2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS) | 979-8-3503-4233-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICTACS59847.2023.10390042

English Language and Linguistic Teaching Strategies Based on Deep Learning

Dr. Mohd Shamim[1], Sourav Choudhury[2], Joyirsiram[3], Dr. Suresh[4], Frederick[4], Akshay Ashok Bannatti[5], L. Rahunathan[6]

[1]*Department of English, Halim Muslim PG College, Chamanganj, Kanpur, UP, India*
[2]*Lovely Professional University, Phagwara, Punjab*
[3]*NITjote, India*
[4]*Department of English, Bishop Heber College, Trichy*
[5]*Sathyabama Institute of Science and Technology*
[6]*Kongu Engineering College*
E-mail: [1]shamimmohd2009@gmail. Com

***Abstract*—This research paper investigates the potential of in corpora ting deep learning techniques into English language and linguistic teaching strategies. With the rapid advancements in deep learning and its proven success in various domains, exploring its application in language education holds significant promise. The paper reviews existing literature on deep learning and language processing, examines case studies of deep learning in educational settings, and proposes practical strategies for implementing deep learning in English language and linguistic instruction. The goal is to assess the benefits, challenges, and implications of using deep learning in language teaching and contribute to the ongoing conversation on innovative approaches to enhance language learning out comes. Deep learning, a subfield of machine learning, has witnessed remarkable advancements and demonstrated exceptional capabilities in various domains. In recent years, re- searchers and educators have started exploring the integration of deep learning technique using to language teaching methodologies. This research paper aims to investigate the potential benefits, challenges, and implementation strategies of using deep learning in English language and linguistic education. The paper reviews existing literature, examines case studies, and presents practical insights to enhance the effectiveness of language instruction through deep learning methodologies.**

Keywords: English Language Teaching(ESL), Deep Learning(DL), Machine Learning(ML), Natural Language Processing(NPL)

I. Introduction

The integration of Artificial Intelligence (AI) into numerous educational disciplines has brought about a transformational shift in teaching methodologies and classroom dynamics. As a consequence of AI's contributions to the improvement of teaching platforms, English language education evaluation, and tutoring, both teaching effectiveness and the learning environment have seen significant improvements. With the advent of AI-powered education systems, teachers may now collaborate on lesson planning regardless of their physical location.

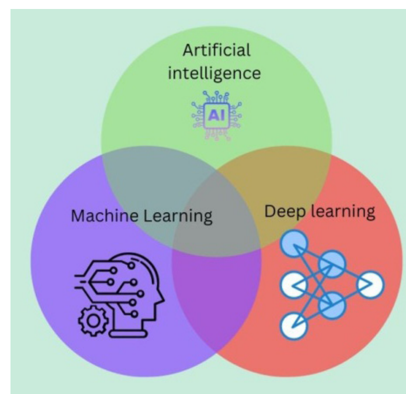

Fig. 1: Deep Learning is the Part of Machine Learning

The need of fair and helpful assessments is also stressed so that problems with the teaching and learning process may be found and fixed as soon as possible. Machine learning is an area of artificial intelligence that helps computers gain knowledge and proficiency via trial and error. These programs can learn broad rules and recognize patterns from structured data by studying massive datasets. These algorithms are skilled in categorizing and differentiating based on learnt qualities from sample data, and they may use this knowledge to anticipate outcomes in actual circumstances. Deep learning is a subfield of machine learning that use layered algorithms to simulate neural networks with the ability to learn and make decisions independently. Fig 1 shows that when trained with big data—a term referring to vast amounts of data—these models perform very well. One of deep learning's defining characteristics is its automatic ability to make sense of input that lacks a predetermined structure[1], [2].

Customized education, as opposed to the one-size-fits-all approach of traditional education, promises to bring forth more innovation and unique contributions to society. It requires teachers to understand students' psychological traits and interests to formulate tailored developmental plans. AI in education aids in this customization by providing a vast database that can suggest various learning plans based on individual students' learning preferences, capabilities, and progress. AI in education represents a synergy between human and machine, where the technology supports educators in meeting instructional demands, yet the teacher remains pivotal to the educational process. For English learners specifically, AI can create a foundational model based on their characteristics and refine their learner profiles for more targeted English instruction. This synergy between AI and education not only streamlines resource utilization but also enhances the overall learning experience, making education more effective and personalized.

