IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Natural Language Generation with Transformer.

MD SAMIR KHAN
BE-CSE (UIE)
CHANDIGARH UNIVERSITY
MOHALI, INDIA

ATEESH KUMAR
BE-CSE (UIE)
CHANDIGARH UNIVERSITY
MOHALI, INDIA

Sandeep Kaur
Assistant professor
CHANDIGARH UNIVERSITY
MOHALI, INDIA

Abstract.

The goal is to study natural language generation (NLG),with the help of transformer which enable's information sharing between the two complimentry activites with the help of Transformers.Transformer allow the automatic generation of high quality human like nautral text from structured data,prompts, or other forms of input.NLG with tranformers leverages the power of deep learning models which include transformer to achieves it goal .NLG models can be integrated into chatbots and virtual assistants to enable more natural and interactive conversations with users.it also provide responses that are contextually relevant and coherent.With this topic we able to see the bridge that build gap between human communication and machine understanding, enabling more natural human-computer interactions.

Introduction.

A branch of artificial intelligence (AI) and natural language processing (NLP) called "natural language generation" (NLG) is concerned with producing text that seems human. The process entails utilizing sophisticated machine learning models, specifically Transformer models, to produce logical and contextually appropriate language. An overview of NLG using Transformer models is given in this introduction, along with a discussion of its importance, uses, and capabilities.

1. The Transformer Revolution: NLP and NLG have undergone a revolution thanks to transformer models. First presented in the 2017 work "Attention Is All You Need" by Vaswani et al., these models have shown unmatched performance in a range of language-related tasks. They are the cornerstone of many cutting-edge natural language processing (NLP) systems due to their capacity to recognize long-range relationships in text through self-attention mechanisms. modest-scale prenatal care in primary care with electronic documentation on a mobile device.

2. Pretraining and Fine-Tuning: Pretraining and fine-tuning are the two steps that NLG using Transformer models usually entails. Large volumes of text data from the internet are used to train models in pretraining. These models learn to anticipate words in sentences, which helps them learn grammar, syntax, semantics, and general knowledge. By fine-tuning, the model is made specifically for NLG activities including content creation, language translation, and text summarization.

3. Flexible Text generating: A variety of text generating jobs, such as the following, can be accomplished with the help of these pretrained Transformer models:

- Content Creation: Generating articles, stories, product descriptions, and more
- Chatbots and Virtual Assistants: Engaging in human-like conversations and providing information or assistance.
- Text Summarization: Condensing lengthy documents into shorter, coherent summaries.
- Language Translation: Translating text fromone language to another.

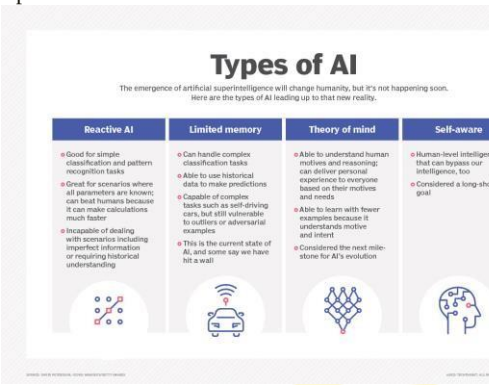
Question Answering: Providing answers to questions based on a given context.

1. Conditional Generation:
NLG models can generate text based on a provided context or input. This means you caninfluence the style, tone, or content of the generated text by giving the model specific
instructions or prompts.

2. Customization and Control:
Users can customize the behaviour of NLG models to align with specific requirements. Youcan steer the generation process to produce
content that fits predefined criteria, ensuring thatthe output meets particular standards and

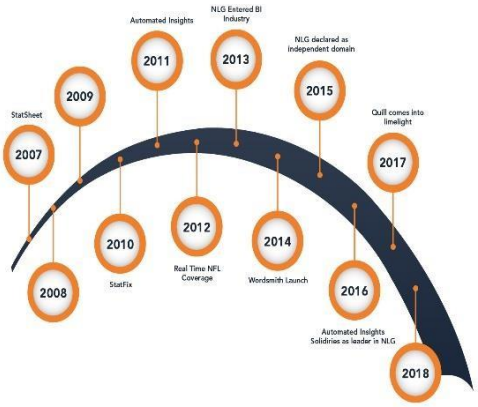
guidelines.

3. The Future of NLG: As NLG technology develops further, we should anticipate even more uses in a range of sectors, such as marketing, healthcare, customer service, and education. Scientists are tackling ethical issues and trying to improve these models' capabilities.



Background.

