Annals of Emerging Technologies in Computing (AETiC)
Vol. 9, No. 3, 2025

Research Article

Investigating the Accuracy of the GPT2 Algorithm in Classifying Identified Targets for an Intelligent Virtual Assistant

Shangying Guo[1] and Jing Zhao[2,*]

[1]School of English Language and Culture, Xi'an Fanyi University, China
Sh.Gu01614314@gmail.com
[2]School of Computer, Wuhan Donghu University, China
zhaojin1112024@163.com
*Correspondence: zhaojin1112024@163.com

Received: 6th October 2024; Accepted: 28th June 2025; Published: 1st July 2025

Abstract: Natural Language Understanding (NLU) is a branch of Natural Language Processing (NLP) that focuses on enabling computers to interpret human language with a level of understanding comparable to humans. NLU encompasses several tasks, including parsing sentences to understand grammatical structure, identifying word and phrase meanings, and determining user intent from natural language inputs. Many AI systems today—such as chatbots and virtual assistants—rely on NLU to accurately interpret and respond to user inputs in real time. This study addresses the challenge of accurately classifying user intents in multilingual intelligent virtual assistants a task critical for enhancing real-time human-computer interaction, by exploring the application of seven GPT-2 based models, leveraging their embedding matrices and tokenizers to design a robust intent-classification framework. The variation in the GPT-2 models in this study lies in the number of final layers and dimensional configurations used for classification. Through a large-scale case study with over one million utterances in 51 languages, the models were evaluated based on key metrics such as Accuracy, Precision, Recall, and F1-Score. Findings indicate that the GPT-256 model consistently achieved the highest values across these metrics, establishing it as the most accurate among the models tested. The GPT-256256 and GPT-128128 models followed closely, both of which showed competitive performance but with slightly lower accuracy than GPT-256. These results underscore the effectiveness of specific model configurations in improving NLU for virtual assistants, particularly in multilingual applications. The findings provide insight into optimizing AI systems for accurate goal classification, enhancing the ability of virtual assistants to understand and respond to diverse user inputs more precisely across languages, making them highly adaptable for global applications.

Keywords: ***Classification; GPT2; Natural Language Processing; Natural Language Understanding; Transformer; Virtual Assistant***

1. Introduction

Natural Language Understanding, or NLU for short, is a sub-field within NLP focused on translating human language into a format more easily readable and understandable by machines [1]. Natural Language Processing develops methods and algorithms to generate, analyze, and interpret speech or text in natural language. Several wide-ranging methods and approaches exist within deep learning, machine learning, and statistical models. Applications involving these methods range from speech recognition to machine translation, sentiment analysis, and text categorization [2]. While the goal of understanding unstructured data is for both, there is a difference between NLP and NLU. NLP studies how computers are coded to understand language and enable natural conversation between computers and people. Meanwhile, NLU deals with machines' ability to understand human language. In other words, first, the computer has to

Shangying Guo and Jing Zhao, "Investigating the Accuracy of the GPT2 Algorithm in Classifying Identified Targets for an Intelligent Virtual Assistant", *Annals of Emerging Technologies in Computing (AETiC)*, Print ISSN: 2516-0281, Online ISSN: 2516-029X, pp. 22-32, Vol. 9, No. 3, 1st July 2025, Published by International Association for Educators and Researchers (IAER), DOI: 10.33166/AETiC.2025.03.002, Available: http://aetic.theiaer.org/archive/v9/v9n3/p2.html.

understand the features of human language, which is called NLU, and then take unstructured text and reconfigure it in a form that is readable by machines, which is called NLP [3-5].

NLU uses computerized techniques to use natural language and analyze and interpret texts or speeches. NLU seeks to enable computers and systems to understand meaningfully, respond effectively, and be helpful to natural language input [6]. Many AI systems today, such as chatbots and virtual assistants, use NLU to translate user entries in natural language form. Also, the techniques of NLU play a vital role in several other sectors like healthcare, finance, and customer service, where interpretation and understanding of natural language must be depicted so that effective communication is enabled with humans [7-9]. NLU generally refers to the capability of a computer to grasp the structure and meaning of human languages, such as Japanese, Spanish, and English. In other words, natural language understanding, NLU, is AI that utilizes computer software to read text and any form of unstructured data. This approach allows an individual to interact with the computer in one's language and using natural phrases. NLU can parse text into machine or computer language and produce output in a form understandable to humans [10]. Such data evaluation can thus be done by computers themselves, automatically, in a matter of seconds using NLU and ML. This saves us time and money when analyzing customer feedback [11]. NLU focuses on understanding the meaning of natural language to have its context by syntactic and semantic analysis. Among the most frequently NLU tasks are, semantic analysis, intent detection, entity recognition, and sentiment analysis are the most prominent. Parsing is a way through which NLU does its job, corrects the structure of sentences, and draws meanings from the text that are exact or taken from dictionaries. On the other hand, semantic analysis deals with the grammatical form of sentences, whereby the arrangement of phrases, words, and sentences is dealt with [7,12]. Transformer-based models such as BERT and GPT-2 have revolutionized natural language understanding (NLU) by significantly improving the accuracy of tasks like intent detection and semantic analysis. These models utilize deep bidirectional or generative architectures to capture complex linguistic patterns and contextual relationships. Despite their success, much of the existing research has predominantly focused on monolingual datasets, especially English, limiting the understanding of their applicability in multilingual settings. Recent work has started to address this by exploring cross-lingual transfer learning techniques that leverage shared representations across languages to improve performance in low-resource languages. However, challenges remain in optimizing model architectures specifically for diverse linguistic contexts and in evaluating these models at scale [13].

