ISSN 2614-0047

Natural Language Processing in Higher Education

Nastiti Susetyo Fanany Putri [a,1,*], Prasetya Widiharso [a,2], Agung Bella Putra Utama [a,3], Maharsa Caraka Shakti [b,4], Urvi Ghosh [c,5]

[a] Universitas Negeri Malang, Malang, Indonesia
[b] BSS High School, Malang, Indonesia
[c] Monash University, Australia
[1] nastiti.susetyo.2005348@students.um.ac.id ; [2] prasetya.widiharso.2005348@students.um.ac.id, [3] agungbpu02@gmail.com, [4] wibawacenter@gmail.com, [5] ugho0001@student.monash.edu
* corresponding author

ARTICLE INFO

Article history
Received December 30, 2021
Revised January 31, 2022
Accepted February 14, 2022

Keywords
NLP
Higher education
Sentiment analysis
Machine translation
Chatbot

ABSTRACT

The application of Natural Language Processing (NLP) in an educational institution is still quite broad in its scope of use, including using NLP on chatterbots for academic consultations, handling service dissatisfaction, and spam email detection. Meanwhile, other uses that have not been widely used are the combination of NLP and Global Positioning Satellite (GPS) in finding the location of lecture buildings and other university facilities. The combination of NLP and GPS is expected to make it easier for new students and visitors from outside the university to find the targeted building and facilities more effectively.

1. Introduction

Digitization has a significant impact on all areas of life. Education is one of the sectors affected. On the other hand, digitalization also triggers a trend towards developing the amount of data or big data, which poses a challenge for educators to make qualitative data analysis effective and efficient. Natural language processing (NLP) answers the challenges and demands of digitization. NLP can help students understand scientific learning. The application of NLP in education is limited to the effectiveness of developing language learning and improving other academic abilities. NLP instruments also help analyze problems and make recommendations by simplifying and accelerating big data processing [1].

NLP is a branch of computational linguistics, artificial intelligence, and computer science aimed at understanding human language automatically [2]. The input is given in human language (natural language) and converted into output that the machine understands. Information extraction, Spam, machine translation, and answer ranking are some of the most common NLP applications [3]. NLP makes it easy for academics to explore the insights contained in big data without the burden of heavy computing. This paper describes the use of NLP in the scope of higher education, which is intended to provide excellent service to students as consumers in the organization of an educational institution. With the achievement of excellent service, it can produce a university that has optimal performance accountability.

2. Natural Language Processing: Overview

Natural Language Processing (NLP) is an area of interest in the artificial intelligence and computer science groups. NLP research comprises theories and approaches that enable successful natural language communication between humans and computers. NLP combines the scientific fields of computer science, linguistics, and mathematics intending to translate human language into commands that computers can execute [4]. Natural language understanding (NLU) and Natural Language Generation (NLG) are two approaches to study in NLP [5].

NLU or linguistics is the study of language and consists of phonology which deals with sound; morphology, with word creation; and syntax, with sentences, semantics, and pragmatics, which deals with understanding [6]. The main goal of NLU is to understand natural languages by analyzing texts and extracting useful information in subsequent assignments [7]. On the other hand, NLG creates texts in natural languages that humans can understand using structured data, text, graphics, audio, and video [8]. NLG is divided into three categories: text-to-text, such as translators and abstracts [9]; *text-to-other*, such as text that produces an image [10]; and other-to-text, like the video to text [11]. The division of understanding of NLU and NLG is depicted in Fig. 1.

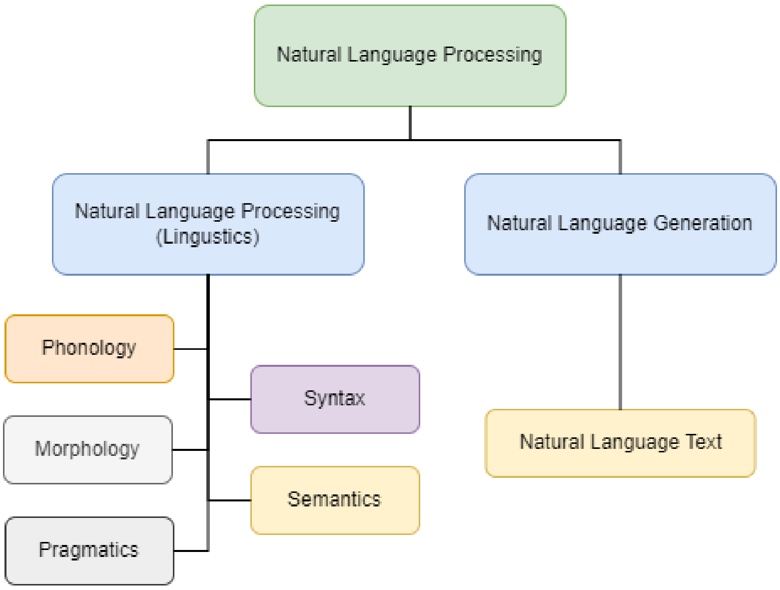

Fig. 1.NLP classification

NLP has undergone four stages of development: an early stage before 1956, a rapid development from 1957 to 1970, a slowing development from 1971-1993, and a period of recovery from 1994 to the present.

The pioneering period or the beginning of NLP began in 1936. Alan Turing introduced the concept of the "Turing Machine". This machine is the theoretical basis for modern computers, the basis for the invention of the electronic computer in 1946. This concept laid the foundation for machine translation and was later called NLP. This project was put forward by Weaver and Booth [12]. Shannon used the probability method of Markov's discrete processes to automate language descriptions in 1948. Then, the notion of thermodynamic entropy was used to calculate the amount of information contained in human language using a probabilistic algorithm [13]. Waver's 1949 memorandum brought the idea of Machine translation (MT) to the world and inspired many other projects. Waver suggests using ideas from cryptography and information theory for language translation. Kleene studied finite automata and regular expressions in the early 1950s, and Chomsky created context-free grammar and applied it to NLP in 1956. As a result of these efforts, two NLP strategies were generated: rule-based and probability. Artificial intelligence (AI) was born in 1956 and supported the development of NLP technology over the next few decades, to improve the technical infrastructure of NLU and NLG.

