

A CNN–SVM Hybrid Model for Learner Adoption Classification in Inclusive E-Learning Systems

S. Shri Tharani¹, Dr.S. Uma², Sonam Delkar³

¹Ph.D. Research Scholar, ² Professor & Head, Department of Computer Science,
Dr. NGP Arts and Science College, Coimbatore

³Scholar, Master of Computing, Curtin University, Western Australia

Abstract

Inclusive e-learning platforms are increasingly adopted to accommodate diverse learners by providing flexible, accessible and personalized learning environments. However, understanding and predicting learner adoption remains a challenge due to the heterogeneity of learner behaviors, emotional feedback and interaction patterns [1], [2]. Traditional machine learning models often fail to capture complex feature representations, while end-to-end deep learning classifiers may suffer from overfitting when applied to limited educational datasets. This paper proposes a hybrid classification framework that integrates a Convolutional Neural Network (CNN) for feature extraction with a Support Vector Machine (SVM) for learner adoption classification. The CNN learns high-level representations from pedagogical, technological and sentiment-based learner interaction data, while the SVM performs robust classification using a maximum-margin decision strategy [8], [9]. Experimental evaluation conducted on a contextual e-learning learner interaction dataset demonstrates that the proposed CNN–SVM model outperforms conventional CNN and classical machine learning classifiers in terms of precision, recall and accuracy. The results indicate that decoupling feature learning and classification improves generalization and provides a reliable early-stage solution for learner adoption prediction in inclusive e-learning environments.

Keywords: Inclusive e-learning, learner adoption, convolutional neural network, support vector machine, educational data mining

1. Introduction

Inclusive e-learning platforms have transformed modern education by enabling learners with diverse abilities, preferences and learning styles to access educational content anytime and anywhere. Features such as adaptive content delivery, accessibility tools and interactive learning resources have significantly enhanced learner engagement [2], [3]. Despite these advancements, a persistent challenge lies in accurately identifying and predicting learner adoption levels within inclusive e-learning systems.

Learner adoption is influenced by multiple factors, including behavioral interactions, emotional responses, pedagogical design and technological accessibility. These factors generate heterogeneous data that are difficult to model using traditional machine learning approaches[2], [4]. While deep learning models such as convolutional neural networks (CNNs) have shown promise in learning complex feature representations, end-to-end deep classifiers often require large datasets and may generalize poorly in educational contexts where data availability is limited [5], [6].

To address these limitations, this paper proposes a hybrid CNN–SVM framework for learner adoption classification. The CNN is employed as a deep feature extractor to capture high-level patterns from learner interaction data, while a support vector machine (SVM) is used as the final classifier to ensure robust decision boundaries. This separation of representation learning and classification improves generalization and reduces overfitting [6], [8].

The main contributions of this work are:

1. A hybrid CNN–SVM architecture for learner adoption classification in inclusive e-learning systems.
2. A feature modeling approach that integrates pedagogical, technological and sentiment-based learner attributes.
3. An experimental evaluation demonstrating the effectiveness of CNN–SVM over conventional CNN and classical classifiers.

2. Related Work

Learner behavior analysis and adoption prediction have been extensively studied using machine learning and deep learning techniques. Traditional machine learning models such as support vector machines, random forests and logistic regression [4], [11], have been applied to learner interaction data for engagement and performance prediction. While these models offer interpretability, they rely heavily on handcrafted features and struggle with high-dimensional data.

Deep learning approaches, particularly CNNs and recurrent neural networks (RNNs) [5], [6], [7], have been used to extract complex representations from learner feedback and interaction logs. CNNs have demonstrated strong performance in sentiment analysis and feature extraction tasks due to their ability to capture local patterns. RNN-based models, including long short-term memory networks, are effective for sequential learning but introduce additional computational complexity [6].

Hybrid models combining deep feature extraction with traditional classifiers have recently gained attention. CNN-SVM architectures have been successfully applied [8], [9], [10], in image classification, text classification and biomedical data analysis. However, their application in inclusive e-learning adoption prediction remains limited [3], [12]. This work bridges this gap by proposing a CNN-SVM hybrid model tailored for educational interaction data.

