Turkish Journal of Computer and Mathematics Education Vol.12 No.13 (2021), 1872-1875

Research Article

Identification of Bloom's Taxonomy level for the given Question paper using NLP Tokenization technique

Dr. G. N. R. Prasad

Sr. Asst. professor,Dept. of MCA, CBIT (A),Gandipet, Hyderabad.

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 4 June 2021

ABSTRACT

Bloom's developed taxonomy level for the questions' to identify the cognitive level of the student. An automation process is required to do the process in a faster manner. Natural Language Processing (NLP) is one of the areas of programming where software processes natural language. Sentiment analysis, language translation, fake news detection, and grammatical error detection are only a few of the applications.Tokenization is the process of splitting down large blocks of text into smaller ones. Tokenization divides the original text into tokens, which are words and sentences. These tokens aid in the comprehension of the context or the development of the NLP model. Through evaluating the series of sentences, tokenization aids in reading the context of the language. Tokenization can be done using a variety of techniques and databases. Any of the libraries that can be used to complete the challenge are NLTK, Gensim, and Keras.This paper provides the solution to the manual process in the identification of bloom's taxonomy levels.

Keywords : Natural language processing, Bloom's Taxonomy; Gensim, Keras.

1.0 INTRODUCTION

Blooms taxonomy is set of three hierarchical models. It is used to categorise the aims of the new educational system in terms of sophistication and detail in order to foster higher ways of thought in engineering education.And also to meet industry's expectations in engineering profession.

Benjamin Bloom developed and suggested the Bloom Taxonomy scheme in 1956. It was a structure for identifying and categorising instructional priorities and targets. [1]. In 2000-01, David Krathwohl, one of the initial team members, and Lorin Anderson, a Bloom student, spearheaded an attempt to update the original cognitive taxonomy. [2]. Bloom's taxonomy is a six-step classification scheme that infers the level of cognitive domain accomplishment based on demonstrated student activity.

2.0 REVIEW OF LITERATURE

Tokenization is the method of dividing sentences into still smaller pieces known as 'tokens.' Natural Language Processing (NLP) is a computer science discipline that involves computer linguistics and artificial intelligence, as well as the integration of human natural languages with computers. Computers are designed to interpret natural language using NLP. Tokenizing data literally involves dividing the text's body. Python text strings are translated to streams of token objects in this method.It should be noted that each token is a separate letter, number, email, punctuation mark, URL/URI, and so on. Tokens are often segmented into streams of sentences with dates and abbreviations in the centre of the sentences. Used NLTK Regular Expressions A regular expression is essentially a character series that assists us in searching for matching patterns in the text that we have. Python's regular expression library is called re, and it comes pre-installed with the Python kit.In most languages, text is composed of words divided by whitespace, where individual words have a semantic meaning. As an example:

- Raw text: What is the definition of Software Engineering.
- Tokenized text: ['What', 'is', 'the', 'definition', 'of', 'software', 'Engineering']

3. BLOOM'S COGNITIVE DOMAIN

Bloom's Taxonomy's cognitive domain (Bloom, 1956) is one of Benjamin Bloom's three domains proposed in the 1950s. During a written exam, this domain is used to evaluate a student's cognitive performance.The famous Bloom's taxonomy consists of six levels i.e. knowledge, comprehension, application,analysis, synthesis and evaluation (Bloom, 1956).

The stages of Bloom's Taxonomy are defined as follows:

Remember	Understand	Apply	Analyze	Evaluate	Create
Describe	Compute	Demonstrate	Characterize	Appraise	Categorize

Enumerate	Explain	Determine	Correlate	Conclude	Collaborate
Identify	Extrapolate	Discover	Dissect	Determine	Compile
Outline	Generalize	Dramatize	Maximize	Interpret	Formulate
Recall	Paraphrase	Express	Prioritize	Predict	Rearrange
Recognize	Restate	Investigate	Relate	Reframe	Reconstruct
Reorganize	Summarize	Predict	Subdivide	Support	Reorganize
			Synthesize	Validate	Role-play

Table: Cognitive Domain verbs in the Bloom's Taxonomy

4. RESEARCH METHODOLOGY

A rule-based approach is used in this work to group the query objects into their respective Bloom's cognitive level. The research products are a list of review questions in the subject of Software Engineering. The training package includes 70 exam questions. With the help of the established laws, the method would automatically classify each question to the corresponding verbs from the Taxonomy.

