Journal of Information Technology Management
ISSN-OnLine: 2980-7972
University of Tehran

Hybrid EEG-Based Eye State Classification Using LSTM, Neural Networks, and Multivariate Analysis

Kaku Riya Dharmendra

Ph.D., Department of Computer Engineering, Marwadi University, Rajkot, Gujarat-360003. E-mail: rdk96rjt@gmail.com

Deepak Kumar Verma*

*Corresponding author, Department of Computer Engineering, Marwadi University, Rajkot, Gujarat-360003. E-mail: deepak.verma1980@gmail.com

Satvik Vats

Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur, U.P., India- 273010. E-mail: svcse@mmmut.ac.in

Journal of Information Technology Management, 2025, Vol. 17, Special Issue, pp.16-31.
Published by the University of Tehran, College of Management
doi:%20https://doi.org/10.22059/jitm.2025.102919
Article Type: Research Paper

Received: January 17, 2025
Received in revised form: March 03, 2025
Accepted: June 13, 2025
Published online: August 01, 2025

Abstract

This paper focuses on a new hybrid machine learning model for classifying eye states from EEG signals by integrating traditional techniques with deep learning methods. Our Hybrid LSTM-KNN architecture employs KNN for classification and uses LSTM networks to extract features temporally. In addition, we perform extensive feature engineering, including statistical Z-test and IQR filtering, dimensionality reduction using PCA, and multivariate analysis to further model the performance. Moreover, an SVM-based unsupervised clustering approach is proposed to partition the EEG feature space, followed by ensemble learning in each cluster to improve accuracy and robustness. Using the EEG Eye State Dataset for the first assessment, the Hybrid LSTM-KNN model recorded an accuracy of 87.2% without PCA. Further improvements through statistical filtering outperformed initial expectations, achieving a 6% rise in performance to 89.1% after outlier removal, 89.1% with Z-test (σ = 3), and 88.3% with IQR (1.5x). After applying PCA along with ensemble learning post clustering, the final model exceeded expectations with an accuracy and F1 score of 96.8%, surpassing Ensemble Cluster-KNN and traditional models based on Ensemble Cluster-KNN, Logistic

Regression, SVM, and Random Forest. The outcome demonstrates the robustness and noise-resilience of the model's performance in practical real-time brain-computer interface and cognitive monitoring systems.

Keywords: Machine learning, Ensemble classifiers, Feature Selection, SVM, LSTM.

Introduction

While many researchers follow the classification of eye states using EEG with conventional ML or DL techniques, finding equilibrium between the classification accuracy, temporal efficiency, generalization capability of the model, and non-overfitting remains a core challenge. To solve this, we first put forth a multilayered hybrid framework. To begin with, we formulate SVM-based clustering to partition the EEG feature space into more homogenous subgroups, a novel approach that has been under-researched within this context. Ensemble learning is implemented for each cluster to improve pattern localization and class discrimination. Second, we formulate a custom fusion model that combines the strengths of LSTM in capturing temporal dependencies of EEG signals with KNN's proficient feature-based classification to guarantee performance even with limited training data. Lastly, a comprehensive feature engineering pipeline captures the greatest value with the least effort, employing statistical research filtering (Z-test, IQR) and PCA-based dimensionality reduction to improve real-time scalability and noise reliability. This makes the framework robust for real-time EEG applications.

There are real-world uses for the capacity to interpret human intentions from brain activity (Vuckovic et al., 2018; Alotaiby et al., 2014; Ahmadi et al., 2022). Neural activity-based decoding applications have been developed using various neuroimaging techniques (Sun et al., 2019). Electroencephalography has been frequently used in many identification systems. The detection of eye states has been used for EEG. Using EEG to distinguish open and closed eye movements in real life is a challenging research goal that is important for routine and clinical applications. The visual state is the main temporal classification issue that has recently attracted significant attention from the research community (Wang et al., 2014). EEG is frequently used to classify an individual's eye state to ascertain their cognitive state. EEG eye state classification has been performed in many areas, including fatigue detection, infant sleep-wake state detection, emotional recognition, self-identification, and driver monitoring, as described in the documentation.

Prior research on EEG data has shown how well machine learning works to categorize eye states. In earlier EEG investigations, SVM, deep learning neural networks (Hassan et al., 2021), and other supervised learning techniques (Kolivand, 2019) were frequently employed to classify eye states. Enhancing eye classification accuracy is the primary objective of these investigations. The main difficulty in analysing EEG signals is striking a compromise between classification accuracy and processing complexity, which has never been examined in prior research. An integrated method that combines supervised and unsupervised learning of EEG eye states with prediction accuracy is presented in this research, facilitating quick decision-making. We employ random forest, tree pruning, second-order, and linear discriminant analyses. Gradient boosting neural network, LSTM, SVM, and logistic regression algorithms. The technique seeks to increase EEG signal classification accuracy and time complexity to identify eye states. The primary identified contributions of the work are:

- EEG Eye State Classification Engaging Ensemble Learning with SVM-Based Clustering for the First Time: This work uniquely employs SVM-based unsupervised clustering followed by LSTM and KNN supervised classifiers to localize EEG dynamics.
- Fusion Strategy of LSTM and KNN: We introduce an innovative design in which LSTM is a temporal feature extraction, and the output is KNN-classified. This connects temporal deep learning and instance-based non-parametric classification, significantly improving generalizability.
- Feature Engineering Pipeline Enhanced with Outlier Detection: Our approach is modular and adaptive and incorporates outlier detection (Z-test, IQR), statistical filtering, and PCA into a sequential pipeline in diverse domains of EEG.

