International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 9 Issue 6 ǁ 2021 ǁ PP. 42-46

Generate and Ranking the Multiple Sentences Based On Context Using Natural Language Generation

Dr.T.M Kiran Kumar[1], LeenaRani A[2*]
[1]*Assistant Professor, Department of MCA, Siddaganga Institute Of Technology, Tumkur*
[2]*Project Student, Department of MCA, Siddaganga Institute Of Technology, Tumkur*

Abstract
In the increasingly popular and diverse research area of Natural Language Generation (NLG), the primary goal is to generate meaningful multiple sentences using unstructured data as words from the neo4j database and ranking based on context. An important aspect of ranking the sentences to choose or select the better meaningful sentences as the response for the chatbot like human communication. While this examination region is likewise a functioning one overall most exertion in NLG is centered around creating great composed content. With the development of one type of module of NLG recurrent neural networks (RNN) various natural language generation tasks have boomed in the past few years such as response generation in conversation and sentence generation. It is very much difficult for human beings to do summarization or create multiple sentences and select the best one of unstructured data manually. This paper aims to present an innovative approach for automatically generating multiple sentences and ranking.

Keywords: *Artificial Intelligence, Automatic Text Generation, Chatbot, Natural Language Generation, Recurrent Neural Networks.*

Date of Submission: 25-05-2021 Date of acceptance: 07-06-2021

I. INTRODUCTION

NLG is usually viewed as an overall term that incorporates a wide scope of errands that take a type of information (e.g., organized info like a dataset or a table, a characteristic language brief, or even a picture) and yield an arrangement of text that is intelligible and justifiable by people. Henceforth, the field of NLG can be applied to an expansive scope of NLP undertakings, for example, producing reactions to client inquiries in a chatbot, interpreting a sentence or a record from one language into another, presenting ideas to help compose a story, or creating outlines of time-concentrated information examination.

The NLG model yield is testing primarily in light of the fact that numerous NLG errands are open-finished. For instance, a discourse framework can create different conceivable reactions for similar client input. A report can be summed up in an unexpected way. Late headways in profound learning have yielded enormous upgrades in numerous NLP assignments. This, thus, presents a requirement for assessing these profound neural organizations (DNN) models for NLG.

NLG—is the sub-space of computational etymology that arrangements with the mechanized creation of excellent spoken or composed substance in human dialects. While the yield of a NLG framework is text, the information can take different structures; now and again, the framework may create text-dependent on other, by and large, human-composed content: applications for this incorporate robotized machine interpretation, outline, and producing rearranged or Reworded adaptations of its information. In different cases, the contribution to the NLG framework is non-etymological, for instance, sports reports, information from natural sensors, or climate, monetary or clinical information. Later applications have included consequently producing text dependent on visual information like pictures or video.

NLG might be utilized not exclusively to produce independent writings, yet in addition to creating phonetic yield to be utilized in an intuitive framework, either in a content-based chatbot, an expressed exchange framework, or an intelligent robot.

The task of NLG has been divided into the 3 sub-tasks - Document planning, Micro planning, Realisation.

The rest of this paper is organized in an accompanying way. Section II: Literature Survey, Section III: Approach, Section IV: Implementation, Section V: Results and Section VI: Conclusion.

II. LITERATURE SURVEY

We have referred some of the papers to understand the uses of NLG module in different way to generate the meaningful sentences and way to use it. Some of them are listed below:-

II-A

Target start to finish information-driven NLG, where we think about 3 unique methodologies. As opposed to NLG techniques assessed in past work, our frameworks can create ungrammatical yield by (a) producing word-by-word, and (b) gaining from uproarious information [1].

II-B

This paper portrays the essential framework put together by the creator of the E2E NLG Challenge on the E2E Dataset Based on the benchmark framework called TGen, the essential framework utilizes REINFORCE to use various references for single Meaning Representation during preparation, while the standard model regarded them as individual preparing instance[2].

II-C

Study to comprehend the issues identified with utilizing NLG to refine clarifications from a famous interpretable AI system called LIME [3].

II-D

The potential for affecting true conduct through naturally created messages has not gotten a lot of consideration. Utilizes NLG and telematics information to make week after week literary criticism [4].

II-E

Narrating is a necessary piece of day-by-day life and a critical piece of how we share data and associate with others. The capacity to utilize Natural Language Generation (NLG) to create stories that are custom-fitted and adjusted to the individual peruse could have a huge effect on a wide range of utilizations. [5].

III. APPROACH

TYPES OF LANGUAGE MODELS

- Markov chain
- LSTM
- RNN
- Transformer

Markov chain

The Markov chain was one of the primary calculations utilized for language age. This model predicts the following word in the sentence by utilizing the current word and thinking about the connection between every special word to figure the likelihood of the following word. You have seen them a ton in prior adaptations of the cell phone console where they were utilized to produce ideas for the following word in the sentence.

