

This article in the formatting stage, it will be change after completing all process.

Recent Trends in Programming Languages

Volume 12 Issue 2, 2025

ISSN: 2455-1821

Review Article

Developing Techniques for Controlling Different Aspects of Text Generation such as Tone and Contents

Rohan Sharma *1, **Ashutosh Naryan**2, **Arun Kumar Rai** 3

^{1,2}Student, Department of Computer Science (Artificial Intelligence and Machine Learning), GNIOT Group of Institutions, Knowledge Park 2, Greater Noida , Uttar Pradesh, India

³Assistant Professor, Department of Computer Science (Artificial Intelligence and Machine Learning), GNIOT Group of Institutions, Knowledge Park 2, Greater Noida , Uttar Pradesh, India

Corresponding author e-mail: Rohan1410sharma@gmail.com

Abstract - Large Language Models (LLMs) have shown excellent text creation quality in Natural Language Processing (NLP). However, LLMs have to satisfy ever-more-complex standards in real-world applications. LLMs are supposed to meet specific user goals, like as mimicking specific writing styles or producing material with poetic richness, in addition to eliminating inaccurate or objectionable content. Controllable Text Generation (CTG) techniques were developed in response to these diverse demands. They guarantee that outputs meet predetermined control conditions, including safety, sentiment, thematic consistency, and linguistic style, while upholding high standards of helpfulness, fluency, and diversity. The aim of this study is to improve the accuracy and flexibility of natural language generation (NLG) models by investigating new methods for managing tone, content, and style, among other elements of text production. Although language models like GPT have made great progress in producing prose that is both fluent and coherent, it is still difficult to regulate some aspects like tone (formal vs. casual) and content (creativity vs. factual accuracy). Using both supervised and reinforcement learning methods, we provide a system that integrates adaptable control mechanisms at several points during the text production process. The framework's fine-grained control over the generated text's tone, voice, and subject matter guarantees that the final product satisfies predetermined objectives. To direct the generating process, our method combines adversarial training, context-aware embeddings, and user feedback loops. Through a series of experiments on various datasets, we assess the efficacy of the suggested methodologies and show that the framework can generate text that satisfies certain tone and content requirements while retaining high linguistic quality. With applications ranging from content creation and tailored communication systems to customer service automation, this study advances the area of natural language generation by offering useful tools for personalizing generated text.

Keywords— Text Generation, Natural language generation, content modification, supervised learning, reinforcement learning, text creation, tone control, embeddings, user feedback, and personalization.

Received Date: 21-Apr-2025

Accepted Date: 2-May-2025

Published Date:

This article in the formatting stage, it will be change after completing all process.

Recent Trends in Programming Languages

Volume 12 Issue 2, 2025

ISSN: 2455-1821

Review Article

Introduction

Recent developments in Natural Language Generation (NLG) have produced extremely powerful systems that can generate text that is fluent, coherent, and pertinent to its context. Numerous industries, such as healthcare, customer service, and content production, have implemented these systems.

Controlling certain elements of text creation, such tone, style, and substance, is still a crucial difficulty in spite of recent developments. Although NLG models are capable of producing grammatically correct and contextually relevant text, their practical applicability in situations needing certain outputs is sometimes limited by their lack of fine-grained control over these qualities. Take the example of an automated email assistant used for customer care, like those found on many e-commerce sites. Composing answers to consumer questions about product problems is the assistant's responsibility. Although the algorithm may be able to provide a technically correct response, it frequently finds it difficult to keep the tone acceptable for various client scenarios. For instance, a reaction to a customer who is upset should be sympathetic and contrite, but an answer to a straightforward question should be uncomplicated and unpretentious. Current NLG systems have a tendency to produce general replies that might not be warm enough for a delicate circumstance or seem too informal in more serious settings [1].

The main goal is to provide a framework that gives users fine-grained control over these elements so they may customize the output to meet particular needs. The architecture suggested here includes a more complex control mechanism that may alter tone and content in response to specific user inputs, in contrast to conventional systems that provide general directions like

"generate formal text" or "be creative." [2]

2. Literature Review

The development of large language models (LLMs) such GPT, BERT, and T5 has made significant progress in the field of text generation in recent years, particularly as it comes to controlling tone and content. The foundational models are the focus of early research in this field, with encouraged the use of natural language processing (NLP) techniques to produce content that was both coherent and appropriate for its context. But as text generation's uses grew, so did the need to improve control over the outputs that were generated [3].