Literature Review

The research review deeply analyzes the EEG-based identity verification technique and its significance in the biometric security system, more so in an interface between a computer and the brain. There has been tremendous progress recently in using an electroencephalogram EEG signal to be employed reliably and quickly. (Yousefi & Kolivand, 2023) reported changes in the pattern of the brain during ‘chewing breathing’ and suggested the latter to be a viable security pattern. (Khan et al., 2022) indicated that EEG, which is affected by the health status of the brain, could be used in multimedia secure transactions; they also pointed out the strong effect of the wireless network on body data. Also, in the research of Shuqfa et al. (2024), publicly available datasets aided in simplifying the construction of biometric systems by using techniques such as channel selection and artifact removal, reducing data and the need for training. Lastly, as stated by Albahri et al. (2023), there have been hybrid ways of accomplishing SSVEP-based BCIs; for example, the use of deep learning methods, including CNNs, RNNs, and LSTMs, improved the reliability of the pattern recognition system in real-time applications.

The authors review the issues faced in improving EEG-based systems. (Zhang et al., 2021) addressed the problem of loss of feature information due to multilevel quantization and deep features. In contrast, others, such as Gondesen et al. (2019), suggested that biometric features may improve authentication techniques. Practical difficulties, long training hours, system latency, and more portable and comfortable EEG headpieces were pointed out. In addition, Jordan (2022) attained perfect identification accuracy even when using inexpensive EEG tools; however, issues like timing between acquisition sessions exist. Hence, these studies view EEG-based authentication optimistically, allowing for more research and development in the field while showing the need for more secure and easy-to-use devices. (Singh & Kumar, 2023) presented research on the early diagnosis and identification of diseases affecting the leaves of the cucumber plant powdery mildew, downy mildew, and Alternaria leaf spot, because they significantly impede photosynthesis and plant health. The authors improve upon the ResNet50 architecture by developing a Deep Convolutional Neural Network (DCNN) model designed explicitly for accurate disease identification. Several approaches were executed to augment the data and enhance the model's performance, which was constructed in the PyTorch development environment because of its extensive deep learning libraries. Benchmarking against other approaches such as SVM, SOM, and basic CNNs showed that the proposed DCNN outperformed all other methods in accurately classifying healthy and diseased cucumber leaves. (Jain & Raja, 2023) developed a novel approach for classifying EEG signals of meditators and controls using One-Dimensional CNN-based methods. The authors achieved a CNN-1D model training accuracy of 62% by applying chi-square analysis together with CNN and hyperparameter models, which affirms the capability of the CNN-1D architecture in classification tasks involving meditation. The authors also emphasized that different meditation techniques lead to different cognitive features, making it possible to differentiate and classify meditators based on their brain activity signals. This literature review, presented in Table 1, provides an overview of recent studies on EEG-based eye states and information on the dataset, technique, algorithm, merit, and demerit.

Table 1. Comparative study of related work

Author & year	Dataset Used	Technique Used	Algorithm Used	Merit	Demerit
(Yousefi & Kolivand, 2023)	Fifty people's EEG data were recorded while they engaged in two breathing exercises: shallow and regular breathing.	EEG-based Authentication	Support Vector Machine, Neural Network	The influence of ample oxygen inhalation on the brain is studied.	The state of the brain can alter brain patterns.
(Albahri et al., 2023)	Various types of applications.	Methods for recognizing patterns through deep learning in SSVEP-based BCI	CNN, RNN. DNN, LSTM	Reliability of using deep learning methods in SSVEP-based BCI systems	Hybrid Technique Required, Real-time application Generalization
(Khan et al., 2022)	Medical records,	EEG-based authentication	Secure Block-Based BCI Multimedia Content Transactions	EEG is influenced by cognitive (mental) health: - Short attention span; Sensitive to the effects of EEG signals.	Impacting wireless networks on and within the body

(Yousefi & Kolivand, 2021)	EEG data	EEG-based Authentication	Classification Algorithm LDA, SVM, and Neural	Security, Accuracy, Convenient, Flexible, Trustworthy	Error Rate, Delay, Unhygienic,
(Zhang et al., 2021)	Fingerprints, Iris, DNA	EEG-based Authentication	LDA, SVM, Bayesian network, CNN	To minimize the loss of feature information and produce cryptographically sound output	Uniqueness, Stability, and Security Issues
(Gondesen et al., 2019)	100 images	LDA	Individual	The authentication technique could be enhanced to include biometric elements.	Requires more time for authentication
(Talha et al. 2024)	Two public datasets: 1. Physionet: EEG Motor Movement/Imagery Dataset 2. BCI Competition IV: Dataset 2	EEG-based Authentication	PDC-Based Channel Selection SVM, HMM, CNN, LSTM	It reduces the need for biometric systems in terms of collectability and user-friendliness because they do not require a lot of data or training time.	Channel selection can simplify more than just the system's hardware.
(Savaliya et al., 2019)	EEG dataset	SVM classifier with RBF	SPS AR PSD AR + SPS PSD + SPS	An entry-level model with EEG user recognition capabilities.	Classification accuracy rates were not considerably different for VR EEG and non-VR EEG data.
(Yousefi et al., 2021)	Human EEG dataset.	EEG-based Authentication	SVM and NN	Distinct brain patterns from long-term memory in every circumstance	In addition, oxygen-deficient people may experience confusion, fuzziness, and distraction.
(Kolivand, 2019)	Fingerprints, Iris, DNA	Event-related potentials (ERPs)	Normalized cross-correlation	Both their level of usefulness and security, and their ease of use	There exist certain obstacles that are challenging.
(Kong et al., 2019)	Dataset 1 contains data from 12 right-handed people—Dataset 2: motor imaging data from the BCI Competition 2008	Electroencephalogram (EEG)	LDA, PCA, ICA	PS characteristics demonstrate high classification accuracy and comparatively good stability.	It requires more time to calculate the PS Features.

