Utilizing Natural Language Processing for Assisting in Writing English Sentences

Mr. Sung Je Bang, Texas A&M University

Sung Je Bang is a PhD student in the Department of Multidisciplinary Engineering at Texas A&M University. He holds a Bachelor of Science and a Master of Science in Computer Engineering from the Department of Computer Science and Engineering at Texas A&M University. During his studies, Sung Je gained industry experience as a software engineering intern. Currently, he serves as a Graduate Research Assistant in the Department of Multidisciplinary Engineering. His research interests include large language models, identity theory, and engineering education.

Dr. Saira Anwar, Texas A and M University

Saira Anwar is an Assistant Professor at the Department of Multidisciplinary Engineering, Texas A and M University, College Station. She received her Ph.D. in Engineering Education from the School of Engineering Education, Purdue University, USA. The Department of Energy, National Science Foundation, and industry sponsors fund her research. Her research potential and the implication of her work are recognized through national and international awards, including the 2023 NSTA/NARST Research Worth Reading award for her publication in the Journal of Research in Science Teaching, 2023 New Faculty Fellow award by IEEE ASEE Frontiers in Education Conference, 2022 Apprentice Faculty Grant award by the ERM Division, ASEE, and 2020 outstanding researcher award by the School of Engineering Education, Purdue University. Dr. Anwar has over 20 years of teaching experience at various national and international universities, including the Texas A and M University - USA, University of Florida - USA, and Forman Christian College University - Pakistan. She also received outstanding teacher awards in 2013 and 2006. Also, she received the ”President of Pakistan Merit and Talent Scholarship” for her undergraduate studies.

Utilizing Natural Language Processing for Assisting in Writing English Sentences

Abstract

Many non-English speaking international students come to the United States to pursue undergraduate engineering programs. However, most of them struggle to learn and use English proficiently. This struggle to learn and use English poses various challenges. For example, such students struggle to describe their plans and thoughts to their college peers and colleagues at work. Also, it is mostly harder for such students to make their place in academic or industry careers. Some of these difficulties arise because students cannot identify sentence structures or differences between various types of sentences in English. Writing in complete sentences is one way to convey ideas effectively in English, and this paper presents the model and its accuracy results for the different types of English language sentences. These types include declarative, imperative, interrogative, exclamative, or invalid. We hypothesize that this model will help students classify written sentences as declarative, interrogative, imperative, exclamative, or invalid. We also discuss the future applications of this model and believe that it can help engineering students correct sentence structure errors according to sentence types. We considered 100 sentences of each sentence type for accuracy and calculated various measures, including precision, F1 score, and recall. Out of 100 declarative sentences, 92 were properly identified as declarative sentences, scoring a high accuracy score of 92%, a precision of 95.8%, a recall of 92%, and an F1 score of 93.9%. Out of 100 interrogative sentences, 77 were correctly classified as interrogative sentences, scoring a moderately high accuracy score of 77%, a precision of 95%, a recall of 92%, and an F1 score of 95.5%. Out of 100 imperative sentences, 55 were correctly classified as imperative sentences, scoring a lackluster accuracy of 55%, a precision of 98.2%, a recall of 55%, and an F1 score of 70.5%. Lastly, out of 100 invalid sentences, 81 were properly determined as invalid, scoring a moderately high accuracy score of 81%, a precision of 50.6%, a recall of 81%, and an F1 score of 62.3%.

Introduction

English is one of the most important natural languages due to its widespread use globally. In the USA, the engineering disciplines use language as a major means of communication [1] . Because of engineers' heavy usage of English, engineering students studying in the United States who wish to pursue their careers in the USA must have a strong grasp of the language in both oral and written form [1] .

However, engineering students from countries that do not have English as their primary language often struggle to comprehend or use English effectively in their work [1][2][3]. One of the problems with students not being able to use English fluently is that when writing in English,

engineering students fail to effectively and accurately utilize sentence structures [4]. The use of the English language is a compulsory requirement for all college degrees, including undergraduate and graduate programs, and for all disciplines, including engineering [4]. However, due to the focus on technical aspects of the field, engineering students often lack appropriate technical communication skills (oral and written) [4][5] and have recognized it as their major perceived need [6]. The importance of the skill is one of the critical ABET criteria for engineering graduates [7]. Further, written communication is considered a critical skill for engineering students [8]. There are two notable reasons in the literature: 1) engineers can think more rigorously about the topic when writing about it and 2) being able to write well helps engineers convey their ideas to their colleagues more effectively [8][9]. Notably, learning to communicate effectively (oral and written) requires correct sentence structure, which is lacking skill and requires assistance.

