International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-12 Issue-2, December 2022

Examining Natural Language Processing Techniques in the Education and Healthcare Fields

Vijaya Kanaparthi

Abstract: ***Natural language processing is a branch of artificial intelligence currently being used to classify unstructured data. While natural language processing is found throughout several fields, these algorithms are currently being excelled in the education and healthcare fields. The healthcare industry has found various uses of natural language processing models. These algorithms are capable of analyzing large amounts of unstructured data from clinical notes, making it easier for healthcare professionals to identify at-risk patients and analyze consumer healthcare perception. In the education field, researchers are utilizing natural language processing models to enhance student academic success, reading comprehension, and to evaluate the fairness of student evaluations. Both fields have been able to find use of natural language model processing models. Some business leaders, however, are fearful of natural language processing. This review seeks to explore the various uses of natural language processing in the healthcare and education fields to determine the benefit and disadvantages these models have on both fields.***

Keywords: ***Artificial intelligence- the development and theory of computer systems completing complex, human tasks (Lucini et. al., 2021 [3]).***

I. INTRODUCTION

Natural language processing is a branch of artificial intelligence "geared toward training computer algorithms to successfully and efficiently recognize and produce human language" (Winless et. al., 2021, p. 523 [1]). It focuses on human or natural language (Mustafina et. al., 2022 [2]), ultimately allowing computers to interpret and understand human speech. First, natural language processing systems organize or "clean" the data into a logical format. This pre-processing stage makes it easier for the natural language processing system to interpret the data. The system then applies algorithms in order to interpret the text (Lucini et. al., 2021 [3]). The two main algorithms used for natural language processing are machine learning models and rule-based systems. Rule-based systems use grammatical rules to interpret text, while machine learning models rely on statistical methods (Goldberg et. al., 2021 [4]). Today, natural language processing is found throughout several fields. However, these algorithms are currently being excelled in the education and healthcare fields. As a newer field of artificial intelligence, natural language processing is currently being used for classifying unstructured data (Garman et. al., 2021 [5]). The healthcare industry has found various uses of natural language processing models.

Manuscript received on 05 October 2022 | Revised Manuscript received on 19 October 2022 | Manuscript Accepted on 15 December 2022 | Manuscript published on 30 December 2022.
***** Correspondence Author (s)
Vijaya Kanaparthi*, Leading Technology Organization, USA. E-mail: datalivesite@gmail.com

These algorithms are capable of analyzing large amounts of unstructured data from clinical notes, making it easier for healthcare professionals to identify at-risk patients and analyze consumer healthcare perceptions (Cho et. al., 2020 [6]). In the education field, researchers are utilizing natural language processing models to enhance student academic success, reading comprehension, and to evaluate the fairness of student evaluations. Both fields have been able to find use of natural language model processing models (Balyan et. al., 2019 [7]). Some business leaders, however, are fearful of natural language processing. For a computer to process text data, the human language must be deconstructed algorithmically, resulting in fear and feelings of unpreparedness for business leaders (Wanless et. al., 2021 [1]). This review seeks to explore the various uses of natural language processing in the healthcare and education fieldsto determine the benefit these models have on both fields. Also, this review seeks to assess various natural language processing models to determine which model is the best for analyzing unstructured data. The following explores various uses of natural language processing and examines the best natural language processing model to analyze unstructured data in the education and healthcare fields.

II. METHODS

The studies found for this systematic review were obtained using searches from EBSCOhost. The search terms utilized for this review were "natural language process", "natural language processing + education", and "natural language processing + healthcare". Only recent studies published within the past five years were utilized for this paper, with the exception of one study that was published in 2015. This study provided ample information regarding the use of natural language processing and the education field, and thus, was selected for this review.

The studies selected for this review were selected based on date and topic. The majority of the studies selected were published within the past two years, providing up-to-date information on the topic. Furthermore, several studies were selected because they compared various natural language processing models. The majority of these studies compared at least three natural language processing models, providing more thorough results. Overall, date and number of models compared were the most important criteria for this review.

III. RESULTS

As mentioned above, the healthcare field has found success in applying natural language processing models to unstructured data.

More specifically, healthcare professionals have found success in applying natural language processing models to clinical notes. The following discusses these applications while also exploring best models for natural language processing in the healthcare field.

IV. HEALTHCARE

The medical field has found use in natural language processing. These uses include developing treatment plans, diagnosing conditions, and optimizing patient experience. As a result, several practitioners and healthcare providers are adopting natural language processing to assist them in understanding the unstructured data in electronic medical records. Natural language processing in life science and health care is expected to grow by 20 percent by the year 2025 (Shiner et. al., 2022 [8]). However, no study has examined which type of natural language processing models are best for analyzing unstructured data in the healthcare field. The first set of studies in this subcategory focus primary on the uses of natural language processing in the healthcare field. The next set of studies examine different natural language processing models that are currently being used to examine unstructured data in electronic medical records. Unstructured health care data can be analyzed using natural language processing. Natural language processing can also be used to diminish the human workload related to interpreting healthcare data. These models can assist with healthcare management, and thus, cut expenses, optimize decision making, and expand access to healthcare. Ionescu (2020) [9] explored the uses of natural language processing in healthcare. The researcher found that natural language processing can be utilized to process data from clinical notes using both machine learning and rule-based systems. By extracting and analyzing data, researchers believe natural language processing is an exceptional tool for healthcare workers to analyze unstructured healthcare data.

Natural language processing algorithms were also applied to Incapacitated with No Evident Advance Directives or Surrogates (INEADS) patients. The prevalence of INEADS patients is unknown because this data is not typically included in electronic health records. In order to overcome this issue, Song et. al. (2022) [10] created a natural language processing algorithm to identify specific information regarding INEADS contained in clinical notes. Researchers utilized a dataset with over 23, 904 adult admissions and 418,393 clinical notes from 2001 to 2012. Researchers also created 17 subcategories indicating elevated or reduced potential for INEADS, then created a language model by applying natural language processing. Any patient with four subcategories was considered part of the high likelihood group. Researchers found that the high likelihood group was significantly more likely to be male and to die during hospitalization. This study shows the benefits associated with utilizing natural language processing in the healthcare field. The results of this study also highlight the potential natural language processing has to identify INEADS patients. Clapp et. al. (2022) [11] examined whether natural language processing could be used to identify patients who are considered at risk of maternal morbidity. The retrospective study consisted of patients admitted to two different hospitals between 2016 and 2020. Natural language processing was used to analyze physical note texts from electronic health records. During the time period, the first hospital had 13,7572 delivery encounters, while the second hospital had 23,397 delivery encounters. Researchers found that natural language processing is capable of predicting severe maternal morbidity from provider documentation at the time of admission. The results of this study can be useful for hospitals, providers, and electronic health record systems by highlighting how artificial intelligences can be used in clinical practice to improve health care. By analyzing physical clinical notes, natural language processing models are capable of assisting healthcare providers decrease maternal morbidity.