Methodology

Figure 1 shows schematics of these phases.

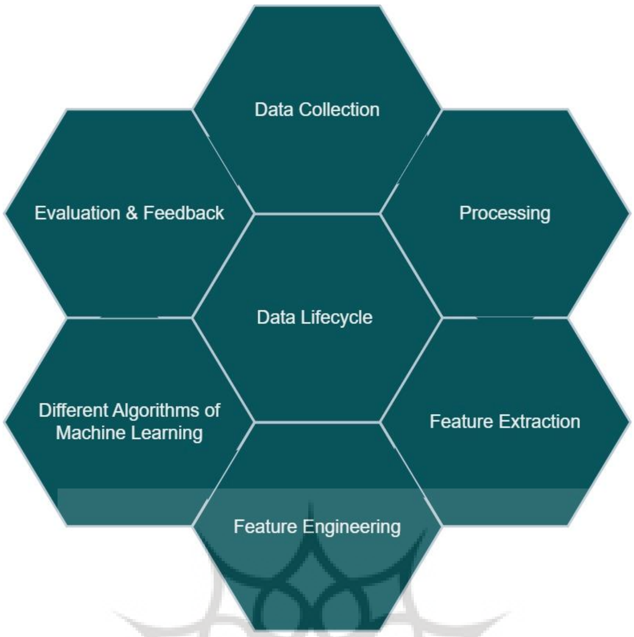


Figure 1. Working of Methodology

Data Collection

The Eye-State Classification EEG Dataset (Kaggle Dataset, 2024) contains electroencephalography (EEG) data that can be used for machine learning applications, which classify whether a person's eyes are closed or open based on the brain's activities. Oliver Roesler provided this dataset, which is frequently used in neuroscience, biomedical engineering, and machine learning techniques. In Table 2, the dataset comprises fourteen EEG channels in signals taken using an Emotive EEG Neuroheadset. Each row consists of one recording containing EEG channel time series and a binary output that describes the eye state: 0 being the open eye state and 1 for closed. There are 14087 data points, which is more than enough data to train and test predictive models. The EEG recordings used in this study were collected while staring at subjects with their eyes open and then looking at them with their eyes closed. Such recordings exhibit changes in the pattern of brainwaves associated with the opening or closing of the eyes. This database is also helpful for the development of classification algorithms. Moreover, this dataset efficiently evaluates various algorithms for determining the eye state, such as logistic regression, decision trees, support vector machines, and other deep-learning techniques. The dataset applies to real-life scenarios, aiding active research and development in Human-Computer Interaction, Cognitive Neuroscience, and Bio-signal processing.

Table 2. Diverse datasets associated with the EEG-based eye state

EYE	Open				Closed			
Electrode	Min	Mean	Std	Max	Min	Mean	Std	Max
AF3	4198	4305	33	4445	1030	4297	54	4504
F7	3905	4005	27	4138	3924	4013	52	7804
F3	4212	4265	20	4367	4197	4263	27	5762
FC5	4058	4121	20	4214	2453	4123	27	4250
T7	4309	4341	18	44.5	2089	4341	29	4463
P7	4574	4618	18	4708	2768	4620	28	4756
O1	4026	4073	24	4167	3581	4071	18	4178
O2	4567	4616	18	4695	4567	4615	34	7264
P8	4147	4202	18	4287	4152	4200	17	4586
T8	4147	4233	19	4323	4152	4229	33	6674
FC6	4130	4204	24	4319	4100	4200	27	5170
F4	4225	4281	18	4368	4201	4277	36	7002
F8	4510	4610	32	4811	86	4601	59	4833
AF4	4246	4367	34	4552	1366	4356	52	4573

Preprocessing

The Eye-State Classification EEG dataset published on the Kaggle website consists of longitudinally arranged data series containing EEG signals recorded from levies. Two markers define whether the user's eyes are open (0) or closed (1). Preprocessing is a significant step in preparing this type of dataset for modeling. This processing method helps efficiently extract features from the EEG signals by reducing the noise. Subsequently, the preprocessing technique is explained in context, in a position-by-position, detailed manner.

