

**STUDENT-CENTERED AUTOMATED FEEDBACK USING  
TRANSFORMER BASED APPROACH**

**Dissertation Phase-2 Report**

*Submitted by*

**Mr. WILLIAM TOM JACOB**

**REG NO : TKM22MEAI17**

*the APJ Abdul Kalam Technological University in partial  
fulfillment for the award of the degree of*

**MASTER OF TECHNOLOGY**

**IN**

**Artificial Intelligence**

**Under the guidance of  
Dr. Sumod Sundar**



**Centre for Artificial Intelligence**

**TKM College of Engineering, Kollam**

**JUNE 2024**

Thangal Kunju Musaliar College of Engineering  
Centre for Artificial Intelligence



C E R T I F I C A T E

This is to certify that, this report titled ***STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH*** is a bonafide record of the **Dissertation Phase-2** presented by **WILLIAM TOM JACOB (TKM22MEAI17)**, under our guidance and supervision, in partial fulfillment of the requirements for the award of the degree, **M. Tech in Artificial Intelligence in APJ Abdul Kalam Technological University**.

Internal Supervisor and Project Coordinator

Head of the Department

Dr. Sumod Sundar  
Associate Professor  
Centre for AI  
TKMCE

Dr. Imthias Ahamed T P  
Professor  
Centre for AI  
TKMCE

## ACKNOWLEDGEMENT

A successful project is a fruitful culmination of efforts by many people, some directly involved and some others indirectly, by providing support and encouragement. Firstly I would like to thank the almighty for giving me the wisdom and grace for making my project a memorable one. I thank him for steering me to the shore of fulfillment under his protective wings.

I express my sincere gratitude to **Dr. T. A. Shahul Hameed**, Principal of T.K.M College of Engineering. I would like to express heartfelt thanks to **Dr. Imthias Ahamed T.P.**, Professor and Head of the Department, Centre for Artificial Intelligence, TKM College of Engineering, Kollam, for his expert guidance, and immense encouragement.

I would also like to thank my supervisor and project coordinator **Dr. Sumod Sundar**, Associate Professor, **Prof. Chinnu Jacob**, Assistant Professor, **Prof. Christy D Ponnai**, Assistant Professor, Centre for Artificial Intelligence(AI), TKM College of Engineering, Kollam for their expert guidance, cooperation, and immense encouragement. I also extend my thanks to the entire faculty and staff members of the Centre for AI, TKMCE, who have encouraged me throughout this work.

I also express my thanks to my loving parents and friends, for their support and encouragement in the successful completion of this work.

William Tom Jacob

## Abstract

Effective communication between teachers and students is essential in today's educational environment. This project presents a transformer model that has been designed for the educational environment. With the use of this model, students can have personalized learning experiences and ask academic questions and while getting immediate answers generated by the model. The personalized feedback is given by using a deep learning model. Due to lack of dataset existing in education domain, a dataset on education domain is developed using web scraping techniques, which scrapes or extracts the relevant data from internet. The data which is obtained through web scrapping is given to the model for training. The deep learning model used is Generative Pre-trained Transformer-2 (GPT-2) a transformer model which then analyses the pattern and gives the customized feedback back to student. The GPT-2 was used for training on the data set. After performing the validation between reference and generated text, the evaluation metrics were obtained. The evaluation metrics showed very low BLEU score and ROUGE score. To increase the value of BLEU score and ROUGE score, data augmentation was performed. Data augmentation technique such as synonym replacement was performed to increase the data size. Parameter efficient fine-tuning techniques such as Low rank Adaption is applied to reduce the size of the model which in turn can produce the results of the model at much less computation time and improved memory efficiency. The integration of this model is expected to transform the educational system by creating an atmosphere in which students are inspired to learn and teachers can skilfully meet each student's unique learning needs.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Literature Survey</b>	<b>4</b>
<b>3</b>	<b>Methodology</b>	<b>10</b>
3.1	Dataset Used . . . . .	10
3.1.1	Dataset-1 . . . . .	10
3.1.2	Dataset-2 . . . . .	11
3.2	Web Scrapping . . . . .	12
3.3	Proposed Framework . . . . .	13
3.4	Technique Used . . . . .	14
3.4.1	GPT-2 . . . . .	15
3.5	PEFT (Parameter Efficient Fine Tuning) . . . . .	19
3.5.1	LoRA (Low-Rank Adaptation of Large Language Models) . . . . .	21
<b>4</b>	<b>Experimental Analysis and Results</b>	<b>23</b>
4.1	Experiments on data pre-processing task . . . . .	23
4.2	Experimental Setup . . . . .	23
4.2.1	Environment Used . . . . .	23
4.3	Data Augmentation . . . . .	24
4.3.1	Synonym replacement . . . . .	25
4.4	Results . . . . .	27
4.4.1	BLEU Score . . . . .	27
4.4.2	ROUGE Score . . . . .	28
<b>5</b>	<b>Conclusion and Future scope</b>	<b>33</b>
	<b>References</b>	<b>35</b>