Shiner et. al. (2022) [8] also examined the use of natural language processing algorithms for clinical notes. According to researchers, natural language processing algorithms may be used to improve quality measurements in clinical practice. The objective of the study was to used structured electronic medical record data to measure quality of care, then supplement measures with data from natural language processing. Researchers examined the quality of care for PTSD through the U.S. Department of Veterans Affairs (VA) across a twenty-year period of time. Using electronic medical record data, researchers measured PTSD care through measurement-based care and evidence-based psychotherapy. Researchers then recalculated the measures using natural language processing of text from clinical notes. Between 2000 and 2019, over two million people were diagnosed with PTSD through the VA. With the structured electronic medical record data, researchers were able to determine approximately 3.2 percent of patents received quality care for PTSD. However, while using natural language processing data, these estimates saw an increase to 4.1 percent. The results of this study highlight how health quality data can be improved through the implementation of natural language processing. With natural language processing, health systems are able to fill documentation gaps when there are barriers in clinical practice or when there are no available structured tools.

Natural language processing was also used by Ioannides et. al. (2022) [12] to examine e-scooter related injuries. Researchers conducted a retrospective review of patients from 2 hospitals and 180 clinics in Los Angeles from 2014 to 2020. Researchers created a natural language processing algorithm to identify injuries. Thirty-six million clinical notes were used for this study. According to researchers, the natural language processing algorithm was able to correctly classify 92 percent of the clinical notes. Researchers discovered that 1,354 individuals had injuries obtained from e-scooters during this time period. Out of these individuals, 30 percent had follow-up visits, 29 percent required advanced imaging, 6 percent required inpatient treatment, and two died. According to this data, researchers were able to concluded there are approximately 115 e-scooter injuries per million trips that were treated in a healthcare setting.

RetrievalNumber:100.1/ijeat.B38611212222
DOI: 10.35940/ijeat.B3861.1212222
Journal Website: www.ijeat.org

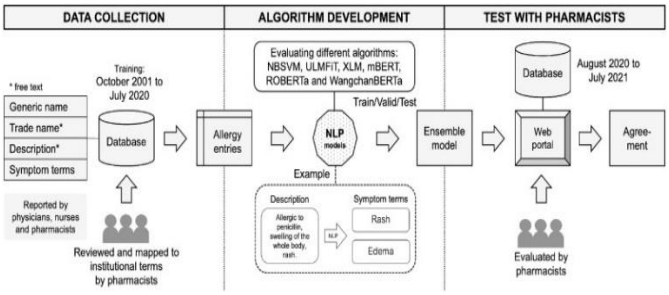
Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Natural language processing algorithms were useful in this study and were able to extract the information necessary for researchers to find an adequate estimate on e-scooter related injuries between 2014 and 2020. Thus, natural language processing algorithms are successful in extracting necessary information from clinical notes.

On the other hand, not all studies have found natural language processing to be beneficial. Miller et. al. (2022) [13] examined the use of natural language processing to automate manual chart review (MCR). Researchers aimed to compare the efficiency and accuracy of natural language processing and manual chart review with surgeon-entered data to determine the rate of gross bile spillage (GBS) during surgery. Bile spillage rates were collected from surgeon-entered data between 2018 and 2019, and then compared to the rates documented in electronic medical record using natural language processing and manual chart review. Altogether, researchers assessed 782 entries and found natural language processing to be 92 percent specific and 27 percent sensitive when it came to detecting bile spillage. In 58 percent of the cases, however, the bile spillage information was ambiguous or missing. Researchers concluded that natural language processing is not able to abstract operative details when they are not clearly documented in the electronic medical records. Thus, in this case, natural language processing is dependent on a surgeon inputting the data. Overall, the results of this study show that while natural language processing can analyze health data, human intervention is required to input the data.

Lastly, Tiyyagura et. al. (2022) [14] examined how natural language processing algorithms can help identify child abuse during infancy. Medically minor findings associated with child abuse, such as bruises, are easily identified by frontline workers. However, the association between these minor findings and child abuse is more difficult to recognize. In order to overcome this issue, researchers developed a natural language processing algorithm to identify injuries in infancy. The algorithm was created to identify ten different injuries associated with infant abuse. The natural language processing algorithm was used on 1,344 provider notes from nine emergency departments over a 3.5-month period of time. Researchers discovered that out of the 1,344 encounters, 3.1 percent had high-risk injuries. The findings of this study show how natural language processing algorithms can be used to identify infants in emergency departments with high-risk injuries, ultimately aiding clinicians to identify infants who may be victims of child abuse. All of the above studies highlight the benefits associated with utilizing natural language processing to analyze unstructured data in the healthcare field. The next set of studies takes a deeper look at natural language processing and analyzes specific models currently being used in the field to determine which is the best model to evaluate unstructured data in the healthcare field.

Chalchulee et. al. (2022) [15] took a specific look at the healthcare field and explored the use of natural language processing in encoding unstructured adverse drug reaction data into institutional symptom terms. According to researchers, 80 percent of all healthcare data is currently unstructured clinical data. Natural language processing could be used to enable automated analysis to "reduce methodological differences in phenotyping clinical data" (p. 3). Thus, natural language processing can be used for a variety of purposes in the healthcare field, such as clinical care and hospital quality improvement. Researchers examined various natural language processing algorithms to encode unstructured adverse drug reaction data in electronic health records into institutional symptom terms.



From Chalchulee, S., Prom chai, C. & Kinwomen, T. (2022). Multi-label classification of symptom terms from free-text bilingual adverse drug reaction reports using natural language processing. *PLo S ONE*, *17*(8), 1-22. http://doi.org/10.1371/journal.pone.0270595

Researchers utilized approximately 80,000 drug allergy recorders and evaluated three different natural language processing techniques: Universal Language Model Fine-tuning (ULM Fit), Naïve Bayes- Support Vector Machine (NB-SVM), and Bidirectional Encoder Representations from Transformers (BERT).

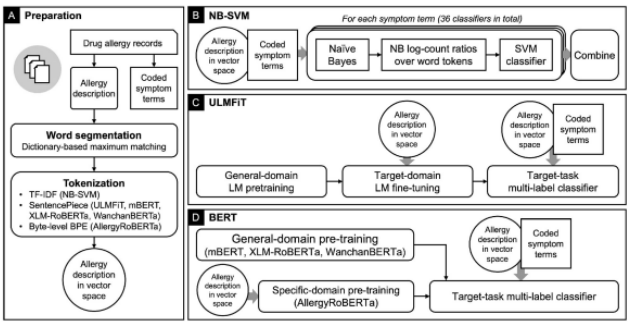


Fig 3. Diagrams outline the steps involved in our methods. (A) Data preparation included word segmentation and tokenization, with each algorithm using different method. (B) NB-SVM involved training multiple pipelines for Naive Bayes feature extraction and SVM classification. (C) ULMFiT involved fine-tuning the pre-trained LM with our target-domain allergy corpus and fine-tuning a classifier for our multi-label classification task. (D) BERT involves fine-tuning a classifier for our multi-label classification task. This study evaluated three pre-trained general-domain BERT models and one target-domain BERT model pre-trained on our allergy corpus.