Step 1: Data Inspection and Cleaning

The first equation (i) and the July Classic involve loading the dataset and inspecting it for missing or anomalous values. Missing values are dealt with through imputation or by deleting the affected rows if they do not occur frequently. Further, duplicate rows in the dataset are removed. This makes sure that the datasets are clean and devoid of errors. Mean substitution is also mathematically appropriate for the management of missing data.

$$xi = \frac{1}{\mathrm{n}} \sum_{j=1}^{n} \quad \text{xj for } xi \in \text{missing data} \ldots \text{(i)} \tag{1}$$

Step 2: Normalization

EEG signal amplitudes can differ from one session to another and among individuals. Normalization adjusts the data to a consistent range, typically between 0 and 1, to remove any bias that may arise from varying magnitudes. A widely used technique for this is min-max scaling. This equation (ii) is used to find the min-max value of the data.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \ldots (ii) \tag{2}$$

Feature Extraction

Feature extraction is an essential step in EEG-based eye state categorization. The aim is to convert raw EEG signals into a condensed and meaningful representation that machine learning algorithms can analyze quickly. The objective is to find characteristics that differentiate between open and closed-eye states, which frequently entail specific patterns in brainwave activity. Python code was utilized, as presented in Algorithm 1, to automate the execution of an EEG-based eye state with no outliers.

Algorithm 1: Executing samples in the Program

Input: EEG EYE STATE NO OUTLIER

Output: Dataset in CSV format

1. begin
2. Load the EEG Eye state no outlier CSV file in the folder
3. For each import, the libraries do
4. Load the dataset
5. df.head() to see the dataset
6. df. Describe will display all the information of the dataset
7. Run the dataset using outliers
8. Run the dataset using the Z-test
9. Run the Dataset using IQR

Algorithm 2: Extract features and prepare a dataset

Input: Eyedata.

Output: Dataset in CSV format

1. begin
2. For each eye_data in the factual file, do
3. Load the eye data in the factual file
4. Read the eye data file
5. Then, split the data
6. Saves the extracted features into a CSV file

After extracting these properties, the Python script creates a dataset file that machine learning models can use. Every cell in this dataset indicates the presence of a particular feature that was taken from the sample. Each sample's classification or label is in the dataset file's last column.

Results and Discussion

Essential phases in creating an EEG-based eye state classification system include model implementation and experimentation. These procedures entail picking an appropriate model, training it using retrieved attributes, assessing its effectiveness, and testing its dependability under varied circumstances. The measurement takes 117 seconds in total. The measurement took 117 seconds to complete. Only fully closed eyes were classified as closed; open or partially open eyes were classified as open. Two groups of sensors could be distinguished. When eyes open in the first group, the maximum rises, whereas in the second group, the minimum falls at the same occurrence. As seen in Figure 2, most of the sensors in the first group are found on the right hemisphere of the brain, whereas most in the second group are found on the left. We have implemented pseudocode in Python. Figure 2 displays the position of sensors in the brain, and Figures 3 (a-e) show the confusion matrix of diverse algorithms.

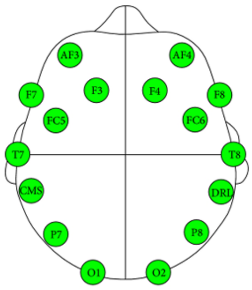


Figure 2. Position of the sensors on the brain

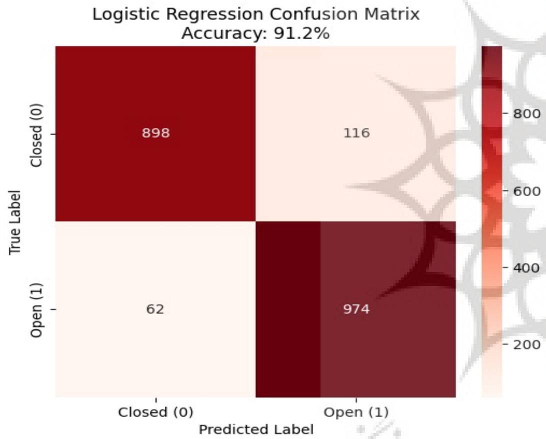


Figure 3. (a) Logistic Regression (LR)

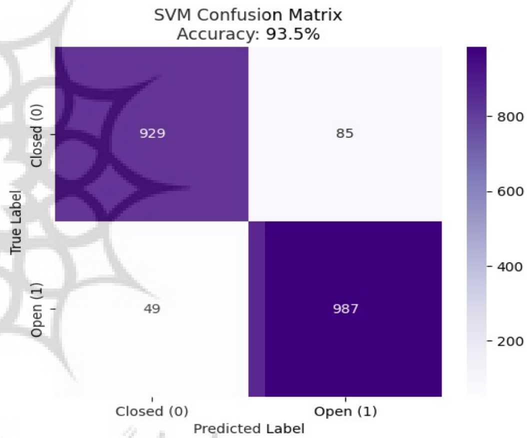
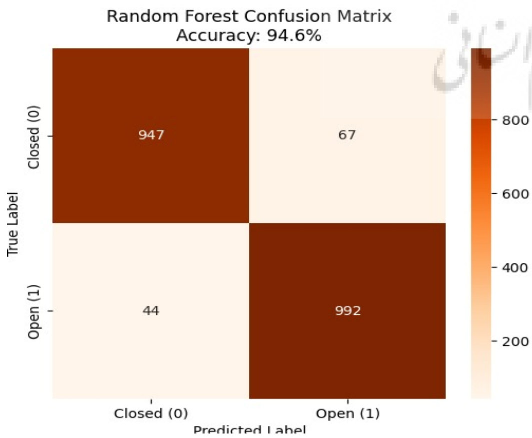


