KnE Social Sciences
ICESRE
6th International Conference on Education and Social Science Research
Volume 2024

Research Article

Development of Mobile Application through the Concept of Artificial Intelligence to Enhance Pronunciation Skill in EFL

Sukma Nur Ardini[1*], Sunarya Sunarya[2], Khoiriya Latifah[3]

[1]English Education Study Program, Universitas PGRI Semarang, 50125, Indonesia
[2]Javanese Education Study Program, Universitas PGRI Semarang, 50125, Indonesia
[3]Informatics Study Program, Universitas PGRI Semarang, 50125, Indonesia

ORCID
Sukma Nur Ardini: https://orcid.org/0000-0001-9746-9005

Abstract.
The utilization of artificial intelligence (AI)-powered mobile applications has demonstrated potential in enhancing pronunciation skills for learners of English as a Foreign Language (EFL). In this study, Natural Language Processing (NLP) was employed for English student learning. Jonglish, an Android mobile application utilizing Machine Translation and Grammarly, served as the platform. Given the novelty of the field in Indonesia, the researchers aimed to investigate the integration of NLP into the creation of Jonglish. To address the research objective, which is to elucidate the development of a mobile application named *Jonglish* through the concept of AI to enhance pronunciation skills in EFL, the researchers utilized the Life Cycle Machine Learning, specifically the Cross Industry Standard Process for Data Mining (CRISP-DM). The results of testing the dataset revealed a 100% success rate in translating the data into Indonesian and English using TextBlob as a translator. Meanwhile, SpellingCheck achieved a 98% accuracy rate for spelling checks. With technological advancements, the collaboration of AI and mobile application will undoubtedly drive further innovation, enhancing convenience, efficiency, and engagement for users around the world. Following the progress report, the subsequent stage is the model testing and deployment phase. In this phase, developers and AI engineers operationalize the concepts and algorithms developed in earlier stages and bring them to life within mobile applications.

Keywords: artificial intelligent, mobile application, pronunciation, EFL

Corresponding Author: Sukma Nur Ardini; email: sukmanurardini@upgris.ac.id

Published 12 March 2024

Publishing services provided by Knowledge E

© Sukma Nur Ardini et al. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Selection and Peer-review under the responsibility of the ICESRE Conference Committee.

OPEN ACCESS

1. Introduction

Artificial Intelligence (AI) is an expanding technological domain that has the potential to change every aspect of our social interactions. In education, AI has begun to produce new teaching and learning solutions, which are now being tested in various contexts [1]. According to one of the computer scientists, AI is the science and engineering of making intelligent machines, especially intelligent computer programs [2]. AI necessitates advanced infrastructure and a thriving innovator ecosystem. In the rapidly evolving

landscape of technology, the fusion of AI and mobile applications has become a transformative force. Artificial intelligence has grown dramatically, and its application in data science, information science, and various platforms is unprecedented [3,4]. The marriage of AI and mobile applications has led to the creation of intelligent, responsive, and personalized experiences for users. [5] In the educational domain, mobile applications are used as tools to facilitate the teaching-learning process in the classroom, reinforce knowledge or skills outside of the classroom (e.g., wikis), replace teacher instruction (e.g., Massive Online Open Courses- MOOCs), or encourage learner motivation and task management skills (e.g., blogs). In addition, Machine learning is a branch of artificial intelligence that has almost become the pronoun of artificial intelligence in many cases. Machine learning systems are used to recognize objects in images, transcribe speech into text, match news items, posts, or products to the interests of users, and select relevant search results [6]. It is also as an existing field of approaches in computer science that perform better with practice [7,8]. In sum, AI encompasses the simulation of human intelligence processes by machines, particularly computer systems.

Mobile applications powered by artificial intelligence (AI) have shown promise in enhancing pronunciation skills in English as a Foreign Language (EFL) learners. These applications leverage automatic speech recognition (ASR) technology to provide learners with real-time feedback and practice opportunities for improving their pronunciation. Several studies have investigated the impact of AI-powered mobile applications on EFL learners' pronunciation skills. Positive result has shown by Zhou [9] in terms of pronunciation improvement. For example, a systematic review of artificial intelligence dialogue systems for EFL students found that chatbots and mobile applications enhanced EFL students' performance in interactional competence.