# List of Figures

3.1	Dataset-1 Sample . . . . .	11
3.2	Dataset-2 Sample . . . . .	11
3.3	Web Scrapper Output . . . . .	12
3.4	Proposed Framework . . . . .	13
3.5	GPT-2 Architecture . . . . .	16
3.6	query,keys and values . . . . .	18
3.7	PEFT . . . . .	20
3.8	LoRA Model . . . . .	21
4.1	Sample of synonym replacement on question . . . . .	25
4.2	Sample of synonym replacement on answer . . . . .	26
4.3	Sample Output 1 . . . . .	29
4.4	Sample Output 2 . . . . .	29
4.5	Sample Output 3 . . . . .	31
4.6	Sample Output 4 . . . . .	31
4.7	Sample Output 5 . . . . .	32

# List of Tables

3.1	Comparison of size before and after applying LoRA . . . . .	22
4.1	Comparison on dataset-1 before and after data augmentation . . . . .	27
4.2	Evaluation metrics 1 . . . . .	29
4.3	Evaluation metrics 2 . . . . .	30
4.4	Evaluation metrics 3 . . . . .	30
4.5	Evaluation metrics 4 . . . . .	30
4.6	Evaluation metrics 5 . . . . .	31
4.7	Evaluation metrics 6 . . . . .	32
4.8	Evaluation metrics 7 . . . . .	32

# Chapter 1

## Introduction

In the rapidly advancing field of educational technology, it is increasingly evident that enhancing communication between teachers and students is crucial. Traditional educational methods are being replaced by more dynamic and interactive approaches, necessitating more efficient communication channels. As educational practices evolve, there is a growing demand for tools that foster greater engagement and participation. Improving these communication routes is essential to ensure that both teachers and students can adapt to and thrive in this new educational landscape. To address this need, we propose developing a state-of-the-art smart device specifically for the educational sector. This advanced tool aims to revolutionize teacher-student interactions through a seamless platform for information sharing. It will foster collaboration and create a more integrated learning environment. By enhancing teamwork and facilitating smoother communication, this device has the potential to significantly improve the educational experience, making learning more coherent and engaging for all participants.

By leveraging the latest technological advancements, we aim to provide a solution that not only meets but exceeds the expectations of teachers and students. Our model, featuring real-time feedback, is designed to surpass traditional communication tools. This enhancement will make the learning process more dynamic and responsive for both educators and students, thereby improving the overall quality of education[2]. Our objective is to create a model that seamlessly adapts to the evolving demands of modern education. As we strive to transform communication within the educational sector, we focus on developing a tool that integrates effortlessly into contemporary educational practices. Our commitment to innovation and flexibility ensures that our solution will remain relevant and effective in fostering improved interactions and learning outcomes. This advanced tool has the inherent capability to revolutionize students interaction with academic material, facilitating easier inquiry on challenging subjects and dispelling any apprehension about voicing questions. Encouraging active participation and confidence among students. This technology bridges the gap between conventional teaching methods and contemporary learning approaches, creating a mutually beneficial relationship between teaching and student involvement.

Real-time feedback is crucial for maintaining the dynamic and flexible nature of training programs. This immediate feedback mechanism ensures that students remain engaged and can quickly resolve any misunderstandings or questions that arise. Instructors can provide

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

instant feedback and constructive criticism due to this timely interaction, which is essential for fostering an effective learning environment. The value of real-time feedback lies in its ability to create a continuous feedback loop between teachers and students. This loop is vital for enhancing the learning process through ongoing adjustments and improvements. Teachers can quickly identify areas where students are struggling and provide targeted guidance to correct misconceptions. This prompt assistance helps students stay on track and make consistent progress.