II. Literature Review

Research in linguistics has explored a variety of statistical methods used in the analysis of second language acquisition. Multiple regression analyses have been compared with ANOVA techniques to evaluate their effectiveness. Additionally, linear mixed-effects models have been investigated for their applicability in psycholinguistics. Issues such as the handling of outliers and the use of exploratory factor analysis in linguistic research have been discussed, with current approaches being summarized. The unsuitability of inferential statistics for English language education research has also been highlighted, alongside a critical examination of significance testing in language studies, advocating for methodological reform. The evolution of research approaches in second language studies has been documented, reflecting on the history and advancements in the field. Reviews have covered the range of quantitative techniques employed in corpus linguistics, including ANOVA and factor analysis, while also addressing practical concerns such as sample size planning. Non-normal data handling via quintile regression with bootstrapping has been described, contributing to the broader understanding of applied linguistics methodologies[3]–[5].

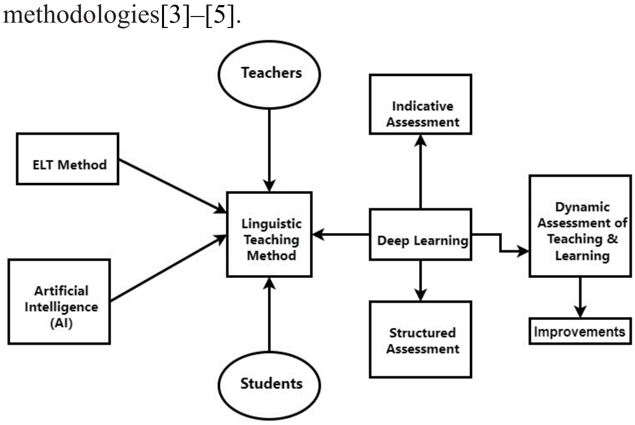

Fig. 2: Teaching Process and assessment

Fig 2 illustrates a flowchart detailing the integration of Artificial Intelligence (AI) into the English Language Teaching (ELT) method. It shows the process beginning with AI influencing linguistic teaching methods, which then informs the interaction between teachers and students, incorporates both structured and indicative assessments, and employs deep learning to facilitate dynamic assessments that lead to improvements in teaching and learning. However, the critique arises from a focus on traditional methods within these studies, prompting a look into the integration of linguistics with deep learning[6]–[8]. The potential synergy between these fields is examined, considering how linguistic theories could inform the development of deep learning models for natural language processing and sentence processing. Moreover, the challenges of describing neural network topologies' inductive biases are discussed, proposing ways to incorporate clear syntactic inductive biases. In practical applications, deep learning has been utilized to create systems like an English to Urdu machine translation system, with performance assessments comparing favorably to existing translation tools. The potential of such systems to overcome language barriers is underscored.

The effect of AI and deep learning on pedagogical methods is investigated as well; this includes the use of mobile apps to teach English to non-native speakers. The application of machine learning to assess and improve students' learning experiences and the use of AR/VR in training and education are also included in this research review. In order to better understand how non-native speakers acquire the language, researchers in the field of English as a Second Language (ESL) have generally concentrated their attention on classroom settings in which English is regularly spoken and utilized. This collection of work seeks to get an understanding of the difficulties and variables that affect language acquisition and competence among students[9]-[13].

III. Proposed Work

The suggested approach uses recurrent neural networks (RNNs) to provide new insights to the study of English. This strategy seeks to improve students' command of the English language by presenting them with a dynamic and engaging learning environment that makes use of language modeling, individualized learning routes, and interactive elements. This approach provides novel opportunities for enhancing language education through the application of AI and NLP methods. RNNs stand out from other neural networks because of their distinctive design, which has a hidden state for each unit that reflects the network's recollection of previous inputs. The network is able to keep a consistent knowledge of the input sequence since this memory is refreshed at each time step. This is especially helpful in language teaching, as students' vocabulary growth is accelerated by drawing on their prior knowledge.

RNNs' capacity to analyze sequences over time is a benefit that's particularly useful for things like time series prediction and increasing the size of the surrounding pixels in image recognition. However, RNNs are notoriously difficult to train and may struggle with extremely lengthy sequences when employing activation functions such as tanh or relu. However, RNNs are useful for tasks such as language modeling, image/face recognition, text-to-speech conversion, and machine translation.Tokenization is a step in natural language processing that reduces large amounts of information to smaller chunks, often individual words. This is a fundamental stage in natural language processing that facilitates the study of simpler forms of more complicated phrases. POS tagging includes categorizing words in a text according to their grammatical function, which is essential for different applications, including text processing and sentiment analysis. Language acquisition relies fundamentally on a sequential representation of facts because of the way in which the order of words transmits meaning. RNNs' ability to retain past inputs, which provides context for comprehending the relationships between words, makes them particularly well-suited to this data. Two natural language processing (NLP) methods that help with linguistic analysis are stemming and lemmatization. The trade-off between speed and linguistic precision in an NLP task will determine which method is used.