Human beings naturally know how to understand a phrase and its context. With machines, it is not easy to understand the meaning behind what is fed into them as input. Virtual assistants, therefore, use these semantic analysis arrangements to define and determine the relationships between independent words and phrases in a particular context. Virtual assistants learn and infer meaning from the structured combination of phrases and words, enabling more accurate user interactions. In NLU, ML models process large datasets of human language. So, to help these models understand patterns in human languages, they are trained using pertinent training data [14]. The training data for NLU models are usually labeled samples of human language, such as chat logs, customer service requests, or other textual data. Preprocessing is the initial step in NLU, cleaning up textual material so that it gets prepared for analysis. Part-of-speech tagging, which would mean labeling each word with its grammatical purpose, and punctuation, which divides the text into distinct words or phrases, could be activities under this category [4]. Accordingly, various ML techniques are applied to understand the text using NLU models. One of those techniques involves the intent detection of the pre-defined text. This refers to how a model identifies the intention behind a particular text. Imagine, for example, if a user sends a message to the NLU model; it recognizes what is being sent based on the demand for good or service information [4]. NLU has allowed individuals and companies to interact organically with machines, unlocking new capabilities. From customer service to data gathering and machine translation, the applications of NLU are changing our lives and workplaces in various ways [11]. With technology still evolving, we can only expect even more complex applications of NLU to make our lives easier. This branch applies NLU to allow computers to automatically understand natural language queries by developing question answering. Speech Recognition and NLP combined and merged the sub-field of question answering. This technology does not respond by merely listing text but tries to give answers in natural human language [15,16].

The following section reviews a few studies related to the research topic. Gašić *et al.* [17] explored the role of ML-based frameworks in human-like conversational virtual assistant systems. The cases examined in this paper were related to clarification in conversations and semantic labeling of sentences. This work proposes a step-by-step approach through NLU inspired by Sam *et al.* [18], which has been used in developing an AI-based virtual assistant to carry out payment processing calculations. The guide analyzes the customer's statement to establish a man-machine interaction. In one review study, Suta *et al.* [19] reviewed the role of ML in chatbot design. The findings of this review study indicated that the inability to process natural language is the main issue related to chatbot design. Moreover, the most critical challenge in answering these correctly is understanding the input. Balsa *et al.* [20] provided a practical example of using an intelligent virtual assistant to care for elderly patients suffering from Type 2 Diabetes Mellitus. The architecture of the virtual assistant proposed works by analyzing the graphic components using behavior change techniques and giving a theoretical framework. Chiu *et al.* [21] proposed an emotionally aware campus virtual assistant drawing on deep neural networks motivated by creating an intelligent environment. Patil *et al.* [22] proposed a virtual personal assistant with speech recognition and text-to-speech capabilities, integrating different NLU platforms, including IBM Watson, Google Dialogflow, and ML. Some researchers, Mekni [23], has proposed an NLP-based virtual assistant to simulate a human conversation using artificial intelligence. This educational virtual assistant, as it has been called, is appropriate for information searching with the convenience of saving time. Do *et al.* [24] (2022) proposed that virtual assistants analyze dangerous signs with separate visual processing capabilities to instruct first responders using AI capabilities. Giachos *et al.* [25], considering the development of virtual assistants, took a systemic approach to research NLP capabilities to create a robot interface from the point of view of virtual assistants. Antonius *et al.* [26] presented a new method to increase NLU and solve challenges caused by user inputs in the real world to improve the efficiency of virtual assistants. The proposed method used transformer-based algorithms, including BERT and RoBERTa. The literature review showed various approaches and techniques for designing virtual assistants, each with specific advantages and characteristics. AI capabilities and ML and DL algorithms can improve and increase efficiency. One usage of AI that makes use of NLU is in virtual assistants. In applications where comprehending and interpreting natural language input is crucial, including virtual assistants, chatbots, and speech recognition systems, natural language understanding (NLU) approaches are frequently employed. Intelligent virtual assistants based on NLU, with the help of voice technologies such as Siri, Cortana, Alexa, and Google Assistant, allow them to infer goals without considering how they are expressed. Therefore, in this study, using GPT2 algorithm-based models, we tried to check the classification accuracy of identified targets for an intelligent virtual assistant [27,28]. Although these studies demonstrate the growing sophistication of virtual assistants, they often rely on either rule-based or standard ML models and do not explicitly evaluate transformer-based models in multilingual settings. Notably, Antonius et al. (2023) and Giachos et al. (2023) began exploring transformer-based solutions (e.g., BERT, RoBERTa), but their focus remained on monolingual or narrowly defined use-cases. This suggests a gap in the systematic evaluation of transformer architectures across different configurations and languages. Previous studies have focused on NLU accuracy improvement but have generally neglected to contrast different GPT-2-based model architectures across different languages, making it impossible to understand the effect of model design on multilingual intent classification. The present study contributes to bridging the existing research gap by systematically evaluating multiple GPT-2 based architectures for multilingual intent classification. By analysing models with varying numbers of layers and neuron dimensions on a dataset comprising over one million utterances across 51 languages, this research provides valuable insights into how architectural choices impact accuracy and computational efficiency. These findings not only advance theoretical understanding but also offer practical guidance for developing more effective and scalable intelligent virtual assistants capable of operating in diverse linguistic environments.

2. Methodology

The present work attempts to verify the classification accuracy of the targets identified for an intelligent virtual assistant using models based on the GPT2 algorithm. Generative Pre-trained Transformer 2, in short GPT2, has been widely used by researchers and programmers for various NLP applications, such as translation, summarization, completion of text, and many others. It is among the most used models within the NLP community and has provided significant strides in research dealing with language production [29].

OpenAI introduced the world to the state-of-the-art GPT2 language model. Its stunning language creation powers had supplanted the earlier version, notorious for attracting much attention. Just like the BERT and Albert models, GPT2 is transformer-based. It comprises many layers consisting of self-attention processes and feed-forward neural networks. The transformer architecture enables GPT2 to catch long-range interactions and generate language that makes sense in the given context. A large corpus of text data is employed for pre-training GPT2 using unsupervised learning. It gains the ability to anticipate the word that will come before it in a phrase by examining the words that came before it. This procedure aids in the deepening of GPT2's grasp of grammar, language, and word semantic linkages [30,31].

To model, first, the pre-processing of the data, including removing punctuations and stopping words, was done. Also, all textual labels were converted to numerical categories. The models used for data classification include seven models based on the GPT2 algorithm with its embedding matrix and tokenizers. The architecture of these models consists of seven different layers. These layers, respectively, include two layers as input: the Modelling TF GPT2 layer, the Global Max Pooling 1d layer, the batch normalization layer, and finally, two dense layers. The difference between the GPT2 models used in this study is in the dimensions and number of the last layers that are used for classification, including GPT-0, GPT-32, GPT-64, GPT-128, GPT-256, GPT-32*32, GPT-64*64, GPT-128*128, and GPT-256*256. In other words, the GPT-32 model means a model with one layer of 32 neurons, and the GPT -32*32 model represents a model with two layers of 32 neurons for classification. Also, the GPT model is the model without the last layer, which performs classification based on the count of default classes.