A significant development occurred from 1957 to 1970 because it was combined with artificial intelligence. During this period, the academia of rule-based and probability-based methods developed rapidly. One of the significant successful projects of this period is the Transformations and Discourse Analysis Project (TDAP) undertaken by the University of Pennsylvania in 1959 and the creation of the Brown American English Corpus. The American psychologist Neisser proposed the concept of cognitive psychology, which directly linked NLP to human cognition in 1967. This period focused more on encoding meaning and developing computationally traceable solutions that previous grammatical theories could not offer. An example of other innovations created from this period is Chomsky's model of linguistic transformation in 1965 [14], Quillian's semantic network [15], and

Schank's conceptual dependency theory, which explains syntactic anomalies and provides semantic representations and other grammar-related innovations [16].

Besides the many developments in the theory of the second period of NLP, it also produced many prototypes. ELIZA by Weizenbaum [17] was built to mimic a conversation between a psychologist and a patient by changing or repeating user input. A robot innovation that can manipulate blocks on a table shows that understanding natural languages is possible for computers, named SHRDLU simulation by Winograd [18]. Wood also developed a system called LUNAR [19], an inference system for a database containing information about moonstone samples using augmentation transitions and semantic procedures.

The second phase of the development of NLP took place between 1971 and 1993. Many projects were undertaken within NLP in the 1970s; for example, the McKeown TEXT discourse designer and McDonald's MUMMBLE response generator use rhetorical predicates to create declarative descriptions in short texts (paragraphs) and TEXT to generate responses that can be understood online. In the second phase, NLP-based applications cannot be completed quickly, and new challenges related to statistical approaches and corpora formation continue to emerge. As a result, many researchers have lost faith in NLP research. Thus, the 1970s put NLP research at its lowest point.

After the mid-1990s, computers became faster and had more storage, sparked NLP research, and allowed speech and language processing technologies to evolve. On the other hand, in 1994, the commercialization of the internet supported the demand for information, and the extraction of natural language-based information was increasing. Joshua Bengio proposed a feed-forward neural network, the first neural language model, in 2001. In 2008 Ronan Colbert introduced multitasking on NLP neural networks. Tomas Mikolov of Google created Word2Vec in 2013, a statistical method for studying word insertion independent of a text corpus using a neural network. Ilya Sutskever proposed a sequence-to-sequence learning model in 2014, a general framework for utilizing neural networks for mapping one sequence to another. Most researchers use statistical models to help machines understand and process human language. Currently, researchers are more focused on updating existing algorithms or creating new methods for NLU and NLG and starting to apply NLP with other research subjects.

3. Natural Language Processing In Education

Educational natural language processing (e-NLP) is a branch of a study investigating the application of natural language processing in Educational settings [20]. This field is an interdisciplinary field of automated text analysis in the context of educational research problems and applications. ENLP can provide educational policymakers with valuable insights into making policies to improve the efficiency and quality of teaching and learning [21]. The three primary responsibilities of NLP in research are assessing, utilizing, and processing languages [22]. In Education, there are many of data in the form of text, so most NLP research in the field of Education focuses on the use of NLP to reveal student behavior (sentiment analysis) [23]–[25], chatbot utilization [26]–[28], machine translation [21], [22], [31] and so on.

3.1. Sentiment Analysis

Sentiment Analysis is discovering, extracting, and studying personal information using NLP and text analysis tools [32]. It is also known as opinion mining because it processes the opinions of several people in a particular domain. In the domain of Education, integrating learners' emotions in the context of teaching and learning and properly handling emotions during the learning process are the focus of two different approaches. In the first approach, emotion is one of the affective domain categories that need to be well developed, while the second approach focuses on integrating the learner's emotions in the teaching-learning context and adequately handling these emotions during the learning process [33]. Sentiment analysis can be described as a non-intrusive, non-invasive, low-cost, design-based tool for a defined emotion-sensing system [34], can serve to increase student profiles with information about themselves. The affective state, through analyzing the traces of his behavior in the teaching and learning environment.

The application of sentiment analysis in education is used to evaluate instruction. Although it has been established that qualitative student feedback highlights variables overlooked by quantitative ratings, student ranking in a Linkert-type questionnaire has become the most popular method of evaluating instruction in higher education [35]. Compared to qualitative feedback, this feedback has a higher level of informativeness and is unstructured, making analysis more difficult. So many studies have switched to using sentiment analysis as an alternative solution for analyzing student responses in an open survey [36]–[38]. Sentiment analysis also can be a reference in an institution's decision-making. The sentiment is used to determine student satisfaction in learning through various comments left on various online platforms [39]. Decision-making regarding education in the pandemic era can also refer to sentiment analysis results [40].

Sentiment analysis is an analytical learning method that can provide valuable research opportunities in the realm of learning analysis and data mining in the field of Education. The latest research adopted a sentiment approach to predict student satisfaction in taking a massive open online course (MOOC) [41]. NLP sentiment analysis is used to analyze unstructured textual data to find positive or negative sentiments contained in the supervising teacher's notes and to predict the probability of dropping out of school.