The roots of NLG can be traced back to the 1960s and 1970s when early computer programs like ELIZA and SHRDLU demonstrated basic text generation capabilities. These systems were rule- based and lacked the sophistication of modern NLG. Neural network-based approaches to NLP began to gain traction in the 1990s and early 2000s. Text generation and other sequence-to-sequence tasks were handled by Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks. The pivotal moment for NLG with Transformers came with the release of the "Attention Is All You Need"
a 2017 study by Vaswani et al. The Transformer architecture changed the field of natural language processing (NLP) by introducing the idea of self-attention, which worked incredibly well for processing sequential data. GPT (Generative Pretrained Transformer) was first released by OpenAI in 2018 with GPT-1. These models could be adjusted for different NLG tasks because they were pretrained on vast volumes of textual data. Due to possible misuse, GPT-2, which was released in 2019, caused controversy and had its release temporarily restricted. Google AI's introduction of BERT (Bidirectional Encoder Representations from Transformers) gave natural language processing (NLP) a fresh viewpoint. Bidirectional context understanding, which BERT introduced in place of just left-to-right or right-to-left context, improved a number of NLG tasks, including text classification and question answering. By employing a permutation-based training strategy, Google Brain and Carnegie Mellon University's XLNet further expanded transformer-based models. It showed gains in tasks related to language generation and comprehension.



With 175 billion parameters, OpenAI unveiled GPT-3, one of the biggest and most potent pretrained language models to date. In numerous NLG tasks, such as text completion, question answering, and text generation, it produced state-of-the-art results. Many uses for NLG with Transformers have been developed, such as sentiment analysis, text summarization, chatbots, content creation, and language translation. The automation and improvement of numerous facets of natural language processing have been made possible by these models. Large-scale pretrained models, such as GPT-3, have sparked debate about ethical problems ranging from potential misinformation, privacy, and biases in generated text. This has sparked conversations about the need for improved model governance and responsible AI. NLG with Transformers research is still in its infancy, with continuous efforts to decrease biases, increase model efficiency, and create new applications in industries like finance, healthcare, and education.

How does NLG work?

Natural language generation (NLG) is a type of software that produces natural language output according to the needed condition .A survey of NLG methods describes NLG as "as the up coming generation artificial intelligence and computational algos that is widely working in construction of computer systems that can produce understandable texts in English or other human languages .
autoated NLG will be compared to the human minds because it will concern enough ton think like humans and than turn the idea into reality >NLG is also can comapred to the translator of Artificial computer languages ,this type os decompiler can also provide readable code generation .
working of NLG has some stages like;
Input of data from user: it can take several inputs from various sources like database, APIs, or manual custom users .
Well Planned Contents: the content should be well manner so that the input can be understand by NLg system so that it can understand the key insight and can give the desire ouput.
NLG have some predefine algos, rule and templates, which read and analyze the input and try to give the closest output. NLG also have a look aver grammatically correct and meaningful sentences, review over the sentence; generated text often undergoes a review process to correct error to ensure it meets quality standard.
output: the stage is to provide the output int he form of text written docuemnt it should be unfiltered and eaasily readable fot the humans.

Future of NLG

The future of NLG is unpredictable ,few years back NLG was unknown to many people but with a course of time it is going to set a big Business in IT industries with will reform and its impact will directly or indirectly on the world economy . let us understand more about NLG, It can extract and process large amounts of data and then convert that information into human language. Effectively, the technology is capable of telling a story in the same way that a person would. Pretty amazing, right? In this post we will show you where NLG shows up in real life, what it can do for companies and what lies in store for the future.

We all known about GPT-3 . GPT-3 is a language model developed by OpenAI . It learns from pre-existing text and is able to complete sentences using different alternatives (similar to predictive text). But GPT-3 doesn't come out of the box like that, you have to train it with thousands and thousands of examples, a task that so far has only been done in English.

I can highlight some trends and potential directions for NLG:

1..Improved Quality and Control:
NLG models are likely to become even more proficient at generating high-quality, coherent, and contextually relevant text. Researchers are actively working on improving control over generated content, allowing users to specify style, tone, and other attributes with greater precision.

2..Multimodal NLG:
Future NLG systems may integrate multiple modes of communication, combining text with images, audio, and video to create more immersive and engaging content. This can enhance storytelling, content creation, and communication in various domains.

3..Conversational AI:
NLG will continue to drive advancements in chatbots and virtual assistants, making them more capable of holding natural, context-aware conversations with users. These systems will become more integral to customer support, education, and various other applications.

4..Translator :
NLG models will become more proficient at generating content in multiple languages, facilitating cross-lingual communication and translation. This will have significant implications for global business and international collaboration.