Moreover, real-time feedback empowers students to actively engage in their learning. Receiving immediate responses to their work allows students to quickly recognize and correct errors, leading to better understanding and retention of the material. This swift corrective process is crucial for building a strong knowledge foundation, as it prevents the reinforcement of incorrect information and promotes deeper comprehension. Beyond improving learning outcomes, immediate feedback also fosters a sense of responsibility and engagement among students. When students know they will receive prompt feedback and guidance, they are more likely to take an active role in their education. This engagement is key to maintaining student interest and enthusiasm, which are essential for effective learning.

Additionally, real-time feedback contributes to creating a supportive and flexible learning environment. Such an environment can be tailored to meet the specific needs of each student, providing personalized support that maximizes the learning experience. Instructors can adjust their feedback to match the individual learning styles and paces of students, ensuring each receives the appropriate level of challenge and support. In summary, real-time feedback is essential for training programs, enhancing instructional effectiveness, encouraging student participation, and fostering a positive learning environment. By enabling prompt identification and correction of errors and maintaining dynamic interaction between instructors and learners, real-time feedback ensures that students can achieve their full potential and meet their educational goals.

Moreover, this immediate communication fosters a classroom environment characterized by vibrancy and dynamism. Seamless interaction between educators and students transforms learning into a collaborative endeavor, with information freely shared between both parties[5]. By ensuring students feel appreciated for their contributions and that their needs are promptly addressed, such engagement cultivates a sense of community and mutual respect. Rather than a static transmission of knowledge, the classroom becomes a dynamic educational setting where information is continuously exchanged. This dynamic engagement enhances students' understanding and retention of subject matter while also fostering critical thinking and active participation. Real-time communication nurtures a collaborative, inclusive learning environment that empowers students and instills a sense of ownership, ultimately enriching the educational experience.

When immediate teacher involvement isn't possible, the device's features become operational. Powered by state-of-the-art AI algorithms, the device effectively addresses various student queries with precise and beneficial answers. What sets this concept apart is its sophisticated AI algorithms designed to identify patterns in student inquiries. By analyzing these trends, the tool offers valuable insights into topics that may require further attention

## **STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH**

---

in future classes. This enables teachers to better accommodate the diverse learning styles of their students and encourages active pursuit of knowledge among pupils. By highlighting areas for deeper exploration, the tool helps educators tailor their teaching to the specific needs of their classroom, ultimately enhancing the overall quality of education.

## Chapter 2

# Literature Survey

In recent years, the development of question-answering systems for conversation agents has been a an area of focus point in the realm of artificial intelligence. Karpagam et al. present a deep learning-based approach for answer selection in question-answering systems, specifically tailored for conversation agents. They leverage word2vector and a biLSTM attentive model to enhance answer selection, utilizing datasets such as Insurance QA, Yahoo! Answer, TREC QA, and bAbI Facebook [1]. While their approach is comprehensive and well-implemented, a notable shortcoming is the lack of a comparative study with other models or approaches that are already in use in question-answering systems. This omission raises concerns about how the performance, accuracy, and efficiency of the suggested model compare to the state-of-the-art methods already in use. It is challenging to evaluate the relative benefits or potential drawbacks of the new model in the absence of such a comparison. Furthermore, more research is required to determine the model's effectiveness and generalizability across various datasets or domains. Understanding the model's resilience and adaptation to various data kinds and real-world situations requires evaluating it in a variety of circumstances. Addressing these aspects would provide a more comprehensive validation of the model, offering deeper insights into its practical applications and potential for widespread adoption in various question-answering tasks.

