STEK: A Supporting Tool to Enhance the English Knowledge

S.I.C. Dias
Department of Computer Science andSoftware Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

D. Vithanage
Department of Computer Science and Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

L.A.P.A. Weerasinghe
Department of Computer Science andSoftware Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

A.P.D. Ranaweera
Department of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

D.I. De Silva
Department of Computer Science and Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

ABSTRACT
English is not just the second most widely spoken language in the world. It is also the language most commonly used to communicate with native English speakers. Sri Lanka is ranked eighty-second on the English Competence Index and has a low level of English proficiency. People who live in rural places in Sri Lanka do not receive the proper education or materials to learn English. There is a near-ubiquity of web-connected devices among language learners in the twenty-first century, and the significant success of mass-market web-based language learning software demonstrates a great need for such resources. After establishing the global demand for web-based digital English learning tools, this article addresses the platforms and programming languages that English educators might employ to construct new online learning activities, particularly for rural locations. STEK, a supporting tool to enhance English knowledge, is the recommended solution. It is built on a web application by employing a combination of machine learning and deep learning approaches. It consists off our components: sample essay generator, pronunciation checker, citation checker, and tense modifier.

Keywords
Machine learning, deep learning, essay generator, pronunciation checker, citation checker, tense modifier

1. INTRODUCTION
Learning the English language is an important task in each and every one life. Because of its pedagogical benefits, learning English as a second language or a foreign language online has been increasingly popular in recent years. Technology integration, according to teachers, policymakers, and education experts, enhances both learners' pedagogical practices and users' learning progress [1]. Technology provides an assortment of interesting teaching tools that learners may adapt and use, making English language learning exciting while allowing people to collaborate andget more engaged in language learning. As a result, there's no denying that technological integration in these settings has become a new, upward trend in language learning.

Skills related the English language is a teacher-oriented task rather than a student-oriented one. First, the topic of writing is specified, and the suggested approach enables a user to generate auto-essays. Natural language processing or artificial intelligence systems can provide users with suitable feedback using employment with the help of the Eleuther AI model. On other hand, Eleuther AI model that already been used in previous works without a high accuracy. The problem is that the output that is generated by the Eleuther AI is not accurate or reliable all the time. The proposed system, STEK, will address the boundaries and the limitations of the Eleuther AI model and provide an output with more reliability and accuracy.

Pronunciation plays a crucial part in language mastery. As one of the most commonly used languages for international communication, learning English with correct pronunciation would make it easy for the conversation partner to understand the message of the non-native speaker. English, as a single language with flawed pronunciation, is frequently mispronounced, particularly by non-native English speakers These English speakers must be shown their mistakes in pronouncing words. The major task entails the capacity to classify or assign a voice in a different cluster. The classification of voice is an important problem in machine-learning systems for applications, such as voice recognition, sound filtering, class summarization, subject identification, and language identification systems. The development of automatic pronouncing systems is essential to tackle pronunciation problems.

The development and expansion of a learner's vocabulary is an important aspect of becoming proficient in any language. This is a portion of language education and grammar instruction that teaches the structure of sentences and the various categories of words from which to choose. Grammar tests are a frequent strategy in language instruction for practicing basic rules of sentence construction and expanding a learner's vocabulary. In many circumstances, a teacher's time is limited, and creating examinations may interfere with other demanding and intricate components of the teacher's everyday work. Grammar and sentence formation are systematic tasks that follow a set of rules, which a computer is generally adept at. An application that can automatically generate grammatical sentences will make the learning of English that much easier for a learner.

Within the academic community, citation parsing is an essential component of search engines, and it is also essential for the protection of intellectual property. Regardless of how they are described, a list of research sources can cause students to suffer alternating feelings of perplexity and frustration. According to a report by Project Information Literacy [4], 41% of polled students reported having difficulty with citations. Personal observation demonstrates that students make regular use of web-based citation tools. Based on reference questions asked and student work in library teaching exercises, it appears that many students rely on the citation tool accessible. While these citations comply to the standard APA, MLA, Harvard, and IEEE standards, they may be missing material, contain incorrect information, or contain formatting problems.

By considering these factors,this paperproposesa web application called STEKtohelp people who are unsatisfied with their fluency in the English language to improve themselves.

2. LITERITURE REVIEW

In past years, several systems have been proposed based different machine learning techniques to learn and improve the English skills of native and non-native English speakers.

