COMPARISON OF SEQUENCE MODELS FOR TEXT NARRATION FROM TABULAR DATA

Mayank Lohani, Rohan Dasari, Praveen Thenraj Gunasekaran, Selvakuberan Karuppasamy and Subhashini Lakshminarayanan

Data and AI, Advance Technology Centers in India,
Accenture, Gurugram, India

ABSTRACT

This paper demonstrates our work on the survey of pre-trained transformer models for text narration from tabular data. Understanding the meaning of data from tables or any other data source requires human effort and time to interpret the content. In this era of internet where data is exponentially growing and massive improvement in technology, we propose an NLP (Natural Language Processing) based approach where we can generate the meaningful text from the table without the human intervention. In this paper we propose transformer-based models with the goal to generate natural human interpretable language text generated from the input tables. We propose transformer based pre-trained model that is trained with structured and context rich tables and their respective summaries. We present comprehensive comparison between different transformer-based models and conclude with mentioning key points and future research roadmap.

KEYWORDS

Survey, NLP (Natural Language Processing), Transformers, Table to Text.

1. INTRODUCTION

In recent years, Natural language processing (NLP) is massively growing field and being used in wide range of applications and features. Table to text generation is a segment, which aims at generating meaningful and descriptive text about defined information in structured data. Few applications in this segment include generating sentences given medical tabular data, descriptions of restaurant menus given meaningful representations, summaries from cricket/football game score tables, generating meaningful text from tables in Wikipedia, converting statistical stock market tables to easily understandable textual form, etc. Existing table to text models has provided a checkpoint for narrating text from tables but those are not completely efficient and reliable as they are exposed to hallucination which means the generated text is meaningful, but it is not related to the source that means it lacks faithfulness.

In this paper we have chosen 3 transformer-based models namely:

- T5
- BART
- mBART

David C. Wyld et al. (Eds): CST, NLMLT, DMS, CLBD, ITCS, VLSIE - 2022
pp. 55-60, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121906

2. LITERATURE REVIEW

Mike Lewis et al., proposes the research work for BART [1]. It introduces BART as a pre-training model approach that learns to map corrupted documents/text to the original one. This paper contains the working architecture as well as the details about the pretrained model and experimental results along with the comparison with other models.

Yinhan Liu et al., proposed the research work for MBART [2]. In this work, authors presented mBART that is a multilingual sequence to sequence model. mBART is a pre-trained by applying the BART to large-scale monolingual corpus on many languages. It contains the detailed analysis on mBART with different ranges and experimental results.

A study on comparison between Bart, t5 and GPT-2

[3] along with experimental conclusions. Its findings show that BART and T5 perform quite better and gives better results than GPT-2 for the chosen task.

Xinyu Xing et al., published the research work for table to text[4]. It proposes the use of STTP model along with the experimental results and comparison with Bart model.

Yang Yang et al., proposed the research work [5] using transformer. It uses the WIKIBIO dataset and proposed a transformer-based model to study several data to text generation tasks.

Tianyu Liu 1 et al., proposed the text narration from table [8] from an entity-centric view. They have used WIKIPERSON dataset contains around 250000, 30000, and 29000 (table, text) pairs in training, validation, and test sets respectively. They have tried decreasing the hallucinated data and increasing faithfulness by evaluating the faithfulness with two entity-centric metrics, both are proven to have good agreement with human perspective. They have also experimented the comparison of transformer and Bart model.

3. BACKGROUND

3.1. Table to Text

A Table is a widely used type of data source on the web, which has a definite structure and contains useful information. Interpreting the meaning of a table and understanding, explaining its content is an important problem in artificial intelligence, with innovative applications like question answering, in search engines and many more. The task of text narration from tables could be used to support many applications, such as conversational agents and various interpretation browsers. As well as the task can be used to generate meaningful sentences for the well-structured tables on the Internet.

3.2. Transformers

Before transformers were introduced, most SOTA (state-of-the-art) NLP systems used to rely on Recurrent Neural Networks (RNN), such as Long Short-Term Memory (LSTM) and Gated recurrent units (GRUs), with the help of added attention mechanisms. Transformers also tends to make use of attention mechanism but, unlike RNN, they do not have a recurrent structure which means that given enough training data, attention mechanisms can match the performance of RNN with attention.