Additionally, research has shown that AI-powered mobile applications can enhance the capabilities of mobile technology for pronunciation learning. Smartphone software developers are increasingly integrating AI technologies to improve the effectiveness of pronunciation training [10,11,12]. However, it is important to note that the effectiveness of AI-powered mobile applications for pronunciation improvement may vary depending on individual learners' motivation, learning styles, and language goals. While these applications can be valuable tools for pronunciation practice, they should be used with comprehensive language learning strategies and instruction.

In conclusion, mobile applications powered by AI can potentially enhance pronunciation skills in EFL learners. These applications offer real-time feedback, personalized learning experiences, convenience, and accessibility. Research evidence supports the

effectiveness of AI-powered mobile applications for pronunciation improvement. However, it is essential to consider individual learners' needs and goals when incorporating these applications into language learning.

Some features are used to create mobile applications. (1) Exploratory Data Analysis (EDA) is a critical phase in the data analysis process that involves examining, summarizing, and visualizing data to gain a deeper understanding of its characteristics, patterns, and potential relationships. EDA plays a pivotal role in uncovering initial insights, identifying anomalies, and guiding the direction of subsequent analyses. It is the foundation for informed decision-making and hypothesis generation in various fields, including statistics, data science, and machine learning. EDA is a well-established statistical tradition that provides conceptual and computational tools for discovering patterns in order to foster hypothesis development and refinement [13]. (2) Text feature extraction is a process in natural language processing (NLP) that involves converting raw text data into a format that machine learning algorithms can understand and work with effectively. Text data is inherently unstructured, and machine learning models require numerical features as input. Feature extraction involves transforming text into a numerical representation that captures the important information and patterns present in the text. Text feature extraction that extracts text information is an extraction to represent a text message, and it is the basis of a large number of text processing [14]. (3) Voice feature extraction is converting raw audio signals, such as spoken language or sound recordings, into numerical features that can be used for various machine learning tasks, including speech recognition, emotion detection, speaker identification, and more. Similar to text feature extraction, voice feature extraction aims to transform unstructured audio data into a structured format that machine learning algorithms can effectively process. (4) NLP stands for Natural Language Processing, a field of AI that focuses on the interaction between humans and computers using natural language. NLP aims to enable computers to understand, interpret, and generate human language in a way that is both meaningful and useful.

According to There are several subfields and techniques within NLP [15], including: (1) Tokenization: this involves breaking down a text into individual words or tokens, making it easier for a computer to process. (2) Part-of-Speech Tagging: it involves categorizing words in a text into their respective grammatical parts of speech, such as nouns, verbs, adjectives, etc. (3) Named Entity Recognition (NER): NER identifies and classifies entities like names of people, places, organizations, dates, and more within the text. (4) Sentiment Analysis: This determines the sentiment or emotional tone of a piece of text, typically as positive, negative, or neutral. (5) Machine Translation: NLP is

used in machine translation systems like Google Translate to translate text from one language to another. (6) Text Generation: NLP models can generate human-like text, which can be used for chatbots, content creation, and more. (7) Topic Modeling: it is a technique to identify and extract topics or themes from a collection of text documents. (8) Text Classification: this involves assigning predefined categories or labels to text documents, like spam detection in emails or topic categorization in news articles. (9) Chatbots and Virtual Assistants: NLP is used to build conversational agents that can interact with users in a human-like manner. (10) Question Answering: NLP models can be trained to answer questions based on a given text or document, as seen in search engines and virtual assistants. These are just a few examples of the various applications and techniques within the field of NLP. NLP has a wide range of real-world applications, from improving customer support through chatbots to analyze social media sentiment and much more.

In this case, we use NLP for English student learning named *Jonglish* (*Jowo*-English). It refers to the study and exploration of Javanese English. It is a unique linguistic phenomenon that combines elements of the Javanese language and English. It is commonly spoken by Javanese people who have a limited proficiency in English. In the current study, *Jonglish* is an android mobile application using Machine Translation and Grammarly. Machine Translation is part of

NLP has a wide range of real-world applications, from improving customer support through chatbots to analyze social media sentiment and much more [16]. In this case, we use NLP for English student learning, *Jonglish* is an android mobile application using Machine Translation and Grammarly. Since the field is new in Indonesia, the researchers wanted to explore how NLP became part of *Jonglish* was created. Therefore, the objective of this research is to describe the development of mobile application named *Jonglish* through the concept of AI to enhance pronunciation skill in EFL.