Xiaocheng Feng, Ming Liu, Jiahao Liu, Bing Qin, Yibo Sun, and Ting Liu conducted research on "Topic-to-Essay Generation with Neural Networks" in [1][2]. They used neural networks to carry out the process in this study. They have created a new network called the "multi-topic-aware long short-term memory (MTA-LSTM)" network because essay generation is difficult and requires a lot of resources. They used a multi-topic coverage vector in this network model, which will learn the weight of the topic automatically and update it sequentially. Finally, the vector is fed into a generator, which produces the essay as a result. Furthermore, Pengcheng Yang, Lei Li, Fuli Luo, Tianya Liu, and Xu Sun's research "Enhancing Topic-to-Essay Generation with External Commonsense Knowledge" is an expanded version of the above "Topic-to-Essay Generation with Neural Networks." Since their previous model lacked common sense, this research was conducted to inject the commonsense component into the generator, resulting in more novel and diverse essays. They used dynamic memory to achieve this task. They also mentioned that adversarial training and more comprehensive evaluation metrics were used to produce a more novel and diverse output.

Craig W. Stanfill research [3][4] was done to solve the inductive inference problem. In this research paper, they used a memory-based reasoning technique. However, it is mentioned in the paper that when the training set is varied and distracting information is added, the output desired is significantly incorrect. Moreover, "Detection of Typical Pronunciation Errors in Non-native English Speech using Convolutional Recurrent Neural Networks" research was conducted by Aleksandr Diment, Semi Fagerlund, Adrian Benfield, and Tuomas Virtanen. In this process, they trained the algorithm with pronunciation errors made by non-native speakers for a selected set of words as well as a selected set of words without errors. Then the algorithm checked the new words fed in and determined the score. The major drawback of this algorithm was that it works only for a selected set of words.

In [5][6], Banging ping conducted research on "English Semantic Feature Processing and Sentence Structure Analysis Based on Hierarchical Network of Concepts. In this study, they have elaborated on the theory of hierarchical networks of concepts for problem-solving using approaches such as function, translator, effect, and relation. They have also addressed the fundamental structure of English semantics and compared the original text to online machine translation tools. In addition, Takuo Matsuzaki and Jun-ichi Tsujii conducted a study titled "Comparative Parser Performance Analysis Across Grammar Frameworks through Automatic Tree Conversion Employing Synchronous Grammars." In this study, a methodology for the comparative performance analysis of parsers designed for distinct grammar frameworks has been explored. Initially, parsing results are converted to a shallow CFG analysis utilizing automatic tree conversion based on synchronous grammar. Various data sets affect performance, as seen by the experiment results.

Niall Martin Ryan conducted a study titled "Citation Data-set for Machine Learning Citation Styles and Entity Extraction from Citation Strings" in [7][8]. They have explored entity extraction from citation strings at length in their study. This method has many problems and limits, as the metadata contains a great deal of noise and irrelevant data and styles. This technique can also greatly reduce precision. In addition, Sehrish Iqbal, Saeed-UI Hassan, Naif Radi Aljihani, Salem Alelyani, Raheel Nawaz, and Lutz Bornmann worked on "A decade of in-text citation analysis based on natural language processing and machine learning techniques: an overview of empirical investigations." As the title suggests, they developed a solution for in-text citation analysis based on natural language processing and machine learning approaches.

3. METHODOLOGY

3.1 System description

Figure 1 shows a block diagram of the proposed system. The application's primary goal is to enable support to learn English as a tool. The web application will aid in the learning of English and assist students who are struggling to speak English, write essays, locate citations, and check grammar.

Different types of algorithms and models were used to examine the preprocess the collected data. After the Machine Learning (ML) model was trained, the binary classifier results were categorized. Python was chosen as the programming language, along with the following libraries: Space for managing array data of tension, Deep Learning for ASR method, Google Collab to manage the ML framework, and Eleuther AI Model to Implementation Using Mesh TensorFlow for distributed computing.

The proposed system consists of four main components:

3.1.1 The sample essay draft generator

This component represents the implementation of users' accurate drafts of essays and sentences that they can use to improve their own writing and evaluate the fluency and lexical complexity of the user-written essays. Moreover, it generates automatic feedback for the essays that are written by the user to help them amend or overcome the mistakes they make.