5. RULES DEVELOPMENT

There are two conditions where the rules will be applied:

- The rules will determine which keyword is appropriate for each query based on its category.
- If a keyword belongs to more than one subclass, this tool will assist you in selecting the right one.

For example, Summarizemay fall under Evaluatingor creating category.

The flow of the tokenization process is depicted in the following picture.

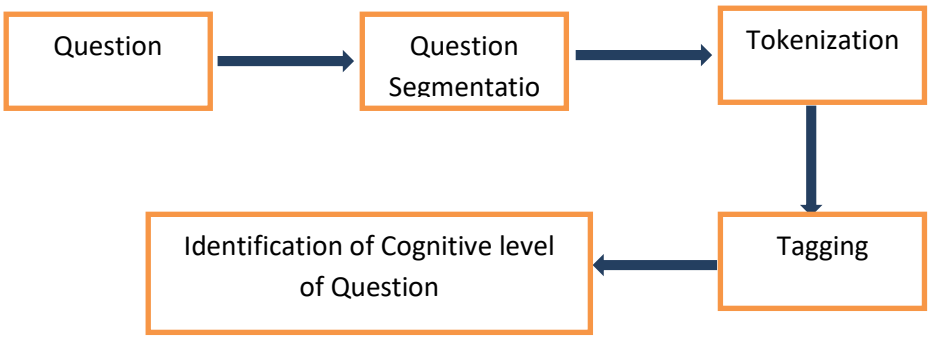

Fig : Flow of the tokenization process

6.0 IMPLEMENTATION

Following is the simple code stub to tokenize the sentence in Python:

```
import nltk.tokenize as nt1
import nltk
Qn="What are the various data mining techniques."
ss1=nt.sent_tokenize(Qn)
token_sent=[nt1.word_tokenize(sent) for sent in ss1]
import re
text_to_search = " How can we find the corresponding strings in this string using the regex library? "
pattern1 = re.compile('\w+')
matches1 = pattern1.finditer(text_to_search)
for match1 in matches1:
```

```
print(match1)

str_a =("define, choose find,how, recall,relate,select,show,spell,tell,what,when,where,which,who,why,
label,list,match,name,omit ")

str_b=("extend,illustrate,infer,interpret,outline,relate,rephrase,show,summarize,translate,
classify,compare,contrast,demonstrate,explain ")

str_c=("apply,build,choose,construct,develop,experimentwith,identify,interview,makeuseof,model,organize,plan,
select,solve,utilize")

str_d=("analyze,assume,categorize,classify,compare,conclusion,contrast,discover,dissect,distinguish,divide,exam
ine,function,inference,inspect,list,motive,relationships,simplify,survey,takepartin,testfor,theme")

str_e=("agree,appraise,assess,award,choose,compare,conclude,criteria,criticize,decide,deuct,defend,determine,di
sprove,estimate,evaluate,explain,importance,influence,interpret,judge,justify,mark,measure,opinion,perceive,prio
ritize,prove,rate,recommend,ruleon,select,support,value")

str_f=("adapt,build,change,combine,compile,compose,construct,create,delete,design,develop,discuss,elaborate,es
timate,formulate,happen,imagine,improve,invent,makeup,maximize,minimize,modify,original,orginate,plan,pred
ict,propose,solution,solve,suppose,test,theory")

str_g=input("Enter Question")
switch(str_g==str_a||str_b||str_c||str_d||str_e||str_f)
{
 case 1:
    print("Remembering category")
 case 2:
   print("Understanding category")
 case 3:
   print("Applying category")
 case 4:
   print("Analyzing category")
 case 5:
    print("Evaluating category")
 case 6:
   print("Creating category")
 default:
   print("The said question does not fall in this categories anymore")
}
```

7.0 CONCLUSION

Tokenization is a very important data pre-processing step in NLP and involves breaking down of a text into smaller chunks called tokens. These tokens can be individual words, sentences or characters in the original text. Tokenization helps in breaking the raw text into small chunks. Through evaluating the series of sentences, tokenization aids in reading the context of the language.Knowledge and development of intellectual skills is at the heart of the cognitive domain of Bloom's taxonomy, The NLP is one of the areas of programming where software processes natural language. This paper given a perfect solution in identifying the cognitive domain level of the given question paper. This allows the trainer to know the exact level of cognitive domain level.