3. Dataset Description and Preprocessing

3.1 Dataset Description

The study utilizes a contextual e-learning learner interaction dataset obtained from the Kaggle repository (url: <https://www.kaggle.com/datasets/zoya77/contextual-e-learning-learner-interaction-dataset>) that captures behavioral, sentiment and demographic information of learners interacting with an inclusive e-learning platform. The dataset includes attributes such as learner sessions, content access patterns, feedback text, device type, engagement indicators and adoption labels. Adoption labels are categorized as high, medium, or low. The dataset consists of approximately 4,500 learner interaction records with multiple behavioral, contextual and sentiment-oriented features collected from inclusive e-learning environments.

Table 1: Description of Learner Interaction Dataset

Attribute	Description
User ID	Unique learner identifier
Session ID	Learning session reference
Clickstream	Interaction events during session
Feedback	Textual learner feedback
Device Type	Desktop, Mobile, or Tablet
Access Time	Timestamp of content access
Engagement Level	Learner engagement intensity
Adoption Label	High, Medium, or Low

3.2 Data Preprocessing

To ensure data quality, consistency and effective feature learning, several preprocessing techniques were applied to the contextual e-learning learner interaction dataset prior to model training.

1. Missing Value Imputation:

The dataset contained incomplete learner interaction attributes and missing engagement information. To address this issue, K-Nearest Neighbor (KNN) imputation was employed. The KNN method estimates missing values based on the similarity between neighboring learner instances using Euclidean distance. Centroid-based estimation and nearest-neighbor filling were used to preserve data consistency and reduce information loss.

2. Dimensionality Reduction:

Singular Value Decomposition (SVD) was applied to reduce redundant and non-informative features from the learner interaction dataset. SVD decomposes the high-dimensional feature matrix into lower-dimensional representations while preserving significant variance information. This dimensionality reduction process improves computational efficiency, minimizes overfitting and enhances the quality of feature learning.

3. Latent Feature Extraction:

Latent Dirichlet Allocation (LDA) was utilized to discover hidden semantic relationships within learner feedback, engagement indicators and contextual interaction attributes. The extracted latent topics were categorized into pedagogical, technological and sentiment-oriented learner representations. These topic-level embeddings enable improved behavioral interpretation and provide semantically enriched feature vectors for the proposed CNN–SVM framework.

4. Feature Representation:

Textual learner feedback and contextual interaction attributes were transformed into semantically enriched numerical representations using LDA and vectorization techniques. Categorical attributes were encoded appropriately and the resulting feature vectors were normalized before being provided to the CNN feature extractor.

The pre-processed and semantically enriched feature vectors were subsequently provided as input to the proposed CNN–SVM hybrid architecture for deep feature extraction and learner adoption classification.

4. Proposed CNN–SVM Hybrid Model

4.1 CNN-Based Feature Extraction

The convolutional neural network serves as a deep feature extractor. A one-dimensional CNN architecture was employed due to its ability to capture localized feature interactions among learner behavioral, pedagogical, technological and sentiment-based representations. Compared to fully connected neural networks, CNNs reduce parameter complexity while effectively learning discriminative local feature patterns from high-dimensional learner interaction data. The convolution operation enables efficient extraction of contextual dependencies among adjacent learner attributes, thereby improving feature generalization and reducing overfitting. It consists of multiple convolutional layers followed by pooling operations that capture local patterns from learner interaction data. Separate convolutional filters learn representations related to pedagogical structure, technological interaction and sentiment information.

Let x denote the input feature vector. The CNN maps the input to a high-level representation as follows:

$$z = f_{CNN}(x)$$

where $f_{CNN}(\cdot)$ represents the convolutional and pooling operations.

4.2 SVM-Based Classification

The output of the final dense layer of the CNN is treated as a discriminative feature vector and provided as input to the SVM classifier. The SVM identifies optimal decision boundaries using a maximum-margin principle, making it suitable for high-dimensional feature spaces.