Besides the many positive values of using sentiment analysis, some requirements must be met beforehand, namely, the need for the availability of large amounts of data in the application of sentiment analysis [42]. However, in reality, there is not too much educational data available, so that in the application of training algorithms often uses non-educational data that has the potential to have differences in the arrangement of patterns and lexicon. The application of sentiment analysis in the real world also has other challenges. (1) The attention-based method applied can lead to a misfocus on syntactically unrelated words. (2) Conventional methods often fail to identify segments with unique sentence structures. (3) Most studies use only one vector to represent the context and target, which limits the results of segmentation analysis because natural languages tend to be complex and complex [43].

3.2. Machine Translation

The main factor in learning is the occurrence of communication, where this communication cannot be separated by language. The understanding process will be very disturbed if there are obstacles to understanding the language. Therefore, there needs to be a bridge to overcoming these language differences. One solution that can be used in developing and applying machine translation in learning. Machine translation (MT) translates from a source text input to text output without human assistance [44].

MT is used to help improve the quality of online education in Africa by leveraging NMT [45]. NMT-based translators are also applied in [46], [47] as a supporting medium in learning foreign languages independently.

Machine translation in education was developed with computer-assisted translation (CAT) in 1930. The most popular software in this field is google translate using a multilingual neural machine translation system that allows zero-shot translation [48], thus being able to translate many languages in 2016. There are two machine translation methods: Statistical Machine Translation (SMT) and Neural Machine Translation (NMT).

Statistical machine translation (SMT) is a new machine translation approach that has recently shown great progress. This approach uses statistics in translating, built on the concept of probability. The SMT system uses a phrase-based model to overcome word-based translation limitations by translating long and varied word sequences [49]. Phrase-based SMT is formulated as a log-linear combination of several statistical models: translation model, Language model, rearrangement model, and word or phrase law [50]. Referring to the opinion of Christopher D Manning and Hinrich Schutze, statistical machine translation, or phrase translation, consists of three components: language model, translation model, and decoder [51]. NLP applications applied in the language model are speech recognition, part-of-speech tagging, and syntactic parsing [52]. SMT excels in terms of flexibility and robustness. However, SMT is difficult to process word mapping correctly, it is difficult to determine the best phrase candidate from input phrases that have different phrase contexts so that the meaning of the translation will be different, it is difficult to predict the derivative structure, and challenging to learn a good language model.

On the other hand, NMT shows better quality than traditional SMT. The most significant advantage of NMT is the gating and attention methods which have proven to be effective in modeling complex remote dependencies and alignment relationships in the translation process, which is a challenge in SMT [53]. NMT is an approach that uses machine learning or deep learning algorithms in its encoder or decoder.

Referring to [54], NMT sometimes has errors in translating due to coverage problems that cause over or under-translation, translation errors due to misinterpreting the natural language of the target sentence so that it does not reflect the original meaning of the input sentence, and UNK problems, and translation quality declines rapidly if the number of UNK words increased. The advantages of SMT include having a mechanism to guarantee each word is translated, treating words as discrete symbols, and explicitly recognizing all translations, including the translation of rare words in UNK in NMT. This makes many researchers combine the two methods in developing machine translation, including in the field of Education.

3.3. Chatbot in Education

A chatbot (Chatterbot) is software that communicates with humans (users) and virtual assistants that can answer several questions correctly [55]. Chatbot technology is widely used in various fields, including industry [56], health [57], marketing [58], tourism [59] even education. Natural language processing or artificial intelligence Markup languages are used by all chatbot algorithms to understand one or more human languag*es.*

Chatbots are text-based and programmed to deal with a limited set of simple questions with answers pre-written by the chatbot developer. The chatbot functions like an interactive FAQ and displays answers like training results. When faced with complex or unexpected questions, the chatbot cannot solve them [60]. Chatbots have evolved to include additional rules and Natural Language processing, allowing users to interact with chatbots in a conversational mode. Because they are faced with many human languages, the latest form of chatbot is contextually aware and able to learn quickly from the input given [61].

The chatbot processes user inquiries and responds appropriately based on applicable mechanisms. Chatbots can be divided into two categories in communicating. Rules-based chatbots and AI chatbots are the two types of chatbots available [62].

Rule-based chatbots, known as pattern matching, use pattern-matching strategies to group words. That way, the bot can provide the correct answer on request. The development of this type of chatbot uses Artificial Intelligent Markup Language (AIML) in constructing the chatbot system [63]. The weakness of this model is that it cannot provide an efficient output if the input pattern is different from the set rules.

AI chatbots use the help of machine learning algorithms that allow chatbots to learn new things. AI chatbots can carry out conversations better than rule-based chatbots because they use machine learning algorithms, NLP, and sentiment analysis [64]. Machine learning chatbots allow identifying user input. Make decisions and learn from previous input. NLP helps chatbots understand how humans communicate and allows chatbots to imitate it. NLP also makes the chatbot understand the conversation context even if the user makes an error in writing the message.

Meanwhile, sentiment analysis makes the chatbot understand the emotions of the user. AI chatbots can understand multiple Languages and read the emotions of the users. However, AI chatbots require much training and must be equipped with a well-defined response to operating correctly.

Some popular chatbot technologies include Apple Siri, Microsoft Carta, Facebook M, and IBM Watson. Recently, the use of chatbots as e-learning learning media has increased [65]. Chatbot technology is a crucial e-learning innovation. It is the most innovative approach to bridging the gap between technology and Education. Including a chatbot enables an engaging learning experience for students, similar to one-on-one interactions with teachers. Adopting chatbots in Higher Education is associated with several benefits, including increasing student motivation and attention, encouraging collaborative learning, promoting communication with friends, and increasing student comfort in learning [66].