References

1. Forehand, M. (2011). Blooms Taxonomy. Emerging perspectives on learning, Teaching and Technology, 1-10.
2. S. IlangoSivaraman and Dinesh Krishna (2015), "Blooms Taxonomy– Application in Exam Papers Assessment" International journal of multidisciplinary sciences and engineering, VOL. 6, NO. 9.
3. Bird, S., Klein, E., &Loper, E. (2009). Natural Language Processing with Python, O'Reilly Media, Inc., 1005 Gravenstein Highway North,Sebastopol, CA 95472.USA.
4. Bloom, B.S. Taxonomy of Educational Objectives – Handbook 1 Cognitive Domain, London.Longman. (1956).

5. Cutrone, L., & Chang, M. (2010). Automarking: Automatic Assessment of Open Questions, 2010 10th IEEE International Conference onAdvance Learning Technologies. 143-147.
6. Johnson C.G. & U. Fuller. (2006). Is Bloom's Taxonomy is Appropriate for Computer Science, Koli Calling 2006, Koli National Park, Finland. 115-118.
7. Jones, K. O., Harland, J., Reid, J. M.V., Barlett, R. (2009) . Relationship Between Examination Questions and Bloom's Taxonomy, 39th
8. ASEE/IEEE Frontiers in Education Conference. October 18-21, 2009. San Antonio, TX. W1G-1 – W1G-6.
9. Lister, R., &Leaney, J. (2003). Introductory Programming, Criterion-referencing and bloom. SIGCSE '03: Proceeding of the 34th SIGCSETechnical Symposium on Computer Science Education, ACM Press. 143-147.
10. Oliver, D., Dobele, T., Greber, M., & Roberts, T. (2004). This course has a Bloom rating of 3.9, Proceedings of the sixth conference on
11. Australasian computing education, Australasian Computing Society Inc– Volume 30, Dunedin, New Zealand, 227-231.
12. Ranganathan, P., &Nygard, K. A. (2010). Bloom's Online Assessment Test (BOAT) to assess student learning outcome in a distance
13. engineering education course. 2010 2nd International Conference on Computer Engineering and Technology.159-161.
14. Scott, T. (2003). Bloom's Taxonomy Applied to Testing in Computer Science Classes. Consortium for Computing Science in Colleges: RockyMountain Conference. (October 2003) 267-274.
15. Starr, C.W., Manaris, B., &Stalvey, R.H. (2008). Bloom's Taxonomy Revisited: Specifying Assessable Learning Objectives in ComputerScience. SIGCSE '08. Portland, Oregon, USA. March 12-15, 261-265.
16. Swart, A.J. (2010). Evaluation of Final Examination Papers in Engineering: A Case Study Using Bloom's Taxonomy, IEEE Transactions onEducation, (May 2010) Vol. 53, No.2 257-264.
17. Thompson, E., Luxton-Reilly, A., Whalley, J. L. Hu, M., P. Robbins. (2008). Bloom's Taxonomy for CS Assessment.Proceeding Tenth
18. Australasian Computing Education Conference (ACE 2008), Wollongong, Australia.155-162.
19. Wen-Chih Chang, Ming-Shun Chung. (2009). Automatic Applying Bloom's Taxonomy to Classify and Analysis the Cognition Level of EnglishQuestion Items.IEEE.727-733.
20. Yusof, N., & Chai, J. H. (2010). Determination of Bloom's Cognitive Level of Question Items using Artificial Neural Network.2010 10thInternational Conference on Intelligent Systems Design and Applications.866-870.
21. Subramaniyan, M., Sampathkumar, A., Jain, D. K., Ramachandran, M., Patan, R., & Kumar, A. (2021). Deep Learning Approach Using 3D-ImpCNN Classification for Coronavirus Disease. Artificial Intelligence and Machine Learning for COVID-19, 141-152.