From Chalchulee, S., Prom chai, C. & Kaewkomon, T. (2022). Multi-label classification of symptom terms from free-text bilingual adverse drug reaction reports using natural language processing. *PLoS ONE*, *17*(8), 1-22. http://doi.org/10.1371/journal.pone.0270595

Overall, all three models were able to analyze unstructured healthcare data. However, when comparing the natural language processing techniques, researchers found that the BERT models outperformed the other models. The results of this study show that the BERT model outperforms the ULM Fit and SN-SVM models when it comes to encoding unstructured adverse drug reaction data in electronic health records. Thus, according to this study, the BERT model is the best model to encode unstructured data found in health records.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Zhou et. al. (2021) [16] looked at natural language processing and its impact on analyzing healthcare data. In this study, researchers examined the use of natural language processing on traditional Chinese medicine. Systematic data mining can be difficult with traditional Chinese medicine because symptoms can have different literal expression but the same meaning. In order to overcome this issue, researchers constructed natural language processing algorithms using four types of traditional Chinese medicine symptom models: bidirectional long short-term memory (BiLSTM), a text classification model based on Bi-LSTM, bidirectional encoder representation from transformers (BERT), and a text classification based on BERT. Similar to Chalchulee et. al. (2022) [15], researchers found the BERT models to outperform the other models. Thus, according to this study, BERT models have superior performance when it comes to normalizing expressions of traditional Chinese medicine symptoms. Similar to Chalchulee et. al. (2022), Zhou et. al. (2022) study highlights the potential benefits BERT models of natural language processing have in the healthcare field.

Chen et. al. (2022) [17] also focused their research on the healthcare industry, but focused their research on emergency departments. Due to the increasing number of emergency department visits, there has been a serious issue with overcrowding across the globe. Emergency department imbalance can cause several issues for a hospital, such as increase risk to patients and burnout amongst medical personal. To help solve the issue, researchers used a natural language processing model to decrease the negative impact associated with overcrowding in emergency departments in Taiwan. Researchers discovered that all natural language processing methods can help with the imbalance in emergency departments. More specifically, the Bi-directional Long Short-term Memory was the best method to help with emergency department imbalance, and thus, can help decrease overcrowding. The results form this study differ from Chalchulee et. al. (2022) and Zhou et. al. (2022) in that they found another natural language processing model to be better analyzing unstructured data in emergency departments.

Natural language processing has also been used to evaluate patient experiences in healthcare. Healthcare experiences are typically collected through the use of both closed and open-ended questions. Natural language processing algorithms can be utilized to automate the responses to these questions, specifically open-ended questions. Van Buc hem et. al. (2022) [18] utilized an Artificial Intelligence Patient-Reported Experience Measure (AI-PREM) tool to analyze the responses to open-ended questionnaires. The AI-PREM tool contained five questions regarding patient experiences in healthcare and was given to 867 patients. Researchers found the AI-PREM to be a comprehensive tool that allows organizing and analyzing of patient feedback. As a result, healthcare professionals will be able to evaluate patient care and prioritize patient experiences through open-ended questions. While close-ended questions are generally easier to analyze, they offer the respondent limited answer options. By utilizing AI-PREM, healthcare professionals are able to assess and analyze responses from open-ended questionnaires, increasing their quality of care. Thus, AI-PREM is an adequate tool to analyze written responses from open-ended questionnaires.

Sutphin et. al. (2020) [19] explored whether natural language processing can assist with adverse drug events. Adverse drug events pose financial and health problems across the globe. Information regarding adverse drug events can help inform healthcare professionals and improve the safety of patients. However, this information is typically buried in texts and requires natural language processing techniques to extract. Researchers examined the effectiveness of three different natural language processing techniques in extracting information regarding adverse drug events. Researchers examined a bidirectional encoder representation from transformers (BERT) model, a bidirectional long short-term memory model, and a conditional random field (CRF). Researchers found all three methods were able to extract the adverse drug event information from narrative text. According to researchers, all three methods had similar performance rates. In other words, one model did not outperform the others.

Yang et. al. (2021) [20] also studied natural language processing in the healthcare field, but focused their research on data obtained from social media. Social media is an effective and rich research for researchers to assess consumers' perception regarding healthcare services. However, social media can be challenging due to the large about of data it contains. Researchers evaluated several different natural language processing algorithms in evaluating this data. Researchers utilized the bidirectional long short-term memory algorithm, support vector machine, shallow neural network, and bidirectional encoder representation from transformers (BERT). Researchers then applied these algorithms to 11,379 collected tweets on Twitter. Researchers discovered that BERT had the highest accuracies out of all of the algorithms (81 percent Corpus 1, 80 percent Corpus 2). According to researchers, BERT is a feasible solution for processing and analyzing data regarding health services from social media. The results of this study highlight the accuracy and efficiency associated with the BERT model.

Han et. al. (2022) [21] explored the use of natural language processing systems to examine social determinants of health. These non-medical factors can have profound impacts on patient outcomes. However, social determinants of health are rarely included in structured electronic medical record data. Researchers conducted a retrospective cohort study on 2,670 clinical notes with 3,500 social related sentences. Researchers then tested three different natural language processing systems: bidirectional encoder representation from transformers (BERT), long short-term memory (LSTM) network, and convolutional neural network (CNN). Researchers sought to determine whether these models could automatically detect eight different social determinants of health categories. All three models were able to accurately detect all social determinants of health categories.

RetrievalNumber:100.1/ijeat.B38611212222
DOI: 10.35940/ijeat.B3861.1212222
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

However, BERT outperformed the other models in the majority of the key metrics (macro-AUC= 0.907, micro-F1= 0.690). Overall, while all three models were able to accurately categories all social determinants of health, BERT outperformed the other models in distinguishing social sentences and logistic regression.