Nuruzzaman and Hussain contribute to the discourse by evaluating IntelliBot, a chatbot designed for the insurance industry. Their work includes a comparative analysis against existing chatbots like RootyAI, ChatterBot, and DeepQA [2]. The research successfully illustrates IntelliBot's advantage in providing thorough and semantically correct information, highlighting its potential as a sophisticated chatbot for advanced question-answering (QA). There are, however, some significant flaws in the assessment. The lack of analysis of the constraints and limitations present in the suggested IntelliBot multi-strategy selection-based QA chatbot system is one major drawback. There are still unsolved concerns regarding the framework's adaptation and flexibility in different situations due to this gap in the analysis. Moreover, the comparison analysis is constrained because it only considers a small portion of chatbots that are currently in use. The study may have missed other strategies that could have provided insightful information or worked better in particular circumstances if a wider range of pertinent models had been taken into consideration. The research successfully illustrates IntelliBot's advantage in providing thorough and semantically correct information, highlighting its potential as a sophisticated chatbot for advanced question-answering (QA).

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

There are, however, some significant flaws in the assessment. The lack of analysis of the constraints and limitations present in the suggested IntelliBot multi-strategy selection-based QA chatbot system is one major drawback. There are still unsolved concerns regarding the framework's adaptation and flexibility in different situations due to this gap in the analysis. Moreover, the comparison analysis is constrained because it only considers a small portion of chatbots that are currently in use. The study may have missed other strategies that could have provided insightful information or worked better in particular circumstances if a wider range of pertinent models had been taken into consideration.

Serban et al. delve into the intricacies of latent representations in dialogue generation models by employing a hierarchical latent variable encoder-decoder model. This innovative technique enhances the model's ability to differentiate between various types of responses and linguistic styles through the utilization of Recurrent Neural Network (RNN) and Variational Autoencoder (VAE) methods [5]. Their approach showcases significant advancements in the field and illustrates the model's aptitude for generating diverse and contextually relevant dialogue. The hierarchical architecture of their model allows it to capture both detailed conversational nuances and broader dialogue context, crucial for generating responses that are meaningful and coherent. This capability is particularly beneficial for creating more naturalistic interactions in automated dialogue systems, a primary objective in the development of intelligent conversational agents. Despite these notable achievements, the study has several limitations that warrant careful consideration.

One major issue is the lack of a comprehensive comparison with current dialogue generation methods. Without conducting thorough benchmarking against established models, it becomes challenging to discern the comparative advantages and disadvantages of the proposed methodology. Such comparisons are essential for understanding the unique contributions of the model and its potential benefits over existing approaches. The model's heavy reliance on human evaluation presents a notable limitation. Human judgments, being inherently subjective, may overlook certain nuances in the model's capabilities, despite offering valuable insights into the naturalness and appropriateness of generated dialogues. This subjectivity can introduce variability in evaluation results, potentially obscuring the model's true performance. To address this issue, incorporating additional objective evaluation measures is essential. Metrics such as BLEU, ROUGE, and perplexity scores provide quantitative assessments of the model's efficacy, scalability, and efficiency. Perplexity evaluates the model's predictive ability, offering insights into its generalization capabilities. ROUGE and BLEU scores, commonly used in assessing machine translation and summarization tasks, measure the quality of the output by comparing it to reference texts.

These metrics offer a more comprehensive and reliable evaluation framework, facilitating comparisons of the model's performance across different studies and applications. The absence of these objective indicators in Serban et al.'s study limits the depth of the evaluation and may obscure the model's true potential. Without these measurements, it is challenging to assess the effectiveness of the proposed approach comprehensively or to identify areas for improvement accurately. Therefore, achieving a comprehensive and accurate assessment of the model's success necessitates a balanced evaluation that integrates both subjective assessments and objective measurements. For a thorough and dependable evaluation of the

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

model's performance in real-world situations, it's crucial to incorporate both subjective and objective assessments. This comprehensive approach ensures that the model's performance is adequately documented from both qualitative and quantitative perspectives, facilitating a deeper understanding of its strengths and weaknesses. Ultimately, this holistic evaluation process would enable more informed decisions in the development and implementation of advanced dialogue generation systems, thereby improving their effectiveness and reliability in practical applications.