The separation percentage of train, test, and validation data equals 0.75, 0.2, and 0.05, respectively. Also, the evaluation indicators used in this study include several trainable parameters, such as time value, train time, *accuracy*, *precision*, *recall*, and *F1 score*. Four probable states, *TP* (true positive), *TN* (true negative), *FP* (false positive), and *FN* (false negative), will likely occur based on possible states for actual and anticipated samples. The evaluation indices are derived using the following equations using these four factors [32-35]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{P} = \frac{TP}{TP + FN} \tag{3}$$

$$F1\ Score = \frac{2 \times Recall\ \times\ Precision}{Recall + Precision} \tag{4}$$

3. Description of the Database

The dataset utilized in this investigation includes more than one million utterances in 51 different languages, with annotations used to investigate NLU issues. Utterances in this dataset include 60 intents and 55 slot types. Also, the classification of this dataset contains 18 different classes with the labels "social," "transport," "calendar," "play," "news," "date time," "recommendation," "email," "IoT," "general," "audio," "lists," "qa," "cooking," "takeaway," "music," "alarm," "weather." Fig. 1 shows the count of samples in each class. According to this figure, the most significant number of examples is related to the calendar, play, qa, email, and IoT classes. Also, the least number of samples is associated with the " cooking " class.

According to the sentence length analysis results, the average length of sentences drawing on words (len/words) equals 6, and the average length of sentences based on characters (len/chars) equals 34.

Fig. 2 shows the correlation between labels, characters, and word count. According to this figure, there is a high correlation between character count and word count. Also, the correlation value between target labels and word count is slightly higher than between target labels and character count. Therefore, the maximum length of sentences based on word count was used to form the embedding layer.

Fig. 3 shows the word count graph by class. This figure indicates that sentences in most classes are less than 40 words long.