Olthof et. al. (2021) [22] examined natural language processing in radiology. According to researchers, natural language processing is able to extract information from radiology reports, resulting in epidemiological research, quality improvement, and monitoring guidelines. Researchers evaluated four different natural language processing models: long short-term memory recurrent neural network (LSTM), fully connected neural network, bidirectional encoder representation from transformers (BERT), and a convolutional neural network (CNN). Researchers analyzed two different datasets. The first dataset consisted of 2,469 annotated radiologist reports. The second dataset consisted of 2,255 computer tomography studies. The natural language processing models were evaluated based on specificity, sensitivity, negative predictive value, and positive predictive value. Researchers found all four models to have high performance (metrics > 0.90). However, the BERT algorithm outperformed LSTM, Dense, and CNN. Despite variations in prevalence and training size, BERT produced stable results when compared to the other models. Lastly, Li et. al. (2019) [23] specifically explored the effectiveness of the bidirectional encoder representation from transformers (BERT) model in natural language processing. Researchers aimed to evaluate BERT-based models for clinical or biomedical entity normalization. Researchers analyzed 1.5 million notes from electronic medical records and focused primarily on indication, medication, and adverse drug reaction. Two normalization systems were used for comparison, MetaMap and DNorm. Not only did researchers found for the BERT model to outperform both Meta Map and DNorm, these models had state-of-the art performance when it came to clinical and biomedical entity normalization. This study also found that BERT models can be fine-tuned for further accuracy. Thus, like the other studies, this study found the BERT model to outperform the other models in regards to natural language processing. Overall, the results of these studies show various uses for natural language processing models in the healthcare field. More specifically, these studies analyze various models that can be used to analyze unstructured data in healthcare. As shown above, the BERT model seems to be the most effective and most used natural language processing model in healthcare. The BERT models seem to outperform any other model in regards to analyzing unstructured healthcare data. However, the BERT model has not been compared to all natural language processing models used to analyze unstructured healthcare data, resulting in a need for additional research on BERT models.

V. EDUCATION

Natural language processing is currently being used for various reasons in the education field. Furthermore, research on its use in the education field is continuing to grow. Researchers are currently examining the use of natural language processing in reading comprehension, testing, and enhancing student success. There is also an interest in using natural language processing models to analyzing answers to open-ended questions. The following studies examine the various uses of natural language processing models in the education field, and discuss the effectiveness of different natural language processing models.

Natural language processing can have several benefits for the education field. It can also be utilized to assess student evaluations. Andrews et. al. (2021) [24] utilized natural language processing algorithms to examine whether there is gender bias in medical resident evaluations. Previous studies indicate that gender bias with medical education evaluations does exists. In order to examine this issue, researchers examined 3,864evaluations between the years of 2012 and 2018 from Yale School of Medicine. Researchers utilized natural language processing algorithms to analyze the written comments in order to determine whether a gender bias exists in resident evaluations. Researchers found that there was not a substantive difference between males and females in terms of negative and positive comments. While researchers did find that specific competencies were more frequently discussed then other competencies, gender did not seem to have an influence. However, it was discovered that female evaluators wrote longer evaluations than their male counterparts. The results of this study highlight how natural language processing can be used to examine educational evaluations.

Natural language processing has also been shown effective when interpreting written dialogue. However, natural language processing algorithms have not been used on group learning conversations. Sullivan et. al. (2019) [25] explored the use of natural language processing algorithms in collaborative learning discussions. Researchers used parts-of-speech (POS) program to transcript verbal dialogue from middle school students. Natural language processing was then used to examine the text. Researchers found natural language processing to be capable of analyzing transcripts of verbal dialogue. From this information, teachers and other individuals in the education field are provided guidance on how to create or construct learning environments and materials. The findings of this study suggest applying natural language processing algorithms to conversation dialogue can increase their effectiveness of their learning environments and materials.

Pugh et. al. (2021) [26] also examined natural language processing on verbal conversations. Researchers investigated the use of natural language processing models to classify problem solving skills from recorded conversation in a noisy setting. Researchers analyzed data from 44 middle school and high school students. These students utilized videoconferencing to collaborate on math and physic problems. Researchers identified seven different social and cognitive skills through 8,600 utterances. Researchers then analyzed the data using the bidirectional encoder representation from transformers (BERT) model.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The AUROC score was 0.80 for the BERT model and 0.90 for human transcripts. The results of this study indicate that the BERT model does not perform as well than humans. However, the noisy setting may have contributed to the effectiveness of the model. The results of this study highlight that the BERT model is not capable of analyzing verbal data in a noisy setting. Further research is needed for these models to be capable of analyzing verbal data in noisy settings.

One challenging task associated with natural language processing is machine reading comprehension. Machine reading comprehension requires algorithms to determine the answer to a given question and passage. In order to overcome this problem, Peng (2021) [27] utilized models using the BERT and ALBERT models. By integrating these models, researchers were able to create a model that was more effective when it comes to machine reading comprehension. While these models still have reading comprehension problems, the results of this studies highlight how the BERT and ALBERT models of natural language processing can be utilized to enhance machine reading comprehension. The findings of this study impact the education field by creating algorithms that are capable of answering questions associated with reading passages.

Allen et. al. (2015) [28] aimed to develop a model with reading comprehension abilities using an intelligent tutoring system, is TART. Prior to the study, student's performance was graded using sentence-level algorithms that did not focus on reading ability. Instead, researchers decided to utilize natural language processing models to create a comprehension ability model based on linguistic properties. Data was collected from 126 students across eight different training sessions. Linguistic properties were calculated with Coh-Metrix. Researchers found that a student's reading comprehension ability could adequately be measured using linguistic indices through natural language processing. Thus, researchers concluded that natural language processing tools and techniques can be useful in improving student models in intelligent tutoring systems.

Johnson et. al. (2015) [29] also examined natural language processing and is TART. Both is TART and Writing Pal are tutoring systems that assess a learner's written responses and provide accurate, immediate feedback. Through natural language processing, both systems are able to grade and respond to student learners. However, researchers proposed a new practice module that is based on past performance data to decrease or increase the difficulty of practice tests. Researchers were able to adopt this model to the Writing Pal tutoring systems. However, researchers noted that both systems need adaptive writing and reading instructions for the natural processing algorithms to successful be able to increase or decrease difficulty of practice tests.

Natural language processing models have also been used to examine reading comprehension. Using a collaborative augmentation and simplification of text system (CoAST), Shardlow et. al. (2022) [30] explored the potential of natural language processing in enhancing and supporting reading comprehension in higher education. CoAST is an online software created to help students with sophisticated academic texts. The software uses natural language processing algorithms to help identify specific words that may be difficult for readers at various reading levels. Researchers utilized a quasi-experimental design to examine 23 undergrad education students and 23 undergrad technology students in the UK. Participants were randomly selected into experimental and control groups. Students were asked to finish synonym matching task, read annotated texts, and complete a reading comprehension task. The experimental group was given access to the annotations in the annotated text. Researchers discovered that students who use CoAST platforms had an increased ability to comprehend key words. Students who had access to annotations also improved more than the control group. The results of this study highlight how CoAST and other natural language processing algorithms can be used provide lecturers with a list of potentially difficult texts to help reduce the time required to identify and teach difficult words to students. The results also show how CoAST can help enhance and mediate the relationships between students, teachers, and difficult theoretical texts (Shardlow et. al., 2022).