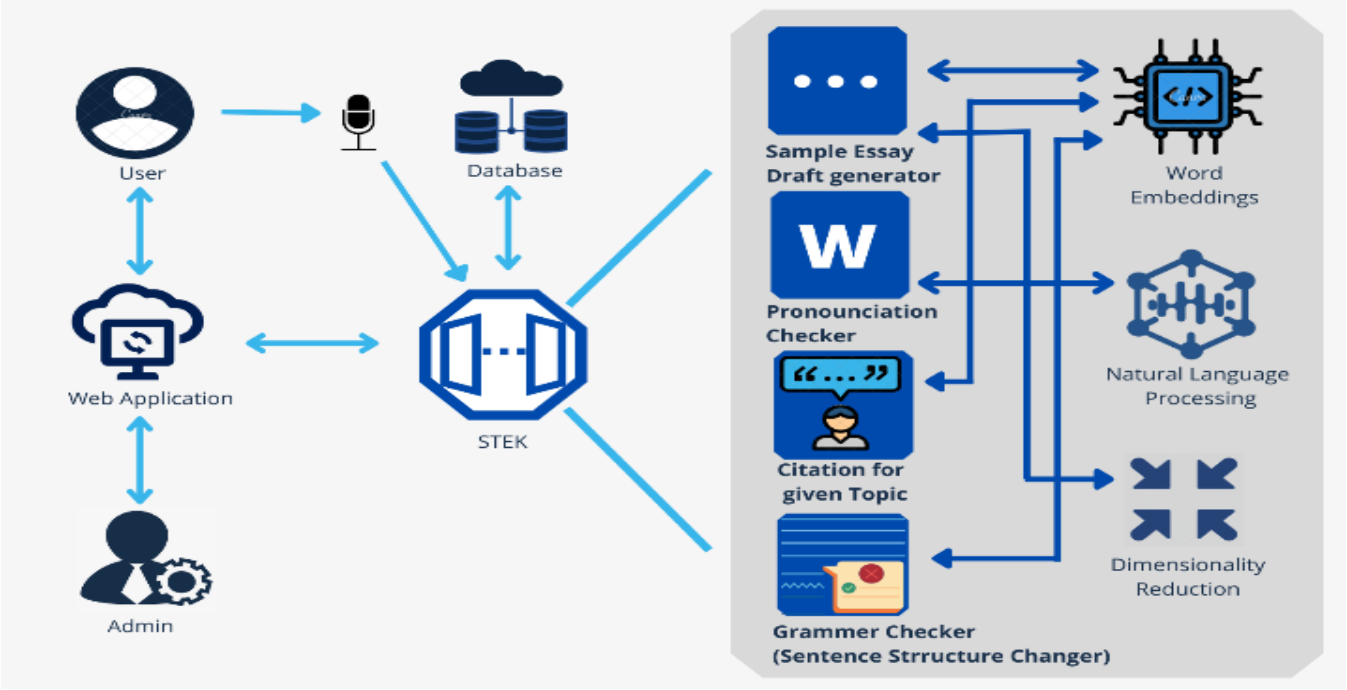

Fig. 1:Overall system architecture

The focus is to auto-generate an essay with a given topic or sentence. The proposed system analyzed the drawbacks of the Eleuther AI Model to overcome the current issue. Eleuther AI Model was used in the machine learning development process. Finally, it writes the sample essay draft. To achieve the desired functionality, several techniques such as grammar correction, paraphrasing, and removing unnecessary characters were used. With the help of the scoring model, the system will be able to compare the score of the generated draft with the draft generated by the authors. This will help to ensure the accuracy and the reliability of the generated draft.

3.1.1.1 Initialization and text preprocessing

Figure 2 shows the system architecture of the essay generator.

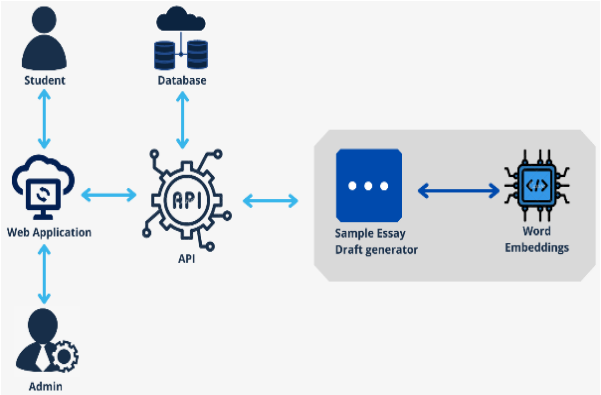

Fig. 2: System architecture of essay generator

This component focuses on essay creation. It takes a topic word as input and generates a well-structured article on the topic. Following the concept of text planning and developing a framework for essay generation is created utilizing the Eleuther AI Model. The framework includes three components: topic comprehension, sentence extraction, and sentence reordering. Multiple statistical methods were utilized and empirically contrasted in terms of qualitative or quantitative analysis for each component. The user inputs a prompt on the front end and then clicks "Generate." Amazon API Gateway receives a POST HTTP request from the web application. API Gateway invokes the database, which in turn contacts the SageMaker endpoint that hosts Eleuther AI. Eleuther AI creates text in response to the prompt. The created text returns to the web application for display.