Repetitive neural organization (RNN)

Neural organizations are models that attempt to emulate the activity of the human cerebrum. RNNs pass everything of the grouping through a feed-forward organization and utilize the yield of the model as a contribution to the following thing in the arrangement, permitting the data in the past advance to be put away. In every cycle, the model stores the past words experienced in its memory and ascertains the likelihood of the following word. For each word in the word reference, the model relegates a likelihood dependent on the past word, chooses the word with the most elevated likelihood, and stores it in memory.

LSTM

To address the issue of long-range conditions, a variation of RNN called Long transient memory (LSTM) was presented. Even though like RNN, LSTM models incorporate a four-layer neural organization. The LSTM comprises four sections: the unit, the information entryway, the yield entryway, and the failed to remember entryway. These permit the RNN to recall or neglect words whenever span by changing the data stream of the unit. At the point when a period is experienced, the Forgotten Gate perceives that the setting of the sentence may change and can overlook the current unit state data. This permits the organization to specifically follow just significant data while additionally limiting the vanishing slope issue, which permits the model to recollect data over a more extended period.

LSTM memory is restricted to a couple of hundred words because of their inalienably unpredictable successive ways from the past unit to the current unit. A similar intricacy brings about high computational necessities that make LSTM hard to prepare or parallelize.

Transformer

A moderately new model has first presented in the 2017 Google paper "Consideration is all you need", which proposes another technique called "self-consideration system." The Transformer comprises a pile of encoders for preparing contributions of any length and another arrangement of decoders to yield the produced sentences. Rather than LSTM, the Transformer performs just a little, steady number of steps, while applying a self-consideration instrument that straightforwardly animates the connection between all words in a sentence. In contrast to past models, the Transformer utilizes the portrayal of all words in setting without packing all the data into a solitary fixed-length portrayal that permits the framework to deal with longer sentences without the soaring of computational prerequisites.

IV. IMPLEMENTATION

The project was implemented using markov chain model of NLG and random module to generate random sentences by extract input from graph database or unstructured dataset.

NLG, a subfield of Artificial Intelligence (AI), is a productive interaction that naturally changes information into a plain-English substance. The innovation can recount a story – precisely like that of a human investigator – by composing the sentences and sections.

NLG is identified with computational phonetics, Natural Language Processing (NLP), and Natural Language understanding (NLU), the spaces of AI worried about human-to-machine and machine-to-human cooperation.

- NLP is when PCs read and transform input text into organized information.
- NLU implies comprehension of the printed/factual information caught by PCs.
- NLG is when PCs transform organized information into a message and compose data in human language.

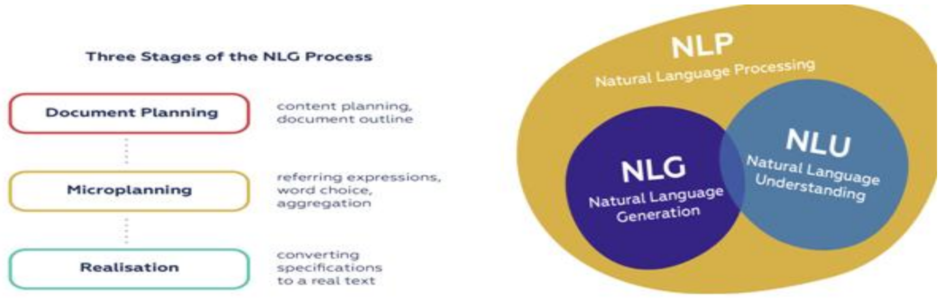


Fig.1 NLG process diagram

Fig.1 explains the process of NLG as three stages document planning, micro-planning, and realization.

In documentation planning, it has concentrated on planning and document the outline. The micro-planning stage it is referring the expression of word choice and aggregation. In the realization stage, it converting specification to a real text.

This section provides an overview of sentences generated randomly and also using the Markov chain method.

IV - A. Randomly sentence generated:-

By using Anaconda software python code to create a sentence by using a module random accept the input from the neo4j - unstructured database edges as verbs and nodes as a noun and we create a list using Python features after creating the list we accept the input from the list and create sentences randomly. Extract data from the graph database and converted that the extracted things into a list and then extract the words from that particular list to make it a noun list verb by identifying nouns and verbs to make a random sentence by using a random module.

IV - B. Markov chain:-

A Markov chain is a describing a sequence of possible words in which the probability of each word or event depends only on the state attained in the previous word. Chain moves in 2 ways - discrete-time Markov chain (DTMC), continuous-time Markov chain (CTMC).