3.2.1. T5

T5 stands for "Text to Text Transfer Transformer". T5 tries to combine all the downstream tasks into a text to text format. T5 is versioned into various type as per sizes:

- t5-small
- t5-base
- t5-large
- t5-3b
- t5-11b.

3.2.2. BART

BART stands for "Bidirectional Auto-Regressive Transformers." BART is a transformer-based Sequence to Sequence model that makes use of corrupted source text. BART can be seen as combination of BERT and GPT2 which tries generalizing Bert due to the bidirectional encoder and GPT2 with the left to right decoder.

3.2.3. mBART

mBART stands for "Multilingual Bidirectional Auto-Regressive Transformers" MBART is a multilingual encoder-decoder sequence to sequence model which is based on transformers primarily used for translation task but not restricted to that. As the model is multilingual it expects the sequences in a different format.

4. ARCHITECTURE

In this paper we have taken transformer-based models which include T5, BART, mBART for carrying out the comparison task. We have provided our explanation and understanding for the same.

4.1. T5

T5 is an encoder-decoder model which is pre-trained on a multitask which is a mixture of both unsupervised as well as supervised tasks and for which each task is converted into a text to text format. T5 works well on a variety of tasks by prepending a different prefix to the input corresponding to each task, e.g., for translation: translate English to German, summarizing text, predicting similarity score among 2 sentences.

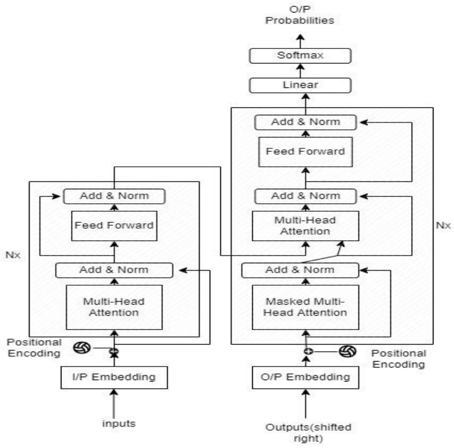

Fig. 1. Architecture of Transformer (T5)

Fig. 1 illustrate the architecture of transformer. The Transformer architecture consists of the Encoder block which is towards the left and the Decoder block which is towards the right.

Encoder: Encoder block consists of a stack of N identical layers. Every layer has a multi-head attention layer.

Decoder: The decoder stack also consists of 6 identical layers. Each decoder layer has 2 multi-head attention layers, followed by a feed forward neural network.

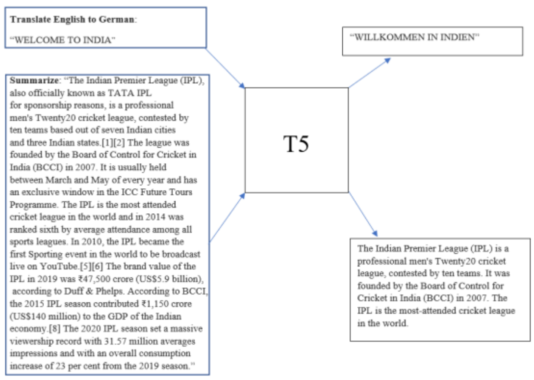

Fig. 2. T5 FLOWCHART

Fig. 2 illustrate the T5 framework. Many tasks can be casted into this framework like language translation, classification task, regression task other sequence to sequence tasks like document summarization for example, summarizing articles from websites, etc.

4.2. BART

BART is a self-supervised Sequence to Sequence auto-encoder model where we send the source data and add some noise to corrupt text, this data we send to the Denoiser which is an encoder decoder model, and then we get the regenerated text at the end. Here Regenerated text gives the feedback back to the Denoiser. i.e., Loss Optimization via Backpropagation. This model works well and gives the best performance when used for Natural Language Generation tasks like translation, summarization but it is also working perfectly for tasks like text classification, Q&A.

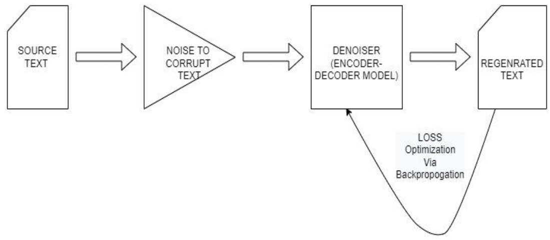

Fig. 3. Architecture of BART

Fig. 3 illustrate the architecture of BART. We have explained this architecture in detail in this paper.