3.1.1.2 Paraphrasing

According to several responses that was received from the Eluther AI model have encountered several problems with the text. One of the problems is the structure of the sentences of the text. The sentence structures that are provided from the Eluther AI model is somewhat unorganized. To address this small issue with the Eleuther AI were trained transformer model that will reorganize the structure of the sentences into a suitable and a more favourable way.

3.1.1.3 Grammar correction

One of the other problems that were encountered is the grammatical correctness of the sentences. After generating several essays drafts by using the Eleuther AI model it is clear to us that the grammar of the sentences that are generated are not grammatically correct all the time. A pre-trained model that will correct the most crucial grammar and spelling mistakes that occur with the Eleuther AI model is used to ensure the highest standards and the accuracy of the essays. By using this model output that will be provided to the user will be more reliable in terms of grammar and spelling mistakes.

3.1.1.4 Sentence to sentence connection

The other major problem that wasfound when using the Eleuther AI model is the connection between sentences. Most of the text generated by the Eluther AI model is not accurate most of the time. The connection between the two sentences or the whole essay is kind of meaningless sometimes. In order to tackle this problem, a transformer model that will allow STEK to determine the connection between the sentences and for the whole essaywas trained. By doing this the accuracy and the reliability of essay generated by STEK will be much higher.

3.1.1.5 Essay score

Even after applying all the above-mentioned solutions one can still say that the text generated by the Elutehr AI model is better than the STEK. To have a standard or a quality assurance for the essays, a model that will score a given essay according to the standards of IELST was trained. The IELTS standards were used to make the score model, as it is well-known and since it gives more reliability.

3.1.2 Pronunciation checker

This component uses voice-based processing in machine learning and a trained deep learning model for speech recognition.Figure 3 shows the system architecture of the pronunciation checker.

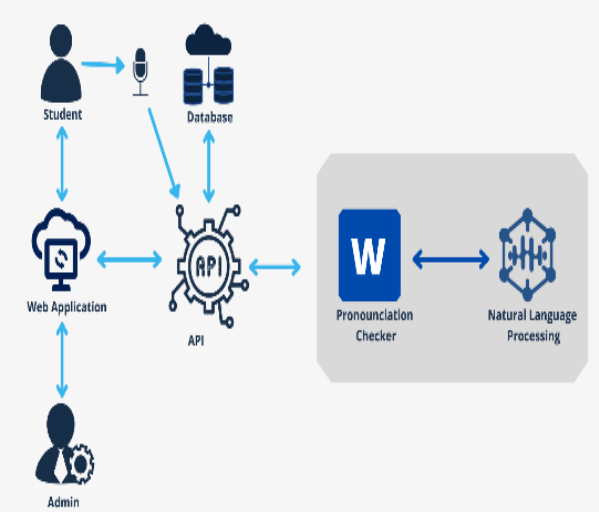

Fig. 3: System architecture of pronunciation checker

The automated Pronunciation Error Detection and Evaluation module performs speech analysis to identify the pronunciation level and errors. A personalized vowel chart is created to identify mispronounced words and mispronounced vowels within the words and provide the goodness of pronunciation value. To create the personalized vowel chart and for the speech analysis part, the system used Audio signal processing methods in a python environment.

To make it easy to understand the flow of the system, the whole process of analyzing human speech and giving feedback can divide into several components as follows:

- Audio preprocessing
- Word segmentation
- Audio Segmentation
- Measure formant frequencies word segment by segment
- Creating a data blog for segmentation
- Find pronounced vowels count within the word segment using a vowel chart
- Search for the vowels that need to be pronounced within the word segment
- Interconnect the vowel pattern and feedback

3.1.2.1 Audio processing

Before performing the analysis part, the data should be preprocessed. In the beginning, the recorded data is in mp3 format. So as the first step, the recorded audio is converted into WAV format with a 16000-sample rate and mono stereotype. The audio converting part is performed using the "pydub" library of the Python pip install package.

3.1.2.2 Word segmentation

The system needs to understand the words that the speaker says. Can achieve that goal by using an ASR model. Present days automatic speech recognition models developed to give a highly accurate word transcript. As the ASR model, the VOSK model [11] was used for this research.

3.1.2.3 Audio segmentation

In the section, the audio is segmented into words with its start and end timestamps. In this part, the audio file will break into segments using the start and end timestamps in the above section.

Once this process is done, there will be a separate word transcript and with related audio segments. In the next section, the system will use created segments by segments to analyze and generate formant frequencies for the prediction.