2.1. Content-Based Approaches

Once it comes to text production, content-based approaches focus on making ensure that the final outcome remain pertinent, coherent, and consistent with the intended theme or subject matter. In contrast to tone control, which modifies the text's emotional or stylistic elements, content control makes sure that its content is in conformity with established standards or subject-specific expertise [4].

2.2. Tone-Based Approaches

The aim of tone-based text-generation approaches is to regulate the created text's formal, stylistic, or emotional features. Models can provide outputs that are suited for particular audiences, situations, or communication goals—like informal chat, formal writing, or persuasive speech—by adjusting tone. This feature is essential for many applications, such as virtual assistants, creating content, and customer support.

3. Scope of The Project

This article in the formatting stage, it will be change after completing all process.

Recent Trends in Programming Languages

Volume 12 Issue 2, 2025

ISSN: 2455-1821

Review Article

This project's goal is centered on creating methods that allow for fine-grained control over various text creation elements, including tone and content. This study attempts to close the gap by providing more adaptable methods that enable the management of textual features, given that natural language generation (NLG) systems have made notable progress in producing coherent, fluent writing. The scope encompasses the creation of a framework that combines many control methods to guarantee that the generated text complies with user-specific objectives in a range of situations [5].

Three main areas are the focus of this project:

1. **Tone Control:** One of the main objectives is to make it possible to change the generated text's emotional undertone, making it more professional, casual, aggressive, or sympathetic. Context-aware embeddings, which enable the application of certain tonal qualities and a detailed comprehension of the emotional context, will be used to accomplish this control.
2. **Content Control:** Making sure the output text complies with specified content criteria is the second main goal. This involves having command of the text's precision, originality, and applicability. The project will investigate techniques like adversarial training and reinforcement learning to make sure that the information produced is in line with predetermined goals, whether they be factual, creative, or a mix of the two.
3. **Customization and Adaptability:** This project will also highlight the system's adaptability, which will allow users to alter the generating process to suit their requirements. The solution will be flexible enough to accommodate a variety of use cases, whether the objective is to modify the tone and substance of marketing emails, customer support answers, or creative writing [6].

3. Proposed Algorithm

The proposed algorithm aims to control both the tone and content of text generation using a hybrid approach that combines content conditioning, reinforcement learning (RL), and attention mechanisms.

The goal is to generate text that adheres to specific content guidelines while maintaining a desired tone.

1. **Input Preprocessing:** The input consists of a prompt that includes both the target content and tone specifications (e.g., "generate a formal, informative text about climate change"). The content is extracted using natural language understanding techniques, and tone-related labels are identified (e.g., formal, casual, persuasive).
2. **Model Conditioning:** A pre-trained language model (e.g., GPT-3,meta) is fine-tuned on a corpus labeled with both content-specific and tone-specific data. The model is conditioned on the prompt, ensuring that the content is relevant to the specified topic and aligned with the target tone.
3. **Reinforcement Learning (RL) for Tone and Content Control:** The model is trained using RL, where rewards are given for generating content that adheres to both content relevance and tone requirements. For example, a formal tone might be rewarded for using polite language, while content relevance is ensured by penalizing off-topic or factually incorrect generation.

This article in the formatting stage, it will be change after completing all process.

Recent Trends in Programming Languages

Volume 12 Issue 2, 2025

ISSN: 2455-1821

Review Article

4. **Output Generation:** The model generates the final output, with both content and tone closely aligned to the specifications. The output is evaluated based on predefined metrics for content coherence and tone accuracy [7].

4. Objective

The main goal of this project is to use Python to create and apply sophisticated methods to handle various text production parameters, particularly tone and content, in natural language generation (NLG) systems. Even while contemporary NLG models like GPT and BERT have shown remarkable talents in producing text that is coherent and contextually relevant, they frequently have trouble exercising fine-grained control over certain characteristics like content alignment, stylistic preferences, and emotional tone. By creating a versatile, adaptable framework that enables users to modify and regulate the generated text's tone and content in accordance with certain user-defined objectives, this study seeks to overcome these constraints [8].

5. Methodology

The methodology for the research focuses on using Pythonbased machine learning techniques to create a strong framework for managing many aspects of text generation, including tone and content. To allow for accurate manipulation of generated text attributes, the system combines deep learning models with natural language processing (NLP) technologies.

In this research, we aim to develop and explore techniques for controlling various aspects of text generation, specifically focusing on tone and content. The methodology is structured to ensure that the generated text aligns with the predefined objectives in terms of tone consistency and content accuracy. We discuss the tools, techniques, and approaches utilized to generate text while maintaining the desired characteristics.