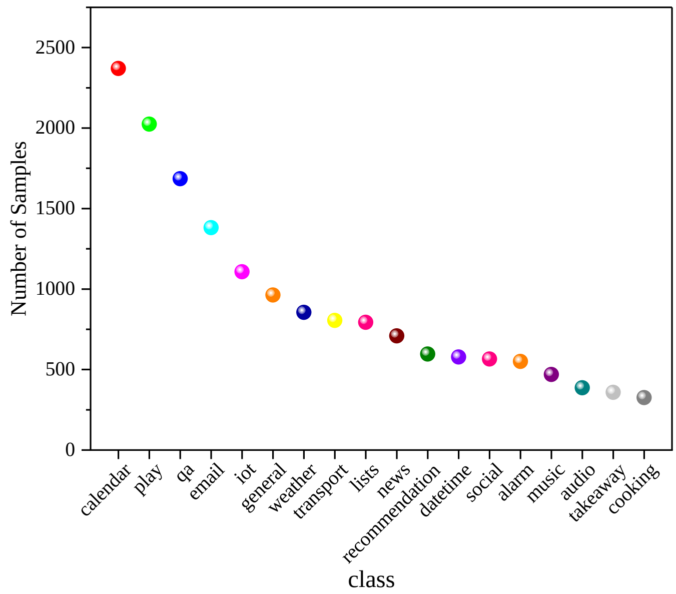

Figure 1. Bar plot related to the frequency of data in each class

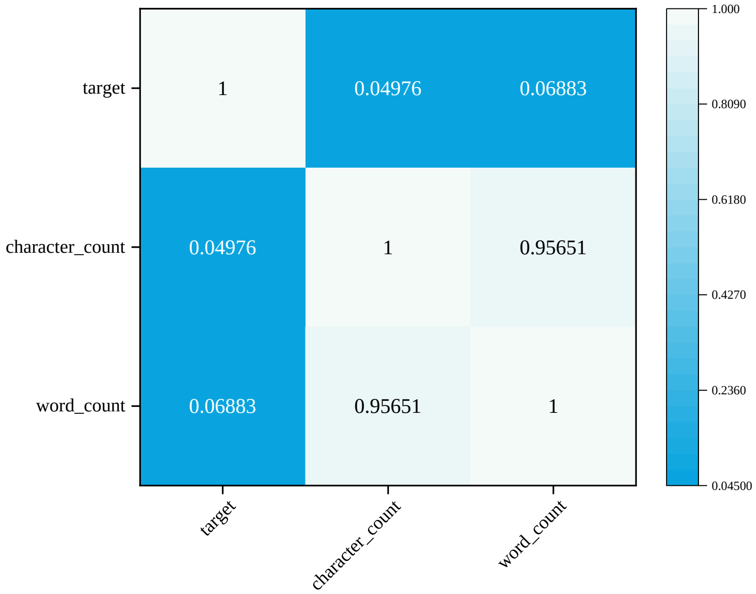

Figure 2. Bar plot related to the frequency of sentence length

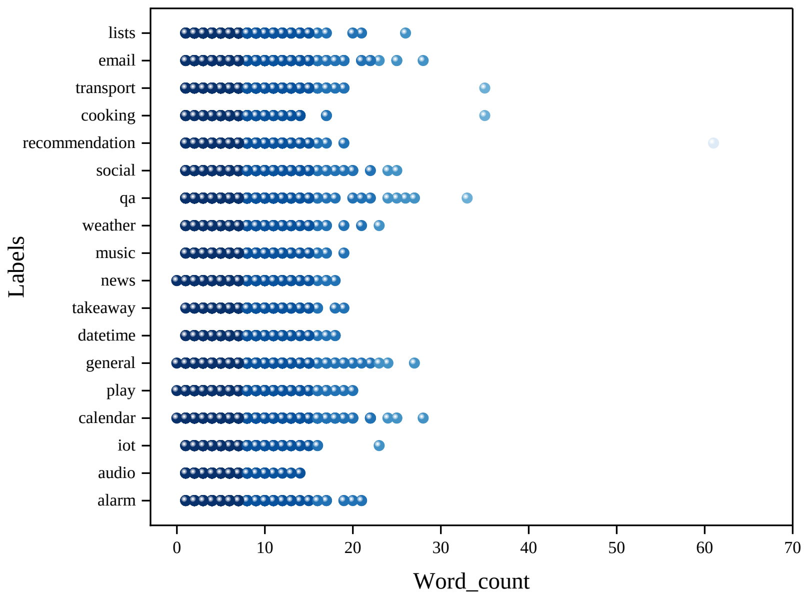

Figure 3. Word count chart by classes

4. Outcomes and Discussion

This section analyzes the accuracy of diverse frameworks based on evaluation indices. Fig. 4 shows a diagram of the count of trainable parameters by model. As this figure shows, GPT-256*256, GPT-256, and GPT-128*128 models have the highest number of trainable parameters, respectively.

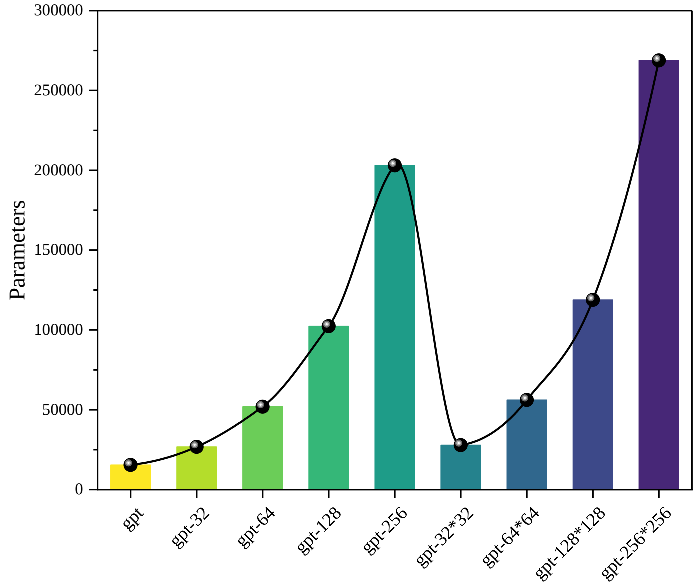

Figure 4. Number of trainable parameters

The values of statistical assessment indices, such as Accuracy, Precision, Recall, and F1-score, for each framework, are displayed in Table 1 and Fig. 5. This figure shows that the model's Accuracy, Precision, Recall, and F1-score indices are respectively, 0.9757, 0.9789, 0.9726, and 0.9755. These values are all greater than those of the equivalent models. As a result, this model performs better in data categorization than other models. After this model, GPT-256*256 and GPT-128 models are in the following ranks of models with the highest accuracy. The results show that the GPT-0 approach has the lowest values of all evaluation indices compared to others. As a result, this model has the least accuracy compared to others.

Table 1. The statistical assessment indices linked to all frameworks

Index	GPT-0	GPT-32	GPT-64	GPT-128	GPT-256	GPT-32*32	GPT-64*64	GPT-128*128	GPT-256*256
Accuracy	0.8944	0.9215	0.9575	0.9727	0.9757	0.9257	0.9484	0.9566	0.9727
Precision	0.8817	0.9175	0.9586	0.9714	0.9789	0.9212	0.9504	0.9622	0.9747
Recall	0.9043	0.9259	0.954	0.9709	0.9726	0.929	0.9487	0.951	0.9687
F1-score	0.8912	0.9211	0.9561	0.971	0.9755	0.9242	0.9485	0.956	0.9714

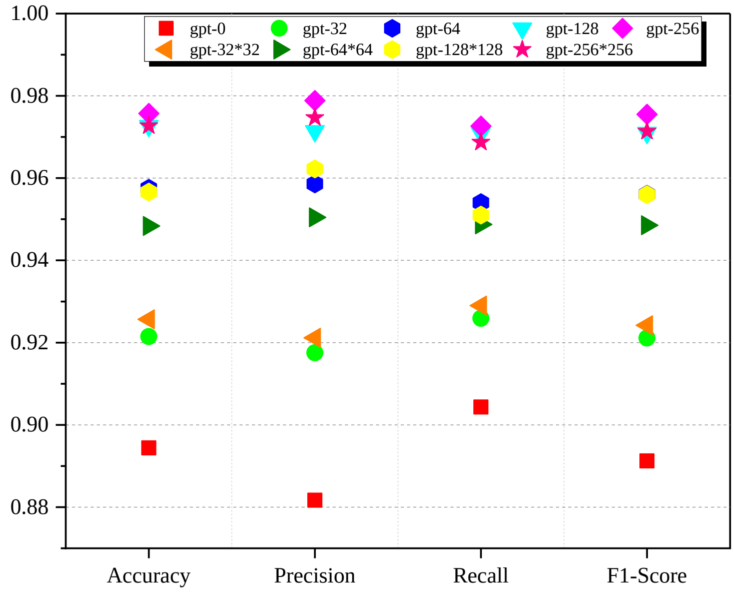

Figure 5. The statistical evaluation indices related to all models

Fig. 6 displays the values of the F1-Score index versus the count of trainable parameters for all models. This figure considers the impact of the count of trainable parameters on processing speed and prediction ability. According to this figure, the highest F1-Score index values are related to the GPT-256 and GPT-256*256 models, which have the highest number of trainable parameters.