3.1.2.4 Measure formant frequencies word segment by segment

Once audio segmentation is done, this step will measure format frequency using the formant position formula. l maximum produced by the acoustic resonance of the human voice tract. It is commonly believed that the first two formants (f1 and f2) relate to the open/close and front/back dimensions of vowels, which are crucial in defining their quality. To extract the f1 & f2 from audio, the 1st step must be the point processing throughout the audio.

A point process object represents a point process, which is a sequence of points it in time, defined on a domain (t min, t max). The index i runs from 1 to the number of points. The points are sorted by time, i.e., t i + 1 >t i. The point process helps to identify the voiced intervals (or glottal pulses) of the given audio file. To detect f1 & f2 formants within the voiced intervals, the filtered voiced intervals from the point process will be re-sampled to a sampling frequency of twice the value of the formant ceiling. Then formants frequencies are measured using LPC coefficients with the algorithm using the Burg method. Then it will give f1 & f2 formants frequencies within the voiced intervals of the given audio file.Figure5 shows the overall system flow of the pronunciation checker.

3.1.2.5 Find pronounced vowels count within the word segment using a vowel chart and feedback

Vowels and consonants are the two main classes of speech sounds. Creating a personalized vowel chart is shown in Figure 4. System Generated vowel chart can help to avoid accent, gender, and age. f1 of monophthongal vowel sound against f2 were used to create the vowel chart as shown in Figure 4.