The SVM decision function is defined as:

$$y = \arg \max (w \cdot z + b)$$

where w and b represent the learned weight vector and bias term, respectively.

Multi-class classification is handled using a one-vs-rest strategy.

Table 2: CNN Architecture for Feature Extraction

Layer	Configuration
Input Layer	Learner feature vector
Convolution Layer 1	64 filters, kernel size 3
Activation	ReLU
Max Pooling	Pool size 2
Convolution Layer 2	128 filters, kernel size 3
Activation	ReLU
Global Pooling	Feature aggregation
Dense Layer	128 units (feature output)

The Dense Layer output is passed to the SVM.

Algorithm CNN–SVM Learner Adoption Classification

Input: Dataset D

Output: Adoption class C

1: Preprocess D

- Impute missing values using KNN
- Normalize numeric features
- Encode categorical features
- Apply SVD for feature reduction

2: Initialize CNN model

3: Train CNN on preprocessed data

4: Extract feature vectors Z from CNN dense layer

5: Initialize SVM classifier

6: Train SVM using feature vectors Z and labels Y

7: For each test instance:

- Extract CNN features z_i
- Predict adoption class C_i using SVM

8: Return predicted adoption classes C

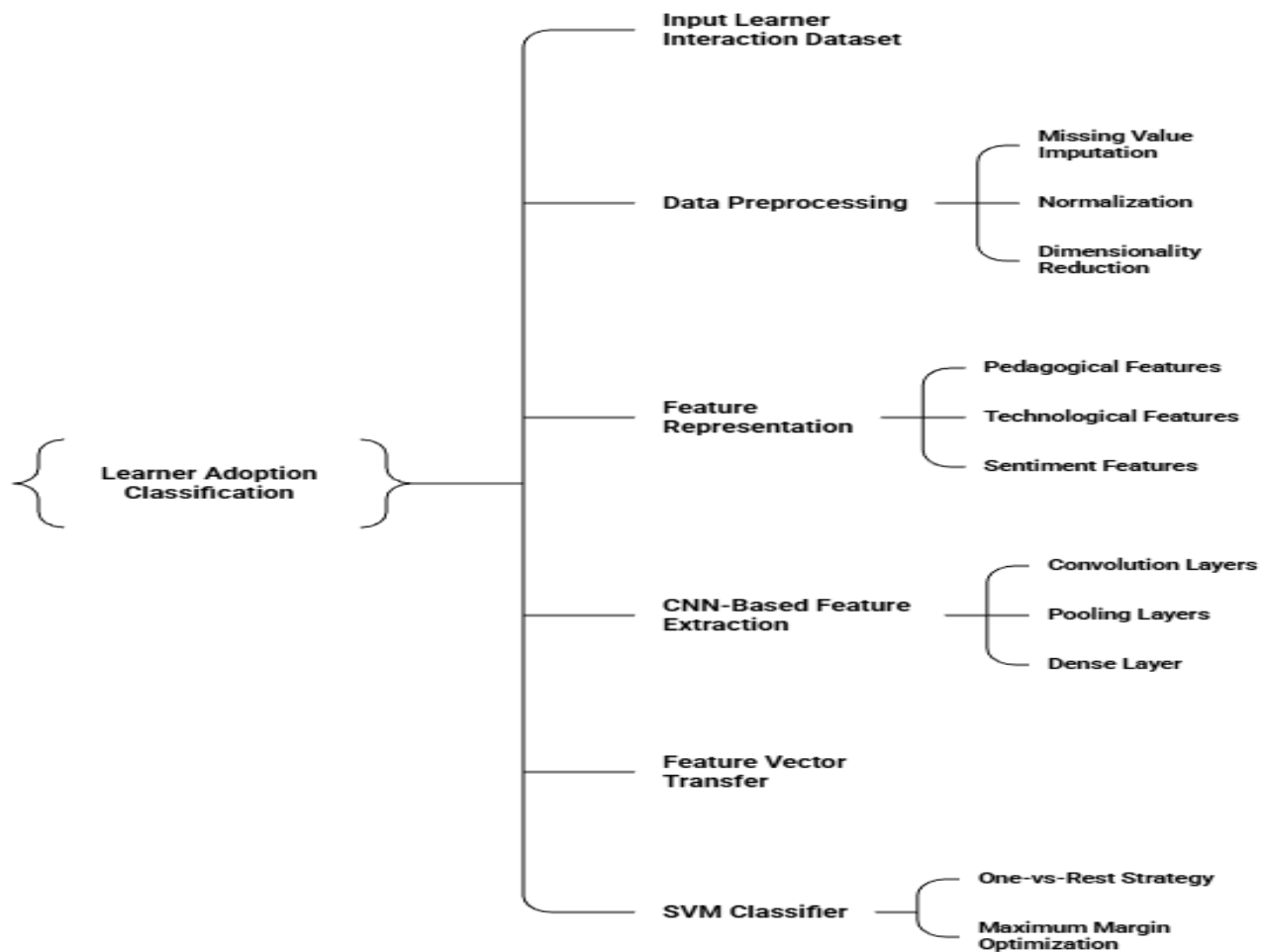


Figure. 1 illustrates the workflow of the proposed CNN–SVM learner adoption classification framework. Initially, learner interaction data are preprocessed through missing value imputation, normalization and dimensionality reduction. The refined features are then categorized into pedagogical, technological and sentiment-based representations. A convolutional neural network is employed to extract high-level feature embeddings, which are subsequently passed to a support vector machine classifier. The SVM performs adoption classification using a maximum-margin decision strategy. Finally, the model performance is evaluated using standard classification metrics.

Table 3: Computational Complexity of the Proposed CNN–SVM Framework

Component	Operation	Time Complexity	Remarks
Data Preprocessing	KNN Missing Value Imputation	$O(n \cdot k \cdot d)$	Distance-based estimation of missing attributes