Recognizing the root problem of engineering students' technical writing as inappropriate use of sentence structure in English, this paper focuses on creating means to allow students to convey their thoughts using correct sentence structure. Prior literature suggests that complete sentences effectively allow an individual to communicate with another because complete sentences enhance the strength and meaning of the ideas in the writer's mind [10].

We hypothesize that this model will assist students in classifying written sentences as declarative, interrogative, imperative, exclamative, or invalid.

For the consideration of complete sentences and their structure in the English language, in this paper, we are focusing on four types of simple sentences [11], which are:

1) declarative, which is inexpressive, neutral, and relatively objective [12](i.e., "An algorithm is a set of instructions designed to perform a specific task.")
2) interrogative, which shows a speculative characteristic [12](i.e., "How does a microprocessor execute an instruction from memory?")
3) imperative, which exhibits request [12] (i.e., "Measure the voltage difference across the terminals of the battery.")
4) exclamative, which shows affective or sentimental aspects [12] (i.e., "The code works without any errors!")

Researchers within the field of Natural Language Processing (NLP) are continuing to develop tools utilizing Artificial Intelligence (AI) to assist in understanding and writing English sentences [13][14][15].

Previous literature shows that to understand English sentences better, students must better understand sentence structures [16]. This paper serves to develop a tool that can help students better understand sentence structures. Our study presents the Natural Language Processing model

that considers sentence type at its core and helps students understand these types using sentence structure and its components. Additionally, the paper discusses the model's accuracy for each sentence type. More specifically, the overarching research question of the study is: How can detecting input sentence components help determine if a sentence is valid and one of the four types of sentences?

Literature Review

Engineering education researchers have investigated how students with low proficiency in English struggle with the language to help develop better methods of learning it [3][17]. Tsai and colleagues [3] presented one example of determining students' difficulties by utilizing data mining to analyze sentences read by Taiwanese students and document their reading errors. In another study, Grumbine and Furuike [17] developed a directed pedagogical approach to teaching English to students of KOSEN, a Japan-based college, through sentence diagramming. The method stemmed from Chomsky's generative grammar.

To assist people in writing tasks, researchers have focused on determining students' difficulties, directed pedagogies for language proficiency, and NLP-based techniques to help in writing mechanisms [14][18].

For instance, Bickel and colleagues [18] developed an N-gram-based completion model that assists users in completing their sentences on a computer by completing the sentence for the user. This model was developed to assist users in completing writing tasks more efficiently, such as communicating with colleagues through email.

Similarly, Melamud and colleagues [14] developed a model that can predict the word or words that would properly complete a given incomplete sentence using a bidirectional long short-term memory (LSTM) network, allowing the model to understand the context of sentences. This method allows the model to have a deep contextual understanding of given sentences to the point of being used to detect human characteristics in writing, such as sarcasm in incomplete sentences [19].

Wiswall states in his study that to help students understand English sentences more effectively, they must be able to understand how sentences can be structured [16]. To help students in understanding sentence structures, we developed a sophisticated model that can enhance existing scholarship.

Method

We adopted a quantitative research design to answer the paper's overarching question. The model, developed with the structures of the four types of sentences in mind, is tested using a

dataset containing strict classifications. Through this study, we also objectively measured our model's performance on this dataset and determined its success and failure based on our statistical analysis of the model's results.

Model

We developed an NLP-based model that takes in a sentence and determines whether it is a valid declarative, exclamative, imperative, or interrogative sentence. We developed this model using a programming library called spaCy [20], which detects all of the components of a given input sentence. Based on the detected components and their order, our model determines if the given sentence is one of the four sentence types or invalid.