ELT, or English Language Teaching, encompasses the methodologies for teaching English to non-native speakers, aiming to enable effective communication in English. The approach chosen for ELT depends on various factors, including the learners' needs and the cultural context.

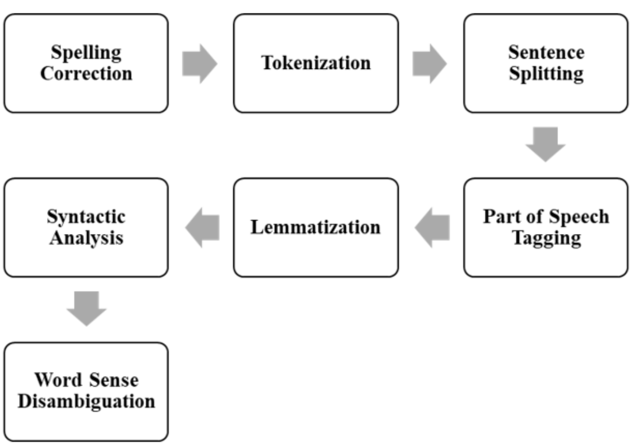

Fig. 3: NPL pipe line used at the English learning method

Fig 3 depicts a sequence of natural language processing (NLP) tasks starting with spelling correction and leading to tokenization, part-of-speech tagging, sentence splitting, and eventually to more complex processes like lemmatization, syntactic analysis, and word sense disambiguation. This flowchart outlines a typical NLP pipeline where text is progressively analyzed and interpreted for computational understanding. Machine learning has greatly assisted in teaching phonetics, particularly in automatic speech recognition and ranking the importance of acoustic features. Communicative Language Teaching (CLT) emphasizes real-world communication skills over rote memorization of grammar and vocabulary, making language acquisition more meaningful for learners.

The importance of machine learning to language education is evident in technology-enhanced learning environments. Machine learning algorithms like decision trees and random forests have been shown to aid in text comprehension and classroom question validation. These algorithms, along with neural networks, have been used to measure student performance and analyze peer evaluation in educational settings. These examples, while diverse, are pertinent to language teaching and learning. Insights from these applications can guide future research in language education and machine learning, potentially enhancing the effectiveness of language teaching methodologies.

IV. Comparative analysis

From the Fig 4we see a line graph that represents the improvement in marks (presumably a measure of student performance or proficiency) as a function of increased linguistic teaching efforts or input.

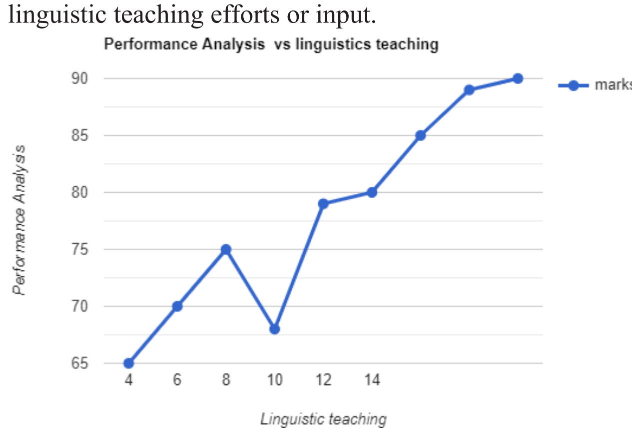

Fig. 4: Performance analysis using Recurrent Neural Network

The 'Linguistic teaching' axis, although not explicitly defined, could be indicative of time, intensity, or a composite index of teaching activities. The trend shown is generally positive, with a significant increase in marks correlating with increased linguistic teaching, despite a noticeable dip in the middle of the graph. This dip might represent a learning challenge or an adjustment period to new teaching strategies, which, once overcome, leads to improved performance. The comparative analysis in Table 1 and Fig 5 of different machine learning algorithms reveals that Random Forest, Recurrent Neural Networks (RNN), and Support Vector Machines (SVM) each have distinct performance metrics and complexities when applied to linguistic tasks.