Data Preprocessing	Dimensionality Reduction (SVD)	$O(\min(n \cdot d^2, d \cdot n^2))$	Reduces redundant and non-informative features
CNN Feature Extraction	Convolution Operations	$O(f \cdot k^2 \cdot n)$	f : filters, k : kernel size
CNN Feature Extraction	Pooling and Dense Layers	$O(n \cdot d')$	d' : extracted feature dimension
Classification	Linear SVM Training	$O(n \cdot d')$	Maximum-margin learning
Classification	SVM Prediction	$O(d')$	Efficient inference

Table 3. summarizes the computational complexity of each component in the proposed CNN–SVM framework, highlighting the scalability of the model for large-scale learner adoption classification.

5. Experimental Setup

5.1 Implementation Details

The proposed model is implemented in Python using TensorFlow, Scikit-learn, NumPy and Pandas libraries. The CNN is trained using a cross-entropy loss function, while the SVM employs a linear kernel for classification. Hyperparameters such as learning rate, batch size and number of convolutional filters were optimized empirically.

Experiments were conducted in the Google Colab environment. The CNN model was trained using the Adam optimizer with a learning rate of 0.001, batch size of 32 and 50 training epochs. GPU acceleration was utilized to improve computational efficiency during training.

Table 4: CNN- SVM Hyperparameter Configuration

Parameter	Value
Kernel Type	Linear
Regularization (C)	1.0
Multi-class Strategy	One-vs-Rest
Optimization	Maximum margin
Learning Rate	0.001
Batch Size	32
Epochs	50
Optimizer	Adam
Activation Function	ReLU

5.2 Cross-Validation Strategy

The proposed CNN–SVM model was evaluated using 10-fold cross-validation to ensure robustness, reduce overfitting and improve statistical reliability. The learner interaction dataset was divided into ten mutually exclusive subsets. During each iteration, nine subsets were utilized for training while the remaining subset was used for testing. The process was repeated ten times so that each subset participated in testing exactly once. The final performance metrics were computed as the average across all validation folds.