There are many approaches related to chatbots in e-learning systems. In [67], the chatbot was developed to address students' complaints from the Benin University computer science faculty, using

Facebook Messenger and Facebook page hosts as a bridge. Other research builds character chatbots that can provide academic information and chatbots about COVID-19. Users can choose whom they want to interact with [68]. Chatbot for international students and academics (CiSA) was developed to facilitate international students about academics and campus life [69].

The use of web bots in developing e-learning chatbots to overcome the problem of delays in responding to student questions has also been developed [70] and designed using Google DialogFlow and implementing Facebook Messenger.

Like all NLP innovations, two sides intersect in every aspect. The benefits of sentiment analysis in education include applying sentiment analysis and structured and insightful knowledge that can be obtained from unstructured text documents, which can be helpful for decision support [39]. In addition, this sentiment analysis provides reasons for academics to visit the school's website so that the information conveyed will spread more quickly. Colleges or schools can also use this method to analyze feedback and comments from surveys to improve their shortcomings [71].

Although it has many positive values, this sentiment analysis also has negative values. Data loss and security issues may arise during analysis [72]. Sentiment analysis also has weaknesses in identifying sarcasm, irony, negation, jokes, and hyperbole expressions, thereby reducing the accuracy of the analysis [73].

Sentiment analysis needs to be done in education, seeing this description. The results of this analysis can be used as a problem-solving tool in improving the quality of education to suit the needs and right on target. Sentiment analysis is also considered to save time and cost for education providers overcoming problems because the main problems can be seen clearly.

Another type of NLP that is often implemented in education is machine translation. This technology is an innovation that is considered to be the most helpful in learning activities. This machine translator is often a bridge for delivering learning materials if the learning resources differ in students' languages. This machine translator also has a relatively short processing time [74]. The development of machine translators is very significant and can translate into many languages using only one tool.

Apart from the many benefits. Machine translators are classified as having a reasonably accurate translation accuracy. However, these results often do not match the grammar and connotations intended by the original text [75]. This weakness will cause a reasonably valuable loss if an error occurs in translating essential financing-related documents. The development of machine translation is considered a threat to the translation profession [76]. This will indirectly lead to an increase in the number of unemployed.

This machine translator's development has made it easier for us in various jobs related to foreign languages. Machine translators are also essential in helping to improve the quality of education because literacy sources and references for learning can be obtained from languages and countries that have different languages from our native language. It is necessary to conduct a final evaluation using a human translator to correct machine translation results to reduce errors in mistranslating.

The world of education cannot be separated from the interaction of humans with humans or humans with machines. The intensity of this interaction tends to be very high, such as in teaching and learning activities, administrative processes, or teacher and student consultation activities. In reality, a teacher or academic staff cannot facilitate it all. So they often use technical assistance, for example, chatbots. This technology was chosen because it can respond quickly compared to humans [77] and is accessible anytime. By implementing this chatbot, educational institutions can reduce expenses because the number of employees they hire can be reduced.

Despite its many benefits, chatbots are devoid of emotions and feelings. This causes interactions with chatbots to tend to be stiff and unable to cope with sudden changes in conversation [78]. Chatbots with AI require continuous optimization, analysis, and maintenance. So educational institutions must frequently update data, so chatbots are more interactive [79]. This ongoing maintenance also costs much money, affecting the school's finances. As already written, the advantages of chatbots are that they can be accessed anytime and limit human interaction.

Chatbots give Education many conveniences, teaching us other methods to interact with the world. This technology has many pros and cons, but Education needs it to carry out administrative activities such as registration and counseling. However, Education cannot entirely rely on chatbots because chatbots do not have emotions, so they can cause errors in providing counseling. So, it is necessary to combine chatbots with human power to make the suggestions given more rational. Chatbots can limit human interaction, but in the current pandemic era, the optimization of chatbots can have other positive effects.

4. Development Directions of NLP in Education

In the previous section, it was explained that NLP had provided a variety of innovations in today's education. However, the current applications are only a few of the advantages NLP, and other artificial intelligence (AI) fields offer for future study. In the future, natural language processing can proliferate from what it is today if researchers analyze the weaknesses of existing technologies and improve them. The following list is some NLP technology development ideas. More ideas will emerge over time

As previously written, it can be seen that both sentiment analysis does not work optimally if there is sarcasm or negation in the sentence. So we need a breakthrough in overcoming this. Several studies use distributed semantic analysis involving word insertion to improve the ability of the sentiment analysis algorithm they developed, as in [80] [81]. In addition, sentiment analysis can also be done using a lexical approach

Future sentiment analysis research in education should also focus more on data generated in natural teaching-learning settings, such as interactions between students on remote collaborative assignments or discussions in cross-semester forums, to develop generalizable and applicable sentiment analysis systems, significantly in the Education scenario.

The main drawback of machine translators and chatbots is that they cannot identify the emotions contained in the input sentences. These deficiencies can be used as a trigger for innovations in developing NLP technology. Innovations that can be created include a combination of chatbots, machine translation, and sentiment analysis. The resulting product can be in the form of applications or robot assistants where the product is designed based on a question-and-answer chatbot about learning life at schools, such as consultation, registration, and other administrative activities. Machine translators facilitate international students so that language barriers do not occur, making it difficult to operate the innovations developed. Meanwhile, sentiment analysis is applied to analyze the user's emotions, with the hope that the user's interaction with the machine will seem more real.

Another development based on e-NLP is combining a chatbot with a global position satellite (GPS). The purpose of making this application is to answer questions from students and prospective students related to the application maker school. GPS is embedded so that answers from bots about specific locations in the campus environment and around campus are more precise and accurate. This development will benefit students and outsiders who will carry out activities on the campus, especially after social distancing and activity restrictions have prevented direct site surveys from being carried out. An illustration of a chatbot that can be developed is shown in Fig. 2.