Random Forest leads in terms of accuracy with a score of 90, followed by RNN with 88, and SVM with 82. In precision, Random Forest again tops the chart with 88, RNN stands at 85, and SVM is at 80. For recall, RNN achieves the highest score at 90, indicating its strength in identifying relevant instances, with Random Forest close behind at 89, and SVM at 79. The F1-Score, which balances precision and recall, is highest for Random Forest at 88, RNN at 87, and SVM at 79.

Table 1: Comparative Analysis

Metric	RNN	SVM	Random Forest
Accuracy	88	82	90
Precision	85	80	88
Recall	90	79	89
F1-Score	87	79	88
Training Time (s)	180	120	200
Model Complexity	High	Medium	Medium

In terms of training time, RNN requires the most with 180 seconds, whereas SVM is the fastest at 120 seconds, and Random Forest takes 200 seconds, which could be attributed to its robustness and the complexity of constructing multiple trees. When considering model complexity, RNN is categorized as high, which may reflect its sophisticated recurrent structure that's capable of handling sequential data like language. Both SVM and Random Forest are rated as medium complexity, indicating they are relatively easier to implement, and tune compared to RNN, but still offer a considerable depth of learning capability.

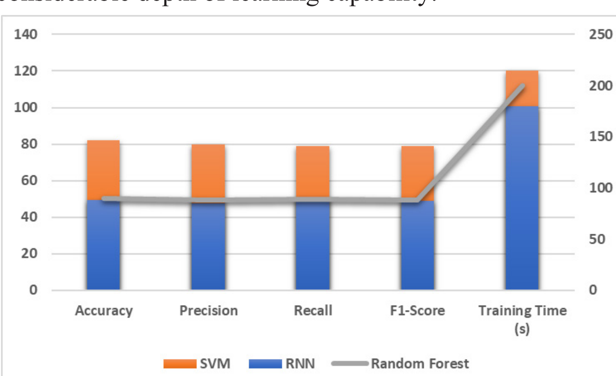

Fig. 5: Comparative Analysis

This analysis suggests that while Random Forest generally provides the highest performance metrics, it does so at the cost of longer training times and moderate complexity. RNN, despite its high complexity, is a strong contender in recall and F1-Score, making it suitable for tasks where understanding the context and sequence is crucial. SVM offers a good balance between speed and performance, particularly where simplicity and fast training are desired.

V. Discussion

The investigation into the effectiveness of various statistical methods for linguistics research, particularly in second language acquisition, has revealed a nuanced landscape. The use of multiple regression analyses and ANOVA techniques has been a staple in this field, providing insights but also inviting scrutiny over their relative merits. The introduction of linear mixed-effects models represents a modern approach that caters to the complexities of psycholinguistics. A notable trend is the critical stance on inferential statistics, which have been deemed less suitable for research in English language education, prompting calls for methodological innovation. Our analysis of artificial intelligence's role in English Language Teaching (ELT) has been synthesized in Figure 2, which encapsulates the transformative potential of AI and deep learning in this domain. The figure underscores the shift towards dynamic assessment methods influenced by AI, which could lead to significant improvements in language teaching efficacy.

The discussion further extends to the realm of Natural Language Processing (NLP), where Figure 3 portrays a standard pipeline of text processing tasks essential for computational language understanding. The sequence begins with spelling correction, proceeds through tokenization and part-of-speech tagging, and culminates in complex operations such as lemmatization and syntactic analysis. The incorporation of these techniques in linguistic teaching methods is poised to enrich the learning experience significantly. The proposed employment of Recurrent Neural Networks (RNNs) innovates upon traditional linguistic teaching methods. With their ability to process sequential data and update their internal state at each time step, RNNs are ideally positioned to advance language education. However, the complexity inherent in training RNNs and their handling of long sequences presents a challenge that must be navigated carefully.

In the comparative analysis, we have juxtaposed RNNs with other machine learning algorithms such as SVMs and Random Forests, as illustrated in Table 1 and Figure 5. Random Forest emerges as the frontrunner in accuracy and F1-Score, though at the expense of longer training times and greater complexity. RNNs, with their high recall rate, are particularly adept at tasks requiring an understanding of context and sequence, while SVMs offer a balanced option with faster training times.

This comparative analysis not only sheds light on the distinct advantages and trade-offs associated with each algorithm but also reflects the evolving nature of linguistic research methodologies. It underscores the potential of leveraging these diverse machine learning approaches to enhance linguistic teaching and assessment strategies, thereby enriching the learning experiences of students and facilitating their mastery of the English language.