Design

Our model aims to take an input sentence and determine if it is one of the four types of simple sentences: declarative, interrogative, imperative, or exclamative. If the model determines the input sentence to be one of the four types of sentences, it outputs its result by categorizing it as such. If the input sentence does not match any of the sentence types, then the sentence is considered incorrect, and the outcome is an invalid sentence, the fifth category.

Our model performs this task by taking the input simple sentence, dividing it into components (i.e., subject, predicate, noun, verb, adjective, etc.), observing what components are present in what order, and using these elements to determine the sentence type.

The model can break a sentence into components with the spaCy module [20]. The spaCy module is a Python-based NLP module that can process an input sentence, tokenize each word, and determine what each token is based on the context in which it is used in the sentence [20]. For example, the spaCy module can split an input sentence into tokens and determine if a token is a subject, a predicate, a verb, a pronoun, and more [20].

Using the detected components, the model observes their usage and the order in which they appear in the sentence to determine what type of sentence it is.

When determining if a sentence is declarative, the model checks for the order of the sentence components: a subject, a predicate, and a period mark. If a subject component appears first, a predicate component with a verb within appears afterward, and a period mark component ends the sentence. The model determines this is a declarative sentence by determining it to be True. If at least one of these criteria is not met, then the model outputs False for declarative. Afterward, the model returns True or False for the input sentence to describe whether it is declarative.

When determining whether a sentence is interrogative, the model checks for formal and colloquial standards. For the formal interrogative standard, if an auxiliary word (i.e., 'are', 'is', 'will') component begins the sentence, a subject component appears second, and a question mark ends the sentence, then the model determines that this is True. For the colloquial interrogative standard, if a subject component appears first, a predicate component carrying a verb component appears second, and a question mark component appears third. The model determines that this is True. If at least one of the two standards is True for the given sentence, then the model determines the input to be an interrogative sentence by returning True.

When determining if a sentence is imperative, the model checks for the order of the following sentence components: a verb, a predicate, and a period mark. If a verb component appears first, a predicate component afterward, and a period mark component ends the sentence, and then the model determines that this is an imperative sentence by determining it to be True. If at least one of these criteria is not met, then the model outputs False for imperative. The model finally returns True or False for the input sentence to describe whether it is imperative.

When determining if a sentence is exclamative, the model checks for the order of the following sentence components: a predicate with a verb and an exclamation mark. If a predicate component containing a verb appears first and an exclamation mark follows it at the end of the sentence, the model determines this is an exclamative sentence by determining it to be True. If at least one of these criteria is not met, the model outputs False for exclamative. The model finally returns True or False for the input sentence to describe whether it is exclamative.

If all four models that check for each sentence type respectively output False, then the given input sentence is determined as an invalid sentence type.

Thus, the model's possible outputs are declarative, interrogative, imperative, exclamative, and invalid.

Figure 1 provides the illustrated design of the model.

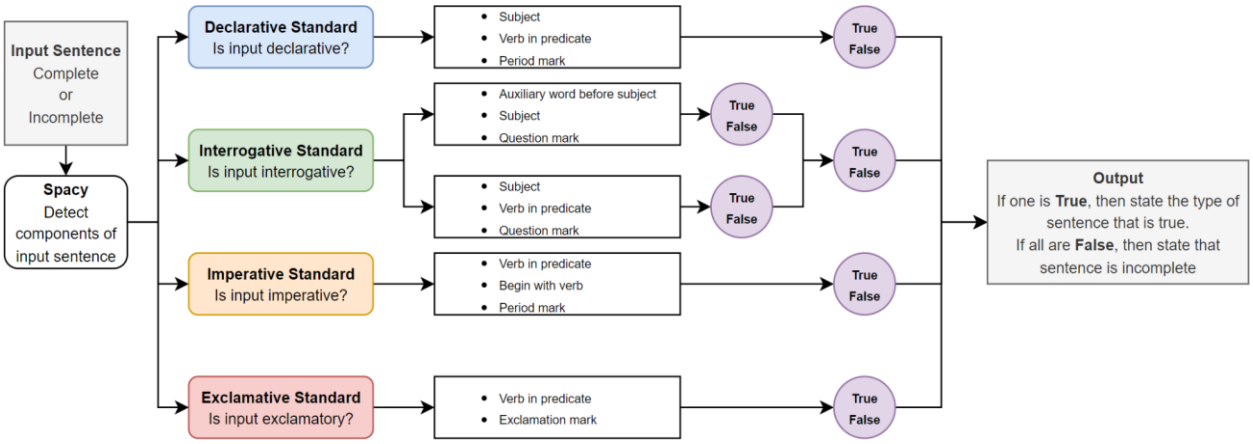

Fig. 1. Sentence type detection model.