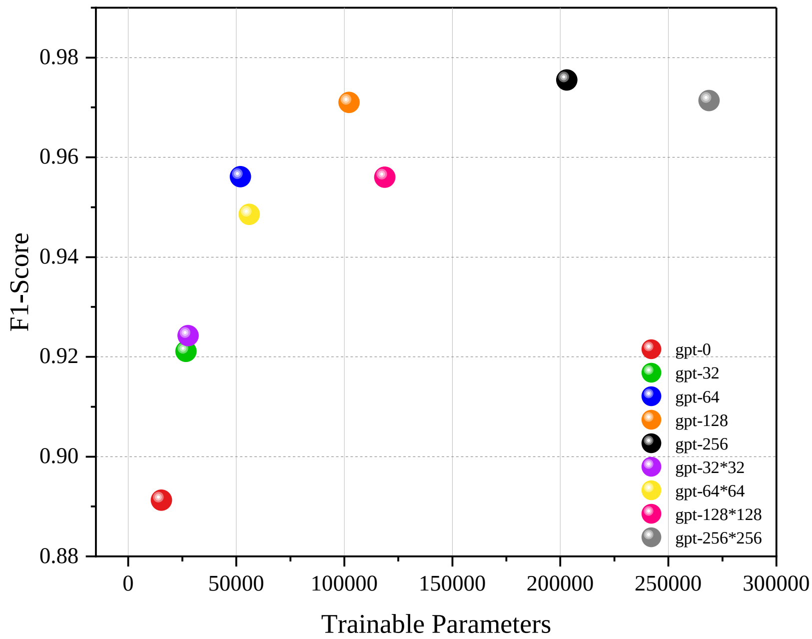

Figure 6. F1-Score index values against the count of trainable parameters

Fig. 7 shows the Time index values against the count of trainable parameters for all models. This chart shows that the maximum number of trainable parameters (GPT-256*256) also has the lowest time. Therefore, it has a higher processing speed than others. Also, the highest value of the Time index corresponds to the GPT-64*64 model.

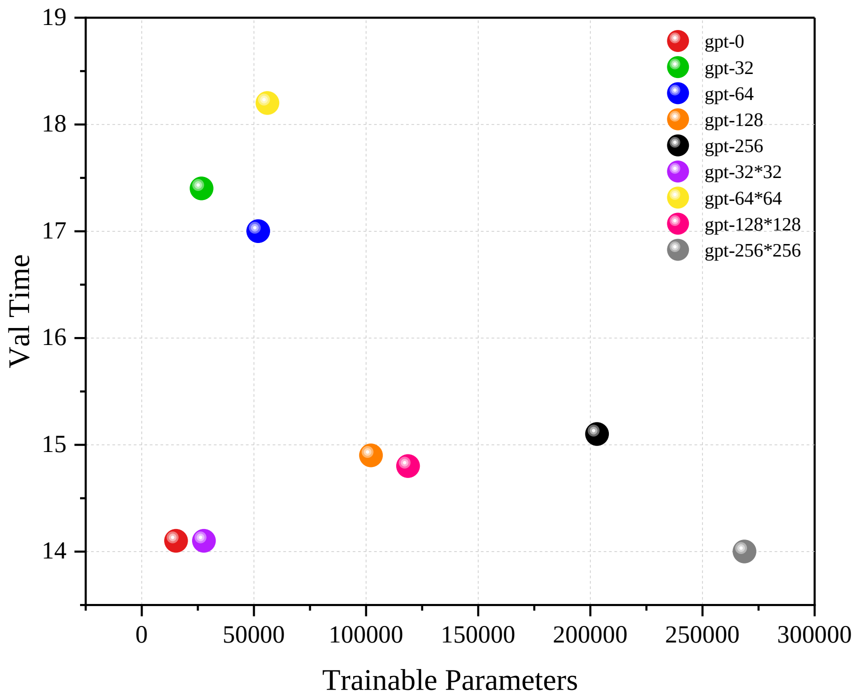

Figure 7. Time index values against the count of trainable parameters

To complete the comparison between the models, Fig. 8 shows the pair plot diagram related to all the evaluation indices, including the count of trainable parameters, Val Time, Train Time, Accuracy, Precision, Recall, and F1-score separately for the frameworks. This figure shows that the GPT-256, GPT-256*256, and GPT-128*128 models have performed better than others.

Figure 8. Number of trainable parameters

To assess the effectiveness of the proposed GPT-based model, we conducted a comparative evaluation against existing state-of-the-art approaches reported in the literature. Table 2 summarizes the accuracy achieved by our best-performing models (GPT-128, GPT-256, GPT-256×256) in comparison with related works.

Table 2. Comparative Analysis

Study	Model Used	Dataset	Accuracy (%)
Patil et al. (2021)	Random Forest + TF-IDF	Multilingual	92.60
Antonius et al. (2023)	RoBERTa	IndicNLP	94.80
Present Study	GPT-2 fine-tuned (128 dim)	Multilingual (1M+, 51lang)	97.27
Present Study	GPT-2 fine-tuned (256 dim)	Multilingual (1M+, 51lang)	97.57
Present Study	GPT-2 wide-deep (256×256)	Multilingual (1M+, 51lang)	97.27

As shown in Table 2, our proposed GPT-256 model outperforms all baseline methods from existing literature, achieving the highest accuracy of 97.57%, which marks a ~5% improvement over RoBERTa and ~6% over traditional ML-based approaches. This performance gain highlights the effectiveness of deeper GPT-based architectures in capturing complex semantic patterns across multilingual datasets. Moreover, while Antonius et al. (2023) utilized pre-trained transformer models, their performance plateaued around 94–95%. Our model surpasses this by leveraging fine-tuned depth-specific architectures optimized for intent classification tasks. These results provide strong empirical evidence of the superiority of our method, particularly in terms of generalizability, model scalability, and multilingual understanding, which are essential for real-world NLU applications.

5. Conclusion

NLU is a branch of natural language processing (NLP) that aims to provide computers with human-like language comprehension. While NLU is a more narrowly focused discipline that focuses on enabling computers to comprehend human language, NLP is a broader field that covers a wide range of strategies for working with natural language. Virtual assistants are one of the AI applications in which NLU is used. Therefore, this investigation presented a framework utilizing transformers to classify the specified goals for an intelligent virtual assistant precisely. The reviewed models for data classification are based on the GPT2 algorithm with its embedding matrix and tokenizers.

The types of GPT2 models used in this study include GPT-0, GPT-32, GPT-64, GPT-128, GPT-256, GPT-32*32, GPT-64*64, GPT-128*128, and GPT-256*256. These frameworks differ in the dimensions and count of the last layers used for classification. The results of a case study on a dataset with more than one million utterances in 51 different languages with annotations showed that the GPT-256, GPT-256*256, and GPT-128*128 models have performed better than others. More specifically, compared to other models, the GPT-256 model has the most significant index values and, thus, the maximum accuracy, according to the analysis of many assessment indices, including Accuracy, Precision, Recall, and F1-Score. After this model, the GPT-256*256 and GPT-128 models are in the following ranks of models with the highest accuracy. The findings demonstrated that, compared to other models, the GPT-0 approach has the lowest accuracy due to its lowest values across all assessment indices.