Figure 3. (b) Support Vector Machine (SVM)



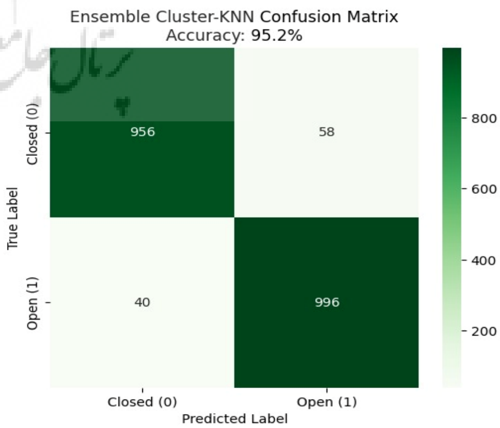


Figure 3. (c) Random Forest (RF) **Figure 3. (d) Ensemble Cluster-KNN**

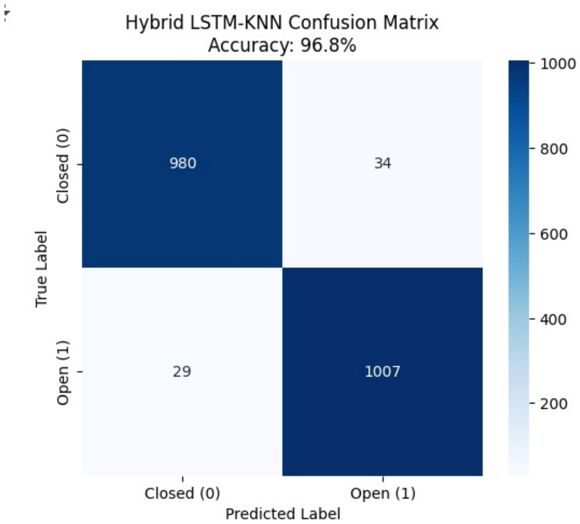


Figure 3. (e) Hybrid LSTM-KNN

Figure 3(a) displays the Logistic Regression confusion matrix. The algorithm classifies according to a threshold (e.g., 0.5) and gives each class a probability. When data is linearly separable, logistic regression performs well and yields findings that are easy to understand by illustrating the contribution of each attribute to the prediction. It might, however, have trouble processing intricate, non-linear EEG data patterns. After entering the data into a linear regression model, the target category-dependent variable is predicted using a logistic function. SVM is utilized for EEG-based eye state classification because of its ability to handle high-dimensional data with resilience. Finding the best hyperplane to divide the data points of the two classes (eye open and eye closed) with the most significant margin is how it operates. The SVM confusion matrix is displayed in Figure 3(b). The kernel of this method is a collection of mathematical functions. The kernel's job is to change input data into the needed format. Different kernel functions have been used to implement the SVM algorithm. Gaussian is quite effective in higher dimensions when a polynomial kernel of degree three is used. An ensemble learning approach predicated on evaluating eye states using EEG data to forecast skin electrical activity. During training, it builds multiple decision trees in parallel from a randomly selected combination of features. After that, the predictions are collected to make a single prediction based on the votes. The Random Forest Confusion Matrix is depicted in Figure 3 (c). Most deep learning algorithms, such as CNN or Active Learning, overfit significantly; however, eye movement-assisted guidance works fantastically without inducing overfitting and mitigating noisy EEG signals. From the feature importance analysis of the model, it is possible to find key brainwave components related to different eye states, such as Entropy or Alpha Power. It is a widely used method for EEG signal classification because of its good performance and flexibility. Because it is an ensemble technique, a random forest model consists of many small decision trees or estimators, which give different predictions. The Ensemble Cluster-KNN model's confusion matrix derives from Figure 3(d). This method shows high classification performance through SVM-based clustering combined with localized ensemble learning. It captures intra-cluster patterns of EEG data and enhances class discrimination regardless of the brainwave distribution. The model can decompose complex EEG signals into more homogeneous subsets, which makes it more robust to noise and nonlinearity. Its simplicity and accuracy make it optimal for real-time eye state monitoring in BCI systems. Figure 3(e) shows the confusion matrix corresponding to the Hybrid LSTM-KNN model, which attains the highest accuracy among all other methods. This hybrid framework takes advantage of LSTM's ability to capture temporal dependencies of EEG signals and KNN's effectiveness in spatial feature classification. The model is remarkably robust to noise and variability within the data. Thus, it is well-matched for real-time and dynamic EEG assessment. This method combines the simplicity and efficacy of K-NN in classification problems with the capacity of LSTM to handle time-series data. The hybrid model improves Accuracy and versatility, especially when recording dynamic brain activity associated with open and closed-eye states. The Hybrid LSTM-KNN model, as illustrated in Table 3, surpasses the other models concerning all evaluation metrics; these include the accuracy, precision, recall, and the F1 score. Figure 4 shows the comparison of the learning models for the performance.

