Scientific Journal of Impact Factor (SJIF): 4.72
e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 06, June -2017

Influence of Natural Language Processing in Education

Kanika Garg[1], Goonjan Jain[2]

[1]School of Computer and Systems Sciences, Jawaharlal Nehru University, Delhi
[2]Department of Applied Mathematics, Delhi Technological University, Delhi

Abstract - *Natural language is the language understandable by humans in which they can interact. Increasing work in Natural Language Processing and education has led the focus of researchers, teachers and students on effects of Natural Language Processing on Education.Although applications of Natural Language Processing differ with educational requirements but this work tried to focus on Natural Language Processing tools for education to provide better knowledge and insight of the subject and topics to students and help teachers in their non-biased evaluation of students. This work will provide an overview in this area of research with the main focus on its applications and challenges.*

Keywords *– Education, Natural Language Processing (NLP), Understanding Language, Teaching Language, NLP Applications*

I. INTRODUCTION

Natural Language Processing (NLP) is a sub-field of artificial intelligence that deals with processing the human understandable language to make it machine understandable. NLP explores how machines can understand and manipulate natural language text and speech just as humans do. NLP has its roots in psychology, linguistics and artificial intelligence etc. In assisting students in their scientific learning NLP can be an effective process. If NLP is implemented in education it will help in creating an efficient language process which will further improve academic performance of a student significantly.Processing of language acquisition along with a scientific approach to learn and understand natural language lays the main basis for NLP techniques.

A lot of work has been done in the field of learning and education, but a step with NLP will focus on new paradigms and improvements in classical approach towards learning especially in the field of linguistics and science. NLP's major role is in providing assistance and guidance in the path of development. It provides effective and novice approaches towards scientific and educational development. Its approaches can be used to create better learning environment by incorporating machines, internet and Web 2.0. It provides a means to language learning, scientific learning by bringing real life experiences to the students to easily understand things their own natural ways.

Developing such software's that provide better educational settings apart from ICT is the main task of NLP. Training and assistance needs to be provided to make teachers and students understand the basic NLP tools for example, e-rater and Text Adaptor. NLP software tools are able to provide language learning environment in the most natural way.This has been helpful in reading research, writing research and teaching of content knowledge since 1950's and 1960's. While methods were available decades ago but computer technology remained a hindrance. But now, developed technology in easily accessible and available to major chunk of population it is easier to implement NLP.It accepts words, sentences and longer texts as input which mimic the human understanding of natural language.For such purpose, it uses various grammatical rules for different languages, and various resources like lexicons, corpuses, morphological analysers and POS taggers [1]. Linguistic approach is useful in enhancing the readability and text quality. In this area further research has been extended using discourse coherence and predicting grade-level of texts [2][3], [4][5]. The various NLP approaches can be used in education system at every level to make better understanding of study material and curriculum to students. The Writing Pal is an intelligent tutoring system. The artificial intelligent system provides writing strategy training with its main characteristics lies in natural language processing algorithms to assess essay quality and guide feedback to students. The natural language algorithms must undertake various linguistic, contextual and rhetorical features as writing in generally distinguished and highly subjective that differs from a person to person [6]. [7]provided with a CIRCSIM-Tutor which is NLP based intelligent tutoring system. This is a dialogue based system that interact with student to help them learn to solve a class of problems in cardio-vascular physiology dealing with the regulation of blood pressure. This system can handle various syntactical structures and lexical items including sentence fragments and misspelled words.

In technology and business strategies, Natural Language Processing finds variety of applications ranging from text analysis, summarization to sentiment analysis etc. but in education it is widely recognized in linguistics and its effectiveness in language learning.It is mostly implemented in English as it is a widely known language but efforts are being made to make its planning and strategies available in different languages. If we talk about our country, there is a dire of need of such systems that will be available in regional languages and some of these tools are available that are discussed below. Experiments done in English language reveals that in educational process natural language learning is effective. NLP has been tested in various countries and Arab is one example [8].

Role of Natural Language in Education[9]

A. Learning technical language and teaching the same
This is teacher's perspective of learning to read, speak and write in a new language. This is possible through Natural Language Processing.

B. Natural language can be used to teach any discipline or subject
Teaching in various disciplines and subjects would be helpful with Language Processing.

C. Processing language to support the needs of students, teachers, researchers
Natural Language MOOC forms and lecture series are useful for education community. Interactive procedures provide better environment for learning and teaching.

