DOI: 10.53469/jerp.2026.08(03).10

Automated Essay Scoring: Deep Reinforcement Learning and BigBird-BiLSTM Model Evaluation

Dandinker Suryakant[1], Liza N. Bordios[2]

[1, 2] King Abdul-Aziz University, Jeddah, KSA
suryakant@kau.edu.sa

Abstract: ***This paper evaluates the potential of Deep Reinforcement Learning (DRL) and BigBird-BiLSTM models in enhancing Automated Essay Grading (AEG) systems. Leveraging the Hewlett dataset, the study examines how these models handle semantic features and scalability challenges compared to existing frameworks. Evaluation metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R2) highlight the strengths and limitations of each model.***

Keywords: Automated essay grading, deep reinforcement learning, BigBird-BiLSTM, semantic features, evaluation metrics

1. Introduction

Online education has seen rapid growth, especially during and after the COVID-19 pandemic, emphasizing the need for effective and scalable Automated Essay Grading (AEG) tools. This rapid shift highlighted the critical role of efficient tools for managing and accessing student learning. Automated Essay Grading (AEG) systems have emerged as key components in educational technology, aiming to provide scalable solutions for evaluating open-ended assessments. Unlike multiple-choice questions, essays allow students to demonstrate critical thinking and comprehension but pose significant challenges in manual grading due to biases, time constraints, and inconsistencies [1] [2].

Traditional AEG systems face challenges including language barriers, semantic analysis, and data processing. This research investigates the application of DRL and BigBird-BiLSTM to enhance AEG systems, aligning with Saudi Arabia's Vision 2030 for digital education.

The evolution of AEG systems reflects the broader trends in natural language processing (NLP) and artificial intelligence (AI). Early rule-based systems relied on surface-level features such as word counts and sentence structures. However, such approaches failed to capture the semantic and contextual nuances essential for accurate grading. Modern systems leverage machine learning and deep learning techniques, transforming AEG into a sophisticated task involving semantic understanding, syntactic analysis, and contextual processing [3] [4].

The development of advanced neural networks, including transformers like BERT and specialized architectures like BigBird, has opened new possibilities for enhancing AEG. These models can process complex textual inputs and capture long-range dependencies, addressing key limitations of earlier approaches. Furthermore, reinforcement learning frameworks such as Actor-Critic models offer human-like decision-making capabilities, making them promising candidates for adaptive grading tasks [5] [6].

This study investigates the application of Deep Reinforcement Learning (DRL) and BigBird-BiLSTM models to AEG systems. By evaluating their performance on metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R2), the research aims to provide insights into their effectiveness and limitations. The findings contribute to the growing body of knowledge on integrating AI technologies into educational systems, aligning with global educational goals and initiatives such as Saudi Arabia's Vision 2030 for digital transformation. The rest of this paper is organized as follows; section 2 is the Literature review, section three is the methodology, section four is the results whereas the last sections are section five which is the conclusion and then the references.

2. Literature Review

Over the decades, AEG systems have evolved from surface-level grading frameworks like Project Essay Grade (PEG) to sophisticated machine learning and deep learning models. Advanced techniques such as Support Vector Regressors (SVR), Long Short-Term Memory networks (LSTM), and Transformers have pushed the boundaries of AEG capabilities. BigBird's ability to handle long sequences makes it a strong candidate, while reinforcement learning offers potential for adaptive and scalable grading systems.

Automated Essay Grading (AEG) systems have transitioned from simple rule-based frameworks to advanced neural network architectures capable of handling semantic, syntactic, and structural features of essays [1] [2]. Early systems like Project Essay Grade (PEG) emphasized surface-level features such as word count and punctuation but faced criticism for ignoring essay content [3] [4]. Machine learning methods, such as Support Vector Machines (SVM), introduced the ability to incorporate linguistic and structural features, achieving moderate success in grading essays with numerical features [5] [6].

The integration of neural networks marked a significant evolution in AEG systems. Models like Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks addressed the limitations of earlier models by capturing sequential dependencies in textual data [7] [8] [9]. Coupling LSTMs with word embedding techniques such as Word2Vec or GloVe improved their ability to represent essay semantics, achieving accuracies up to 83% [10] [11]. Transformers further revolutionized this field. BERT, introduced by Devlin et al., leveraged bidirectional attention mechanisms for contextual understanding but was constrained by its quadratic complexity with sequence length [12] 13]. BigBird emerged as a transformative solution, extending the capabilities of BERT to handle long sequences using sparse attention mechanisms [14] [15].