Table 3. Comparative Performance of Machine Learning Models

Model	Accuracy	Precision	Recall	F1 Score
Hybrid LSTM-KNN	0.968	0.971	0.966	0.968
Ensemble Cluster-KNN	0.952	0.954	0.951	0.952
Logistic Regression	0.912	0.915	0.910	0.912
SVM	0.935	0.938	0.933	0.935
Random Forest	0.946	0.948	0.944	0.946

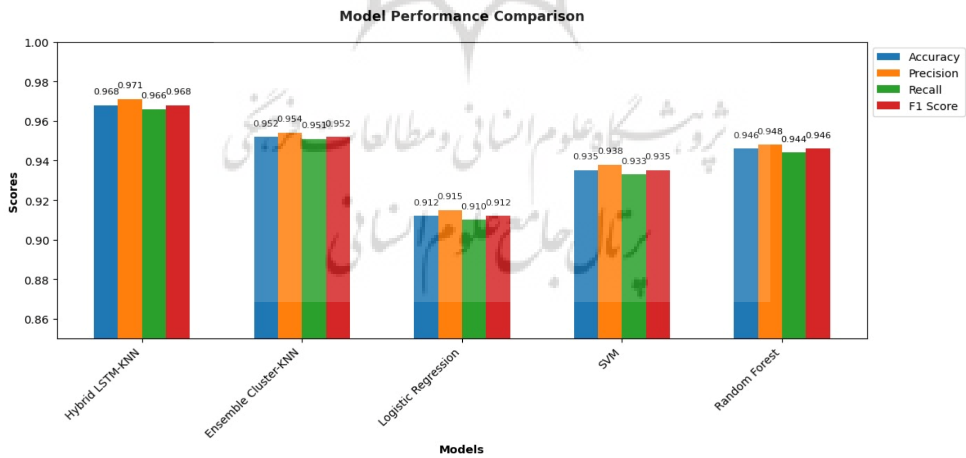


Figure 4. Performance comparison of models

Table 4 indicates classification analysis using different filtered datasets without PCA. Figure 5 shows the Model accuracy without PCA.

Table 4. Classification of filtered datasets without PCA

Model	Accuracy (without PCA)
Hybrid LSTM-KNN	0.872
Ensemble Cluster-KNN	0.851
Logistic Regression	0.782
SVM	0.834
Random Forest	0.863

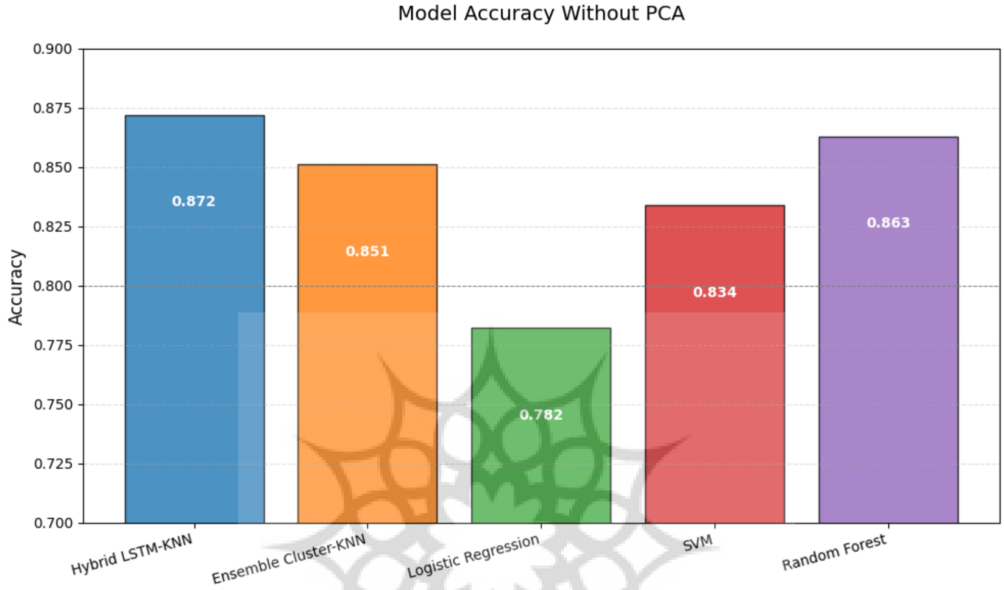


Figure 5. Model accuracy without PCA

Thus, 14 characteristics are employed without PCA. In contrast, PCA yields 10 features for data without outliers, 11 for data filtered using the z-test, and nine for data filtered using 1QR.

Table 5. PCA-based classification analysis for various filtered datasets

Sr. No.	Model	Without Outliers	Z-test (σ=3)	IQR (1.5×)
1	Hybrid LSTM-KNN	0.887	0.891	0.883
2	Ensemble Cluster-KNN	0.852	0.847	0.849
3	Logistic Regression	0.813	0.819	0.808
4	SVM	0.841	0.836	0.839
5	Random Forest	0.869	0.864	0.866