Dataset

The dataset used to test this model contains 500 sentences, all generated by ChatGPT, an NLP-powered chatbot [21]. Every sentence within the dataset has been verified to be correctly classified as its respective type of sentence. All sentences were designed to be simple, with no compound or complex sentences. Out of the 500 sentences, 100 sentences were simple declarative sentences, 100 sentences were interrogative, 100 were imperative, and 100 were exclamative. The last 100 sentences were invalid, incomplete sentences that were none of the four types of sentences. The created model took the complete data set, read every row of the dataset, and determined what type of sentence each row was.

Results - Accuracy

Table 1 presents the results of the model's accuracy, precision, recall, and F1 scores based on correct and incorrect identification. Accuracy, calculated with the number of correct classifications divided by the total number of classifications made, describes the model's ability to respond to the input correctly. Precision, calculated with the number of classifications properly marked as correct divided by the sum of the number of classifications properly marked as correct and the number of classifications incorrectly marked as correct, describes how effective the model is in determining correct classifications as correct. Recall, calculated with the number of classifications properly marked as correct divided by the total number of classifications properly marked as correct and the number of classifications incorrectly marked as incorrect, describes how effective a model is in determining specific classifications in a dataset. Lastly, the F1 Score, calculated with the harmonic mean of precision and recall scores, describes how balanced the model is regarding precision and recall.

TABLE I

EVALUATION OF THE OVERALL PERFORMANCE OF THE SENTENCE DETECTION MODEL

Sentence Types	**Correct Count**	**Total**	**Percent-Accuracy**	**Precision**	**Recall**	**F1 Score**
Declarative	92	100	92%	95.8%	92%	93.9%
Interrogative	77	100	77%	95.0%	96%	95.5%
Imperative	55	100	55%	98.2%	55%	70.5%
Exclamative	96	100	96%	88.5%	77%	82.4%
Invalid	81	100	81%	50.6%	81%	62.3%

The model was tested on a dataset containing 500 sentences generated by ChatGPT. Out of 100 declarative sentences, 92 were correctly identified as declarative sentences, resulting in a high accuracy score of 92%, a precision of 95.8%, a recall of 92%, and an F1 score of 93.9%. The high accuracy, precision, recall, and F1 score indicate that the model effectively detects declarative statements. Out of 100 interrogative sentences, 77 were correctly classified as interrogative sentences, resulting in a moderately high accuracy score of 77%, a precision of 95%, a recall of 92%, and an F1 score of 95.5%. This set of results means that the model is generally effective in detecting interrogative statements despite not being perfect. Out of 100 imperative sentences, 55 were correctly classified as imperative sentences, resulting in a lackluster accuracy of 55%, a precision of 98.2%, a recall of 55%, and an F1 score of 70.5%. These results indicate that while 98.2% of imperative sentences were correctly identified as such, the recall score of 55% indicates that many imperative sentences were falsely identified as non-imperative. Out of 100 exclamative sentences, 96 were properly determined as exclamative sentences, resulting in a high accuracy score of 96%, a precision of 88.5%, a recall of 77%, and an F1 score of 82.4%. These results show that, while highly effective in detecting exclamative sentences properly, the recall score of 77% indicates that the model incorrectly classifies some exclamative sentences as non-exclamative. Out of 100 invalid sentences, 81 were properly determined as invalid, resulting in a moderately high accuracy score of 81%, a precision of 50.6%, a recall of 81%, and an F1 score of 62.3%. These results suggest that while the model effectively determines invalid sentences as invalid sentences, the precision score of 50.6% indicates that the model classifies a significant number of valid sentences as invalid.