2. Method

The current research was part of the second year of R&D research, which included (1) exploratory data analysis, (2) data preprocessing, (3) text feature extraction, (4) voice feature extraction, (5) model building. These were Life Cycle Machine Learning called CRISP-DM [17], seen in Figure 1.

CRISP-DM is a model of a process that forms the foundation of a data science procedure. There are six successive phases: (1) Business understanding; what is required to understand business? (2) Understanding data; what information do we have or need?

Is it spotless? (3) Data organization; how should the data be set up for modeling? (4) What modeling techniques ought to be used? (5) Evaluation; which model best satisfies the company's goals? (6) How do stakeholders obtain the results after deployment? CRISP-DM provides a systematic and structured approach to solving complex data problems, helping organizations derive actionable insights and value from their data. It emphasizes the importance of understanding the business context and the data.

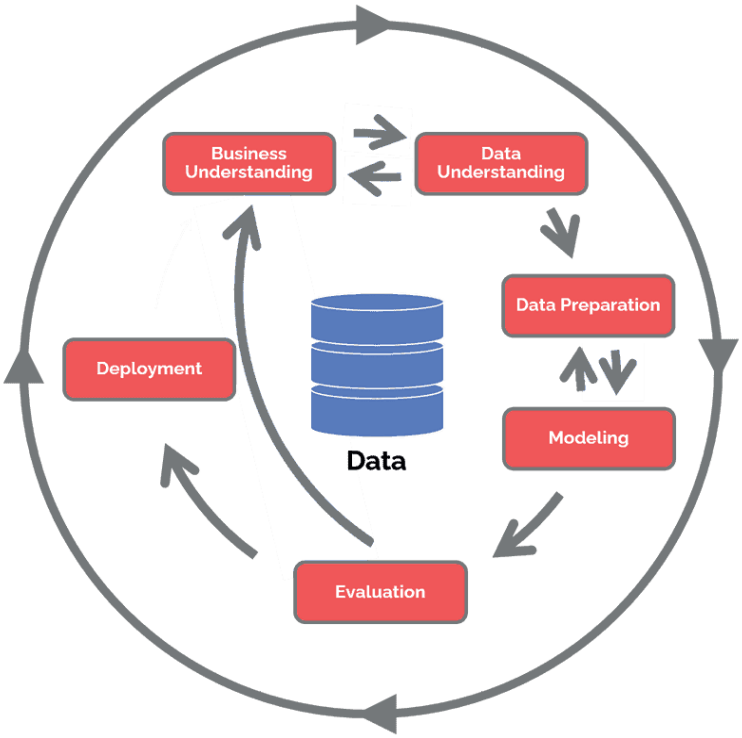

Figure 1: CRISP-DM Diagram.

Then, the researchers employed descriptive qualitative design, where the qualitative method investigates human phenomena [18]. It was to explain and describe the results found in this research. The object of the research was an android mobile application called *Jonglish,* which was created and designed by the researchers. Therefore, the application has not yet been widely known or documented in any sources.

3. Result and Discussion

The study aimed to develop a mobile application named Jonglish through AI to enhance pronunciation skills in EFL. *Jonglish* is an Android mobile-based application for English learning media for students whose mother tongue is Javanese. This application was built using the Python programming language with the TextBlob library. We can use

TextBlob to process text and perform stemming or lemmatization based on specific NLP needs [19,20]. It is a handy tool for text preprocessing and analysis in Python.

TextBlob is a Python library for processing textual data, and it includes a feature for stemming and lemmatization, which can be used to convert words to their base forms. Figure 2 explains the process of developing *Jonglish* by word recognition or searching for basic words using the TextBlob library with stemming and lemmatization to reduce words to their root or base form. The steps included stemming, lemmatization, and word sense. Stemming and lemmatization are two common techniques used in NLP for word normalization. While both techniques aim to reduce words to their base or root form, there are some important differences between them. Stemming is a rule-based process that removes prefixes and suffixes from words to obtain the base or root form. Stemming is faster than lemmatization because it applies simple rules to truncate words without considering the context or part of speech. Stemming may result in the creation of non-words or words that do not exist in the dictionary. Stemming is commonly used in applications like sentiment analysis, where word variations may not significantly affect the analysis.