CRediT Author Contribution Statement

Shangying Guo: Methodology, Software, Validation; Jing Zhao: Writing – Original Draft, Conceptualization, Supervision.

Acknowledgments

This work was supported by the multidisciplinary project of Xi'an Fanyi University (Project Number: 25XYH002).

References

[1] Navin Sabharwal, Amit Agrawal, Navin Sabharwal and Amit Agrawal, "Introduction to Natural Language Processing", In *Hands-on Question Answering Systems with BERT: Applications in Neural Networks and Natural Language Processing*, Apress, Berkeley, CA: Springer Nature, 2021, Print ISBN: 978-1-4842-6663-2, Online ISBN: 978-1-4842-6664-9, ch. 1, pp. 1–14, DOI: 10.1007/978-1-4842-6664-9_1, Available: https://link.springer.com/chapter/10.1007/978-1-4842-6664-9_1.

[2] Sowmya Vajjala, Bodhisattwa Majumder, Anuj Gupta and Harshit Surana, *Practical natural language processing: a comprehensive guide to building real-world NLP systems*, 1st ed. Sebastopol, California, USA: O'Reilly Media, 2020, Available: https://www.oreilly.com/library/view/practical-natural-language/9781492054047/.

[3] Irene Li, Jessica Pan, Jeremy Goldwasser, Neha Verma, Wai Pan Wong *et al.*, "Neural natural language processing for unstructured data in electronic health records: a review", *Computer Science Review*, Print ISSN: 1574-0137, Online ISSN: 1876-7745, p. 100511, Vol. 46, November 2022, Published by: Elsevier, DOI: 10.1016/j.cosrev.2022.100511, Available: https://www.sciencedirect.com/science/article/abs/pii/S1574013722000454.

[4] Roman Egger and Enes Gokce, "Natural language processing (NLP): An introduction: making sense of textual data", In *Applied Data Science in Tourism: Interdisciplinary Approaches, Methodologies, and Applications*, Cham, Switzerland: Springer Nature, 2022, ISBN: 978-3-030-88389-8, ch. 15, pp. 307–334, DOI: 10.1007/978-3-030-88389-8_15, Available: https://link.springer.com/chapter/10.1007/978-3-030-88389-8_15.

[5] Saskia Locke, Anthony Bashall, Sarah Al-Adely, John Moore, Anthony Wilson *et al.*, "Natural language processing in medicine: a review", *Trends in Anaesthesia and Critical Care*, Print ISSN: 2210-8440, Online ISSN: 2210-8467, pp. 4–9, Vol. 38, June 2021, Published by Elsevier, DOI: 10.1016/j.tacc.2021.02.007, Available: https://www.sciencedirect.com/science/article/abs/pii/S2210844021000411.

[6] Dastan Hussen Maulud, Siddeeq Y. Ameen, Naaman Omar, Shakir Fattah Kak, Zryan Najat Rashid *et al.*, "Review on natural language processing based on different techniques", *Asian Journal of Research in Computer Science*, Online ISSN: 2581-8260, pp. 1–17, Vol. 10, No. 1, 21st June 2021, Asian Society of Research & Development (ASRD), DOI: 10.9734/AJRCOS/2021/v10i130231, Available: https://journalajrcos.com/index.php/AJRCOS/article/view/184.

[7] Matteo Zubani, Luca Sigalini, Ivan Serina and Alfonso Emilio Gerevini, "Evaluating different natural language understanding services in a real business case for the italian language", *Procedia Computer Science*, Online ISSN: 1877-0509, pp. 995–1004, Vol. 176, 2020, Published by Elsevier, DOI: 10.1016/j.procs.2020.09.095, Available: https://www.sciencedirect.com/science/article/pii/S1877050920319955.

[8] Kinjal Basu, Sarat Chandra Varanasi, Farhad Shakerin, Joaquin Arias and Gopal Gupta, "Knowledge-driven natural language understanding of english text and its applications", In *Proceedings of the AAAI Conference on Artificial Intelligence*, 18th May 2021, California USA, Online ISSN 2374-3468, Print ISSN 2159-5399, DOI: 10.1609/aaai.v35i14.17488, pp. 12554–12563, Published by AAAI Publications, Available: https://ojs.aaai.org/index.php/AAAI/article/view/17488.

[9] Valmir Oliveira Dos Santos Júnior, Joao Araújo Castelo Branco, Marcos Antonio De Oliveira, Ticiana L Coelho Da Silva and Lívia Almada Cruz, "A natural language understanding model COVID-19 based for chatbots", In

Proceedings of the 2021 IEEE 21st International conference on bioinformatics and bioengineering (BIBE), IEEE, 25th-27th October 2021, Kragujevac, Serbia, Online ISBN: 978-1-6654-4261-9, Print ISBN: 978-1-6654-4262-6, pp. 1–7, IEEE, DOI: 10.1109/BIBE52308.2021.9635248, Available: https://ieeexplore.ieee.org/abstract/document/9635248.

[10] K. R. Chowdhary, "Natural language processing", In *Fundamentals of Artificial Intelligence*, New Delhi: Springer Nature, 2020, Print ISBN: 978-81-322-3970-3, Online ISBN: 978-81-322-3972-7, ch. 19, pp. 603–649, DOI: 10.1007/978-81-322-3972-7_19, Available: https://link.springer.com/chapter/10.1007/978-81-322-3972-7_19.

[11] Benfeng Xu, Licheng Zhang, Zhendong Mao, Quan Wang, Hongtao Xie, "Curriculum learning for natural language understanding", In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, July 2020, USA, Online ISBN: 978-1-952148-25-5, Print ISBN: 978-1-952148-32-2, Published by Association for Computational Linguistics, pp. 6095–6104, DOI: 10.18653/v1/2020.acl-main.542, Available: https://aclanthology.org/2020.acl-main.542/.