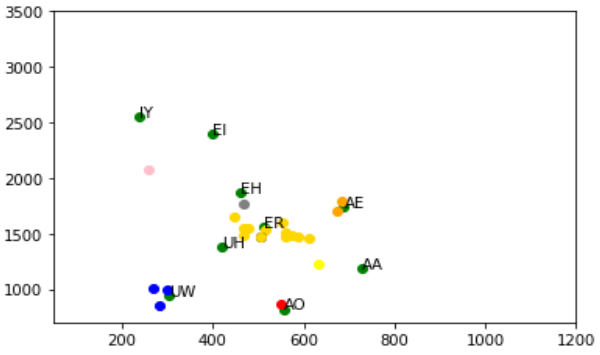

Fig. 4: Vowel categorization

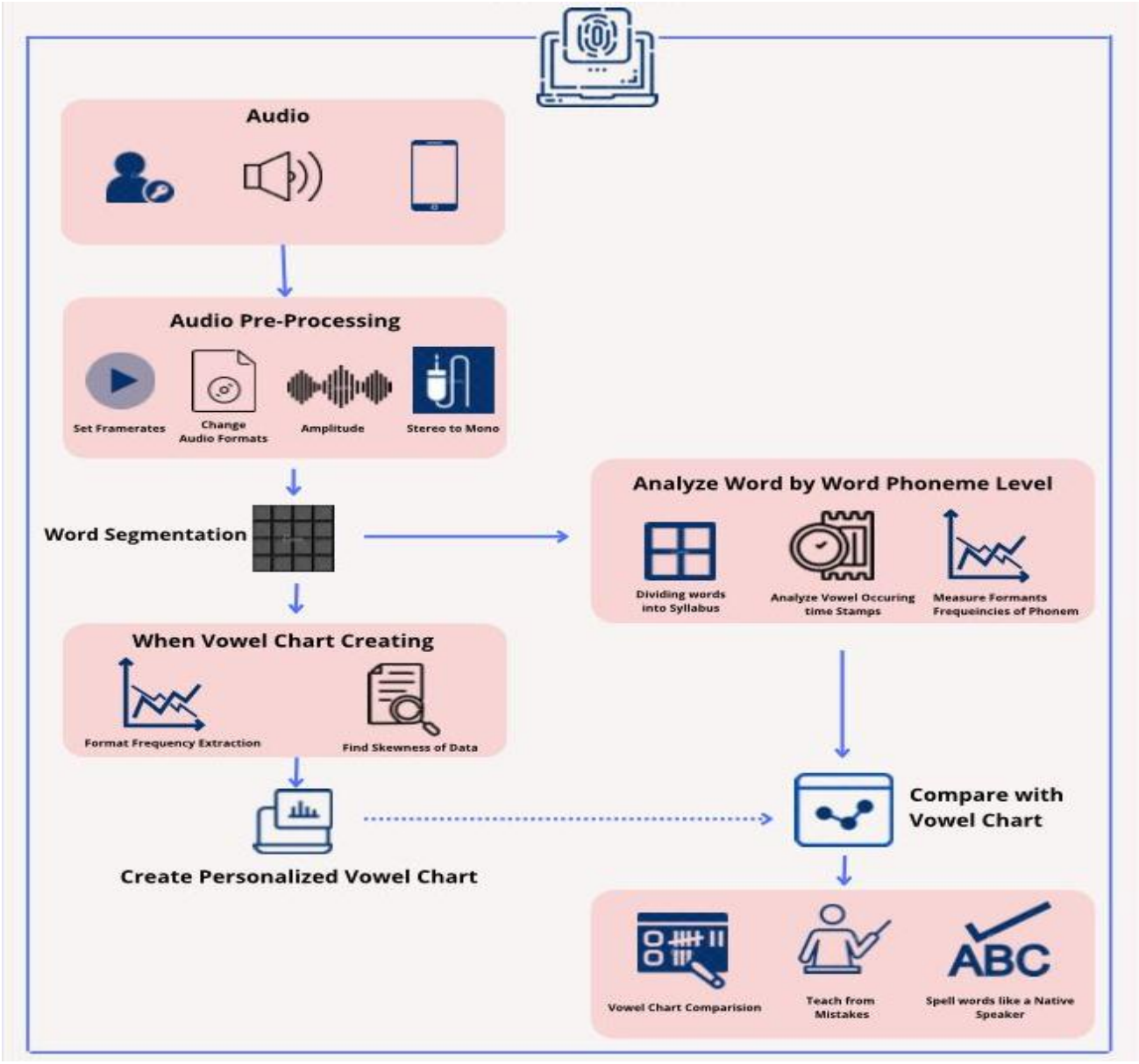

Fig.5: Overall system flowof the pronunciation checker

3.1.3 A grammar changer

To provide a robust and usable online grammar checking facility this component uses a Transformer model with a number of pre-processing techniques such as the byte pair encoding algorithm, tokenization, and spellchecker. As the significant target of preprocessing is to extract the first and rectified versions of each section in the information, as well as the alterations that change the previous into the last option, it is believed that this is an effective method for developing an online grammar checker system.Figure 6 shows the system architecture of the grammar checker.

3.1.3.1 Transformer model

The only connection between the encoder and decoder in the encoder-decoder architecture is a fixed-length semantic vector C. In other words, the encoder will compress the entireinformation sequence into a vector of defined length. There are two negative aspects to this. One is that the semantic vector cannot adequately convey the sequence's information. The second is that the data conveyed by the initial information will

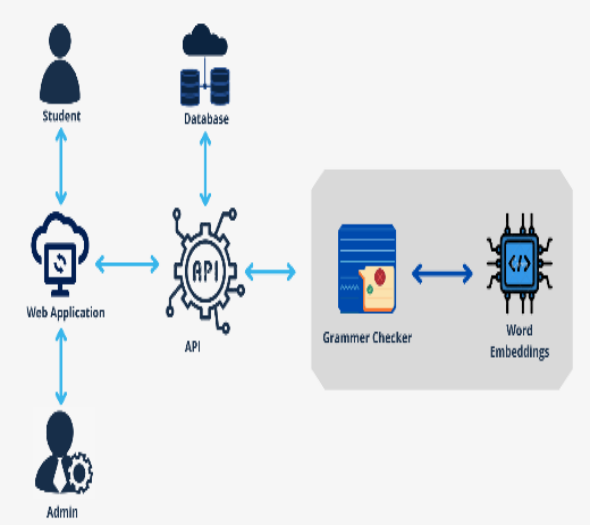

Fig.6: System architecture of grammar checker

be diluted or covered by ensuing sources of info. This makes it difficult to accumulate adequate data about the information sequence at the start of translating; subsequently, the decoding accuracy will be diminished. The Transformer model's central focus mechanism was developed to address the problem.

3.1.3.2 Rule-Based systems

A working memory or knowledge base, a rule base, an inference engine, and an execution engine are all components of a rule-based system. The facts and conditions are described in the knowledge base or working memory. The relationship between the premises and the conclusions is described by the rule basis. An agenda lists all relevant and applicable rules, and the inference engine has a pattern matcher that applies rules based on the fact. Given the input, the execution engine determines which rules to apply. Forward chaining and backward chaining are the two basic techniques used to infer a rule-based system. Forward chaining is a top-down strategy in which the facts are used to draw conclusions or trigger an action based on the fact. Backward chaining takes a bottom-up method, starting with a hypothesis or set of goals and searching the rule space for rules that could be used to prove the observed hypothesis, while also defining new sub-goals to prove as the process progresses.