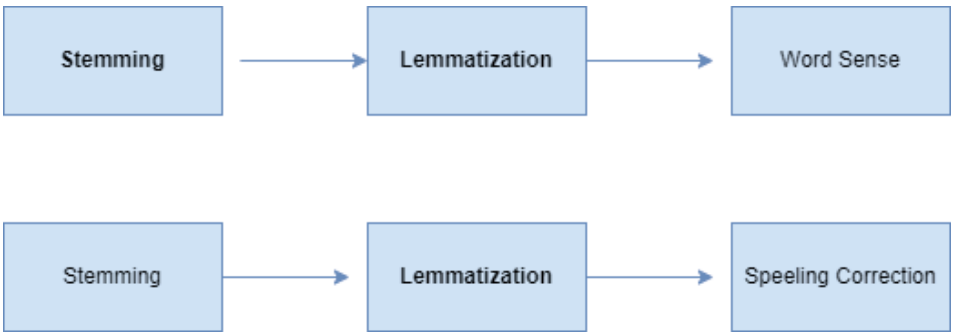

Figure 2: Techniques used in Natural Language Processing (NLP) in *Jonglish*.

On the other hand, lemmatization is a more sophisticated process that considers the context and part of speech of words to derive their base or dictionary form. Lemmatization takes more time than stemming because it finds meaningful word representation by using a vocabulary and performing morphological analysis. Lemmatization produces real dictionary words, which can be beneficial for tasks that require accurate word representations. Lemmatization is commonly used in applications like chatbots and question-answering systems, where the accuracy of word representation is crucial.

WordNet is a lexical database and resource that is often used in NLP. It is a valuable tool for researchers and developers working with text and language-related applications. WordNet is a large lexical database of English. It groups English words into sets of synonyms, called synsets, and provides a brief definition and semantic relationships for

each word. WordNet is used in various NLP tasks such as text classification, information retrieval, and sentiment analysis, and in this case, we use WordNet to make word sense disambiguation. It helps understand the meanings of words and their relationships, which is crucial in NLP applications.

One common use of WordNet in NLP is for Word Sense Disambiguation. This is determining which sense of a word is used in a particular context. WordNet provides different senses (or synsets) for many words, helping NLP algorithms disambiguate word meanings. WordNet is a crucial resource in the field of NLP for understanding word meanings and their relationships. It plays a significant role in various NLP tasks, contributing to developing more accurate and context-aware natural language processing applications.

TextBlob uses WordNet to provide features related to word sense disambiguation, semantic similarity, and more. Synsets help TextBlob determine the correct sense of a word in a given context. For example, the word “bank” = [can refer to a financial institution or the side of a river], “Believe” =[’capable of being believed’]. By identifying the synset associated with the word, TextBlob can improve its spellcheck’s accuracy by suggesting contextually appropriate corrections.

The second feature is to translate Indonesian into English and vice versa from English into Indonesian using the Text Blob Library. There is code that will translate the text to Bahasa Indonesia using TextBlob. After those phases were completed, the researchers displayed the features of *Jonglish,* as seen in Figure 3.

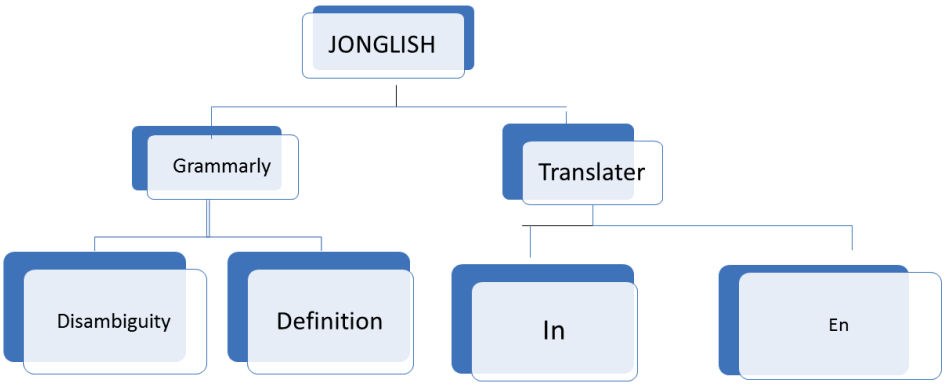

Figure 3: The Features of *Jonglish* Application.