[12] Alessandro Lenci, "Understanding natural language understanding systems", *Sistemi Intelligenti*, Print ISSN: 9550-1120, Online ISSN: 1293-1972, pp. 277–302, Vol. 35, No. 2, May 2023, Published by Società Editrice Il Mulino, DOI: 10.1422/107438, Available: https://www.rivisteweb.it/doi/10.1422/107438.

[13] Trieu Phong Nguyen, Heungwoo Nam and Daehee Kim, "Transformer-based attention network for in-vehicle intrusion detection", *IEEE Access*, Online ISSN: 2169-3536, pp. 55389–55403, Vol. 11, 1st June 2023, Published by IEEE, 10.1109/ACCESS.2023.3282110, Available: https://ieeexplore.ieee.org/abstract/document/10141599.

[14] Ivano Lauriola, Alberto Lavelli and Fabio Aiolli, "An introduction to deep learning in natural language processing: Models, techniques, and tools", *Neurocomputing*, Print ISSN: 0925-2312, Online ISSN: 1872-8286, pp. 443–456, Vol. 470, 22nd January 2022, Published by Elsevier, DOI: 10.1016/j.neucom.2021.05.103, Available: https://www.sciencedirect.com/science/article/abs/pii/S0925231221010997.

[15] Yu Meng, Jiaxin Huang, Yu Zhang and Jiawei Han, "Generating training data with language models: Towards zero-shot language understanding", *Advances in Neural Information Processing Systems*, 28th November-9th December 2022, New Orleans, Louisiana, USA, Print ISSN: 1049-5258, Vol. 35, pp. 462–477, Published by NeurIPS Foundation, DOI: 10.5555/3600270.3600304, Available: https://dl.acm.org/doi/abs/10.5555/3600270.3600304.

[16] Dariusz Mrozek, Szymon Kwaśnicki, Vaidy Sunderam, Bożena Małysiak-Mrozek and Krzysztof Tokarz, "Comparison of Speech Recognition and Natural Language Understanding Frameworks for Detection of Dangers with Smart Wearables", In *Proceedings of the International Conference on Computational Science*, 16th–18th June 2021, Kraków, Poland, Online ISBN: 978-3-030-77970-2, Print ISBN: 978-3-030-77969-6, pp. 471–484, Published by Springer Nature, DOI: 10.1007/978-3-030-77970-2_36, Available: https://link.springer.com/chapter/10.1007/978-3-030-77970-2_36.

[17] Milica Gašić, Dilek Hakkani-Tür and Asli Celikyilmaz, "Spoken language understanding and interaction: machine learning for human-like conversational systems", *Computer Speech & Language*, Print ISSN: 0885-2308, Online ISSN: 1095-8363, pp. 249-251, Vol. 46, November 2017, Published by Elsevier, DOI: 10.1016/j.csl.2017.05.006, Available: https://www.sciencedirect.com/science/article/abs/pii/S0885230817301353.

[18] Albert P. Sam, Brijesh Singh and Ananda Swarup Das, "A robust methodology for building an artificial intelligent (ai) virtual assistant for payment processing", In *Proceedings of the 2019 IEEE Technology & Engineering Management Conference (TEMSCON)*, 12th-14th June 2019, Atlanta, GA, USA, Online ISBN: 978-1-7281-1139-1, Print ISBN: 978-1-7281-1140-7, pp. 1–6, Published by IEEE, DOI: 10.1109/TEMSCON.2019.8813584, Available: https://ieeexplore.ieee.org/abstract/document/8813584.

[19] Prissadang Suta, Xi Lan, Biting Wu, Pornchai Mongkolnam and Jonathan H. Chan, "An overview of machine learning in chatbots", *International Journal of Mechanical Engineering and Robotics Research*, Online ISSN: 2278-0149, pp. 502–510, Vol. 9, No. 4, April 2020, Published by IJMERR, DOI: 10.18178/ijmerr.9.4.502-510, Available: https://www.ijmerr.com/show-176-1358-1.html.

[20] João Balsa, Isa Félix, Ana Paula Cláudio, Maria Beatriz Carmo, Isabel Costa e Silva *et al.*, "Usability of an intelligent virtual assistant for promoting behavior change and self-care in older people with type 2 diabetes", *Journal of Medical Systems*, Online ISSN: 1573-689X, pp. 1–12, Vol. 44, 13th Jun 2020, Published by Springer Nature, DOI: 10.1007/s10916-020-01583-w, Available: https://link.springer.com/article/10.1007/s10916-020-01583-w.

[21] Po-Sheng Chiu, Jia-Wei Chang, Ming-Che Lee, Ching-Hui Chen and Da-Sheng Lee, "Enabling intelligent environment by the design of emotionally aware virtual assistant: A case of smart campus", *IEEE Access*, Online ISSN: 2169-3536, pp. 62032–62041, Vol. 8, 30th March 2020, Published by IEEE, DOI: 10.1109/ACCESS.2020.2984383, Available: https://ieeexplore.ieee.org/abstract/document/9050793.

[22] Jaydeep Patil, Atharva Shewale, Ekta Bhushan, Alister Fernandes and Rucha Khartadkar, "A voice based assistant using Google dialogflow and machine learning", *International Journal of Scientific Research in Science and Technology*, Print ISSN: 2395-6011, Online ISSN: 2395-602X, pp. 6–17, Vol. 8, No. 3, 5th May 2021, Technoscience Academy, DOI: 10.32628/IJSRST218311, Available: https://ijsrst.com/home/issue/view/article.php?id=IJSRST218311.