5.3 Evaluation Metrics

Model performance is evaluated using precision, recall and accuracy derived from the confusion matrix. These metrics provide a comprehensive assessment of classification effectiveness across adoption categories.

Performance Metrics

The performance of the proposed CNN–SVM learner adoption classification model is evaluated using precision, recall and accuracy, which are derived from the confusion matrix. The confusion matrix consists of four fundamental components: true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN).

Precision

Precision measures the proportion of correctly predicted positive instances among all instances predicted as positive.

$$Precision = TP / TP + FP$$

Recall

Recall measures the proportion of correctly predicted positive instances among all actual positive instances.

$$Recall = TP / TP + FN$$

F1-Score

F1-score represents the harmonic mean of precision and recall and provides a balanced evaluation metric for multi-class learner adoption classification. It is particularly effective when class distributions are imbalanced.

$$F1 - Score = (2 \times Precision \times Recall) / (Precision + Recall)$$

A higher F1-score indicates improved balance between precision and recall, reflecting stronger learner adoption classification performance.

Accuracy

Accuracy represents the overall correctness of the classification model by measuring the proportion of correctly classified instances among all instances.

$$Accuracy = TP + TN / TP + TN + FP + FN$$

Table 5: Performance Evaluation Metrics

Metric	Description
Precision	Correctly classified adoption samples among predicted samples
Recall	Correctly classified adoption samples among actual samples
Accuracy	Overall classification correctness
F1-Score	Harmonic mean of precision and recall

6. Results and Discussion

Experimental results demonstrate that the proposed CNN–SVM model achieves superior performance compared to baseline models, including CNN with softmax, standalone SVM and random forest classifiers. The hybrid approach consistently improves precision and recall across all adoption classes.

The improved performance can be attributed to the CNN’s ability to learn discriminative feature representations and the SVM’s robustness in classification. The results confirm that decoupling feature learning and classification is effective for learner adoption prediction in inclusive e-learning environments.

Table 6: Accuracy Comparison of Learner Adoption Classification Models

Model	High Adoption (%)	Medium Adoption (%)	Low Adoption (%)	Overall Accuracy (%)
CNN (Softmax)	90.4	89.6	89.1	89.7
Support Vector Machine (SVM)	88.9	88.1	87.6	88.2
Random Forest	91.2	90.5	90.1	90.6

Proposed CNN-SVM	98.1	98.2	98.0	98.1
------------------	------	------	------	------

Table 6 presents the accuracy comparison of the proposed CNN-SVM model against baseline classifiers. The hybrid CNN-SVM framework achieves the highest overall accuracy across all adoption categories, demonstrating the effectiveness of combining deep feature extraction with maximum-margin classification.