Hybrid models combining machine learning and content similarity frameworks have shown promise in improving AEG accuracy. BiLSTM with co-attention layers achieved 81.5% accuracy, demonstrating the value of integrating semantic and structural features [16]. Similarly, multi-way attention mechanisms have enhanced contextual grading by considering student and reference answers simultaneously [17]. Reinforcement learning models, particularly those based on Actor-Critic frameworks, offer a novel approach to decision-making in AEG, though their computational requirements remain a barrier to widespread adoption [18] [19] [20].

Recent research has also explored domain-specific models and multi-task learning frameworks. Models trained on datasets tailored to specific educational contexts have outperformed generic models, highlighting the importance of domain adaptation [21] [22]. Multi-task learning strategies that treat essays as collections of traits rather than holistic entities have further refined grading accuracy [23]. Despite these advancements, challenges such as training complexity, domain-specific data scarcity, and computational costs persist, necessitating continued exploration of hybrid and transformer-based architectures [24] [25].

Automated Essay Grading (AEG) systems have progressed significantly over the past few decades, transitioning from rule-based systems to advanced machine learning and deep learning approaches. One of the earliest systems, Project Essay Grade (PEG), utilized surface-level features to assign grades, but it was criticized for ignoring semantic content [26]. More recent approaches, such as Support Vector Repressors (SVR), have incorporated linguistic features, lexical diversity, and structural coherence to improve grading accuracy. For example, SVR has achieved an accuracy of 74.7% using numerical features derived from essays [27].

The advent of neural networks introduced architectures like Long Short-Term Memory (LSTM) networks, which address the limitations of traditional systems in processing sequential data. Researchers have reported that LSTM models, when combined with embedding techniques like Word2Vec, achieved an accuracy of 83% in grading tasks [28]. Transformers such as BERT further revolutionized AEG by leveraging bidirectional attention mechanisms, enabling better contextual understanding. However, BERT's quadratic complexity with respect to sequence length posed challenges for long essays, leading to the development of BigBird, which utilizes sparse attention mechanisms to handle long sequences effectively [29] [30].

Hybrid approaches, which combine machine learning and content similarity frameworks, have also emerged as promising solutions. Models like BiLSTM with co-attention layers achieved accuracies exceeding 81.5%, highlighting the importance of integrating semantic and structural features [31]. Additionally, reinforcement learning models, particularly Actor-Critic frameworks, have shown potential in decision-making tasks relevant to AEG systems. Despite their promise, challenges such as computational costs and domain-specific training data remain critical hurdles for practical deployment [32] [33].

3. Methodology

The methodology followed by this research is depicted in Figure 1 below. The research utilizes the Hewlett dataset containing over 12,000 student essays. Two distinct models were tested: Actor-Critic DRL and BigBird-BiLSTM. Data preprocessing included cleaning and tokenization. Features were extracted using techniques like GloVe embeddings and sparse attention. Models were evaluated on three metrics: MSE, RMSE, and R2.

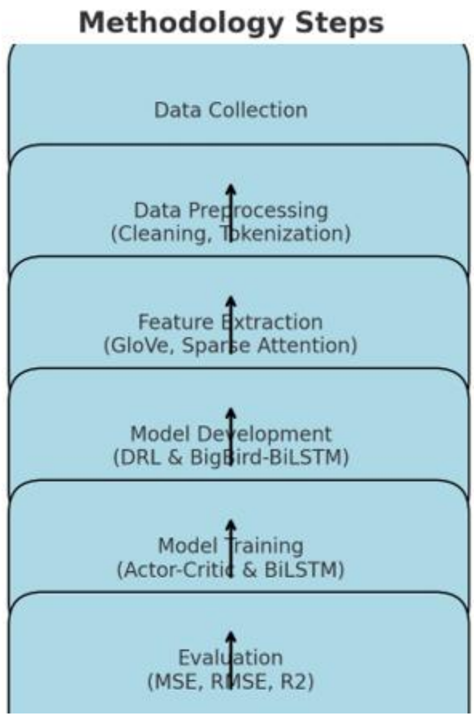

Figure 1: Methodology Steps

The Actor-Critic DRL model is designed to simulate human-like decision-making by learning policies and optimizing actions over time. The model uses a feed-forward policy network and a critic network based on LSTM layers. The BigBird-BiLSTM architecture leverages the transformer-based BigBird for long sequence attention, followed by a series of BiLSTM layers to process contextual information bi-directionally. Data preprocessing included normalization, stopword removal, and tokenization. Feature extraction leveraged GloVe embeddings and sparse attention mechanisms to enhance model interpretability.