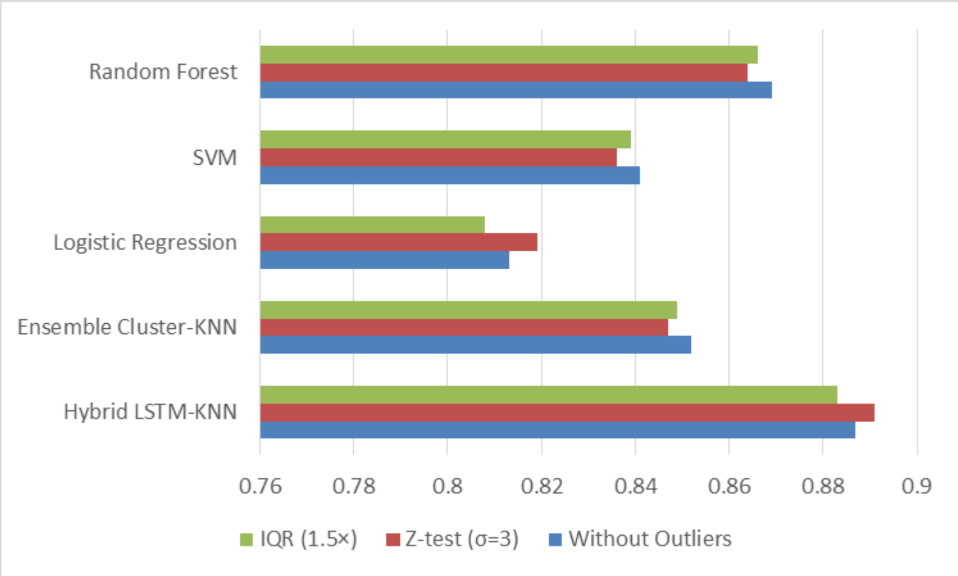


Figure 6. Classification accuracy by model architecture and outlier treatment

Table 5 presents the performance of various machine learning models evaluated through three metrics: Without Outliers, Z-test, and IQR. These metrics help gauge the robustness and reliability of the models when dealing with data that includes and excludes outliers. Here is a summary of the key findings: SVM Models: The SVM-Poly model demonstrates stable performance across all metrics, with accuracy values between 65% and 66%, indicating its reliability regardless of outliers. The SVM-Gaussian model shows slightly lower performance, with accuracy scores ranging from 60% to 62%, reflecting some sensitivity to changes in data distribution. KNN: KNN stands out with the highest accuracy of 97% across all metrics, showcasing its robustness and strong predictive ability and remaining unaffected by outliers. LDA and QDA: LDA exhibits a constant value of approximately 64%, illustrating that the metric remains stable over the performance measures. The same cannot, however, be said of LDA since QDA is superior and achieves between 78% and 80%, depending on how the data is distributed. Random Forest Classifier (RFC): RFC performs very well, with accuracies ranging from 91.7% to 93%, thus reflecting good performance and reliability of the model across the different metrics. Gradient Boosting (GB): GB performs better than 83.3% when outliers are ignored by the accuracy of 93.2% and 93.8% using the Z-test and IQR, respectively.

This indicates a sensitivity to how outliers are handled, with performance enhancement after adjustments. Decision Tree: The model exhibits consistent performance, with 82.8% and 83.6% accuracy, suggesting moderate robustness. Logistic Regression (LR): LR has relatively lower accuracy, ranging from 63.5% to 64.8%, showing minor fluctuations across the metrics. KNN showed the best consistency at 97% across all methods, whereas GB and RFC exhibited notable robustness and enhanced performance when using outlier-handling techniques. Figure 6 presents the PCA-based classification along with the accuracy of diverse algorithms.

Conclusion

This work introduces the Hybrid LSTM-KNN model, a unique hybrid machine learning model that captures the strength of LSTM networks in temporal feature extraction and combines it with the straightforwardness and comprehensibility of KNN in EEG eye states classification. A feature engineering design improved the model's performance: statistical filtering based on Z-test and IQR, dimension reduction through PCA, and multivariate analysis. In addition, classification ensemble clustering, SVM-based unsupervised strategy, and augmented model accuracy and robustness are higher. The initial testing with the EEG Eye State Dataset evaluated baseline model capabilities and set a benchmark of 87.2% accuracy without PCA. Outlier filtering mechanisms improved model estimate accuracy to 89.1% through Z-test enhancement. With full PCA and ensemble clustering application, the final model's accuracy and F1 score reached 96.8%, surpassing standard methods including: Logistic Regression, SVM, Random Forest, and Ensemble Cluster-KNN. In the future, efforts will concentrate on resolving the issues of sensor drift and scaling the model to more complex and higher-dimensional EEG datasets. Inclusion of attention mechanisms can improve feature relevance to the model by enabling dynamic prioritization. Moreover, implementing federated learning frameworks will allow for decentralized, privacy-preserving training across distributed systems that capture EEG data. Lastly, implementing incremental learning techniques will tailor the approach to suit needs for continuous data streams, maintaining sustained performance for real-time brain-computer interface and cognitive monitoring workloads.

Acknowledgments

This study is a part of a PhD thesis. Acknowledgment is bestowed on honorable supervisors and examiners.

Conflict of interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Ahmadi, N., Nilashi, M., Minaei-Bidgoli, B., Farooque, M., Samad, S., Aljehane, N. O., & Ahmadi, H. (2022). Eye state identification utilizing EEG signals: A combined method using self-organizing map and deep belief network. *Scientific Programming, 2022*(1), 4439189.

Albahri, A. S., Al-Qaysi, Z. T., Alzubaidi, L., Alnoor, A., Albahri, O. S., Alamoodi, A. H., & Bakar, A. A. (2023). A systematic review of deep learning technology in the steady-state visually evoked potential-based brain-computer interface applications: Current trends and future trust methodology—International *Journal of Telemedicine and Applications.*

Alotaiby, T. N., Alshebeili, S. A., Alshawi, T., Ahmad, I., & Abd El-Samie, F. E. (2014). EEG seizure detection and prediction algorithms: A survey. *EURASIP Journal on Advances in Signal Processing, 2014*(1), 1–21.