NLP is a field of artificial intelligence that deals with the interaction between computers and human languages. It encompasses various tasks and applications, including grammar correction, machine translation, disambiguation, and definition extraction. As well as *Jonglish,* which focuses on Grammar Correction (e.g., Grammarly): Grammarly is a popular application that uses NLP to help users improve their writing by identifying and

correcting grammatical errors, spelling mistakes, punctuation issues and suggesting better word choices. NLP models in Grammarly analyze the context of the text and provide real-time feedback to enhance the overall readability and correctness of the content. Machine Translation (e.g., Google Translate): Machine translation is automatically translating text or speech from one language to another. NLP plays a central role in machine translation systems like Google Translate. NLP models analyze the structure and meaning of sentences in the source language and generate equivalent sentences in the target language, considering grammar, syntax, and context. Disambiguation: NLP techniques are used to resolve ambiguities in language. Natural languages often contain words or phrases with multiple meanings, and disambiguation aims to determine the correct interpretation of a word or phrase based on the context in which it appears. This is essential for accurate language understanding and meaningful communication. Definition Extraction: NLP can automatically extract definitions or meanings of words and phrases from large text corpora or dictionaries. NLP models analyze the context in which words are used to provide concise and accurate definitions. This can help build language resources and provide instant access to word meanings for users.

NLP achieves these tasks through various techniques and technologies, including Tokenization: Breaking down text into individual words or tokens. Part-of-Speech (POS) Tagging: Assigning grammatical categories (e.g., noun, verb) to each word in a sentence. Syntactic Parsing: Analyzing the grammatical structure of sentences to understand relationships between words. Named Entity Recognition (NER): Identifying and categorizing named entities such as names of people, places, and organizations. Word Embeddings: Representing words in a continuous vector space to capture semantic relationships. Machine Learning: Training models on large datasets to perform specific NLP tasks, such as language translation, sentiment analysis, and more. Overall, NLP technologies like Grammarly and machine translation services are powerful examples of how artificial intelligence can enhance language-related tasks, assisting with writing, communication across languages, and improving the clarity and accuracy of text. Disambiguation and definition extraction are fundamental components of NLP that contribute to the refinement and precision of language understanding and communication.

In addition, from the results of testing the data set, it was obtained that 100% of the data was successfully translated correctly using TextBlob as a translator into Indonesian and English. Meanwhile, for spelling check, SpellingCheck has an accuracy of 98 %. Therefore, the testing results of *Jonglish* dataset have demonstrated the effectiveness of using TextBlob as a translator for both Indonesian and English, achieving a remarkable 100% accuracy rate in accurately translating the provided data. Furthermore, the

SpellingCheck tool exhibited high accuracy, with a 98% success rate in identifying and correcting spelling errors within the dataset. These findings underscore the reliability and proficiency of TextBlob as a translation tool and the SpellingCheck tool for spelling correction, emphasizing their potential as valuable resources for language-related tasks and enhancing overall data quality and communication.

4. Conclusion

The evolution of mobile applications through the concept of artificial intelligence has revolutionized user experiences. From personalized recommendations to efficient automation, AI has reshaped how we interact with our devices. As technology advances, the collaboration between AI and mobile apps will undoubtedly lead to further innovation, enhancing convenience, efficiency, and engagement for users worldwide. AI has revolutionized user experiences in various ways, and its collaboration with mobile application is poised to bring even more innovation.

The next stage after the progress report is the model testing and deploying phase. In the model testing and deploying phase, developers and AI engineers take the concepts and algorithms developed in earlier stages and bring them to life within mobile applications. In conclusion, the model testing and deploying phase is critical in realizing the potential of AI-powered mobile applications. It involves rigorous testing, data integration, optimization, security considerations, user feedback, and ongoing maintenance to ensure that the AI enhances the user experience and provides value to users worldwide. As technology advances, this collaboration between AI and mobile applications will lead to exciting new developments and improved user experiences.