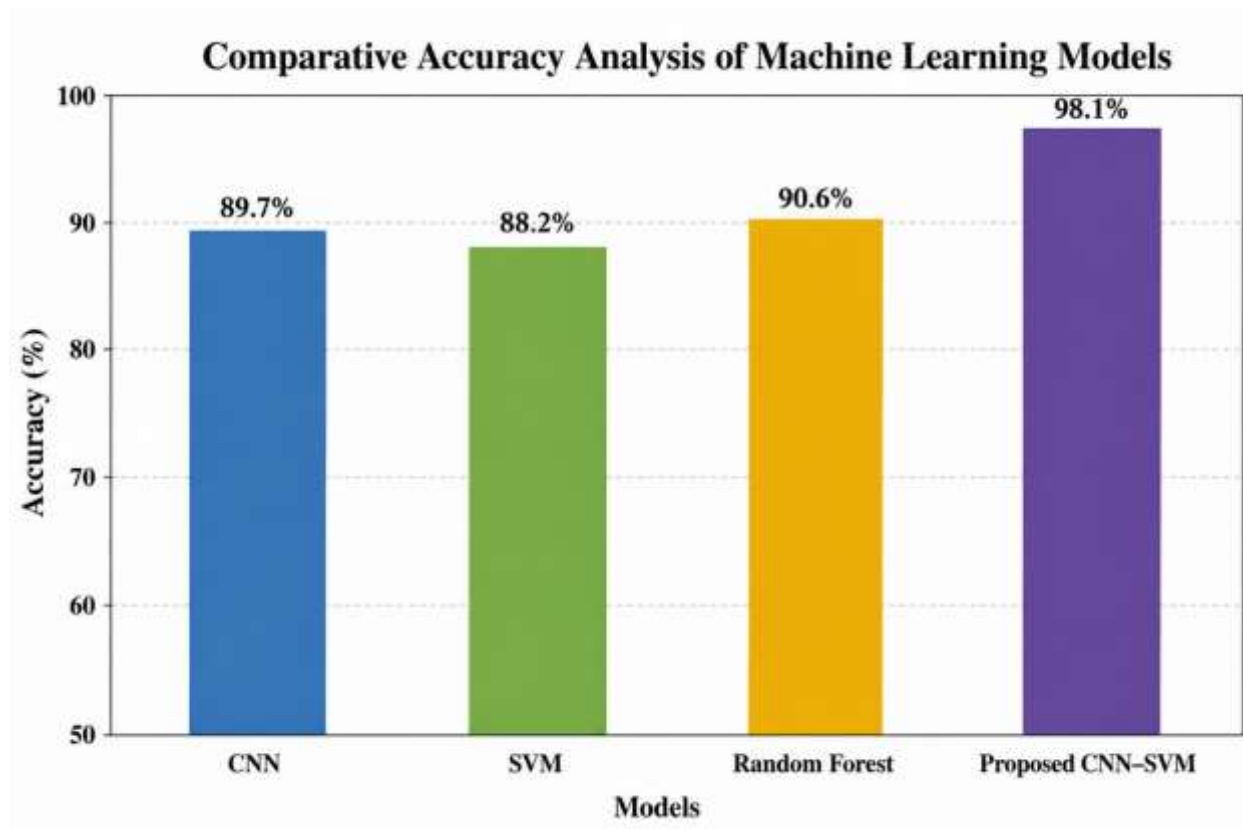


Figure 4: Comparative Accuracy Analysis of Machine Learning Models

The proposed CNN-SVM framework demonstrates superior learner adoption classification accuracy compared to conventional machine learning models.

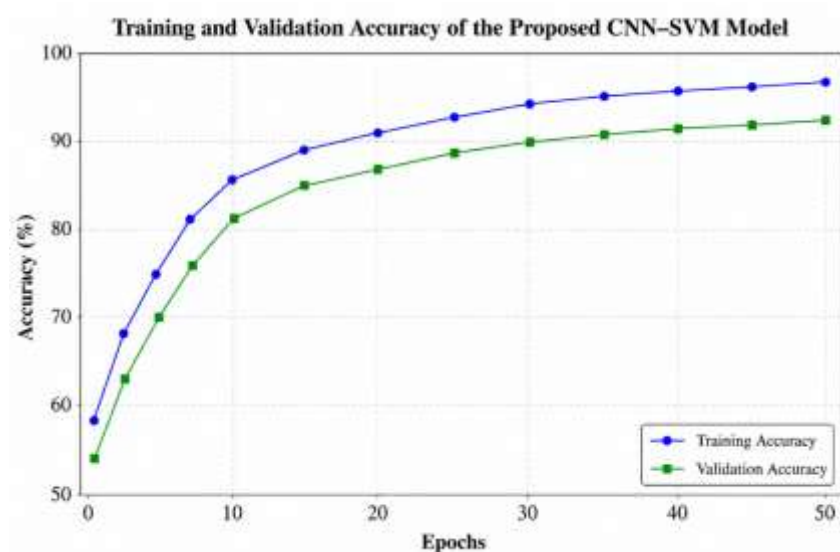
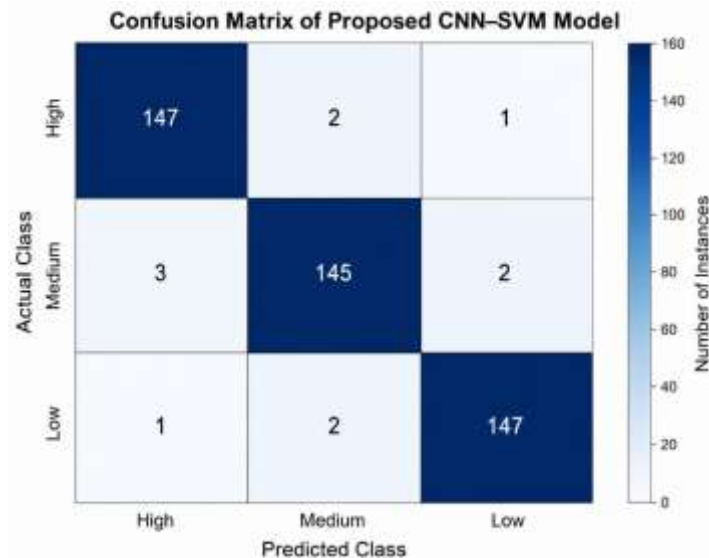