5.1 Defining the Aspects of Text Generation

The first step in developing a methodology for controlling text generation is clearly defining the aspects that need to be controlled:

- **Tone Control:** Tone refers to the emotional quality or attitude expressed in the generated text. It can be formal, informal, neutral, persuasive, or analytical, among other variations. The tone is crucial because it dictates how the message is perceived by the reader.
- **Content Control:** Content control involves ensuring that the generated text is relevant, coherent, and accurate. This includes organizing the text, maintaining logical flow, and ensuring that the facts, data, or arguments presented are aligned with the research question or subject matter[9][10].

5.2 Data Collection and Preprocessing To develop effective techniques for tone and content control, high-quality data is necessary. The data for this study is collected from multiple sources:

- **Text Corpora:** We use diverse corpora that include academic articles, technical papers, blogs, and news reports to analyze how tone and content are maintained across different writing styles.
- **Annotation and Labeling:** Text samples are annotated based on tone (e.g., formal, persuasive, neutral) and content type (e.g., descriptive, analytical, explanatory). This labeling helps in training and validating the text generation models.

This article in the formatting stage, it will be change after completing all process.

Recent Trends in Programming Languages

Volume 12 Issue 2, 2025

ISSN: 2455-1821

Review Article

To make sure the data is appropriate for machine learning applications, it goes through pretreatment stages include text cleaning (noise removal), tokenization (word or phrase breakdown), and lemmatization (word reduction to their base forms).

5.3 Model Selection and Training

To control tone and content in text generation, we adopt advanced natural language generation (NLG) models. The models selected are fine-tuned and optimized to ensure precision in controlling both aspects

- **Pre-trained Language Models:** Transformer-based models, such GPT (Generative Pretrained Transformer), are selected because of their capability to generate language that is both coherent and fluid. To learn the subtleties of tone and content control, these models are improved using our annotated corpora.
- **Tone Control Mechanisms:** We implement fine-tuning strategies that allow the model to generate text in the desired tone. This is achieved by conditioning the model's output based on specific tone-related tokens or tags provided during training. For example, we might label the dataset with tone labels like "formal," "neutral," or "persuasive" to guide the model's response.
- **Content Control Mechanisms:** We use topic-specific prompts and keywords to make sure the information is relevant and in line with the study's goals. In order to preserve factual accuracy relevance, and coherence, the model is trained to concentrate on content production methods such information retrieval, summarization, and content adaptation.

5.4 Tone Control Techniques

The process of controlling tone involves several steps:

- **Tone Tags and Prompts:** During training, specific tone tags (e.g., "informal," "neutral," "technical") are introduced to guide the tone of the generated content. By embedding these tone markers within the input prompts, the model learns to adjust its language output accordingly.
- **Reinforcement Learning for Tone Consistency:** Reinforcement learning (RL) techniques are utilized to refine the tone across generated text. An RL agent is employed to continuously evaluate the tone of the output and provide feedback that reinforces the tone consistency through reward systems. For example, if the generated text matches the target tone, the model receives a higher reward.
- **Human-in-the-Loop Feedback:** In cases where automatic methods fall short, human evaluators provide feedback on tone quality, ensuring the generated text adheres to the required emotional stance. This feedback loop helps further refine the model's tone output.

5.5 Content Control Techniques

Controlling content is a more challenging aspect of text generation, as it requires a deeper understanding of context, accuracy, and relevance. The following techniques are employed to maintain content quality:

- **Topic-Specific Training:** The model is adjusted using domain-specific data to make sure it produces information that is pertinent to particular study areas. Scientific publications,

This article in the formatting stage, it will be change after completing all process.

Recent Trends in Programming Languages

Volume 12 Issue 2, 2025

ISSN: 2455-1821

Review Article

essays, or datasets pertaining to the topic of interest can all be included in this data. The model is trained to guarantee that the information it generates is based on the chosen field of study.

- **Fact-Checking Mechanisms:** We use fact-checking technologies that compare created content with reliable databases or sources since accuracy is essential. An external verification module that indicates content that may include errors or inconsistencies is incorporated into this process.

5.6 Evaluation and Quality Assurance

Once the model has been trained and fine-tuned, we evaluate the effectiveness of the text generation system using both automated and manual methods:

- **Automated Metrics:** The generated content's coherence, relevance, and fluency are assessed using metrics like BLEU, ROUGE, and METEOR. These metrics assess how accurately and stylistically the generated text matches references published by humans.
- **Human Evaluation:** Human evaluators assess the tone and content quality based on a set of predefined criteria, such as: *Tone Appropriateness:* Does the text match the desired tone (formal, persuasive, etc.)?

Content Relevance: Is the generated text aligned with the research topic and context?

Coherence and Structure: Does the text maintain logical flow and clarity?