Fig. 2. Chatbot and GPS Development conversation illustration

In addition to the three categories mentioned above, NLP has another category: detecting spam in emails based on text classification. This category can be applied in education as a student permit detector. This project combines spam detection with sentiment analysis, functioning as a detector for the choice of words used in letters. So that it can be known whether the letter was written by the student's parent or guardian or by the student's friend.

The future of NLP also includes enhanced intelligent search, one of the highlights of expert systems. It can be embedded in the school's website to allow users to search for documents or content using natural language. This search engine can also be inserted into the chatbot so that the chatbot application is faster in finding answers for users from the database. This semantic search can also be embedded in the chatbot, making it possible for the user to use the voice-to-text feature and making the chatbot more varied and efficient because the user can give commands by voice rather than typing through the device.

5. Conclusion

NLP is widely used in the higher education sector as a support for independent academic activities such as machine translators, decision-makers with sentiment analysis, and administrative activities using chatbots. The latest utilization that has not been widely applied in NLP and GPS in the search for buildings and educational facilities at universities. With the increasing use of NLP, it is hoped that it can support improving service quality to optimize a university's performance.

References

[1] E. A. Olivetti *et al.*, "Data-driven materials research enabled by natural language processing and information extraction," *Appl. Phys. Rev.*, vol. 7, no. 4, 2020, doi: 10.1063/5.0021106.

[2] A. Rokade, B. Patil, S. Rajani, S. Revandkar, and R. Shedge, "Automated Grading System Using Natural Language Processing," *Proc. Int. Conf. Inven. Commun. Comput. Technol. ICICCT 2018*, no. Icicct, pp. 1123–1127, 2018, doi: 10.1109/ICICCT.2018.8473170.

[3] P. Gupta, J. Bigham, Y. Tsvetkov, and A. Pavel, "Controlling Dialogue Generation with Semantic Exemplars," pp. 3018–3029, 2021, doi: 10.18653/v1/2021.naacl-main.240.

[4] H. Al-khalifa, H. T. Al-natsheh, H. Bouamor, K. Bouzoubaa, V. Cavalli-sforza, and S. R. El-beltagy, "AND HAMDY MUBARAK Survey of Natural Language Processing in the Arab World". doi: 10.48550/arXiv.2011.12631.

[5] S.-Y. Su, C.-W. Huang, and Y.-N. Chen, "Towards Unsupervised Language Understanding and Generation by Joint Dual Learning," pp. 671–680, 2020, doi: 10.18653/v1/2020.acl-main.63.

[6] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: state of the art, current trends and challenges," *Multimed. Tools Appl.*, vol. 82, no. 3, pp. 3713–3744, Jan. 2023, doi: 10.1007/S11042-022-13428-4.

[7] J. Shabbir, "NUBOT : Embedded Knowledge Graph With RASA Framework for Generating Semantic Intents Responses in Roman Urdu". doi: 10.48550/arXiv.2102.10410.

[8] L. Leppänen, M. Munezero, S. Sirén-heikel, and M. Granroth-wilding, "Finding and Expressing News From Structured Data Full Paper", doi: 10.1145/3131085.3131112.

[9] A. Belz and A. Shimorina, "The ReproGen Shared Task on Reproducibility of Human Evaluations in NLG : Overview and Results," no. September, pp. 249–258, 2021. Available at : aclanthology.org.

[10] S. García-méndez, M. Fernández-gavilanes, E. Costa-montenegro, J. Juncal-martínez, F. J. González-castaño, and E. Reiter, "A System for Automatic English Text Expansion," vol. 7, 2019, doi: 10.1109/ACCESS.2019.2937505.

[11] A. L. Review, "Natural Language Processing (NLP) in Management Research A Literature Review Natural Language Processing (NLP) in Management Research : A Literature Review," 2020, doi: 10.1080/23270012.2020.1756939.

[12] D. Zhang, J. Wang, and M. Sun, "Made Towards Human-level AI," vol. 3, pp. 38–47, 2020, doi: 10.23977/jaip.2020.030107.

[13] E. Calvanese Strinati and S. Barbarossa, "6G networks: Beyond Shannon towards semantic and goal-oriented communications," *Comput. Networks*, vol. 190, no. October 2020, p. 107930, 2021, doi: 10.1016/j.comnet.2021.107930.

[14] J. W. Mwakapina, "What is the Nature of Language? How does it Behave? What is Language Learning then? A Review Paper in Applied Linguistics," *Randwick Int. Educ. Linguist. Sci. J.*, vol. 2, no. 4, pp. 603–618, 2021, doi: 10.47175/rielsj.v2i4.357.

[15] N. M. Hewahi and Y. S. A. Khateeb, "Neural Networks Representation for Semantic Networks," *2020 Int. Conf. Innov. Intell. Informatics, Comput. Technol. 3ICT 2020*, 2020, doi: 10.1109/3ICT51146.2020.9311964.

[16] J. Savage *et al.*, "Semantic reasoning in service robots using expert systems," *Rob. Auton. Syst.*, vol. 114, pp. 77–92, 2019, doi: 10.1016/j.robot.2019.01.007.

[17] S. Natale, "If software is narrative: Joseph Weizenbaum, artificial intelligence and the biographies of ELIZA," *New Media Soc.*, vol. 21, no. 3, pp. 712–728, 2019, doi: 10.1177/1461444818804980.

[18] J. Ott, "The Pretense of Knowledge : On the insidious presumptions of Artificial The Pretense of Knowledge : On the insidious presumptions of Artificial Intelligence," no. October, 2020, doi: 10.13140/RG.2.2.31551.46245.