Figure 5: Training and Validation Accuracy of the Proposed CNN-SVM Model

The proposed CNN-SVM framework demonstrates stable convergence with increasing training and validation accuracy across training epochs while maintaining minimal overfitting.

Table 7: Confusion Matrix of Proposed CNN–SVM Model

Actual / Predicted	High	Medium	Low
High	147	2	1
Medium	3	145	2
Low	1	2	147


Figure 6: Confusion Matrix of the Proposed CNN–SVM Framework

The confusion matrix analysis demonstrated that the proposed CNN–SVM framework achieved consistent classification performance across high, medium and low learner adoption categories with minimal misclassification.

6.1 Ablation Study

An ablation study was conducted to evaluate the contribution of individual components within the proposed CNN–SVM framework. The performance of standalone CNN, standalone SVM, CNN with Softmax classifier and the proposed CNN–SVM hybrid model were compared.

The standalone CNN model demonstrated effective feature extraction capability but showed comparatively lower classification generalization. The standalone SVM classifier achieved stable decision boundaries but lacked deep representation learning. The hybrid CNN–SVM architecture achieved superior performance by combining discriminative CNN feature extraction with robust maximum-margin classification.

The results confirm that decoupling feature extraction and classification improves learner adoption prediction performance in inclusive e-learning environments.

7. Limitations

Although the proposed CNN–SVM framework demonstrated strong learner adoption classification performance, the study has several limitations. The dataset is limited to contextual learner interaction data obtained from a single repository and may not fully represent inclusive e-learning environments. In addition, the proposed model does not explicitly capture temporal learner behavior evolution. Future work will focus on large-scale multi-institutional datasets and temporal deep learning architectures for dynamic learner adoption prediction.

8. Conclusion and Future Work

This paper presented a CNN–SVM hybrid framework for learner adoption classification in inclusive e-learning systems. By leveraging CNN-based feature extraction and SVM-based classification, the proposed approach effectively models complex learner interaction data while maintaining robust generalization. Experimental evaluation demonstrated improved performance over conventional deep learning and machine learning models.

Future work will focus on incorporating temporal modeling techniques and attention mechanisms to capture evolving learner engagement patterns and further enhance adoption prediction accuracy.