Balyon et. al. (2020) [7] also examines how natural language processing can be used to examine text difficulty. Researchers used natural language processing to predict language difficulty using intelligent tutoring systems. Human rates approximated the text difficulty of 262 texts in two different data sets. Researchers found that natural learning processing was able to increase accuracy by over 10 percent when compared to classic readability metrics. The findings from this study highlight the importance of integrating natural language processing algorithms to text difficulty as well as the potential of natural language processing in developing text difficulty classification.

It can be challenging for computers and software to analyze answers to open-ended questions since it requires examining complex sentences. However, Smith et. al. (2020) [31] investigated how natural language processing algorithms can be used to grade open-ended questions in eBooks. These questions are used to motivate children and increase reading comprehension. Computer science students developed natural language processing algorithms to grade answers to the open-ended questions. Researchers discovered that the best natural language processing algorithm was capable of grading 85 percent of the questions correct or partly correct. While the natural language processing algorithm was in Slovenian, the algorithm mode was successfully implemented in a language art class for fourth graders. Thus, natural language processing algorithms are capable to grade open-ended questions at least partially correct.

Open-ended questions were also explored by Wulff et. al. (2022) [32]. According to researchers, open-ended questions can be troublesome for natural language processing models due to the ambiguity associated with language. Researchers utilized a pretrained language model to uncover patterns to help decrease language ambiguity. Researchers also examined the challenges and potentials associated with pretrained language model and clustering.

RetrievalNumber:100.1/ijeat.B38611212222
DOI: 10.35940/ijeat.B3861.1212222
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Researchers analyzed open-ended questions from 75 preservice physics teachers to determine whether a pretrained language model can cluster specific sentences. Researchers found that pretrained language models can assist with clustering, and thus, decrease the ambiguity associated with language. Researcher conclude that a pretrained language model can be used to assess open-ended questions. The findings from this study highlight the potential of natural language processing in examining open-ended questions.

Some researchers have found success in their own natural language processing models. Mao et. al. (2021) [33] developed their own natural language processing model called Temporal-ASTNN. This model was developed to examine student learning progression. Temporal-ASTNN combines long-short term memory (LSTM) with abstract syntactic trees (AST). LSTM is responsible for handing the temporal nature of progression, while AST is responsible for the linguistic nature. The effectiveness of the model was compared against others, including Code2Vec, iSnap, and Java. Other temporal models were also examined. Researchers found that Temporal-ASTNN outperforms all other models and can achieve the best performance within four minutes. According to the findings from this study, Temporal-ASTNN can be a beneficial natural language processing algorithm to examine student learning progression.

Natural language processing can also be utilized to increase student academic success. Using Course MIRROR mobile system, Menekse (2020) [34] examined the reflection-informed learning and instruction (RILI) model and its impact on student academic success. The researcher hypothesized that the RILI model can be used to provide students with reflection information, ultimately helping teachers to address student difficulty, resulting in enhanced academic success. According to the researcher, students need self-monitoring activities that can help them in their understanding and identification of confusing concepts. The researcher examined 153 undergraduate engineering classes and found that students under the RILI condition were able to perform significantly better than the control group (Cohen's "d" = 0.82). Furthermore, a reflection analysis highlighted that exam performance was significantly associated with both quantity and quality of reflections. Lastly, a survey was given to the participants in the study. The results of the survey suggest that uses of the RIFI model highly valued the model and were more likely to utilize Course MIRROR in other classes. The results of this study show how natural language processing models can be used to enhance academic success in students.

Student academic success can also be measured and analyzed through peer assessments. Jia et. al. (2021) [35] examined the use of natural language processing on peer assessments. Various academic fields have applied peer assessments over the last couple of decades. While they have been found to be effective, peer assessments are dependent on high-quality peer reviews to be successful. Previous studies have found that high-quality peer reviews typically include problems, suggestions, and a positive tone. Natural language processing can be used to detect these features in peer assessments. Researchers utilized a multi-task learning (MTL) model to evaluate these reviews and compared the results to a pretrained language model BERT. Researchers found the BERT model to outperform the other models when it came to evaluating and assessing the peer-review comments. More specifically, the results of this study show that the BERT model outperformed the MLT model when analyzing peer-review comments.

Overall, the results of this study highlight various uses for natural language processing in the education field. According to the results of this study, natural language processing models can help assist with reading compression, answering open-ended questions, analyzing small group discussions, and enhance student academic success. When it comes to the best model, the findings of these studies show that the best model is dependent on the situation. For instance, Temporal-ASTNN was effective in examining student learning progression, while the RILI model and BERT model were effective in measuring and enhancing student academic success. However, these studies did not compare their models to the same natural language processing models. Regardless, the results of this study highlight how specific natural language processing models are better depending on what it is attempting to analyze.

VI. DISCUSSION

The results of these studies highlight the current use and future advantages natural language processing offers. Overall, natural language processing seems to be a useful tool in both the healthcare and education fields. Ionescu (2020), Song et. al. (2022), and Clapp et. al. (2022) all highlight the abilities and benefits of using natural language processing in the healthcare field. By using natural language processing models, Ionescu (2022) found that natural language processing can be utilized to process data from clinical notes using both machine learning and rule-based systems. Song et. al. (2022) also found success with using natural language processing to identify INEADS patients. This study shows the benefits associated with utilizing natural language processing in the healthcare field. Furthermore, Clapp et. al. (2022) found that natural language processing is capable of predicting severe maternal morbidity from provider documentation at the time of admission. The results from Shiner et. al. (2022), Ioannides et. al. (2022), and Tiyyagura et. al. (2022) also show how natural language processing can be useful in analyzing clinical notes from electronic medical records. Shiner et. al. (2022) highlight how health quality data can be improved through the implementation of natural language processing. With natural language processing, health systems are able to fill documentation gaps when there are barriers in clinical practice or no available structured tools. Similar, Ioannides et. al. (2022) found natural language processing algorithms to be useful in extracting information from clinical notes to find an adequate estimate on e-scooter related injuries. Lastly, Tiyyagura et. al. (2022) examined how natural language processing algorithms can help identify child abuse during infancy.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The findings of this study show how natural language processing algorithms can be used to identify infants in emergency departments with high-risk injuries, ultimately aiding clinicians to identify infants who may be victims of child abuse.

Only one of the studies, Miller et. al. (2022) found natural language processing to not be successful in analyzing data from clinical notes. In this study, researchers compared the efficiency and accuracy of natural language processing and manual chart review with surgeon-entered data to determine the rate of gross bile spillage during surgery. Researchers concluded that natural language processing is not able to abstract operative details when they are not clearly documented in the electronic medical records. This study is important because it highlights the dependence natural language processing has on human input. Without human input, the natural language processing algorithm is not fully capable of analyzing data. Thus, accurate human input is required for natural language processing to be successful.