[19] K. Divya, B. Siddhartha, N. Niveditha, and B. Divya, "An Interpretation of Lemmatization and Stemming in Natural Language Processing," *J. Univ. Shanghai Sci. Technol.*, vol. 22, no. 10, pp. 350–357, 2020. Available at : just.org.

[20] P. G. Naik, S. G. Patil, and G. R. Naik, "Natural Language Interface for Querying Hardware and Software Configuration International Journal of Computer Sciences and Engineering Open Access Natural Language Interface for Querying Hardware and Software Configuration of a Local Area Network," no. March, 2019, doi: 10.26438/ijcse/v7i2.949963.

[21] A. Peña-Ayala, "Educational data mining: A survey and a data mining-based analysis of recent works," *Expert Syst. Appl.*, vol. 41, no. 4 PART 1, pp. 1432–1462, 2014, doi: 10.1016/j.eswa.2013.08.042.

[22] D. Litman, "Natural Language Processing for Enhancing Teaching and Learning," pp. 4170–4176, 2013. doi: 10.1609/aaai.v30i1.9879

[23] S. Rani and P. Kumar, "Cover Feature Advances in Learning Technologies," *Adv. Learn. Technol.*, pp. 36–43, 2017. doi : 10.1109/MC.2017.133.

[24] R. Hajrizi and K. P. Nuçi, "Aspect-Based Sentiment Analysis in Education Domain," 2020. doi: 10.48550/arXiv.2010.01429.

[25] G. S. Chauhan, P. Agrawal, and Y. K. Meena, *Aspect-based sentiment analysis of students' feedback to improve teaching–learning process*, vol. 107. Springer Singapore, 2019. doi: 10.1007/978-981-13-1747-7_25.

[26] A. J. Moraes Neto and M. A. Fernandes, "Chatbot and conversational analysis to promote collaborative learning in distance education," *Proc. - IEEE 19th Int. Conf. Adv. Learn. Technol. ICALT 2019*, pp. 324–326, 2019, doi: 10.1109/ICALT.2019.00102.

[27] E. Kasthuri and S. Balaji, "A chatbot for changing lifestyle in education," *Proc. 3rd Int. Conf. Intell. Commun. Technol. Virtual Mob. Networks, ICICV 2021*, no. Icicv, pp. 1317–1322, 2021, doi: 10.1109/ICICV50876.2021.9388633.

[28] A. S. Sreelakshmi, S. B. Abhinaya, A. Nair, and S. Jaya Nirmala, "A Question Answering and Quiz Generation Chatbot for Education," *2019 Grace Hopper Celebr. India, GHCI 2019*, no. November 2019, 2019, doi: 10.1109/GHCI47972.2019.9071832.

[29] I. Orife *et al.*, "Masakhane -- Machine Translation For Africa," pp. 1–4, 2020. Available at : arxiv-vanity.com.

[30] V. Kordoni *et al.*, "PDF hosted at the Radboud Repository of the Radboud University Nijmegen Enhancing Access to Online Education : Quality Machine Translation of MOOC Content Overview of TraMOOC," 2021. Available at : core.ac.uk.

[31] N. S. Khan, A. Abid, and K. Abid, "A Novel Natural Language Processing (NLP)–Based Machine Translation Model for English to Pakistan Sign Language Translation," *Cognit. Comput.*, vol. 12, no. 4, pp. 748–765, 2020, doi: 10.1007/s12559-020-09731-7.

[32] C. Zuheros, E. Martínez-Cámara, E. Herrera-Viedma, and F. Herrera, "Sentiment Analysis based Multi-Person Multi-criteria Decision Making methodology using natural language processing and deep learning for smarter decision aid. Case study of restaurant choice using TripAdvisor reviews," *Inf. Fusion*, vol. 68, no. October 2020, pp. 22–36, 2021, doi: 10.1016/j.inffus.2020.10.019.

[33] F. S. Dolianiti, D. Iakovakis, S. B. Dias, S. Hadjileontiadou, J. A. Diniz, and L. Hadjileontiadis, "Sentiment analysis techniques and applications in education: A survey," *Commun. Comput. Inf. Sci.*, vol. 993, pp. 412–427, 2019, doi: 10.1007/978-3-030-20954-4_31.

[34] S. Caballe, "Towards a Multi-modal Emotion-Awareness e-Learning System," *Proc. - 2015 Int. Conf. Intell. Netw. Collab. Syst. IEEE INCoS 2015*, pp. 280–287, 2015, doi: 10.1109/INCoS.2015.88.

[35] A. Tsang, "The relationship between tertiary-level students' self-perceived presentation delivery and public speaking anxiety: A mixed-methods study," *Assess. Eval. High. Educ.*, vol. 45, no. 7, pp. 1060–1072, 2020, doi: 10.1080/02602938.2020.1718601.

[36] K. Van Nguyen, V. D. Nguyen, P. X. V. Nguyen, T. T. H. Truong, and N. L. T. Nguyen, "UIT-VSFC: Vietnamese Students' Feedback Corpus for Sentiment Analysis," *Proc. 2018 10th Int. Conf. Knowl. Syst. Eng. KSE 2018*, pp. 19–24, 2018, doi: 10.1109/KSE.2018.8573337.

[37] D. Marcu and M. Danubianu, "Sentiment analysis from students' feedback," *15th Int. Conf. Dev. Appl. Syst.*, pp. 204–209, 2020. doi : 10.3389/frai.2021.728708.

[38] Z. Nasim, Q. Rajput, and S. Haider, "Sentiment analysis of student feedback using machine learning and lexicon based approaches," *Int. Conf. Res. Innov. Inf. Syst. ICRIIS*, pp. 1–6, 2017, doi: 10.1109/ICRIIS.2017.8002475.