References

- [1] R. S. Baker and P. S. Inventado, "Educational data mining and learning analytics," in *Learning Analytics: From Research to Practice*, J. A. Larusson and B. White, Eds. Berlin, Germany: Springer, 2014, pp. 61–75, doi: 10.1007/978-1-4614-3305-7_4.
- [2] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 6, pp. 601–618, Nov. 2010, doi: 10.1109/TSMCC.2010.2053532.
- [3] O. Zawacki-Richter, V. I. Marín, M. Bond and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education," *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, pp. 1–27, 2019, doi: 10.1186/s41239-019-0171-0.
- [4] S. B. Kotsiantis, I. D. Zaharakis and P. E. Pintelas, "Machine learning: A review of classification and combining techniques," *Artificial Intelligence Review*, vol. 26, no. 3, pp. 159–190, 2006, doi: 10.1007/s10462-007-9052-3.
- [5] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [6] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [7] X. Zhang, J. Zhao and Y. LeCun, "Character-level convolutional networks for text classification," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2015, pp. 649–657.
- [8] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1007/BF00994018.
- [9] V. N. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, 1998.
- [10] Z. Liu, J. Tang, Y. Han, M. Chen and J. Wang, "A hybrid CNN–SVM approach for text classification," *Knowledge-Based Systems*, vol. 163, pp. 852–861, Jan. 2019, doi: 10.1016/j.knosys.2018.10.015.
- [11] W. Xing, R. Guo, E. Petakovic and S. Goggins, "Participation-based student final performance prediction model through interpretable genetic programming," *Journal of Educational Data Mining*, vol. 7, no. 2, pp. 1–36, 2015.
- [12] O. C. Santos and J. G. Boticario, "User-centred design and educational data mining support during the recommendations generation process," *Expert Systems with Applications*, vol. 42, no. 8, pp. 4214–4226, May 2015, doi: 10.1016/j.eswa.2015.01.041.
- [13] B. Tang, H. He, P. M. Baggenstoss and S. Kay, "A Bayesian classification approach using class-specific features for text categorization," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 9, pp. 2158–2171, Sept. 2014, doi: 10.1109/TKDE.2013.150.
- [14] S. Wang and R. M. Summers, "Machine learning and radiology," *Medical Image Analysis*, vol. 16, no. 5, pp. 933–951, July 2012, doi: 10.1016/j.media.2012.02.005.
- [15] C. Xing, H. Hu and S. Mai, "Deep learning-based engagement analysis in online education," *IEEE Access*, vol. 8, pp. 102560–102572, 2020, doi: 10.1109/ACCESS.2020.2998692.
- [16] J. Kim, J. Kim and S. Kim, "A CNN-based hybrid model for learning engagement prediction in e-learning systems," *Computers & Education*, vol. 151, pp. 103–118, 2020, doi: 10.1016/j.compedu.2020.103865.
- [17] M. Khalil and M. Ebner, "Learning analytics: Principles and constraints," *Journal of Learning Analytics*, vol. 2, no. 3, pp. 178–203, 2015.

- [18] R. D. Baker, "Stupid tutoring systems, intelligent humans," *International Journal of Artificial Intelligence in Education*, vol. 26, no. 2, pp. 600–614, 2016.
- [19] P. Brusilovsky and E. Millán, "User models for adaptive hypermedia and adaptive educational systems," *The Adaptive Web*, Springer, pp. 3–53, 2007.
- [20] M. F. Naeem, R. Gohar and A. Khan, "Student engagement prediction using deep learning in online learning environments," *IEEE Access*, vol. 9, pp. 166053–166064, 2021, doi: 10.1109/ACCESS.2021.3135779.
- [21] S. M. Jayaprakash, E. W. Moody, E. J. M. Lauría, J. R. Regan and J. D. Baron, "Early alert of academically at-risk students," *Journal of Learning Analytics*, vol. 1, no. 1, pp. 6–47, 2014.
- [22] Y. B. Moon, "The learning machine: Predicting learning outcomes using machine learning," *Decision Support Systems*, vol. 124, pp. 113–126, 2019, doi: 10.1016/j.dss.2019.113126.
- [23] A. Peña-Ayala, "Educational data mining: A survey and a data mining-based analysis of recent works," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1432–1462, 2014.
- [24] S. L. Lim, J. W. Park and S. Y. Park, "Deep learning-based learner behavior analysis in MOOCs," *IEEE Transactions on Learning Technologies*, vol. 13, no. 4, pp. 728–740, 2020.
- [25] H. Drachsler and W. Greller, "Privacy and analytics – It's a DELICATE issue," *Journal of Learning Analytics*, vol. 3, no. 1, pp. 89–98, 2016.