5..Domain-Specific NLG:
NLG models will be increasingly tailored to specific industries and domains. For example, NLG systems specialized in healthcare, finance, legal, and other sectors will generate content specific to their respective fields.

NLG Text Generation

It involves the use of algorithms and models to convert structured data or input into coherent and contextually appropriate language output. NLG finds applications in various domains such as chatbots, virtual assistants, content generation, data reporting, and more.

To understand NLG text generation in detail, let's break down the process into its key components:

1. Data Analysis: NLG starts with analyzing the input data or structured information. This could include numerical data, facts, or any other form of structured content. The analysis involves understanding the relationships between the data points and identifying patterns.

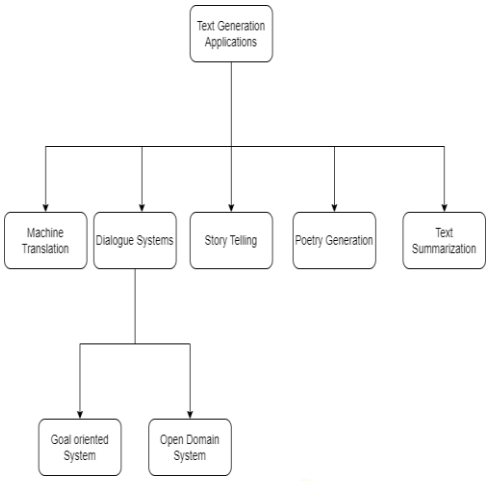
2. Content Determination: After analyzing the data, NLG determines what information is relevant and should be included in the generated text. It decides which details to emphasize and how to structure the content to convey the intended message effectively.

3. Text Planning: Once the relevant content is determined, NLG focuses on organizing the information into a coherent narrative.

4. Sentence Generation: NLG generates sentences based on the structured content and the planned narrative. This involves selecting appropriate vocabulary, grammar, and syntax to construct meaningful and grammatically correct sentences. Depending on the application, NLG systems may use rule-based systems, template-based systems,

machine learning, or deep learning techniques for sentence generation.

5. Surface Realization: In this final step, NLG converts the generated sentences into the final textual output. It takes care of formatting, punctuation, and any language-specific conventions to produce the text in a natural and readable format. Surface realization ensures that the generated text is contextually appropriate and aligns with human language patterns.



Rule-based systems use predefined grammar rules and templates to generate text, while template-based systems employ pre-designed templates that can be filled with dynamic data.

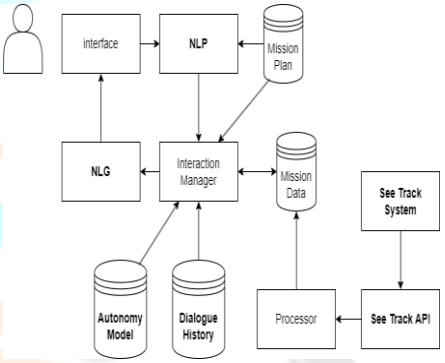
NLG has various applications across industries. In customer service, NLG can be used to automate responses and provide personalized recommendations. In journalism, it can generate news articles based on structured data. In e-commerce, NLG can create product descriptions or personalized marketing content. NLG is also used in data reporting to convert complex data sets into understandable narratives.

Transformer Architecture

Text creation from computer data is known as natural language generation, or NLG. It acts as a translator, converting digital data into representations in natural language. In this instance, the data gathered and user input are used to generate a conclusion or sentence. In natural language processing, it is the challenge of producing natural language from a machine representation system. Natural Language Understanding and Natural Language Generation are incongruent in certain aspects. Whereas in natural language generation, the system must choose how to put an idea into words, in natural language understanding, the system must resolve the input sentence to produce the machine representation language.

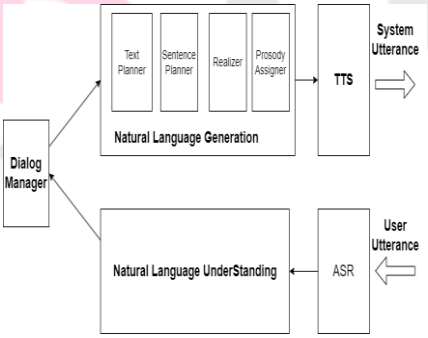
Having a list of pre-written phrases could be the easiest way to create content. material that has been cut and pasted. In straightforward applications like horoscope machines or personalized business, consequences can be satisfying.

generators of letters. However, in order to produce language that seems natural and does not become repetitive, a complex NLG system needs to go through stages of planning and information merging. One of the hardest problems in artificial intelligence (AI) is thought to be conversational AI. The creation of dialogue systems that can converse rationally and in-depth with people is a goal shared by many AI researchers. Task-oriented and open-domain dialogue systems are the two categories into which dialogue systems fall. Task-oriented dialogue systems interact with users to accomplish a predetermined goal, like helping clients purchase online or reserve a ticket. Through a series of discussions, an open-domain dialogue system, on the other hand, can help users ask questions, share knowledge, and develop social etiquette. Single-turn conversations from ten distinct domains make up our dataset.