VI. Conclusion

In conclusion, the research has illuminated the intricate relationship between statistical methods, machine learning, and the field of linguistics, especially in the context of second language acquisition. It has been demonstrated that while traditional statistical techniques like ANOVA and regression analyses have their place, the integration of advanced computational methods, particularly AI and NLP, presents a compelling advancement in linguistic education.

Our findings indicate that deep learning, specifically through the use of RNNs, holds significant promise for enhancing language instruction. The capacity of RNNs to process sequential data and learn from context aligns well with the complexities of language learning. Despite the challenges in training such models, the benefits in terms of personalized and dynamic educational experiences are substantial.

The comparative analysis of machine learning algorithms, as presented in our results, showcases the strengths of Random Forest in accuracy and precision, while also highlighting the agility of SVMs for tasks requiring quicker computational times. It is crucial to note, however, that the choice of algorithm should be tailored to the specific linguistic objective at hand, considering the trade-offs between performance metrics, training time, and model complexity.

In the broader scope, the integration of AI into the ELT method, as visualized in Figure 2, and the step-by-step breakdown of the NLP pipeline in Figure 3, underscore the transformational impact these technologies can have on the educational landscape. These advancements not only provide a more structured and indicative assessment of students' language abilities but also pave the way for dynamic and continuous improvement in teaching methodologies.

Overall, the intersection of machine learning and linguistic education offers a fertile ground for innovation, promising more effective and engaging language learning experiences. Future research should continue to refine these technologies and explore their application in diverse linguistic settings to fully harness their potential for language teaching and learning.

References

[1] J. Shi, "Deep Learning for College English Education Evaluation," *Mobile Information Systems*, vol. 2022, 2022, doi: 10.1155/2022/3558558.

[2] J. Xu, Y. Liu, J. Liu, and Z. Qu, "Effectiveness of English Online Learning Based on Deep Learning," *Comput Intell Neurosci*, vol. 2022, 2022, doi: 10.1155/2022/1310194.

[3] C. Song, "An Evaluation Method of English Teaching Ability Based on Deep Learning," *Security and Communication Networks*, vol. 2022, 2022, doi: 10.1155/2022/8339137.

[4] Y. Zhong and X. Yue, "On the correction of errors in English grammar by deep learning," *Journal of Intelligent Systems*, vol. 31, no. 1, 2022, doi: 10.1515/jisys-2022-0013.

[5] G. Singh, V. Garg, P. Tiwari, and R. Goel, "Emergence of IoT and Big Data: A Study in the Healthcare Industry," in *Applications of Deep Learning and Big IoT on Personalized Healthcare Services*, IGI Global, 2020, pp. 42–54.

[6] N. Kaur and G. Singh, "Role of Brokers/Institutional Investors to Induce Investment in Indian Stock Markets," *Int J Adv Res (Indore)*, vol. 3, no. 7, pp. 1545–1562, 2015.

[7] S. Sahai and G. Singh, "Partial Confirmatory Factor Analysis for E-Service Delivery Outcomes Using E-Tools Provided by the Government," in *Advances in Computing and Data Sciences: Second International Conference, ICACDS 2018, Dehradun, India, April 20-21, 2018, Revised Selected Papers, Part II 2*, 2018, pp. 463–470.

[8] C. Krishnan, K. Thapliyal, and G. Singh, *Global Higher Education and the COVID-19 Pandemic: Perspectives, Challenges, and New Opportunities*. CRC Press, 2023.

[9] G. Stobbe, *Just Enough ENGLISH GRAMMAR*, vol. 4, no. 1. 2013.

[10] H. Ding, Y. Chen, and L. Wang, "College English Online Teaching Model Based on Deep Learning," *Security and Communication Networks*, vol. 2021, 2021, doi: 10.1155/2021/8919320.

[11] X. Tao, "Ways to Promote Students' Deep Learning in English Teaching Based on Computer Technology," in *Journal of Physics: Conference Series*, 2021. doi: 10.1088/1742-6596/1881/2/022042.

[12] J. Huang, "Personalized College English Learning Based on Deep Learning under the Background of Big Data," *Comput Intell Neurosci*, vol. 2022, 2022, doi: 10.1155/2022/7361746.

[13] Y. Cho and J. Kim, "Production of mobile english language teaching application based on text interface using deep learning," *Electronics (Switzerland)*, vol. 10, no. 15, 2021, doi: 10.3390/electronics10151809.