3.1.3.3 SpaCytokenization

The BEA-2019 shared task text was tokenized using spaCy, an open-source Python package. encore web sm is the name of the tokenization core model, which is an English perform multiple tasks A convolutional neural network prepared on OntoNotes. The genre of text intended for processing is web-based including blogs, news, remarks, and text, which are the main fields where the internet-based syntax checker framework needs to be utilized. The model utilized by the spaCy library has a tokenization accuracy score of 99.76, which is completely reliable for an online grammar checker system. The spaCy Python library offers excellent documentation, excellent interoperability with Pycharm and neural networks, and an industrial focus. Figure 7 depicts the flowchart of the grammar checker.

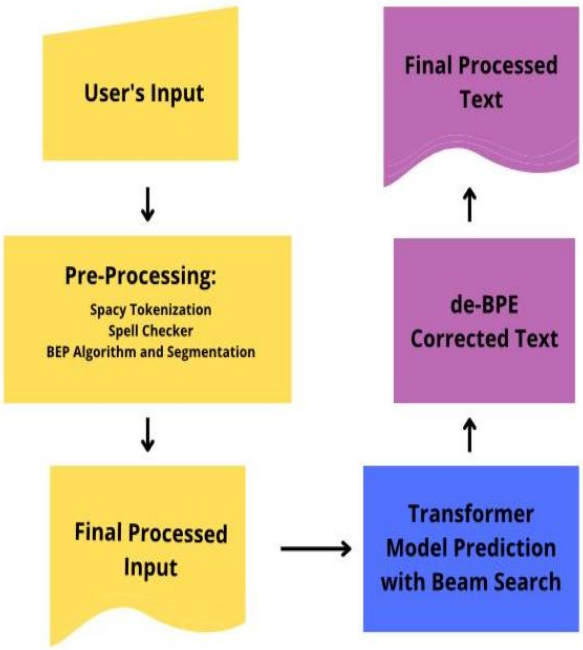

Fig. 7: SpaCy tokenization

3.1.3.4 Spell and Tense checker

Numerous modern GEC systems have a spellchecker that not only corrects "TENSE"-type grammatical errors but also deals with unknown words caused by typos or misspellings. Restrained the popular open-source spellchecker hunspell to circumvent its constraint of operating solely at the word level without taking context into account.

3.1.4 Citation generator

With the help of the citation generator, the user will be able to generate the citation automatically while inputting their idea and the topic. Once the topic is identified by the tool, the STEK will find out similar projects related to the given citation topic. Finally, a report will be able to be generated.Figure 8 shows the system architecture of the citation checker.

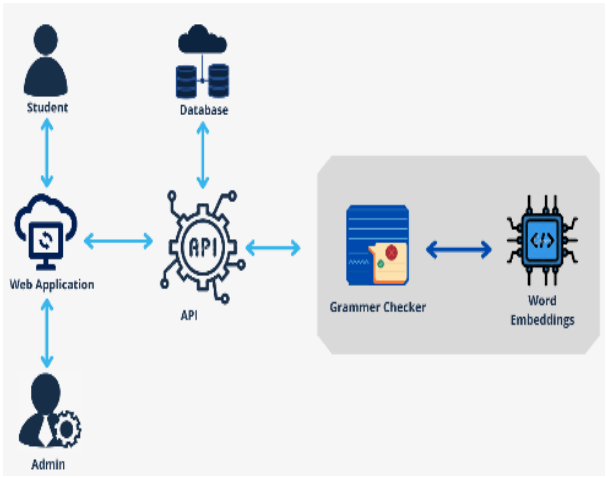

Fig. 8: System architecture of citation checker

3.1.4.1 Web scraping

A web scraping tool is a technology solution for quickly, efficiently, and automatically extracting data from websites and providing it in a more structured and user-friendly way" It is possible to extract unstructured data from the World Wide Web using classic copy-and-paste since some websites restrict access to automated machines. However, this strategy is extremely inefficient for larger projects. Occasionally, websites or online services provide APIs to retrieve or interact with data. However, it is not uncommon for APIs to be unavailable or for available solutions to fall short of user requirements. Using APIs also demands some programming competence. If APIs are unavailable or insufficient for the task, the web scraping approach might be utilized.

3.1.4.2 Feature extraction

Text feature extraction plays a significant function in text classification and directly affects classification accuracy. Based on the vector space model (VSM), a text is represented as a point in N-dimensional space. Each dimension of the dot's data represents a (digitized) text characteristic. And text features typically employ a keyword set. It means that, based on a group of predetermined keywords, the weights of the words in the text is computed using specific methods and then construct the text's feature vector, a digital vector. Existing methods for text feature extraction include filtration, fusion, mapping, and clustering, which are briefly described below.