The Markov chain was one of the primary calculations utilized for language age. This model predicts the following word in the sentence by utilizing the current word and thinking about the connection between every special word to figure the likelihood of the following word. You have seen them a ton in prior adaptations of the cell phone console where they were utilized to produce ideas for the following word in the sentence.

V. PSEUDO CODE

Author would try to give sample code snippet for random sentence generation using methods and random module.

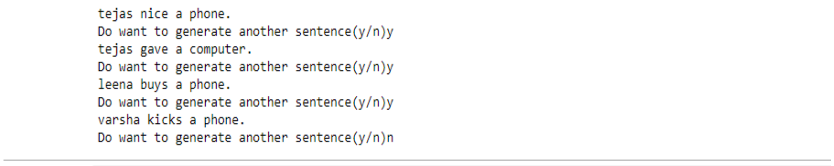
```
#randomly sentence generation
from random import randint
name = ['leena','tejas','varsha','govinda','sumanth','sangeetha']
verb = ['buys','kicks','codes','rides','gave','saw','nice']
noun = ['computer','tv','phone','bike']
ch="y"
def select(word):
   words=leyn(word)
   picked =randint(0,num_words-1)
   picked=word[picked]
   return picked
while(ch=="y"):
   print(select(name), select(verb),'a', select(noun),end='.\n')
   ch=input("Do want to generate another sentence(y/n)")
```

VI. SIMULATION RESULTS

Fig2 and Fig3 shows the result of randomly sentence generation one as with loop condition another as menu driven based – until you press 'n'.

```
her Computer treated Our Man
his Fore ran her Computer
her Computer ran her Edge
his Ball coded The Heart
her Heart polished Our Edge
Their Computer treated her Monkey
Our Edge healed Their Heart
her Man Hit Our Edge
Their Ball healed her Monkey
The Monkey treated Our Computer
```

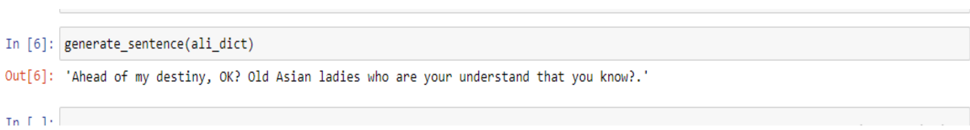
Fig.2 Output diagram of randomly sentence generation



```
tejas nice a phone.
Do want to generate another sentence(y/n)y
tejas gave a computer.
Do want to generate another sentence(y/n)y
leena buys a phone.
Do want to generate another sentence(y/n)y
varsha kicks a phone.
Do want to generate another sentence(y/n)n
```

Fig.3 Output diagram of randomly sentence generation

Fig4 shows the result of sentence generation using markov chain method and for train the model using pandas package in python.



```
In [6]: generate_sentence(ali_dict)
Out[6]: 'Ahead of my destiny, OK? Old Asian ladies who are your understand that you know?.'
```

Fig.4 Output diagram of sentence generation using markov chain approach

VII. CONCLUSION

In this paper author made an attempt to generate sentences using unstructured data from graph database and ranking based on context.Natural language generation are now commonly used for sentence generation or response generation for chatbot.NLG continues to evolve, it will become more diversified and will provide effective communication between us and computers in a natural way. As well as significant research from the related area of spoken dialogue systems that using natural language generation on the output side can also have a significant effect on use a subjective opinion of the system. We just plan for future enhancement more training the model based on particular situations or context for generating meaningful sentences and ranking based on the context.NLG coupled with NLP are the core of chatbot and other automated chats and assistants that provide us with everyday support.

REFERENCES

[1]. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 2231-2242, Copenhagen, Denmark, September 7-11, 2017.https://arxiv.org/abs/1707.06875

[2]. H Gong - E2E **NLG** Challenge System Descriptions, 2018 - macs.hw.ac.uk

[3]. Forrest , J , Sripada , S , Pang , W & Coghill , G 2018 , Towards making NLG a voice for interpretable Machine Learning . in E Krahmer , A Gatt & M Goudbeek (eds) , Proceedings of The 11th International Natural Language Generation Conference . , W18-6522 , Association for Computational Linguistics (ACL) , pp. 177-182 , 11th International Conference on Natural Language Generation (INLG 2018) , Tilburg , Netherlands , 5/11/18.

[4]. Braun, Daniel; Reiter, Ehud and Siddharthan, Advaith (2018). SaferDrive: an NLG-based Behaviour Change Support System for Drivers. *Natural Language Engineering*, 24(4) pp. 551–588.

[5]. Stephanie M. Lukin, Marilyn A. Walker (**2019-02-22**). A Narrative Sentence Planner and Structurer for Domain Independent, Parameterizable Storytelling**.** 10 NO 1 (2019).