[39] P. Adinolfi, E. D'Avanzo, M. D. Lytras, I. Novo-Corti, and J. Picatoste, "Sentiment Analysis to Evaluate Teaching Performance," *Int. J. Knowl. Soc. Res.*, vol. 7, no. 4, pp. 86–107, 2016, doi: 10.4018/ijksr.2016100108.

[40] M. Umair and A. Hakim, "Sentiment Analysis of Students' Feedback before and after COVID-19 Pandemic Sentiment analysis of Students Feedback before and after COVID-19 Pandemic View project," no. July, 2021. doi : 10.2139/ssrn.4324732

[41] K. F. Hew, X. Hu, C. Qiao, and Y. Tang, "What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach," *Comput. Educ.*, vol. 145, p. 103724, 2020, doi: 10.1016/j.compedu.2019.103724.

[42] I. Lazrig and S. L. Humpherys, "Using Machine Learning Sentiment Analysis to Evaluate Learning Impact," vol. 20, no. February, pp. 13–21, 2022. doi : 10.47750/pnr.2022.13.04.047

[43] B. Zhang, X. Li, X. Xu, K. C. Leung, Z. Chen, and Y. Ye, "Knowledge Guided Capsule Attention Network for Aspect-based Sentiment Analysis," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 9290, no. 3, pp. 1–13, 2020, doi: 10.1109/TASLP.2020.3017093.

[44] K. N. Dew, A. M. Turner, Y. K. Choi, A. Bosold, and K. Kirchhoff, "Development of machine translation technology for assisting health communication: A systematic review," *J. Biomed. Inform.*, vol. 85, pp. 56–67, 2018, doi: 10.1016/j.jbi.2018.07.018.

[45] J. Z. Abbott and L. Martinus, "Towards Neural Machine Translation for African Languages," 2018. doi : 10.48550/arXiv.1811.05467.

[46] A. Schmidhofer and N. Mair, "Machine Translation in Translator Education," vol. 4, no. 2, pp. 163–180, 2018. doi: 10.14201/clina201842163180

[47] J. Moorkens, "What to expect from Neural Machine Translation: a practical in-class translation evaluation exercise," *Interpret. Transl. Train.*, vol. 12, no. 4, pp. 375–387, 2018, doi: 10.1080/1750399X.2018.1501639.

[48] M. Johnson *et al.*, "Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation," *Trans. Assoc. Comput. Linguist.*, vol. 5, pp. 339–351, 2017, doi: 10.1162/tacl_a_00065.

[49] D. Banik, A. Ekbal, and P. Bhattacharyya, “Machine Learning Based Optimized Pruning Approach for Decoding in Statistical Machine Translation,” *IEEE Access*, vol. 7, pp. 1736–1751, 2019, doi: 10.1109/ACCESS.2018.2883738.

[50] M. Artetxe, G. Labaka, and E. Agirre, “An effective approach to unsupervised machine translation,” *ACL 2019 - 57th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf.*, pp. 194–203, 2020, doi: 10.18653/v1/p19-1019.

[51] C. Park, Y. Yang, K. Park, and H. Lim, “Decoding strategies for improving low-resource machine translation,” *Electron.*, vol. 9, no. 10, pp. 1–15, 2020, doi: 10.3390/electronics9101562.

[52] A. Seker, E. Bandel, D. Bareket, I. Brusilovsky, R. S. Greenfeld, and R. Tsarfaty, “AlephBERT:A Hebrew Large Pre-Trained Language Model to Start-off your Hebrew NLP Application With,” no. 2005, 2021. doi : 10.48550/arXiv.2104.04052.

[53] F. Stahlberg, “Neural Machine Translation: A Review,” *J. Artif. Intell. Res.*, vol. 69, pp. 343–418, Oct. 2020, doi: 10.1613/jair.1.12007.

[54] X. Wang, Z. Lu, Z. Tu, H. Li, D. Xiong, and M. Zhang, “Neural machine translation advised by statistical machine translation,” *31st AAAI Conf. Artif. Intell. AAAI 2017*, pp. 3330–3336, 2017. doi : 10.1609/aaai.v31i1.10975.

[55] M. Lombardi, F. Pascale, and D. Santaniello, “An application for Cultural Heritage using a Chatbot,” *2nd Int. Conf. Comput. Appl. Inf. Secur. ICCAIS 2019*, pp. 1–5, 2019, doi: 10.1109/CAIS.2019.8769525.

[56] M. Nuruzzaman and O. K. Hussain, “IntelliBot: A Dialogue-based chatbot for the insurance industry,” *Knowledge-Based Syst.*, vol. 196, no. xxxx, p. 105810, 2020, doi: 10.1016/j.knosys.2020.105810.

[57] C. Blanc *et al.*, “FlauBERT vs. CamemBERT: Understanding patient’s answers by a French medical chatbot,” *Artif. Intell. Med.*, p. 102264, 2022, doi: 10.1016/j.artmed.2022.102264.

[58] A. K. Kushwaha and A. K. Kar, “MarkBot – A Language Model-Driven Chatbot for Interactive Marketing in Post-Modern World,” *Inf. Syst. Front.*, 2021, doi: 10.1007/s10796-021-10184-y.

[59] R. Pillai and B. Sivathanu, “Adoption of AI-based chatbots for hospitality and tourism,” *Int. J. Contemp. Hosp. Manag.*, vol. 32, no. 10, pp. 3199–3226, 2020, doi: 10.1108/IJCHM-04-2020-0259.

[60] T. Zemčík, “Failure of chatbot Tay was evil, ugliness and uselessness in its nature or do we judge it through cognitive shortcuts and biases?,” *AI Soc.*, vol. 36, no. 1, pp. 361–367, 2021, doi: 10.1007/s00146-020-01053-4.