In their groundbreaking paper, Vaswani et al. introduce the Transformer model, which revolutionizes the field of neural networks by eliminating the need for complex recurrent or convolutional architectures. Instead, the Transformer exclusively employs attention mechanisms, marking a significant departure from traditional models [9]. This shift towards attention mechanisms brings several advantages. Firstly, the Transformer model enhances efficiency. Unlike traditional models that often require substantial training time, the Transformer is designed to be more time-efficient. This is primarily because it allows for parallelization, where multiple processes can run simultaneously, speeding up the training process significantly. Traditional models, particularly those relying on recurrent neural networks (RNNs), struggle with this as they often require sequential processing. Furthermore, the performance of the Transformer in machine translation tasks is superior to that of traditional models. By leveraging fully connected layers and stacked self-attention mechanisms, the Transformer can effectively capture and model global dependencies between input and output sequences. This capability is crucial in translation tasks, where understanding the context and relationships between words across long sequences is vital.

The Transformer model's performance surpasses that of previous models, including those using ensemble techniques, which combine multiple models to improve performance. Achieving state-of-the-art results on translation tasks, such as the WMT 2014 English-to-German and English-to-French benchmarks, highlights its effectiveness. Additionally, the architecture of the Transformer model underscores its scalability and robustness. Its design facilitates faster and more efficient training, making it well-suited for large-scale translation tasks and other potential applications in natural language processing and beyond. The model's success in these areas demonstrates its promising future and the potential for broader applications in various domains.

The paper's introduction of the Transformer model is primarily centered around sequence transduction tasks, such as machine translation, and does not directly address challenges that arise when dealing with other data modalities like images, audio, or video. This focus means that the specific issues and adjustments required for effectively applying the Transformer to these types of data are not covered in the paper. For instance, while the Transformer leverages Multi-Head Attention to help capture relationships between different parts of the input, its method of averaging attention-weighted positions can lead to a reduction in effective resolution. This can impact the model's ability to learn dependencies between distant points within the input data, potentially limiting its effectiveness when applied to non-sequential data or tasks requiring high-resolution details.

Despite the impressive performance of the Transformer in achieving state-of-the-art re-

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

sults for translation tasks, the paper does not extend its evaluation to other types of tasks or datasets. This leaves a gap in understanding how well the Transformer model might perform in different contexts beyond machine translation. Additionally, the paper does not delve into the particular challenges or modifications necessary for scaling the Transformer model to handle tasks with substantial input and output sizes. For example, managing the complexity of lengthy documents or high-resolution images requires adaptations that are not discussed in the publication. These scenarios present unique difficulties, such as maintaining coherence over long texts or preserving fine-grained details in images, which the standard Transformer architecture might not be fully equipped to handle without further refinement or additional techniques. Therefore, while the Transformer model shows great promise, further research and experimentation are needed to explore its applicability and performance across a broader range of tasks and data types.

Grenander et al. focus on enhancing student learning gains in the realm of intelligent tutoring systems by utilizing automated, customized feedback methods. Their innovative approach involves deconstructing student responses and mapping them into a relational network using advanced neural discourse segmentation and classification algorithms [6]. This technique aims to provide detailed and contextually relevant feedback that can adapt to each student's specific needs, thereby personalizing and improving the learning process. The study by Grenander et al. successfully highlights several benefits of this proposed methodology. A significant advantage is the system's ability to deliver personalized feedback that aligns closely with each student's unique learning requirements and progress. This tailored interaction can lead to better understanding and retention of the material, as students receive guidance that directly addresses their current knowledge level and misconceptions.

However, the study also identifies a notable drawback of this approach. Specifically, there are instances where the model generates poor or even nonsensical results. Such occurrences can undermine the effectiveness of the feedback system, as the quality and consistency of the feedback are crucial for fostering student learning and trust. These erroneous outputs underscore the model's limitations in consistently producing logical and coherent responses. This issue highlights a significant challenge in the development and deployment of intelligent tutoring systems: ensuring the accuracy and consistency of automated feedback. While neural discourse segmentation and classification show great potential for enhancing educational outcomes, the model's inconsistent performance indicates that further refinement is necessary. Addressing these shortcomings is essential for creating a robust and reliable intelligent tutoring system that can consistently support and enhance student learning.