While the first set of research focused on uses of natural language processing in the healthcare field, the second set of research focused on the best natural language processing model for analyzing unstructured data. Both Chalchullee et. al. (2022) and Zhou et. al. (2021) found the BERT models to be successful when analyzing unstructured data in the healthcare field. Chalchulee et. al. (2022) explored the use of natural language processing in encoding unstructured adverse drug reaction data into institutional symptom terms. Researchers evaluated three different natural language processing techniques: Universal Language Model Fine-tuning (ULM Fit), Naïve Bayes- Support Vector Machine (NB-SVM), and Bidirectional Encoder Representations from Transformers (BERT). Researchers found that the BERT models outperformed the other models. Zhou et. al. (2021) had similar results. In this study, researchers examined the use of natural language processing on traditional Chinese medicine. Similar to Chalchulee et. al. (2022), researchers found the BERT models to outperform the other models. The results of these two studies highlights the potential benefits BERT models of natural language processing have in the healthcare field.

Yang et. al. (2021), Han et. al. (2022), Olthof et. al. (2021), and Li et. al. (2019) also found BERT-based models to be best for analyzing unstructured healthcare data. Yang et. al. (2021) focused their research on social media and consumer perceptions regarding healthcare services. Researchers utilized the bidirectional long short-term memory algorithm, support vector machine, shallow neural network, and bidirectional encoder representation from transformers (BERT). Researchers found BERT had the highest accuracies out of all of the algorithms. Furthermore, in Han et. al. (2022), researchers tested three different natural language processing systems: bidirectional encoder representation from transformers (BERT), long short-term memory (LSTM) network, and convolutional neural network (CNN). While all models were able to detect eight different social determinants of health categories, BERT outperformed the other models in the majority of the key metrics, distinguishing social sentences, and logistic regression.

Olthof et. al. (2021) also found the BERT model to be useful in radiology. Researchers evaluated four different natural language processing models: long short-term memory recurrent neural network (LSTM), fully connected neural network, bidirectional encoder representation from transformers (BERT), and a convolutional neural network (CNN). All four models had high performance metrics, but BERT outperformed the other three models and produced more stable results. Lastly, Li et. al. (2019) specifically examined the effectiveness of the BERT model by comparing it to two other models: Meta Map and DNorm. Not only did the BERT model outperform the other two models, the BERT model had state-of-the-art performance when it came to clinical and biomedical entity normalization. This study also found that BERT models can be fine-tuned for further accuracy.

Chen et. al. (2022) and Sutphin et. al. (2020), on the other hand, had different findings. Researchers used a natural language processing model to decrease the negative impact associated with overcrowding in emergency departments in Taiwan. In this study, researchers found the Bi-directional Long Short Term Memory was the best method to help with emergency department imbalance, and thus, can help decrease overcrowding. Furthermore, in Sutphin et. al. (2020), the BERT model did not outperform any other model when examining adverse drug events. Researchers examined a bidirectional encoder representation from transformers (BERT) model, a bidirectional long short-term memory model, and a conditional random field (CRF). Researchers found all three methods were able to extract the adverse drug event information from narrative text. According to researchers, all three methods had similar performance rates. The findings of both these studies are important because they contradict the findings of the above studies. These findings, however, were only found in two of the studies. There were substantially more studies that found the BERT model to be superior when analyzing unstructured healthcare data. Thus, according to the above results, the BERT model is the best for analyzing unstructured healthcare data.

There were similar results in the education field. Natural language processing can have several benefits for the education field. Both Sullivan et. al. (2019) and Pugh et. al. (2021) showed how natural language processing can be effective when interpreting written and verbal dialogue. Sullivan et. al. (2019) explored the use of natural language processing algorithms in collaborative learning discussions. Researchers found natural language processing to be capable of analyzing transcripts of verbal dialogue. Pugh et. al. (2021) also examined natural language processing on verbal conversations. More specifically, researchers analyzed whether the BERT model can analyze verbal data based on social and cognitive skills. Researchers found the BERT model capable of analyzing verbal data, but not as effectively as written data. Furthermore, Peng (2021) utilized the BERT and ALBERT models to answer questions to written passages.

While these models still have reading comprehension problems, the results of this studies highlight how the BERT and ALBERT models of natural language processing can be utilized to enhance machine reading comprehension. Overall, however, both Peng (2021) and Pugh et. al. (2021) found relative success in applying the BERT model to the education field.

Both Allen et. al. (2015) and Johnson et. al. (2017) examined natural language processing and the is TART tutoring system. Allen et. al. (2015) focused their research on reading comprehension. Researchers decided to utilize natural language processing models to create a comprehension ability model based on linguistic properties. Researchers found that a student's reading comprehension ability could adequately be measured using linguistic indices through natural language processing. Johnson et. al. (2017) also examined natural language processing and is TART. Researchers had different findings in that adaptive writing and reading instructions are needed for the natural processing algorithms to be successful in increasing or decreasing difficulty of practice tests.

Reading comprehension and open-ended questions were explored in Shardlow et. al. (2022), Balyon et. al. (2020), Smith et. al. (2020), and Wulff et. al. (2022). Shardlow et. al. (2022) explored the potential of natural language processing in enhancing and supporting reading comprehension in higher education through CoAST. Researchers discovered that students who use CoAST platforms had an increased ability to comprehend key words. Balyon et. al. (2020) also examines how natural language processing can be used to examine text difficulty. Researchers found that natural learning processing was able to increase accuracy by over 10 percent when compared to classic readability metrics. The findings from these studies highlight the importance of integrating natural language processing algorithms to text difficulty as well as the potential of natural language processing in developing text difficulty classification.

Natural language processing and open-ended questions were also examined in the above studies. Smith et. al. (2020) investigated how natural language processing algorithms can be used to grade open-ended questions in eBooks. Researchers discovered that the best natural language processing algorithm was capable of grading 85 percent of the questions correct or partly correct. Thus, natural language processing algorithms are capable to grade open-ended questions at least partially correct. Furthermore, Wulff et. al. (2022) utilized a pretrained language model to uncover patterns to help decrease language ambiguity in open-ended questions. Researchers found that pretrained language models can assist with clustering, and thus, decrease the ambiguity associated with language. Researcher conclude that a pretrained language model can be used to assess open-ended questions. The findings from both these studies highlight the potential of natural language processing with examining open-ended questions.

Models were also examined in the education field. Mao et. al. (2021) developed their own natural language processing model called Temporal-ASTNN. This model was developed to examine student learning progression. The effectiveness of the model was compared against others, including Code2Vec, is nap, and Java. Researchers found that Temporal-ASTNN outperforms all other models and can achieve the best performance within four minutes. Furthermore, Menekse (2020) examined the reflection-informed learning and instruction (RILI) model and its impact on student academic success. Researchers found the RILI model to enhance academic success in students. On the other hand, Jia et. al. (2021) examined the use of natural language processing on peer assessments. Researchers utilized a multi-task learning (MTL) model to evaluate these reviews and compared the results to a pretrained language model BERT. Researchers found the BERT model to outperform the other models when it came to evaluating and assessing the peer-review comments.