Find further extraction of 64 features based on the methodology of four cutting-edge procedures. Using the Extra-Trees classifier,29 best features were chosen. Then the

Random Forest and Support Vector Machine classifiers were applied to all selected techniques. Using the Random Forest classifier, the supervised model outperforms the industry standard by 11.25%, with an area under the Precision-Recall curve of 89%. Finally, the deep-learning model of the proposed system, which uses all 64 features to distinguish between significant and insignificant citations, reports an accuracy of 92.57%.

4. CONCLUSION

The most frequently spoken language in the world now is English. It has become the international language of business, communication, and even entertainment. In these examinations, non-native English speakers from all over the world struggle, particularly Sri Lankans. As a result, individuals encounter substantial obstacles in achieving their career and academic objectives, and they miss out on a number of crucial opportunities. There are a number of limitations to be aware of despite the availability of numerous English-learning software options. As part of this research, a web-based intelligent STEK application was created to address the deficiencies of currently available programs. Within the scope of this study, a web-based application was created for intelligent e-learning that addresses the deficiencies of existing programs.

This study focuses on deploying a software solution to increase non-native speakers' English language competency and enhance their English learning experience. Non-native English language learners can generate any type of essay draft, which will aid to increase their writing skills. Non-native English speakers' speakinglevels will increase as a result of automated speaking and pronunciation error detection evaluations, which will also teach them how to pronounce words correctly. Cutting-edge technologies such as Audio Signal Processing and Automatic Speech Recognition are used to create the suggested system. In comparison to traditional methodologies and existing solutions, STEK has been able to give more accurate and dependable findings. Aid to improve the grammar accuracy, which can lead to a proper grammatical sentence, and finally, students will not cheat on their topic at the same time. They may find similar research papers to get over their research.

5. REFERENCES

[1] B. Qin, D. Tang, X. Geng, D. Ning, J. Liu, and T. Liu, "A Planning based Framework for Essay Generation," arXiv:1512.05919 [cs], Jan. 2016, Accessed: Jun. 15, 2022. [Online]. Available: https://arxiv.org/abs/1512.05919.

[2] P. Yang, L. Li, F. Luo, T. Liu, and X. Sun, "Enhancing Topic-to-Essay Generation with External Commonsense Knowledge," in *Proc. 57th Annual Meeting of the Association for Computational Linguistics*, July 2019, pp. 2002 – 2012.

[3] C. Stanfill, "Memory-Based Reasoning Applied to English Pronunciation," in *Proc. sixth National conference on Artificial intelligence*, vol. 2 , July 1987, pp. 577–581.

[4] A. Diment, E. Fagerlund, A. Benfield, and T. Virtanen, "Detection of Typical Pronunciation Errors in Non-native English Speech Using Convolutional Recurrent Neural Networks,*"International Joint Conference on Neural Networks (IJCNN*), July 2019, pp. 1-8.

[5] T. Matsuzaki and J. Ichi Tsujii, "Comparative Parser Performance Analysis across Grammar Frameworks through Automatic Tree Conversion using Synchronous Grammars*,"* in*Proc. 22nd International Conference on Computational Linguistics*, vol. 1, Aug. 2008, pp: 545–552.

[6] L. Kallmeyer, "Extraposed relative clauses in Role and Reference Grammar. An analysis using Tree Wrapping Grammars," *Journal of Language Modelling*, vol. 9, no. 2, Feb. 2022.

[7] S. Iqbal, S.-U. Hassan, N. R. Aljohani, S. Alelyani, R. Nawaz, and L. Bornmann, "A Decade of In-text Citation Analysis based on Natural Language Processing and Machine Learning Techniques: An overview of empirical studies," arXiv:2008.13020 [cs], Aug. 2020, Accessed: Jun. 15, 2022. [Online]. Available: https://arxiv.org/abs/2008.13020.

[8] B. Bruce and J. K. Peyton, "A New Writing Environment and an Old Culture: A Situated Evaluation of Computer Networking to Teach Writing*,"Interactive Learning Environments*, vol. 1, no. 3, pp. 171–191, Sep. 1990.

[9] N. Humble, "Developing A Web Application For Auto-Generating Grammar Tests*,"12th Annual International Conference on Education and New Learning Technologies*, Jul. 2020, pp. 7196-7201.

[10] L. Homol, "Web-based Citation Management Tools: Comparing the Accuracy of Their Electronic Journal Citations," *The Journal of Academic Librarianship*, vol. 40, no. 6, pp. 552–557, Nov. 2014.

[11] E. Dikici, M. Semerci, M. Saraclar, and E. Alpaydin, "Classification and Ranking Approaches to Discriminative Language Modeling for ASR," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 2, pp. 291–300, Feb. 2013.