II. UNDERSTANDING LANGUAGE

NLP is used in languageassessment to evaluatespoken or written instances with respect to linguistic dimension(s). *Syntactic* analysis has been used to detect and potentially correct writing errors such as incorrect preposition usage for populations such as ESL (English as Second Language) or deaf students [10], [11].Teachers and students face various problems due to language barriers, in that case, use of various linguistic tools such as grammar, syntax, and textual patterns are very effective for learning and assessment of text.

Various levels of language that needs to be understand by any individual are:

A. *Phonology*: This deals with the interpretation of speech sounds within and across the words. It has rules for sounds within the word, variations in sounds when words are spoken together and for variation and stress across a sentence.

B. *Morphology*: This deals with basic construct of a word which is made up of terms called morphemes. E.g. *unlawfully* is a word made from prefix "un", root word "law" and suffix "fully". NLP systems can infer the meaning of a word using each individual morpheme and can help others to learn about it. Morphological analyser is created by IIT-B for Hindi[1]. English morphological analyser[2] is also available.

C. *Lexical*: Here meaning of individual word is being interpreted. In a sentence, each word has a different label and meaning. To understand this, we have a Part-of-Speech (POS)[3] tagger which will tag individual words in a sentence for part of speech. This can help learners to better understand the sentence construct. POS taggers are available for various languages like Bengali, Hindi, English, Tamil and Telugu.

D. *Syntactic*: This helpsuncover the grammar that has been used in a sentence. This provides a representation of a sentence that shows structural dependency of each individual word upon each other in a sentence. E.g. Krishna runs faster than Ram is totally different from Ram runs faster than Krishna as they are similar but differs in syntactical construct.

E. *Semantic*: This deals with the possible meanings of a sentence by concentrating on the meaning of each individual word in a sentence. E.g Bat flies all over the ground. Here, bat can be an object which is used in cricket or a mammal that can fly. But going with the sentence it seems to be a mammal that can fly all over the ground. This level helps in semantic disambiguation which in itself is a problem in NLP but at preliminary level can be used for some basic examples.

F. *Discourse*: This level deals with more than one sentence or we say, a bunch of sentences. This provides overall meaning of a paragraph comprising of some sentences. Automatic Text Analyser[4] use discourse level to analyse the given text. This can be used to identify the author of the article, main events occurred etc.

G. *Pragmatic*: This level of language requires world knowledge. This is concerned with the constructive use of language by understanding intentions and goals. This may require the use of NLP knowledge bases. E.g. Parents refused children to go on trek as they fearedinjury. Here, *they* is implicit for parents. In another sentence, Parents refused children to go on trek as they are not old enough. In this sentence *they* referred to children.

Various NLP resources are available like language corpora, various linguistic tools and WordNet which helps in effective process of language learning and understanding. It provides deeper insights of language to understand every aspect of it.

[1] http://www.cfilt.iitb.ac.in/~ankitb/ma/
[2] http://nlpdotnet.com/services/Morphparser.aspx
[3] https://nlp.stanford.edu/software/tagger.shtml
[4] https://www.online-utility.org/text/analyzer.jsp

Corpora provide a large amount of written or spoken data which is active and can be used for experimental computations. Text-to-speech and Speech processing are also a task of NLP that can be very useful in education for regional areas. Students can learn and understand in their own language. For example, Hindi WordNet[5] provides a way to define lexical and syntactical relationship among different Hindi words and provides information about the vocabulary usage. The large collection of information provides sufficient data regarding the usage of words, which assist enhancing the information and academic skills of the students. Some NLP tools available are POS Tagger and Stemmer.

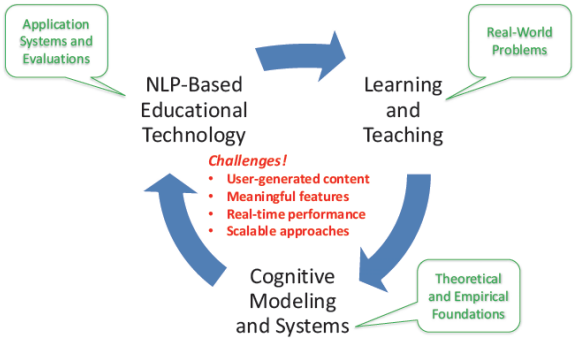

Figure 1. Education Research lifecycle using NLP [9]

III. TEACHING LANGUAGE

If we define ways of studying, typically there can be 3 ways: one-on-one study with human tutors, group study in a class with one human tutor and computer assisted study. It has been seen that in case 1 students generally score higher than case 2 students, with both typically scoring higher than case 3 students [9]. Human tutors has an advantage over automated tutors that humans can participate in natural language dialogues with their students and can go up to any extent by quoting different plots and exampleswhich has led to a belief that human tutors are more effective mainly because of theircommunication. So, to minimize the performance gap between human tutoring and computer tutoring, dialogue-based intelligent tutoring systems have been introduced and become ubiquitous. Further, making such systems socially aware and socially intelligent will make teaching easy and more realistic. It enables such systems to understand student's emotional state of mind and will be able to counsel them and make them understand.