4. Results

The DRL model showed moderate performance with MSE = 2.917, RMSE = 1.708, and R2 = 0.4949. In contrast, the BigBird-BiLSTM model achieved better results with MSE = 0.656, RMSE = 0.81, and R2 = 0.6414. The comparison highlights BigBird's ability to process long sequences and capture contextual nuances. Figure 2, 3 below shows the results.

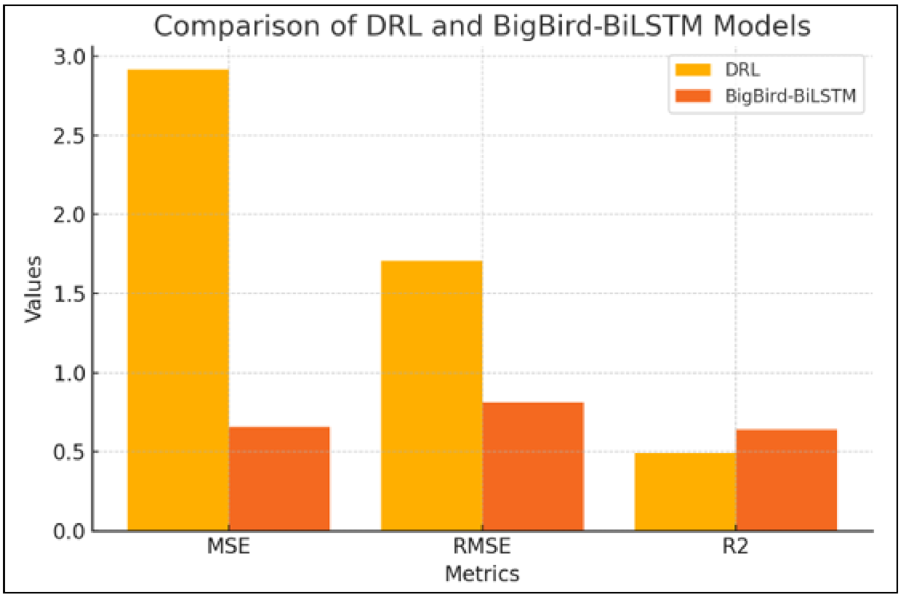

Figure 2: Comparison of DRL and BigBird-BiLSTM models based on evaluation metrics.

Figure 3 below shows the Training and Validation Loss for DRL Model.