Additionally, the study incorporates multiple RoBERTa models, a choice that introduces fresh challenges alongside their considerable potential. One significant concern revolves around the substantial memory requirements of these models, potentially constraining their scalability and practicality, particularly in environments with limited computational resources. The utilization of multiple RoBERTa instances raises concerns regarding the system's memory usage, potentially restricting its usefulness and effectiveness in real-world scenarios. Moreover, this setup may exacerbate processing challenges, leading to increased computational costs and delays in delivering feedback. These obstacles not only impact the system's efficiency but also compromise its responsiveness and overall performance. Address-

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

ing these issues is crucial to enhancing the viability and effectiveness of the feedback system in practical applications, where response accuracy and resource efficiency are pivotal factors. Implementing strategies to optimize memory usage and streamline processing workflows is essential to ensure the system's continued viability and its ability to furnish users with accurate and timely feedback across diverse contexts.

Song et al. delve into conversational aspect sentiment analysis (CASA), providing fine-grained sentiment information for dialogue planning and understanding. Their work showcases strong results for in-domain and out-of-domain datasets using Bidirectional Encoder Representations from Transformers (BERT) [7]. A notable constraint of the research is the absence of any comparative analysis with alternative models or techniques in the Conversational Agent Sentiment Analysis (CASA) domain. It is challenging to evaluate the relative inventiveness and efficacy of the suggested strategy because of this omission. Furthermore, a thorough examination of the possible disadvantages or difficulties associated with identifying sentiment expression in conversational dialogues is not done in this work. It is vital to comprehend these difficulties because sentiment expression can differ significantly amongst speakers and circumstances, which could have an impact on the correctness and dependability of the model. The effect of domain adaptation on CASA systems' performance is another important factor that the study fails to take into consideration. Domain adaptation can significantly influence how well a sentiment analysis model generalizes to new, unseen data, and its neglect may lead to an overestimation of the model's capabilities. Addressing these gaps would provide a more comprehensive and realistic evaluation of the model's performance and applicability in diverse real-world scenarios.

In conclusion, these studies encompass a wide range of AI-related topics, particularly focusing on chatbots, sentiment analysis, conversation generation, and question-answering systems. Each research contributes novel and innovative elements to these domains. For instance, improvements in sentiment analysis allow systems to better understand and respond to human emotions, while advancements in chatbot technology enhance user interactions by making conversations more natural and intuitive. Question-answering systems are becoming more proficient at providing accurate and relevant answers based on extensive datasets, and dialogue production techniques are being refined to generate more coherent and contextually appropriate responses. Collectively, these advancements expand the capabilities of AI in understanding and generating human language, paving the way for more sophisticated and user-friendly applications.

Despite these advancements, further research and comparative studies are necessary to verify the effectiveness and applicability of the proposed models across various settings and datasets. While each model performs well in the specific context for which it was designed, its performance can vary significantly in different contexts or with different data. Comprehensive comparative evaluations enable researchers to identify the strengths and weaknesses of various models, leading to improvements and adaptations that ensure these AI systems are robust and versatile. Ongoing research is essential for advancing AI technologies and ensuring their reliable deployment in real-world applications, maximizing their potential benefits across diverse fields.

## **STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH**

---

The primary goal of this project is to offer students immediate, personalized answers to their questions, significantly enhancing their learning experience. As the number of students and the volume of educational content grow, it becomes increasingly difficult for instructors to address each student's queries individually. This issue is exacerbated by limited time and instructor availability. In contrast, an AI model can provide instant responses at any time, ensuring that learning resources are accessible 24/7. This enables students to get the support they need outside of regular classroom hours, promoting a more continuous and flexible learning environment. Ultimately, this approach enhances timely problem-solving and encourages student independence.

## Chapter 3

# Methodology

The primary goal is to develop a deep learning model that allows students to receive instant, personalized answers to their questions, thereby enhancing their overall educational experience. Students will input their queries into the model, which will analyze and generate relevant feedback. This method aims to understand the context and content of each student's question to deliver accurate and customized responses that address their specific needs. Utilizing advanced natural language processing techniques, the model will be capable of comprehending complex queries and providing precise and helpful answers. This approach not only facilitates immediate support but also enhances student engagement and interaction, making education more effective and accessible for learners.

### 3.1 Dataset Used

Due to the lack of a specific educational dataset, the GPT-2 model was first trained on dataset-1. Later, dataset-2 was created through web scraping to better meet specific educational needs. This new dataset aimed to improve the GPT-2 model's accuracy and effectiveness in educational applications by providing more relevant and focused instructional content, thereby enhancing the model's overall performance.