[61] G. Alexakis, S. Panagiotakis, A. Fragkakis, E. Markakis, and K. Vassilakis, “Control of smart home operations using natural language processing, voice recognition and iot technologies in a multi-tier architecture,” *Designs*, vol. 3, no. 3, pp. 1–18, 2019, doi: 10.3390/designs3030032.

[62] A. Veglis and T. A. Maniou, “Chatbots on the Rise: A New Narrative in Journalism,” *Stud. Media Commun.*, vol. 7, no. 1, p. 1, 2019, doi: 10.11114/smc.v7i1.3986.

[63] S. Singh and H. K. Thakur, “Survey of Various AI Chatbots Based on Technology Used,” *ICRITO 2020 - IEEE 8th Int. Conf. Reliab. Infocom Technol. Optim. (Trends Futur. Dir.*, pp. 1074–1079, 2020, doi: 10.1109/ICRITO48877.2020.9197943.

[64] Q. Zheng, Y. Tang, Y. Liu, W. Liu, and Y. Huang, “UX Research on Conversational Human-AI Interaction: A Literature Review of the ACM Digital Library,” pp. 1–34, 2022. doi : 10.1145/3491102.3501855.

[65] E. E. Program and U. N. Malang, “Web Site:,” vol. 14, no. 2, pp. 179–187, 2021. Available at : urj.uin-malang.ac.id.

[66] S. Studente, S. Ellis, and S. F. Garivaldis, “Exploring the Potential of Chatbots in Higher Education: A Preliminary Study,” *Int. J. Educ. Pedagog. Sci.*, vol. 14, no. 9, pp. 768–771, 2020. Available at : publications.waset.org.

[67] F. O. Chete and G. O. Daudu, “An Approach towards the Development of a Hybrid Chatbot for Handling Students ’ Complaints,” *J. Electr. Eng. Electron. Control Comput. Sci.*, vol. 6, no. 22, pp.

29–38, 2020. Available at: jeeeccs.net.

[68] W. El Hefny, A. El Bolock, C. Herbert, and S. Abdennadher, "Applying the Character-Based Chatbots Generation Framework in Education and Healthcare," *HAI 2021 - Proc. 9th Int. User Model. Adapt. Pers. Human-Agent Interact.*, pp. 121–129, 2021, doi: 10.1145/3472307.3484172.

[69] S. Gupta, Y. Chen, R. C. Gupta, Y. S. Inclusive, and Y. Chen, "Information Systems Education Supporting Inclusive Learning Using Chatbots ? A Chatbot- Led Interview Study Supporting Inclusive Learning Using Chatbots ? A Chatbot- Led Interview Study Sambhav Gupta," vol. 33, no. March, 2022. Available at : aisel.aisnet.org.

[70] M. Deng *et al.*, *Interface Design for Boccia Robot Considering Operation Characteristic*, vol. 11786 LNCS. 2019. doi: 10.1007/978-3-030-30033-3_30.

[71] Z. Kamisli Ozturk, Z. I. Erzurum Cicek, and Z. Ergul, "Sentiment analysis: An application to Anadolu University," *Acta Phys. Pol. A*, vol. 132, no. 3, pp. 753–755, 2017, doi: 10.12693/APhysPolA.132.753.

[72] S. Sharma and A. Jain, "Role of sentiment analysis in social media security and analytics," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 10, no. 5, 2020, doi: 10.1002/widm.1366.

[73] D. Li, R. Rzepka, M. Ptaszynski, and K. Araki, "HEMOS: A novel deep learning-based fine-grained humor detecting method for sentiment analysis of social media," *Inf. Process. Manag.*, vol. 57, no. 6, 2020, doi: 10.1016/j.ipm.2020.102290.

[74] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, "Machine Learning for Fluid Mechanics," *Annu. Rev. Fluid Mech.*, vol. 52, pp. 477–508, 2020, doi: 10.1146/annurev-fluid-010719-060214.

[75] D. Emelin, I. Titov, and R. Sennrich, "Detecting word sense disambiguation biases in machine translation for model-agnostic adversarial attacks," *EMNLP 2020 - 2020 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, pp. 7635–7653, 2020, doi: 10.18653/v1/2020.emnlp-main.616.

[76] L. Nunes Vieira and E. Alonso, "Translating perceptions and managing expectations: an analysis of management and production perspectives on machine translation," *Perspect. Stud. Transl. Theory Pract.*, vol. 28, no. 2, pp. 163–184, 2020, doi: 10.1080/0907676X.2019.1646776.

[77] X. Luo, M. S. Qin, Z. Fang, and Z. Qu, "Artificial Intelligence Coaches for Sales Agents: Caveats and Solutions," *J. Mark.*, vol. 85, no. 2, pp. 14–32, 2021, doi: 10.1177/0022242920956676.

[78] C. H. Hong *et al.*, "Audrey: A Personalized Open-Domain Conversational Bot," pp. 1–20, 2020. doi: 10.48550/arXiv.2011.05910

[79] M. Adam, M. Wessel, and A. Benlian, "AI-based chatbots in customer service and their effects on user compliance," *Electron. Mark.*, vol. 31, no. 2, pp. 427–445, 2021, doi: 10.1007/s12525-020-00414-7.

[80] M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub, Y. Jararweh, and B. Gupta, "Deep recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews," *J. Comput. Sci.*, vol. 27, pp. 386–393, 2018, doi: 10.1016/j.jocs.2017.11.006.

[81] V. G. Djokic, J. Maillard, L. Bulat, and E. Shutova, "Decoding brain activity associated with literal and metaphoric sentence comprehension using distributional semantic models," *Trans. Assoc. Comput. Linguist.*, vol. 8, pp. 231–246, 2020, doi: 10.1162/tacl_a_00307.