Similar to the healthcare field, the BERT model tends to outperform the other models in the education field. However, unlike the healthcare field, other models were considered successful as well. For instance, Temporal-ASTNN was effective in examining student learning progression, while the RILI model and BERT model were effective in measuring and enhancing student academic success. The results of this review show that, while the BERT model may have more favorable results, the best natural language processing model in the education field is dependent on situation. However, the BERT model did tend to perform better the most out of any other model in both the healthcare and education field.

VII. CONCLUSION

As seen from above, natural language processing models can be beneficial for both the healthcare and education fields. These models can analyze unstructured data in both fields, making it easier for humans to interpret large amounts of data. Natural language processing algorithms can be used to analyze clinical notes and identify at-risk patients. By extracting and analyzing data, researchers believe natural language processing is an exceptional tool for healthcare workers to analyze unstructured healthcare data. These algorithms can also be utilized to answer open-ended questions, measure reading comprehension, and enhance student academic success. While several natural language processing models were discussed throughout this review, the BERT model seemed to have the most favorable results for both the healthcare and education fields.

There are several implications associated with the results of this review. First, there are several uses of natural language processing that can be useful to individuals in the healthcare and education field. Natural language processing can be used to adequately analyze clinical notes from electronic medical records. Furthermore, natural language processing can be used to analyze consumer healthcare perceptions through social media. It can also be used to identify at-risk patients. Two of the studies found natural language processing models to be effective at finding adverse drug reactions and at-risk patients by analyzing clinical notes. Prior to natural language processing, computers were not capable of this ability.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

In regards to the education field, natural language processing can assist teachers and educators with multiple tasks. They can be used to monitor reading comprehension and student academic success. In the education field, natural language processing can help enhance the performance of both the student and the educator, ultimately enhancing the entire educational field.

There are also limitations associated with this review. First, only a handful of studies were selected for this review. However, thousands of articles are written on both subjects. The number of studies makes it difficult to properly analyze the topic. This is especially true with the natural language processing models. Without properly reviewing all the literature on the topic, it may be premature to determine the BERT model to be the most successful in analyzing unstructured data. There may be other studies that found another model to be more successful than the BERT model. On the other hand, the studies for this review were randomly selected, and the majority of these randomly selected articles found the BERT model to be the most efficient at analyzing unstructured data.

Future research should be conducted on the BERT model and the healthcare field. More specifically, there needs to be additional research comparing the BERT model to various types of models for analyzing unstructured healthcare data. While these studies did compare the BERT models to other models, not all of the studies compared the BERT model to the same model. Furthermore, two of the studies, Sutphin et. al. (2020) and Chen et. al. (2022), did not find the BERT model to outperform the other models it was compared to. Further research needs to be conducted to determine the actual efficiency of the BERT model. There also needs to be further research on the use of natural language processing and open-ended questions. While natural language processing algorithms seem to be able to analyze and answer open-ended questions, it only has an 85 percent accuracy rate. Thus, additional research is required to increase the accuracy of natural language processing algorithms with open-ended questions.

REFERENCES

1. Wanless, L., Seifried, C., Bouchet, A. &Valenant, A. (2021). The diffusion of natural language processing in professional sports. *Sport Management Review*, *25*(1), 1-24. http://www.doi.org/10.1080/14413523.2021.1968174.
2. Mustafina, H., Suwarsito, S., Pinandita, T. (2022). Natural language processing for mapping exam questions to the cognitive process dimension. *iJet*, *17*(13), 4-16. https://doi.org/10.399/ijet.v17i13.29095 [CrossRef]
3. Lucini, F., Krewulak, K., Fiest, K. & Bagshaw, S. (2021). Natural language processing to measure the frequency and mode of communication between healthcare professionals and family members of critically ill patients. *Journal of the American Medical Informatics Association*, *28*(3), 541-548. http://doi.org/10.1093/jamia/ocaa263 [CrossRef]
4. Goldberg, S., Flemotomos, N. & Martinez, V. (2020). Machine learning and natural language processing in psychotherapy research: Alliance as example use case. *Journal of Counseling Psychology*, *67*(4), 438-448. http://doi.org/10.1037/cou0000382.supp [CrossRef]
5. Garman, A., Erwin, T., Garman, T. & Kim, D. (2021). Developing competency frameworks using natural language processing: An exploratory study. *Journal of Competency-Based Education*, *6*(3), 1-4. http://doi.org/10.1002/CBE2.1256 [CrossRef]
6. Cho, I., Lee, M. & Kim, Y. (2020). What are the main patient safety concerns of healthcare stakeholders: A mixed-method study of web-based text. *International Journal of Medical Informatics*, *140*, 104-162. http://doi.org/10.1016/j.ijmedinf.2020.104162 [CrossRef]
7. Balyan, R., Crossley, S. & Brown, W. (2019). Using natural language processing and machine learning to classify health literacy from secure messages: The ECLIPPSE study. *PLoS ONE*, *14*(2), 1-17. http://doi.org/10.1371/journal.pone.0212488 [CrossRef]
8. Shiner, B., Levis, M., Dufort, V., Patterson, O. & Watts, B. (2022). Improvements to PTSD quality metrics with natural language processing. *Journal of Evaluation in Clinical Practice*, *28*(4), 520-530. http://doi.org/10.1111/jep.13587 [CrossRef]
9. Ionescue, D. (2020). Deep learning algorithms and big health care data in clinical natural language processing. *Linguistic and Philosophical Investigations*, *19*, 86-92. http://doi.org/10.22381/lpi1920204 [CrossRef]
10. Song, J., Topaz, M. Landau, A. & Klitzman, R. (2022). Using natural language processing to identify acute care patients who lack advance directives, decisional capacity, and surrogate decision makers. *PLoS ONE*, *17*(7), 1-12. http://doi.org/10.1371/journal.pone.0290220 [CrossRef]
11. Clapp, M., Kim, E., James, K. & Perlis, R. (2022). Natural language processing of admission notes to predict severe maternal morbidity during the delivery encounter. *Am J Obstet Gyncol*, *227*(3), 511-519. http://doi.org/10.1016/j/ajog.2018.04.015 [CrossRef]
12. Ioannides, K., Wang, P., Kowsari, K., & Kojima, N. (2022). E-scooter related injuries: Using natural language processing to rapidly search 36 million medical notes. *PLoS ONE*, *17*(4), 1-16. http://doi.org/jounral.pone.0266097 [CrossRef]
13. Miller, B., Fafaj, A., Tastaldi, L. & Alkhatib, H. (2022). Capturing surgical data: Comparing aquality improvement registry to natural language processing and manual chart review. *Journal of Gastrointestinal Surgery*, *26*(7), 1490-1494. http://doi.org/10.1007/s11605022-05282-4 [CrossRef]
14. Tiyyagura, G., Asnes, A., Leventhal, J. & Shapiro, E. (2022). Development and validation of anatural language processing tool to identify injuries in infants associated with abuse. *Academic Pediatrics*, *22*(6), 981-988. http://doi.org/10.1016/j.acap.2021.11.004 [CrossRef]
15. Chalchulee, S., Promchai, C. &Kaewkomon, T. (2022). Multi-label classification of symptom terms from free-text bilingual adverse drug reaction reports using naturallanguage processing. *PLoS ONE*, *17*(8), 1-22. http://doi.org/10.1371/journal.pone.0270595 [CrossRef]
16. 16. Zhou, L., Lie, S., Li, C. & Sun, Y. (2021). Natural language processing algorithms for normalizing expressions of synonymous symptoms in traditional Chinese medicine. *Evidence-Based Complementary & Alternative Medicine*, 1-12. \http://doi.org/10.1155/2021/6676607 [CrossRef]
17. Chen, T., Chen, J., Chang, W. & Tsai, W. (2022). Imbalanced prediction of emergency department admission using natural language processing and deep neural network. *Journal of Biomedical Informatics*, *133*, 104-171. http://doi:10.1016/j/jbi.2022.104171 [CrossRef]
18. Van Buchem, M., Neve, O., Kant, I. &Steyerberg, E. (2022). Analyzing patient experiences using natural language processing: Development and validation of the artificialintelligence patient reported experience measure (AI-PREM). *BMC Medical Informatics and Decision Making*, *15*(22), 183. http://doi.org/10.1186/s12911-022-01923-5 [CrossRef]
19. Sutphin, C., Lee, K., Yepes, A. &Uzuner, O. (2020). Adverse drug event detection using reason assignments in FDA drug labels. *Journal of Biomedical Informatics*, *110*. http://doi.org/10.1016/j.bi.2020.103552 [CrossRef]
20. Yang, Y., Al-Garadi, M., Bremer, W. & Zhu, J. (2021). Developing an automatic system for classifying chatter about health services on Twitter: Case study for Medicaid. *Journal of Medical Internet Researcher*, *23*(5), 1-5. http://doi.org/10.2196/26616. [CrossRef]
21. Han, S., Zhang, R., Shi, L. & Richie, R. (2022). Classifying social determinants of health from unstructured electronic health records using deep learning-based natural language processing. *Journal of Biomedical Informatics*, *127*. http://doi.org/10.1016/j.jbi.2021.103984 [CrossRef]
22. 22. Olthof, A., Van Ooijen, P. & Cornelissen, L. (2021). Deep learning-based natural language processing in radiology: The impact of report complexity, disease prevalence, dataset size, and algorithm type on model performance. *Journal of Medical Systems*, *45*(10), 116. http://doi.org/10.1016/j.cmpb.2021.106304 [CrossRef]