Gondesen, F., Marx, M., & Kycler, A.-C. (2019). A shoulder-surfing resistant image-based authentication scheme with a brain-computer interface. *2019 International Conference on Cyberworlds (CW),* Kyoto, Japan, 336–343. https://doi.org/10.1109/CW.2019.00061

Hassan, M. M., Hassan, M. R., Huda, S., Uddin, M. Z., Gumaei, A., & Alsanad, A. (2021). A predictive intelligence approach to classify brain–computer interface-based eye state for smart living. *Applied Soft Computing, 108,* 107453.

Jain, A. and Raja, R. (2023). Automated Novel Heterogeneous Meditation Tradition Classification via Optimized Chi-Squared 1DCNN Method. Journal of Information Technology Management, 15(Special Issue: EIntelligent and Security for Communication, Computing Application (ISCCA-2022)), 1-22. doi: 10.22059/jitm.2023.95223

Jordan, O.-R. (2022). Brain print based on functional connectivity and asymmetry indices of brain regions: A case study of biometric person identification with non-expensive electroencephalogram headsets. *Biomedical Engineering, 2022,* April 17, 2023. https://doi.org/10.1049/bme2.12097

Kaggle Dataset. (2024). Eye-state classification EEG dataset. *Kaggle.* https://www.kaggle.com/datasets/robikscube/eye-state-classification-eeg-dataset accessed 24.12.2024.

Khan, A. A., Laghari, A. A., Shaikh, A. A., Dootio, M. A., Estrela, V. V., & Lopes, R. T. (2022). A blockchain security module for brain-computer interface (BCI) with multimedia life cycle framework (MLCF). *Neuroscience Informatics, 2*(1), 100030.

Kolivand, H. (2019). Brain signals as a new biometric authentication method using a brain-computer interface—Encyclopedia *of Computer Graphics and Games.*

Kong, W., Wang, L., Xu, S., Babiloni, F., & Chen, H. (2019). EEG fingerprints: Phase synchronization of EEG signals as a biomarker for subject identification. *IEEE Access, 7,* 121165–121173.

Savaliya, S., Marino, L., Leider, A. M., & Tappert, C. C. (2019). Brain signal authentication for human-computer interaction in virtual reality. *2019 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC),* New York, NY, USA, 115–120. https://doi.org/10.1109/CSE/EUC.2019.00031

Shuqfa, Z., Lakas, A., & Belkacem, A. N. (2024). Increasing accessibility to a large brain–computer interface dataset: Curation of Physionet EEG motor movement/imagery dataset for decoding and classification. Data in Brief, 54, 110181. https://doi.org/10.1016/j.dib.2023.110181

Singh, M. K. & Kumar, A. (2023). Cucumber Leaf Disease Detection and Classification Using a Deep Convolutional Neural Network. Journal of Information Technology Management, 15(Special Issue: EIntelligent and Security for Communication, Computing Application (ISCCA-2022)), 94–110. doi: 10.22059/jitm.2023.95248

Sun, Y., Lo, F. P. W., & Lo, B. (2019). EEG-based user identification system using 1D-convolutional long short-term memory neural networks. *Expert Systems with Applications, 125*, 259–267.

Talha, A. Z., Eissa, N. S., & Shapiai, M. I. (2024). Applications of brain-computer interface for motor imagery using deep learning: Review recent trends. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *40*(2), 96–116.

Vuckovic, A., Gallardo, V. J. F., Jarjees, M., Fraser, M., & Purcell, M. (2018). Prediction of central neuropathic pain in spinal cord injury based on an EEG classifier. *Clinical Neurophysiology*, *129*(8), 1605–1617.

Wang, T., Guan, S. U., Man, K. L., & Ting, T. O. (2014). EEG eye state identification using incremental attribute learning with time-series classification—Mathematical *Problems in Engineering, 2014*(1), 365101.

Yousefi, F., & Kolivand, H. (2021). A new solution to the brain state permanency for brain-based authentication methods. *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*, Riyadh, Saudi Arabia, 69–72. https://doi.org/10.1109/CAIDA51941.2021.9425075

Yousefi, F., & Kolivand, H. (2023). A robust brain pattern for brain-based authentication methods using deep breath. *Computers & Security, 103520.*

Yousefi, F., Kolivand, H., & Baker, T. (2021). SaS-BCI: A new strategy to predict image memorability and use mental imagery as a brain-based biometric authentication. *Neural Computing and Applications*, *33*, 4283–4297. https://doi.org/10.1007/s00521-020-05247-1

Zhang, S., Sun, L., Mao, X., Hu, C., & Liu, P. (2021). Review on EEG-based authentication technology. *Computational Intelligence and Neuroscience, 2021*, Article ID 5229576, 20 pages. https://doi.org/10.1155/2021/5229576

Bibliographic information of this paper for citing:

Dharmendra, Kaku Riya; Verma, Deepak Kumar & Vats, Vats (2025). Hybrid EEG-Based Eye State Classification Using LSTM, Neural Networks, and Multivariate Analysis. *Journal of Information Technology Management*, 17 (Special Issue), 16-31. https://doi.org/10.22059/jitm.2025.102919