We can say BART is a modified version of Sequence-to-Sequence model made to work as an auto-encoder. Only difference in BART architecture is the use of GELU instead of RELU activation layer. If we compare BART with BERT, BART doesn't make use of a feed-forward network at the top for word prediction while BERT does. BART uses just approximately 10% more parameters compared to equivalent BERT. BART achieves better performance for language generation tasks compared to BERT.

Applications of BART:

We can fine-tune BART achieving better performance for the following tasks:

- Sequence Generation
- Token Classification
- Sequence Classification
- Machine Translation

4.3. mBART

mBART is a sequence to sequence denoising auto-encoder model for pretraining a complete sequence to sequence model by denoising full texts in multiple languages, while earlier approaches have concentrated only on the encoder, decoder, or reconstructing parts of the text. The best part about mBART model is that it learns some structure of the languages during pretraining, and this structure goes beyond linguistic borders and allows intrinsic knowledge transfer between languages. This language-transfer significantly outperform fine-tuning on the target language pair and can be the turning point for some applications.

5. INFERENCE

In this paper we did a comparative survey on three transformer-based models for text generation from tabular data.

- One of the advantages of using T5 model against other models is that it does not provide a label or a span of the inputs as an output to the provided input sentence, but instead it generates the output as a string formatted text.
- BART is extremely flexible and can account for nonlinearities and interactions without overfitting due to the Bayesian priors. In addition, BART's default tuning parameters are effective in many cases.

- In short, we can describe mBART as a multilingual model with encoder-decoder primarily used for translation task but not limited to that. Being multilingual in nature it expects the sequences in a different format.

6. Conclusion & Future Work

In this paper, we have done the survey on three transformer based models for Table to text generation and approaches for the same. As an extension to this research work, we are planning to implement transformer based T5 model as our use case. We will be using the ToTTo dataset which is published by google and train the model on maximum possible epochs and optimize the solution by creating dynamic interface where we can provide our input and predict the meaningful sentence as an output from the model.

References

[1] Mike Lewis*, Yinhan Liu*, Naman Goyal*, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, Luke Zettlemoyer-BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension-BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension (aclanthology.org)

[2] Yinhan Liu*, Jiatao Gu*, Naman Goyal*, Xian Li, Sergey Edunov Marjan Ghazvininejad, Mike Lewis, Luke Zettlemoyer-Multilingual Denoising Pre-training for Neural Machine Translation-2001.08210.pdf (arxiv.org)

[3] Auto-regressive Text Generation with Pre-Trained Language Models: An Empirical Study on Question-type Short Text Generation-Anonymous ACL submission-pdf (openreview.net)

[4] Xinyu Xing and Xiaojun Wan Wangxuan Institute of Computer Technology, Peking University The MOE Key Laboratory of Computational Linguistics, Peking University-Structure-Aware Pre-Training for Table-to-Text Generation-Structure-Aware Pre-Training for Table-to-Text Generation (aclanthology.org)

[5] Yang Yang Communication University of China Juan Cao Communication University of China Yujun Wen (wenyujun@cuc.edu.cn) Communication University of China Pengzhou Zhang Communication University of China-Table-to-Text Generation with Accurate Content Copying-https://www.nature.com/articles/s41598-021-00813-6

[6] Google AI Blog: Exploring Transfer Learning with T5: the Text-To-Text Transfer Transformer (googleblog.com)

[7] Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer-https://openreview.net/pdf?id=5Jcma3gBao

[8] Tianyu Liu 1 *, Xin Zheng 2 3 *, Baobao Chang 1 4, Zhifang Sui 1- Towards Faithfulness in Open Domain Table-to-text Generation from an Entity-centric View- 2102.08585.pdf (arxiv.org)

[9] Ankur P. Parikh, Xuezhi Wang, Sebastian Gehrmann, Manaal Faruqui, Bhuwan Dhingra, Diyi Yang, Dipanjan Das https://arxiv.org/abs/2004.14373

[10] Sha, L., Mou, L., Liu, T., Poupart, P., Li, S., Chang, B., & Sui, Z. Order-planning neural text generation from structured data. *Proc. AAAI.* Vol. 32, No. 1(2018, April)