Discussion

The model effectively detects what type of sentence an input sentence is, excluding imperative sentences in terms of accuracy and recall and invalid sentences in terms of precision. Regarding F1 scores, the model effectively determines all four types of sentences except invalid ones,

which helps determine if the sentence fully conveys the writer's thought. With this tool, users, such as engineering students with English as their second language, can write their sentences and determine if they wrote the correct type of sentence.

The results of the model must be viewed with some limitations. Firstly, the model can only determine an input sentence to be declarative, interrogative, imperative, exclamative, or invalid if the input is a simple sentence, meaning that the model does not consider more complicated sentence structures. The model cannot properly observe and determine the right classifications for compound and complex sentences. Second, the model struggles with prepositions. When a preposition is added to a sentence, the model struggles to consider how the sentence structure is affected compared to when there is no preposition in the same sentence. Third, the model struggles when observing sentences containing heteronyms because the spaCy module struggles to determine the proper sentence component type when encountering heteronyms. For example, when observing the sentence "Wind it", the spaCy module may detect "Wind" as a noun rather than a verb. When a word that can be a noun or a verb based on its definition is added to the sentence, the module sometimes incorrectly labels the word with the wrong component type. The spaCy module's struggle to determine the right sentence component of heteronyms plays a significant role in the model's inaccuracy on imperative statements, where the beginning verb was often incorrectly classified as a noun. The module's struggle with heteronyms in imperative sentences is one of the causes of inaccuracy in specific cases. Lastly, the dataset only contains sentences generated from ChatGPT. Although every sentence has been verified, they are not extracted from real conversations between students or any other humans. Because of this, the model remains untested in real-world situations.

Conclusion

Despite being an important aspect of engineering education, non-native-speaking students struggle to write in English, which prevents them from having more thorough thoughts on the engineering topics they write about. Non-native-speaking students' struggle to write in English also hampers their ability to effectively express their thoughts to their colleagues.

To assist students in conveying their thoughts fully by writing in full sentences, students should be able to write using the four sentence types: declarative, interrogative, imperative, and exclamative.

Our model is designed to address an important aspect of scholarship by taking the input sentence and determining what type of simple sentence it is or if it is invalid. Such a model can be a valuable tool for engineering education research studies. Some potential uses of this model include assistance in writing research papers, as it will help the students write complete sentences by aiming to write with specific types of sentences, such as declarative sentences.

When students in group work must communicate through written form with each other while sharing ideas or thoughts related to their project, they can use our model to receive help in writing complete sentences to fully convey their thoughts. It will allow students to write specific types of sentences that will help better convey their thoughts and ideas to other group members.

Future work can focus on eliminating the limitations discussed. The algorithm can be improved to classify compound and complex sentences as valid or invalid properly. The model can also be improved to address sentences with prepositions properly. Another step future work can take is to improve the model to observe and understand heteronyms in sentences. Lastly, the model can be further tested on a larger dataset with sentences from real humans. Additionally, conducting a usability study with students and getting their feedback on the model's efficacy for improving their writing skills will be important.

References

[1] R. M. O. Pritchard and A. Nasr, “Improving reading performance among Egyptian engineering students: principles and practice,” *Engl. Specif. Purp.*, vol. 23, no. 4, pp. 425–445, Jan. 2004, doi: 10.1016/j.esp.2004.01.002.

[2] Eun Gyong Kim and A. Shin, “Seeking an Effective Program to Improve Communication Skills of Non-English-Speaking Graduate Engineering Students: The Case of a Korean Engineering School,” *IEEE Trans. Prof. Commun.*, vol. 57, no. 1, pp. 41–55, Mar. 2014, doi: 10.1109/TPC.2014.2310784.

[3] Y.-R. Tsai, C.-S. Ouyang, and Y. Chang, “Identifying Engineering Students’ English Sentence Reading Comprehension Errors: Applying a Data Mining Technique,” *J. Educ. Comput. Res.*, vol. 54, no. 1, pp. 62–84, Mar. 2016, doi: 10.1177/0735633115605591.

[4] L. R. Cox and K. G. Lough, “The importance of writing skill to the engineering students,” *Univ. Ark.-Fort SmithUniversity Mo.-Rolla*, 2007.