Tutoring generally differs depending upon its application area especially for which spoken intelligent dialogue systems have been developed. For example, tutorial dialogue versus airline information systems are evaluated on different criteria. E.g. airline information system need shorter dialogues rather than longer ones. There is a difference in emotional states of humans in airline systems such as anger, but in tutorial it will be boredom and differs in types of user goals such as learn physical chemistry rather than find a flight from Delhi to Bangalore.
Dialogue tutors and computer tutors are similar with respect to tutoring research where main concentration is on knowledge domain and its correctness regarding a subject e.g., biology (Evens and Michael 2006), computer science (Boyer et al. 2008), electricity and electronics (Dzikovska et al. 2010), physics (VanLehn et al. 2007), and elementary school science (Ward et al. 2011). Natural language instructions can be added to quantitative problem-solving tutors so they can help in better conceptual learning by using dialogues with students to impart conceptual knowledge directly into them. Some of these systems allow students to speak rather than type their answers, which also supports handsfree conversation that could be useful during lab work. Tutorial dialogue technology is just starting to be applied to more ill-defined domains than STEM, e.g., teaching second language learners how to chat.

NLP applications in Education

A. E – Rater

It is automatic essay rater [12]that examined lexical, syntactic and discourse-related text features and Intelligent Essay Asessor analyzed content through vocabulary usage with latent semantic analysis.

B. Text Adaptor

This is basically introduced for English language learners [12]. It helps in modification of text to better understand difficult language or vocabulary. It supports elaboration of text, text summarization etc.

[5] http://www.cfilt.iitb.ac.in/wordnet/webhwn/

C. Intelligent tutoring systems

Intelligent tutoring systems e.g. [6] are associated with support and evaluation of content knowledge acquisition [13]–[15]. The goal of intelligent tutoring is to help students work through problem sets in various domains (e.g, physics). This is another area that has involved increasingly more NLP for the purpose of evaluating students' responses as they work through problem sets in a subject area. Intelligent tutoring applications make use of systems that use propositional information in responses to identify correct knowledge, given a particular problem [12].

D. Finding grammatical errors

NLP provides tools with minimal error for some Indian languages that can be used as assessment for students in a learning phase. Teachers and students are at ease with this NLP tool that can be used for learning as well as evaluation. Students can check their text for any grammatical mistake and can easily correct them. This tool can be created for many other regional languages so as to benefit the students where teachers are scarcely available.

E. Autocorrecting grammar

NLP tools are available to autocorrect the sentences having grammatical mistakes. If a student is hesitant to publish its work or to write to someone can easily use this tool.Some online services are available like ginger[6], grammarly[7], bhashagiri[8] for Hindi etc. to auto correct grammar.

IV. CHALLENGES

A. Equal access to learning for all:

Institute has to ensure equal access of such NLP systems to all. Particularly to disable students. There must be some sizeable students who are struggling with their studies and need a help. This must be given to all not only to the one who can afford costly things rather it should be a part of institution.

B. Lack of Instructional Design methodology:

Projects related to NLP require instructional designer. They are the subject expert who can author the content and teaches how content should be written and understood.

C. Time and effort constraints:

These things need efforts and time for understanding them and having a hands-on. Generally, teachers are busy with their teaching and organizing schedule. They have a lot of work apart from teaching like checking answer sheets, maintaining student's records,managing and creating forms, quizzes etc. for brainstorming, making question paper, interacting with students, conducting seminars creating assignments etc. So, within a limited time frame, they need to learn paradigm first and then prepare the course content according to the instructions and the various assessment documentations as per the formats compatible to NLP tool.

D. Lack of Motivation:

The engagement of both faculty and student engagement is closely linked to the quality of the learning and both of them can lose motivation if the quality is skipped. Faculty need to be motivated with financial and non-financial motivational strategies in order to maintain a positive attitude.