23. Li, F., Jin, Y., Lui, W. & Rawat, B. (2019). Fine-tuning bidirectional encoder representations from transformers (BERT) - Based models on large-scale electronic health record notes: An empirical study. *Journal of Medical Internet Research, 21*(9), 1-6. http://doi.org/10.2196/14830 [CrossRef]
24. Andrews, J., Chartash, D. & Hay S. (2021). Gender bias in resident evaluations: Natural language processing and competency evaluation. *Medical Education, 55*(12), 1383-1342. http://doi.org/10.1111/medu.14593 [CrossRef]
25. Sullivan, F. & Keith, P. (2019). Exploring the potential of natural language processing to supportmicrogenetic analysis of collaborative learning discussions. *British Journal of Educational Technology, 50*(6), 3047-3063. http://doi.org/10.111/bjet.12875 [CrossRef]
26. Pugh, S., Subburaj, S. & Rao, A. (2021). Say what? Automatic modeling of collaborative problem-solving skills from student speech in the wild. *International Educational Data Mining Society, 14*, 1-13.
27. Peng, Y., Li, X., Song, J. & Luo, Y. (2021). Verification mechanism to obtain an elaborate answer span in machine reading comprehension. *Neurocomputing, 466*, 80-91. http://doi.org/10.1016/j.neucom.2021.08.084 [CrossRef]
28. Allen, L., Snow, E. & McNamara, D. (2015). Are you reading my mind? Modeling students' reading comprehension skills with natural language processing techniques. *International Conference on Learning Analytics.* [CrossRef]
29. Johnson, A., McCarthy, K., Kopp, K. & Perret, C. (2015). Adaptive Reading and Writing Instruction in iStart and W-Pal. *FLAIRS.* https://aaai.org/ocs/index.php/FLAIRS/FLAIRS17/paper/view/15513/15002
30. Shardlow, M., Sellar, S. & Rousell, D. (2022). Collaborative augmentation and simplification oftext (CoAST): Pedagogical applications of natural language processing in digital learning environment. *Learning Environments Research, 25*(2), 399-421. https://doi.org/10.1007/s10984-021-09368-9 [CrossRef]
31. Smith, G., Haworth, R. & Zitnik, S. (2020). Computer science meets education: Natural language processing for automatic grading of open-ended questions in eBooks. *Journal of Educational Computing Research, 58*(7), 1227-1255. http://doi.org/10.1177.0735633120927486 [CrossRef]
32. Wulff, P., Buschhuter, D., Westphal, A. & Mientus, L. (2022). Briding the gap between qualitative and quantitative assessment in science education research with machine learning – A case for pretrained language models- based clustering. *Journal of Science Education & Technology, 31*(4), 490-513. http://doi.org/10.1007/s10956-022-09969-w [CrossRef]
33. Mao, Y., Shi, Y. & Marwan, S. (2021). Knowing "when" and "where": Temporal-ASTNN forstudent learning progression in novice programming task. *International Education Data Mining Society, 14*, 1-11.
34. Menekse, M. (2020). The reflection-informed learning and instruction to improve students' academic success in undergraduate classrooms. *Journal of Experimental Education, 88*(2), 183-199. http://doi.org/10.1080/00220973.2019.1620159 [CrossRef]
35. Jia, Q., Cui, J. & Xiao, Y. (2021). All-in-on: Multi-Task learning BERT models for evaluating peer assessments. *International Educational Data Mining Society, 14*, 1-8.

AUTHORS PROFILE

Vijaya Kanaparthi, I hold a bachelor's degree in commerce& Stats. I am passionate about learning and building AI &Machine learning technologies that empower the world to do more, accelerate the decision-making process and improve customer experience. Have 15+ experience in commerce and different IT technologies. The passion for automation led to a new chapter in my life. Currently working for a leading technology organization in USA, where every day is a new beginning. I'm also interested in many other things: entrepreneurship, alternative energy, science fiction, and politics.