Information about Bixby, the user, rivals, emotions, emergencies, greetings, entertainment, sensitive devices, and events.

This dataset was produced by skilled annotators with a background in linguistics and related fields. With an average of 7.31 and 14.44 words for utterance and response, respectively, the total dataset consists of 184,849 pairs of utterances and responses. First, we divided the data in a 95:5 ratio into a train and test set. Next, we create the validation set using 5% of the training set. An attention strategy is used by a Transformer, a deep learning architecture, to process text sequences. Transformers can capture long-term relationships in text because they do not rely on sequential connections like standard models based on recurrent neural networks do.



Transformer blocks and attention are the two main operating components of a transformer.

By allocating different weights to terms within a sequence, the model can concentrate its attention on the most pertinent portions of the literature. To help the model learn linguistic structures and patterns, transformation blocks are layers that apply nonlinear transformations to input representations.

NLG Application

Transformers are widely used in natural language processing software. Examples include automated summarization, text classification, question answering, sentiment analysis, machine translation, and text production.

A notable example of a Transformer in NLP is GPT-4 (Generative Pre-trained Transformer 4), created by OpenAI and currently dominating the large language model landscape based on numerous evaluations from both humans and machines. GPT-4 and its predecessor GPT-3.5 (better known as chatGPT) took the world by storm with their capacity to produce reliable and engaging content under a range of conditions.

These models have found application in the production of articles, essays, and even programming code in tasks such as automatic text synthesis. They have also been used in the creation of chatbots, virtual assistants, and customized recommendation engines. It should be noted, though, that human evaluation and interaction are crucial because the models have a propensity to "hallucinate," or make up information or facts.

Another noteworthy example is Google's Bard, a conversational generative AI chatbot that was first developed using the PaLM big language model family and later on the LaMDA family of big language models. At the moment, Bard can only speak English, Japanese, and Korean. Nevertheless, this model is currently unavailable in a number of European Union countries, including Spain, as a result of the General Data Protection Regulation (GDPR). Google is attempting to confirm, though, that Bard satisfies the requirements for use in Spanish-speaking territory.

Apart from GPT-4 and Bard, there are a number of other transformer-based large language transformers. models that focus on particular tasks and domains and are easier to use. For instance, LLaMa models have enhanced quality and accuracy in applications like information extraction, summarization, and machine translation.

As mentioned earlier, these NLP Transformer models require a significant amount of computer power to develop and run because they are trained on enormous volumes of data. But their effectiveness and adaptability have led to important advances in natural language processing.

In summary, Transformers have revolutionized language modeling by offering improved comprehension and reliable text production across an array of uses. They are helpful natural language processing tools because of their ability to capture long-term interdependence, versatility, and adaptability to varied tasks. Research and development are moving forward, and these

REFERENCES.

1. Galderisi, S., Heinz, A., Kastrup, M., Beezhold, J., & Sartorius, N. (2015). Toward a new definition of mental health. *World psychiatry*, *14*(2), 231.
2. How mental health care should change as a consequence of the COVID-19 pandemic, The Lancet Psychiatry,Volume 7, Issue 9, 2020, Pages 813-824, ISSN 2215-0366.
3. Caldeira, C., Chen, Y., Chan, L., Pham, V., Chen, Y., & Zheng, K. (2017). Mobile apps for mood tracking: an analysis of features and user reviews. In *AMIA Annual Symposium Proceedings* (Vol. 2017, p. 495). American Medical Informatics Association. . https://appadvice.com/app/mood-balanceself-caretracker/1471631164.
4. https://www.geeksforgeeks.org/a-completeguide-to-learn-xml-for-android-appdevelopment/
5. https://XAMP//SQL.
6. Attention Is All You Need" (Transformer)

   Authors: Vaswani.
7. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding"-

   Authors: Devlin.
8. GPT-3: Language Models are Few-Shot Learners"

   Authors: Brown.
9. Turing-NLG: A 17-billion-parameter Language Model by Microsoft-

   Authors: Kitaev
10. Generating Long Sequences with Sparse Transformers"

    Authors: Child
11. DialoGPT: Large-Scale Generative Pre-training for Conversational Response Generation"

    Authors: Zhang