[23] Mehdi Mekni, "An artificial intelligence based virtual assistant using conversational agents", *Journal of Software Engineering and Applications*, Print ISSN: 1945-3116, Online ISSN: 1945-3124, pp. 455–473, Vol. 14, No. 9, September

2021, Published by Scientific Research, DOI: 10.4236/jsea.2021.149027, Available: https://www.scirp.org/journal/paperinformation?paperid=111666.
[24] Vickie Do, Alexander Huyen, Federick J. Joubert, Mina Gabriel, Kyongsik Yun *et al.*, "A virtual assistant for first responders using natural language understanding and optical character recognition", In *Proceedings of the Pattern Recognition and Tracking XXXIII*, 27th May 2022, Orlando, Florida, United States, Online ISBN: 978-1-5106-5079-4, Print ISBN: 978-1-5106-5078-7, pp. 94–104, Published by SPIE, DOI: 10.1117/12.2620729, Available: https://dataverse.jpl.nasa.gov/dataset.xhtml?persistentId=hdl:2014/56166.
[25] Ioannis Giachos, Evangelos C. Papakitsos, Petros Savvidis and Nikolaos Laskaris, "Inquiring natural language processing capabilities on robotic systems through virtual assistants: A systemic approach", *Journal of Computer Science Research*, Online ISSN: 2630-5151, pp. 28–36, Vol. 5, No. 2, April 2023, Published by Bilingual Publishing Group, DOI: 10.30564/jcsr.v5i2.5537, Available: https://journals.bilpubgroup.com/index.php/jcsr/article/view/5537.
[26] Franciskus Antonius, Purnachandra Rao Alapati, Mahyudin Ritonga, Indrajit Patra, Yousef A. Baker El-Ebiary *et al.*, "Incorporating Natural Language Processing into Virtual Assistants: An Intelligent Assessment Strategy for Enhancing Language Comprehension", *International Journal of Advanced Computer Science and Applications*, Print ISSN: 2158-107X, Online ISSN: 2156-5570, Vol. 14, No. 10, January 2023, DOI: 10.14569/IJACSA.2023.0141079, Available: https://thesai.org/Publications/ViewPaper?Volume=14&Issue=10&Code=IJACSA&SerialNo=79.
[27] Junhong Zhao, Christopher James Parry, Rafael dos Anjos, Craig Anslow and Taehyun Rhee, "Voice interaction for augmented reality navigation interfaces with natural language understanding", In *Proceedings of the 2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ)*, 25th-27th November 2020, Wellington, New Zealand, Online ISBN: 978-1-7281-8579-8, Print ISBN: 978-1-7281-8580-4, pp. 1–6, Published by IEEE, DOI: 10.1109/IVCNZ51579.2020.9290643, Available: https://ieeexplore.ieee.org/abstract/document/9290643.
[28] Pin Ni, Yuming Li, Gangmin Li and Victor Chang, "Natural language understanding approaches based on joint task of intent detection and slot filling for IoT voice interaction", *Neural Computing and Applications*, Print ISSN: 0941-0643, Online ISSN: 1433-3058, pp. 16149–16166, Vol. 32, 13th March 2020, Published by Springer Nature, DOI: 10.1007/s00521-020-04805-x, Available: https://link.springer.com/article/10.1007/s00521-020-04805-x.
[29] Márk Lajkó, Viktor Csuvik and László Vidács, "Towards javascript program repair with generative pre-trained transformer (gpt-2)", In *Proceedings of the Third International Workshop on Automated Program Repair*, 19th May 2022, Pittsburgh Pennsylvania, United States, Online ISBN: 978-1-4503-9285-3, pp. 61–68, Published by ACM Digital Library, DOI: 10.1145/3524459.3527350, Available: https://dl.acm.org/doi/abs/10.1145/3524459.3527350.
[30] C. R. Dhivyaa, K. Nithya, T. Janani, K. Sathis Kumar and N Prashanth, "Transliteration based generative pre-trained transformer 2 model for Tamil text summarization", In *Proceedings of the 2022 International Conference on Computer Communication and Informatics (ICCCI)*, 25th-27th January 2022, Coimbatore, India, Online ISBN: 978-1-6654-8035-2, Print ISBN: 978-1-6654-8036-9, DOI: 10.1109/ICCCI54379.2022.9740991, pp. 1–6, Published by IEEE, Available: https://ieeexplore.ieee.org/abstract/document/9740991.
[31] Jia-Hong Huang, Luka Murn, Marta Mrak and Marcel Worring, "Gpt2mvs: Generative pre-trained transformer-2 for multi-modal video summarization", In *Proceedings of the 2021 International Conference on Multimedia Retrieval*, 21st-24th August 2021, Taipei, Taiwan, Online ISBN: 978-1-4503-8463-6, pp. 580–589, Published by ACM Digital Library, DOI: 10.1145/3460426.3463662, Available: https://dl.acm.org/doi/abs/10.1145/3460426.3463662.
[32] Mehdi Ghadiri, Amir Abbas Rassafi and Babak Mirbaha, "The effects of traffic zoning with regular geometric shapes on the precision of trip production models", *Journal of Transport Geography*, Online ISSN: 1873-1236, Print ISSN: 0966-6923, pp. 150–159, Vol. 78, June 2019, Published by Elsevier, DOI: 10.1016/j.jtrangeo.2019.05.018, Available: https://www.sciencedirect.com/science/article/abs/pii/S0966692317308566.
[33] Amir Rastgoo and Hamed Khajavi, "A novel study on forecasting the airfoil self-noise, using a hybrid model based on the combination of CatBoost and Arithmetic Optimization Algorithm", *Expert Systems with Applications*, Print ISSN: 0957-4174, Online ISSN: 1873-6793, p. 120576, Vol. 229, 1st November 2023, Published by Elsevier, DOI: 10.1016/j.eswa.2023.120576, Available: https://www.sciencedirect.com/science/article/abs/pii/S0957417423010783.
[34] Hamed Khajavi and Amir Rastgoo, "Improving the prediction of heating energy consumed at residential buildings using a combination of support vector regression and meta-heuristic algorithms", *Energy*, Print ISSN: 0360-5442, Online ISSN: 1873-6785, p. 127069, Vol. 272, 1st June 2023, Published by Elsevier, DOI: 10.1016/j.energy.2023.127069, Available: https://www.sciencedirect.com/science/article/abs/pii/S0360544223004632.
[35] Hamed Khajavi and Amir Rastgoo, "Predicting the carbon dioxide emission caused by road transport using a Random Forest (RF) model combined by Meta-Heuristic Algorithms", *Sustainable Cities and Society*, Print ISSN: 2210-6707, Online ISSN: 2210-6715, p. 104503, Vol. 93, June 2023, Published by Elsevier, DOI: 10.1016/j.scs.2023.104503, Available: https://www.sciencedirect.com/science/article/abs/pii/S2210670723001142.