#### 3.1.1 Dataset-1

The figure 3.1 shown above is the InsuranceQA dataset for the task of insurance-related question answering. The dataset is obtained from kaggle which is in csv format. There is a total of 1000 – rows and 3 – columns in the dataset. The three different columns are

- ID
- Question
- Answer

# STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

id	question	answer
0	Is Disability Insurance Required By Law?	Not generally. There are five states that require most all employers carry short term disability insurance on their employees. These states are: t
1	Can Creditors Take Life Insurance After Death?	If the person who passed away was the one with the debt, creditors generally cannot take the life insurance proceeds left as long as the benefi
2	Does Travelers Insurance Have Renters Insurance?	One of the insurance carriers I represent is Travelers and yes, you can purchase Renters insurance through Travelers. I would look for a local ag
3	Can I Drive A New Car Home Without Insurance?	Most auto dealers will not let you drive the car off of the lot unless you have proof of insurance if there is a loan on the car. It is never a good i
4	Is The Cash Surrender Value Of Life Insurance Taxable?	Cash surrender value comes only with Whole Life Insurance, never with Term. It is the amount you can terminate your policy and receive. That
5	How Is Annuity Income Reported?	You should news is that annuity income is usually reported by the insurance company issuing the annuity. When they make a disbursement they a
6	What Does AAA Home Insurance Cover?	AAA Home insurance, like all other major carriers, covers a wide variety of claims, including fire, theft, vandalism, and many other items. Howes
7	What Is A Simple Retirement Plan?	What is a SIMPLE retirement plan? A Simple is an inexpensive opportunity for a small business to have a retirement plan that doesn't break the l
8	What Does Social Security Disability Insurance Cover?	Social Security Disability Insurance pays a monthly benefit to people who cannot work due to a disability. The benefit is paid in cash to the reci
9	Is Car Insurance Prepaid?	Yes, automobile insurance is typically paid in advance. Normally no less than thirty days at a time. Each carrier sets their own requirements as t
10	What Does Medicare Part B Cover?	Medicare Part B covers the doctor services, outpatient hospital services, medical services and supplies. There is a monthly cost charged to the t
11	Can Veterans Get Life Insurance?	Unless a service person gets Veterans Group Life Insurance (VGLI) within a year from separating from the service, their Servicemembers Group
12	Does My Homeowners Insurance Cover Lost Wedding Ring?	Great question. Generally no, your Homeowners' policy will not cover for anything that is lost. This is referred to as "Mysterious Disappearance
13	How Does Assigned Risk Auto Insurance Work?	Assigned risk in California which is what I am familiar with is a processing center by which people having trouble buying auto insurance are conne
14	Can My Boyfriend Add Me To His Car Insurance?	If you are living in the same household then he should put you on the list of drivers. The insurance follows the car and that is linked to a househ
15	Do I Need A Police Report To Submit A Claim For A Car Accident On Private Property?	Actually, though the local police might prepare an accident report for you, in many cases, because the incident occurred on private property, it i
16	Does Full Coverage Auto Insurance Cover Repairs?	By "Full Coverage" I am assuming you mean, does carrying Comprehensive and Collision coverage cover repairs? If you have a covered claim i.e
17	Is Life Insurance Acceptable In Islam?	Do to its use of interest and chance there are those who believe that insurance is not permitted in Islam (i.e. that it is haram or sinful). Due to t
18	Will Health Insurance Cover A Tubal Reversal?	Most insurance plans will not cover the reversal of a voluntary sterilization or any treatments designed to overcome infertility. There are fifteen
19	How Much Life Insurance Can You Get If You Have Been Diagnosed With Breast Cancer?	Traditionally, you can get up to 20-30x your income level of life insurance. The medical history does not necessarily limit how much you can get
20	What Do I Need To Register For Medicare?	If you have paid taxes into the system then enrollment is automatic. You should receive your Medicare card and a packet of information from I
21	Does Short Term Disability Insurance Work?	Short-term Disability does work if you have a qualifying disability. When you are applying for short or long-term disability, you need to look at t
22	Will Homeowners Insurance Cover Pool?	In short, the answer is yes. But it needs to be clarified whether you are asking for property damage to the pool itself or liability coverage. Both
23	When Do I Have To Sign Up For Medicare?	You should sign up for Medicare in the three months before the month of your 65th birthday, the month of your 65th birthday, or in the three n
24	Will An MIP Affect My Car Insurance?	That is a great question! If you are the minor in question with the possession charge, you better bet it will. If the minor was someone in your ca
25	Does Blue Cross Blue Shield Have Life Insurance?	Blue Cross / Blue Shield is the name of the network association for a number of health insurance companies (includes Anthem and CareFirst). N
26	What Are The Advantages And Disadvantages To Long Term Care Insurance?	The advantages of coverage in your senior years can be significant based on the average senior couple spending \$250,000 in medical costs dur
27	Will Medicare Pay For Life Alert?	Medicare covers items that are considered medically necessary. This includes diagnosis and treatment of medical conditions caused by acciden