E. Broadband Speed:

Developing countries like India, education system is striving with different problems. It may be improper infrastructure, imprecise skilled manpower, and higher level of corruption. As a developing economy, India is far behind the other developed countries in engineering education and research because teachers are resistant towards drift in the education system, there's a little technical support for handling such expensive hardware.

F. Pedagogical Issues:

Teachers can bring a potential change into the education system of any country by integrating technology with their wisdom. Teachers should possess knowledge about relevant NLP tools, and how to use them in their teaching process effectively.

ACKNOWLEDGEMENT

This work was supported in part by the Council of Scientific and Industrial Research (CSIR), New Delhi with file no. 09/263(1049)/2015-EMR-I.

[6] http://www.gingersoftware.com/
[7] https://www.grammarly.com/
[8] https://bhashagiri.com/

V. CONCLUSION

In this work, NLP tools have been studied for their applications in education. NLP can be helpful in learning new language from scratch, gaining technical knowledge in some different language, communicate across different platforms to different countries with different accent and language. It not only helps in learning new language but also in linguistic evaluation, teaching and testing already lerant things. Many NLP tools have already been developed for the research community that can be used and tested with education community. Till this date, only intrinsic evaluation has been undertaken for NLP tools in education and there is a need of extrinsic evaluation of such scheme at a larger level to actually understand its effectiveness. Although extrinsic studies are also present but their number is really small. For this we needed technological parameters in schools and educational institutions. With the growing research in NLP, we will soon be able to get educational systems which are tolerant to NLP errors and can be generalized across students, teachers, courses, schools, etc., and for fostering richer collaborations between NLP and educational researchers.

REFERENCES

[1] R. Flesch, "A new readability yardstick.," *J. Appl. Psychol.*, vol. 32, no. 3, pp. 221–233, 1948.
[2] W. H. MacGinitie and R. Tretiak, "Measures of sentence complexity as predictors of the difficulty of reading materials.," in *Proceedings of the 77th Annual Convention of the APA*, 1969.
[3] S. E. Schwarm and M. Ostendorf, "Reading Level Assessment Using Support Vector Machines and Statistical Language Models," no. June, pp. 523–530, 2005.
[4] P. Deane, K. M. Sheehan, J. Sabatini, Y. Futagi, and I. Kostin, "Structure, Differences in text Readers, and its implications for assessment of struggling," *Cient. Stud. Read.*, vol. 10, no. 3, pp. 257–275, 2006.
[5] R. Barzilay and M. Lapata, "Modeling local coherence: An entity-based approach," *Comput. Linguist.*, vol. 34, no. 1, pp. 1–34, 2008.
[6] D. S. Mcnamara, S. A. Crossley, and R. Roscoe, "Natural language processing in an intelligent writing strategy tutoring system," pp. 499–515, 2013.
[7] C. W. W. S, J. A. Michael, and A. A. Rovick, "CIRCSIM-Tutor: An Intelligent Tutoring System Using Natural Language Dialogue*," pp. 13–14.
[8] N. Y. Habash, "Introduction to Arabic natural language processing," *Synth. Lect. Hum. Lang. Technol.*, vol. 3, no. 1, pp. 1–187, 2010.
[9] D. Litman, "Natural Language Processing for Enhancing Teaching and Learning," pp. 4170–4176, 2013.
[10] L. Michaud and K. McCoy, "Capturing the Evolution of Grammatical Knowledge in a CALL System for Deaf Learners of English," in *Proceedings of IJAIED*, 2006, pp. 65–97.
[11] J. Tetreault and M. Chodorow, "The ups and downs of preposition error detection in ESL writing," in *In Proceeding of COLING 2008*, 2008.
[12] J. Burstein, "Opportunities for Natural Language Processing Research in Education," pp. 6–27, 2009.
[13] J. S. Brown, R. R. Burton, and A. G. Bell, "SOPHIE: A sophisticated instruction environment for teaching electronic troubleshooting (An example of AI in CAI).," *Int. J. Man-machine Stud.*, vol. 7, no. 5, pp. 675–696, 1975.
[14] J. . Carbonell, "AI in CAI: An artificial intelligence approach to computer-assisted instruction," *IEEE Trans. Man-Machine Syst.*, vol. 11, no. 4, pp. 190–202, 1970.
[15] Burton, R. Richard, and J. S. Brown, "An investigation of computer coaching for informal Activities.," *Int. J. Man. Mach. Stud.*, vol. 11, no. 1, pp. 5–24, 1979.