Acknowledgements

The authors would like to thank the Indonesian Ministry of Education, Culture, Research, and Technology (Kemendikbudristek) for the support through DRTPM funding.

References

[1] Pedro F, Subosa M, Rivas A, Valverde P. Artificial intelligence in education: Challenges and opportunities for sustainable development. France: United Nation Educational, Scientific, and Cultural Organization; 2019.

[2] Christopher M. Artificial intelligence definitions. Human-Centered Artificial Intelligence (HAI). Stanford University. September 2020. Available from: https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf

[3] Pikhart M. Intelligent information processing for language education: the use of artificial intelligence in language learning apps. Procedia Comput Sci. 2020;176:1412–9.

[4] Semenenko D, Nazarov Y. Exhibition environment for visitors with visual impairments. Questions of Expertise in Culture, Arts and Design [Internet]; 2019 Jun 6-7; Yekaterinburg, Russia. Dubai: KnE Social Sciences; 2020 [cited 2021 Jun 3]. 6 p. Available from: https://doi.org/10.18502/kss.v4i11.7558.

[5] Sánchez-Morales LN, Alor-Hernández G, Rosales-Morales VY, Cortes-Camarillo CA, Sánchez-Cervantes JL. Generating educational mobile applications using UIDPs identified by artificial intelligence techniques. Comput Stand Interfaces. 2020 Jun;70:103407.

[6] Liang H, Sun X, Sun Y, Gao Y. Text feature extraction based on deep learning: a review. EURASIP J Wirel Commun Netw. 2017;2017(1):211.

[7] Manikanthan SV, Padmapriya T, Hussain A, Thamizharasi E. Artificial intelligence techniques for enhancing smartphone application development on mobile computing. 2020;14(17):4. https://doi.org/10.3991/ijim.v14i17.16569.

[8] Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science. 2015 Jul;349(6245):255–60.

[9] Zhou A. Investigating the impact of online language exchanges on second language speaking and willingness to communicate of Chinese EFL learners: a mixed methods study. Front Psychol. 2023 May;14:1177922.

[10] Rogerson-Revell PM. Computer-Assisted Pronunciation Training (CAPT): Current Issues and Future Directions. RELC J. 2021;52(1):189–205.

[11] Ardini SN, Faradilla PE, Budiman TC. The Effectiveness of Discovery Listening to Improve Listening Competence in Higher Education. E-Structural (English Studies on Translation, Culture, Literature, and Linguistics). 2023 May 22;6(01):29-40. https://doi.org/10.33633/es.v6i01.7823.

[12] Zakiyyah F, Setyaji A, Ardini SN. The analysis of pronunciation application based on the concept of artificial intelligence. In UNCLLE (Undergraduate Conference on Language, Literature, and Culture) 2022 Jul 2 (Vol. 2, No. 01, pp. 559-569).

[13] Behrens JT. Principles and procedures of exploratory data analysis. Psychol Methods. 1997 Jun;2(2):131–60.

[14] Singh V, Kumar B, Patnaik T. Feature extraction techniques for handwritten text in various scripts: a survey [IJSCE]. International Journal of Soft Computing and Engineering. 2013 Mar;3(1):238–41.

[15] Pustejovsky J, Stubbs A. Natural language annotation for machine learning. Cambridge: O'relly; 2012.

[16] Daniel S. Unleashing the power of natural language processing (NLP) advancements. August 4, 2023. https://www.linkedin.com/pulse/unleashing-power-natural-language-processing-nlp-advancements-s

[17] Hotz N. What is CRISP DM? January 19, 2023. https://www.datascience-pm.com/crisp-dm-2/

[18] Creswell JW, Creswell JD. Research design: Qualitative, quantitative and mixed methods approaches. 5th ed. Lincoln: University of Nebraska; 2019.

[19] Lutkevich B. Natural language processing (NLP). Tech Accelerator: A guide to artificial intelligence in the enterprise. TechTarget. January 2023. Available from: https://www.techtarget.com/searchenterpriseai/definition/natural-language-processing-NLP

[20] Spencer EL. Natural language processing (NLP) techniques & Examples. Revuze. January 13, 2023. Available from: https://www.revuze.it/blog/natural-language-processing-techniques/