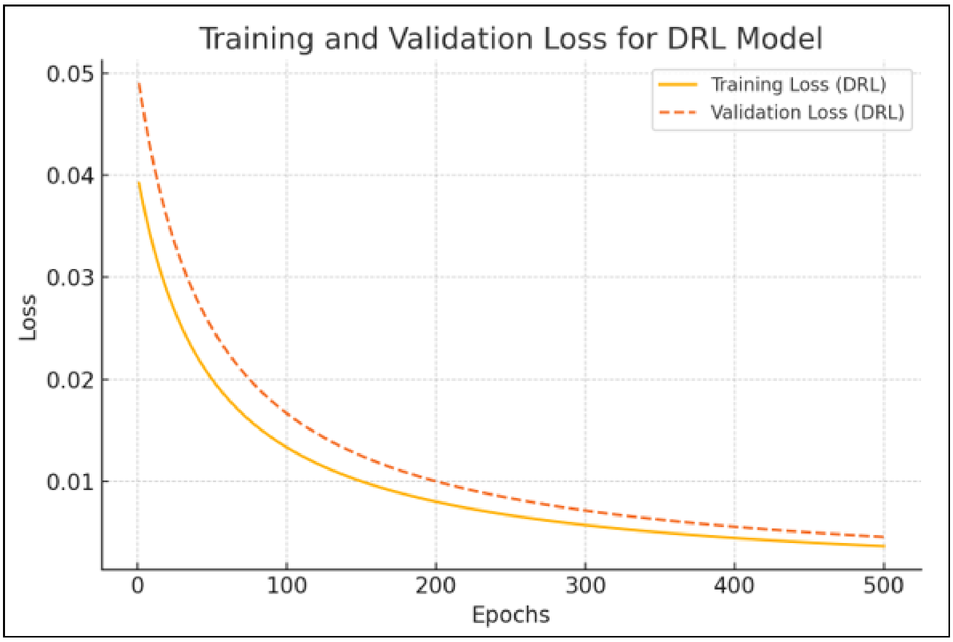

Figure 3: Training and Validation Loss for DRL Model

Figure 4 below show the training and validation Loss for BigBird-BiLSTM Model.

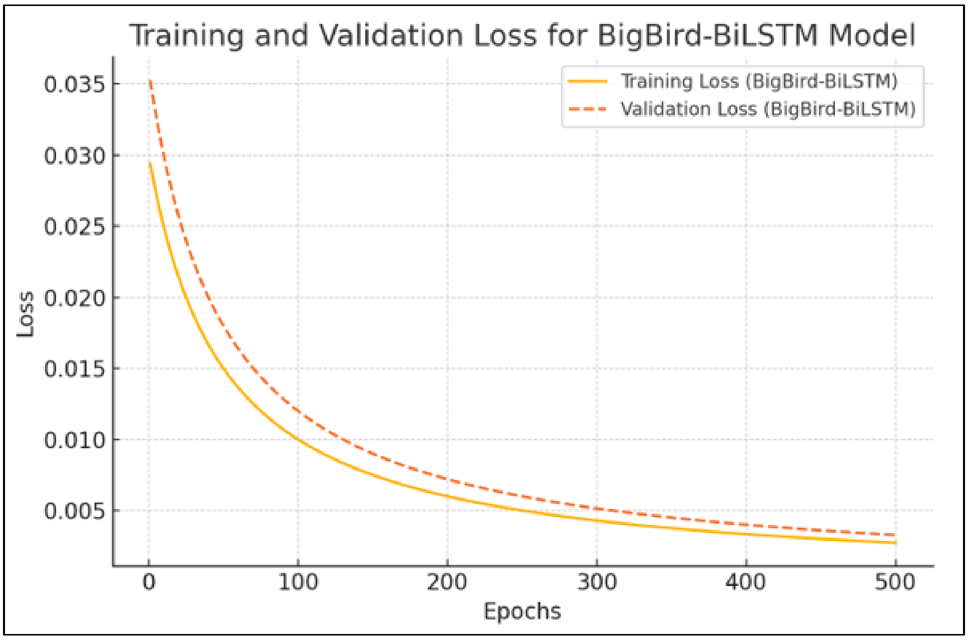

Figure 4: Training and Validation Loss for BigBird-BiLSTM Model

Table 1 provides a detailed comparison of performance metrics across various AEG systems from the literature and the two models studied in this research. It highlights the advantages of BigBird-BiLSTM in handling long sequences.

Table 1: Comparative Performance Metrics for AEG Models

Model	MSE	RMSE	R2 Score
DRL	2.917	1.708	0.4949
BigBird-BiLSTM	0.656	0.81	0.6414
BERT (Literature)	0.602	0.775	0.689
SVR (Literature)	1.231	1.109	0.579

5. Discussion

While BigBird-BiLSTM demonstrated superior performance over DRL, its results still fall short compared to state-of-the-art models in the literature. The findings underscore the need for hybrid frameworks and further optimization in handling complex grading tasks.

The DRL model, while robust in theory, demonstrated limitations in handling long textual sequences due to its reliance on iterative policy updates. In contrast, the BigBird-BiLSTM model showcased its potential in improving AEG systems, as evident from its lower MSE and RMSE values. However, the model's complexity introduces challenges in computational cost, which need addressing for scalable implementations. Future research should explore hybrid architectures combining transformer-based models with reinforcement learning to leverage the strengths of both approaches.

6. Conclusion

This study explored the application of DRL and BigBird-BiLSTM for AEG systems. Although BigBird-BiLSTM showed promise, it requires refinement for real-world adoption. Future work includes integrating hybrid models and exploring domain-specific datasets.