[5] K. Kozan, M. Menekse, and S. Anwar, “Exploring the Role of STEM Content, Professional Skills, and Support Service Needs in Predicting Engineering Students’ Mid-College Academic Success,” *Int. J. Eng. Educ.*, vol. 37, no. 3, pp. 690–700, 2021.

[6] M. Menekse, X. Zheng, and S. Anwar, “Computer science students’ perceived needs for support and their academic performance by gender and residency: An exploratory study,” *J. Appl. Res. High. Educ.*, vol. 12, no. 5, pp. 1025–1044, Feb. 2020, doi: 10.1108/JARHE-07-2019-0194.

[7] “Criteria for Accrediting Engineering Programs, 2022 - 2023.” ABET, Jan. 05, 2022. [Online]. Available: https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2022-2023/

[8] E. Wheeler and R. L. McDonald, “Writing in Engineering Courses,” *J. Eng. Educ.*, vol. 89, no. 4, pp. 481–486, Oct. 2000, doi: 10.1002/j.2168-9830.2000.tb00555.x.

[9] J. Miller and R. Weinert, *Spontaneous spoken language: Syntax and discourse*. Oxford University Press: Oxford, 1998.

[10] M. Demirezen, “The Recognition of Extended Simple Sentences as a Teaching Writing Problem,” *Procedia - Soc. Behav. Sci.*, vol. 70, pp. 560–566, Jan. 2013, doi: 10.1016/j.sbspro.2013.01.093.

[11] P. Collins, “Clause Types,” in *The Handbook of English Linguistics*, 1st ed., B. Aarts, A. McMahon, and L. Hinrichs, Eds., Wiley, 2020, pp. 131–144. doi: 10.1002/9781119540618.ch8.

[12] L. Jiang and M. Ren, “Declarative, Exclamatory, Interrogative and Imperative Sentences,” in *A General Theory of Ancient Chinese*, in Qizhen Humanities and Social Sciences Library. , Singapore: Springer Nature Singapore, 2023, pp. 201–234. doi: 10.1007/978-981-16-6042-9_11.

[13] D. Cer *et al.*, “Universal Sentence Encoder.” arXiv, Apr. 12, 2018. Accessed: Apr. 25, 2024. [Online]. Available: http://arxiv.org/abs/1803.11175

[14] O. Melamud, J. Goldberger, and I. Dagan, “context2vec: Learning Generic Context Embedding with Bidirectional LSTM,” in *Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning*, Berlin, Germany: Association for Computational Linguistics, 2016, pp. 51–61. doi: 10.18653/v1/K16-1006.

[15] R. Gao, H. E. Merzdorf, S. Anwar, M. C. Hipwell, and A. R. Srinivasa, “Automatic assessment of text-based responses in post-secondary education: A systematic review,” *Comput. Educ. Artif. Intell.*, vol. 6, p. 100206, Jun. 2024, doi: 10.1016/j.caeai.2024.100206.
[16] Z. E. Wiswall, “Improvement in Sentence Structure,” *Engl. J.*, vol. 16, no. 7, p. 518, Sep. 1927, doi: 10.2307/803380.
[17] R. Grumbine and S. Furuike, “Sentence Diagramming as an Effective Tool for Teaching English to Engineers”.
[18] S. Bickel, P. Haider, and T. Scheffer, “Learning to Complete Sentences,” in *Machine Learning: ECML 2005*, vol. 3720, J. Gama, R. Camacho, P. B. Brazdil, A. M. Jorge, and L. Torgo, Eds., in Lecture Notes in Computer Science, vol. 3720. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 497–504. doi: 10.1007/11564096_47.
[19] A. Joshi, S. Agrawal, P. Bhattacharyya, and M. J. Carman, “Expect the Unexpected: Harnessing Sentence Completion for Sarcasm Detection,” in *Computational Linguistics*, vol. 781, K. Hasida and W. P. Pa, Eds., in Communications in Computer and Information Science, vol. 781. , Singapore: Springer Singapore, 2018, pp. 275–287. doi: 10.1007/978-981-10-8438-6_22.
[20] “spaCy · Industrial-strength Natural Language Processing in Python.” [Online]. Available: https://spacy.io/
[21] “Introducing ChatGPT.” [Online]. Available: https://openai.com/blog/chatgpt