6. RESULT AND DISCUSSION

Through a series of checks using Python-based models, the suggested methods for managing various elements of text generation—particularly tone and content—were assessed. The purpose of these studies was to evaluate the efficacy of the hybrid strategy that combines attention processes, reinforcement learning (RL), and content conditioning.

Evaluation of Content Control: The refined models showed notable gains in content relevance when conditioned on domain-specific cues. The models effectively followed the necessary content requirements while creating text on predetermined themes (such as technology or healthcare), regularly delivering material that was in line with the desired subject matter. By incorporating reinforcement learning, content accuracy was further improved, as the models were rewarded more for producing outputs that were factually accurate and pertinent to the topic.

Evaluation of Tone Control: Models trained with labels specific to a certain tone demonstrated encouraging outcomes in terms of modifying the tone of generated text. In contrast to informal or amusing tones, the model produced more structured and professional language when it was asked to generate text in a formal tone. Additionally, reinforcement learning was essential in adjusting the tone because it taught the models to maximize for the intended stylistic or emotional traits, including assertiveness or empathy.

7. CONCLUSION

This paper provides an overview of recent advancements in the area of Controllable Text Generation (CTG) for Large Language Models (LLMs), focusing on aspects such as tone and content. It systematically outlines the core concepts, while addressing the control criteria and the quality standards for generated text. Additionally, the paper introduces a new approach for

This article in the formatting stage, it will be change after completing all process.

Recent Trends in Programming Languages

Volume 12 Issue 2, 2025

ISSN: 2455-1821

Review Article

classifying CTG tasks, distinguishing between content control (also known as linguistic or hard control) and attribute control (referred to as semantic or soft control).

This study has provided a thorough framework for managing several facets of text production, including tone and content, in Python-based natural language generation (NLG) systems. Even though NLG technologies have advanced significantly, current models frequently lack fine-grained control over the generated text's substance, style, and emotional tone. These drawbacks have prevented NLG from being widely used in fields like marketing, customer service, and creative writing that need for specialized outputs. The experimental findings demonstrate enhanced flexibility and accuracy in managing tone and content across a range of use scenarios, confirming the framework's efficacy.

Finally, this study identifies current research difficulties and suggests future development paths by offering a thorough discussion of the fundamental ideas, technological techniques, assessment strategies, and real-world applications in the field of controlled text production. Its purpose is to provide as a methodical guide and reference for study into controlling text production.

REFERENCES

1. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. *Advances in neural information processing systems*. 2017;30:1-11.
2. Ranzato MA, Chopra S, Auli M, Zaremba W. Sequence level training with recurrent neural networks. *arXiv preprint arXiv:1511.06732*. 2015 Nov 20:1-16.
3. Keskar NS, McCann B, Varshney LR, Xiong C, Socher R. Ctrl: A conditional transformer language model for controllable generation. *arXiv preprint arXiv:1909.05858*. 2019 Sep 11:1-18.
4. Liu Y, Su Y, Shareghi E, Collier N. Plug-and-Play Recipe Generation with Content Planning. *arXiv preprint arXiv:2212.05093*. 2022 Dec 9:223–234.
5. Liu Y, Liu X, Zhu X, Hu W. Multi-aspect controllable text generation with disentangled counterfactual augmentation. *arXiv preprint arXiv:2405.19958*. 2024 May 30.
6. Controllable Text Generation for Large Language Models: A Survey
7. Yu D, Yu Z, Sagae K. Attribute alignment: Controlling text generation from pre-trained language models. *arXiv preprint arXiv:2103.11070*. 2021 Mar 20.
8. Jiang Q, Chen L, Xu R, Ao X, Yang M. A challenge dataset and effective models for aspect-based sentiment analysis. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP) 2019 Nov* (pp. 6280-6285).
9. Abdali S, Anarfi R, Barberan CJ, He J. Decoding the ai pen: Techniques and challenges in detecting ai-generated text. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining 2024 Aug 25* (pp. 6428-6436).
10. Shen C, Cheng L, Zhou R, Bing L, You Y, Si L. Mred: A meta-review dataset for structure-controllable text generation. *arXiv preprint arXiv:2110.07474*. 2021 Oct 14.
11. Shin T, Razeghi Y, Logan IV RL, Wallace E, Singh S. Autoprompt: Eliciting knowledge from language models with automatically generated prompts. *arXiv preprint arXiv:2010.15980*. 2020 Oct 29..

This article in the formatting stage, it will be change after completing all process.

Recent Trends in Programming Languages
Volume 12 Issue 2, 2025
ISSN: 2455-1821
Review Article