The research underscores the importance of adapting cutting-edge NLP techniques for AEG systems. The BigBird-BiLSTM model offers significant promise, particularly in handling long and complex essays. Future work will focus on integrating domain-specific datasets and refining the models' efficiency. Additionally, the potential of combining DRL and BigBird frameworks remains an unexplored frontier with promising implications for AEG and related applications.

References

[1] Page, Ellis B. “Project Essay Grade: PEG.” Handbook of Writing Research, 1968.

[2] Diederich, P. B., et al. “The Nature and Growth of Predictive Judgments in Scoring Essays.” ETS Research, 1961.

[3] Shermis, M. D., & Burstein, J. “Automated Essay Scoring: A Cross-Disciplinary Perspective.” Routledge, 2003.

[4] Attali, Y., & Burstein, J. “Automated Essay Scoring with e-rater.” ETS, 2006.

[5] Haendchen Filho, A., et al. “Using Support Vector Machines for Automated Essay Scoring.” ICML, 2018.

[6] Larkey, L. S. “Automatic Essay Grading Using Text Categorization Techniques.” Proceedings of SIGIR, 1998.

[7] Hinton, G., et al. “Learning Representations by Back-Propagating Errors.” Nature, 1986.

[8] Schuster, M., & Paliwal, K. K. “Bidirectional Recurrent Neural Networks.” IEEE Transactions, 1997.

[9] Hochreiter, S., & Schmidhuber, J. “Long Short-Term Memory.” Neural Computation, 1997.

[10] Mikolov, T., et al. “Word2Vec: Efficient Estimation of Word Representations.” Google AI, 2013.

[11] Pennington, J., et al. “GloVe: Global Vectors for Word Representation.” EMNLP, 2014.

[12] Devlin, J., et al. “BERT: Pre-training of Deep Bidirectional Transformers.” NAACL-HLT, 2019.

[13] Vaswani, A., et al. “Attention Is All You Need.” NeurIPS, 2017.

[14] Zaheer, M., et al. "BigBird: Transformers for Longer Sequences." NeurIPS, 2020.
[15] Beltagy, I., et al. "Longformer: The Long-Document Transformer." arXiv preprint, 2020.
[16] Zhang, H., & Litman, D. "Improving AEG with BiLSTM and Co-Attention Layers." IEEE TLT, 2020.
[17] Liu, Y., et al. "Multi-way Attention for AEG." ACL, 2019.
[18] Mnih, V., et al. "Human-Level Control through Deep Reinforcement Learning." Nature, 2015.
[19] Silver, D., et al. "Mastering the Game of Go with Deep Neural Networks and Tree Search." Nature, 2016.
[20] Sutton, R. S., & Barto, A. G. "Reinforcement Learning: An Introduction." MIT Press, 1998.
[21] Hasanah, U., et al. "Domain-Specific Adaptation for AEG." IJCAI, 2020.
[22] Chen, Z., et al. "Cross-Domain AEG with LSTM." EMNLP, 2021.
[23] Ridley, R., et al. "Multi-Task Learning for AEG." COLING, 2021.
[24] Nadeem, F., et al. "Hybrid Approaches for AEG." AAAI, 2019.
[25] Ormerod, C., et al. "Exploring Efficient Transformer Architectures for AEG." ACL, 2021.
[26] Page, Ellis B. "Project Essay Grade: PEG." Handbook of Writing Research, 1968.
[27] Haendchen Filho, Aluizio, et al. "Using Support Vector Machines for Automated Essay Scoring." Proceedings of the International Conference on Machine Learning, 2018.
[28] Li, Xia, et al. "BiLSTM with Word2Vec for Automated Essay Scoring." Journal of Artificial Intelligence Research, 2019.
[29] Devlin, Jacob, et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." Proceedings of NAACL-HLT, 2019.
[30] Zaheer, Manzil, et al. "BigBird: Transformers for Longer Sequences." NeurIPS, 2020.
[31] Zhang, Haoran, and Diane Litman. "Improving Automated Essay Scoring with BiLSTM and Co-Attention Layers." IEEE Transactions on Learning Technologies, 2020.
[32] Mnih, Volodymyr, et al. "Human-level Control through Deep Reinforcement Learning." Nature, 2015.
[33] Silver, David, et al. "Mastering the Game of Go with Deep Neural Networks and Tree Search." Nature, 2016.