Yu., "MI-EEG classification using Shannon complex wavelet and convolutional neural networks", Applied Soft Computing, Vol. 130, 2022, p. 109685.

[9] J. S. Kirar, R. K. Agrawal, "Composite kernel support vector machine based performance enhancement of brain computer interface in conjunction with spatial filter", Biomedical Signal Processing and Control, Vol. 33, 2017, pp. 151-160.

[10] J. Chen, P. K. Kudjo, S. Mensah, S. A. Brown, G. Akorfu, "An automatic software vulnerability classification framework using term frequency-inverse gravity moment and feature selection", Journal of Systems and Software, Vol. 167, 2020, p. 110616.

[11] M. S. Zulfiker, N. Kabir, A. A. Biswas, T. Nazneen, M. S. Uddin, "An in-depth analysis of machine learning approaches to predict depression", Current Research in Behavioral Sciences, Vol. 2, 2021, p. 100044.

[12] V. Vapnik, "The nature of statistical learning theory", Springer Science & Business Media, 2013.

[13] D. Madroñal, R. Lazcano, R. Salvador, H. Fabelo, S. Ortega, G. M. Callicó, E. Juarez, C. Sanz., "SVM-based real-time hyperspectral image classifier on a manycore architecture", Journal of Systems Architecture, Vol. 80, 2017, pp. 30-40.

[14] S. Selim, M. M. Tantawi, H. A. Shedeed, A. Badr, "A CSP\AM-BA-SVM Approach for Motor Imagery BCI System", IEEE Access, Vol. 6, 2018, pp. 49192-49208.

[15] A. Singh, S. Lal, H. W. Guesgen, "Small sample motor imagery classification using regularized Riemannian features", IEEE Access, Vol. 7, 2019, pp. 46858-46869.

[16] Y. Park, W. Chung, "Optimal channel selection using covariance matrix and cross-combining region in EEG-based BCI", Proceedings of the 7th International Winter Conference on Brain-Computer Interface, Gangwon, Korea, 18-20 February 2019, pp. 1-4.

[17] K. Belwafi, S. Gannouni, H. Aboalsamh, H. Mathkour, A. Belghith, "A dynamic and self-adaptive classification algorithm for motor imagery EEG signals", Journal of Neuroscience Methods, Vol. 327, 2019, p. 108346.

[18] A. Hekmatmanesh, H. Wu, F. Jamaloo, M. Li, H. Handroos, "A combination of CSP-based method with soft margin SVM classifier and generalized RBF kernel for imagery-based brain computer interface applications", Multimedia Tools and Applications, Vol. 79, No. 25, 2020, pp. 17521-17549.

[19] B. Blankertz et al. "Results of the BCI Competition III", BCI Meeting, 2005, p. 30.

[20] Z. Qiu, J. Jin, H.-K. Lam, Y. Zhang, X. Wang, A. Cichocki, "Improved SFFS method for channel selection in motor imagery based BCI", Neurocomputing, Vol. 207, 2016, pp. 519-527.

[21] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, "Scikit-learn: Machine learning in Python", Journal of Machine Learning Research, Vol. 12, 2011, pp. 2825-2830.

[22] C. D. Whitmire, J. M. Vance, H. K. Rasheed, A. Missaoui, K. M. Rasheed, F. W. Maier, "Using machine learning and feature selection for alfalfa yield prediction", AI, Vol. 2, No. 1, 2021, pp. 71-88.

[23] F. Thabtah, F. Kamalov, S. Hammoud, S. R. Shahamiri, "Least Loss: A simplified filter method for feature selection", Information Sciences, Vol. 534, 2020, pp. 1-15.

[24] I. Dagher, "Quadratic kernel-free non-linear support vector machine", Journal of Global Optimization, Vol. 41, No. 1, 2008, pp. 15-30.

[25] H. Ling, C. Qian, W. Kang, C. Liang, H. Chen, "Combination of Support Vector Machine and K-Fold cross validation to predict compressive strength of concrete in marine environment", Construction and Building Materials, Vol. 206, 2019, pp. 355-363.

[26] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, R. Budiarto, "Evaluating trust prediction and confusion matrix measures for web services ranking", IEEE Access, Vol. 8, 2020, pp. 90847-90861.

[27] W. Castro, J. Oblitas, M. De-La-Torre, C. Cotrina, K. Bazán, H. Avila-George, "Classification of cape gooseberry fruit according to its level of ripeness using machine learning techniques and different color spaces", IEEE Access, Vol. 7, 2019, pp. 27389-27400.