Figure 3.1: Dataset-1 Sample

The dataset's two primary features used in implementation are the question and the response. Each row contains a question in one column and its corresponding answer in the adjacent column. This structure simplifies linking queries with their specific answers, making it ideal for training and evaluating models in automated question-answering tasks.

## 3.1.2 Dataset-2

A	B	C
id	question	answer
1	1 What was the first European trading company to establish a foothold in India?	The British East India Company, arriving in 1600 with permission from Mughal Emperor Akbar.
2	2 What key Battle helped solidify British influence in India?	The Battle of Plassey (1757), where a smaller British force strategically defeated the Nawab of Bengal.
3	3 Who was a key figure in Indian social reform movements during British rule?	Raja Ram Mohan Roy, advocating against Sati (widow burning) and founding the Brahmo Samaj.
4	4 What was the Sepoy Mutiny of 1857?	A large-scale rebellion by Indian soldiers against the British East India Company, ultimately leading to British Crown rule.
5	5 What policy aimed to assimilate Indians into British culture?	The English Education Act of 1835, promoting English language education and Western ideas.
6	6 What was the significance of the Indian National Congress (INC)?	A major political party formed in 1885, spearheading the Indian independence movement.
7	7 Who were two prominent leaders of the Indian independence movement?	Mahatma Gandhi, advocating for non-violent resistance, and Bhagat Singh, known for his revolutionary actions.
8	8 What was the impact of the Rowlatt Act of 1919?	It granted the British broad emergency powers, leading to protests like the Jallianwala Bagh Massacre.
9	9 What was the significance of the Government of India Act of 1935?	It introduced limited self-government in India, but also solidified the separation of British India and Princely States.
10	10 What was the Quit India Movement?	A nationwide campaign launched by Gandhi in 1942, demanding an immediate British withdrawal from India.
11	11 What role did World War II play in the Indian independence movement?	It weakened British power and boosted Indian nationalism, with movements like the Azad Hind Fauj.
12	12 What were the key figures behind the partition of India in 1947?	Muhammad Ali Jinnah, leader of the Muslim League, and Jawaharlal Nehru, representing the INC.
13	13 What were the consequences of the partition of India?	Mass displacement and violence along religious lines, creating the independent nations of India and Pakistan.
14	14 What economic impact did British rule have on India?	Exploitation of resources for British benefit, while neglecting Indian infrastructure and development.
15	15 What social and cultural changes occurred in India during British rule?	A rise in Western education, changes in dress and customs, and the emergence of new social movements.
16	16 How did the British justify their rule in India?	Often through the concept of the "White Man's Burden," claiming to bring civilization and progress to India.
17	17 What are some lasting legacies of British rule in India?	The English language, a parliamentary system of government, and a complex social and religious landscape.
18	18 How do historians view the British Raj (rule) in India?	With a critical eye, acknowledging both modernization efforts and the exploitation of India's resources and people.
19	19 What are some sources for studying Indian-British history?	Colonial records, Indian literature and poetry, and archaeological evidence.
20	20 Why is understanding Indian-British history important?	country.
21	21 colonial rule?	Indo-Gothic, seen in structures like Mumbai's Chhatrapati Shivaji Maharaj Terminus.
22	22 What was the significance of the Bengal Renaissance?	An intellectual and cultural movement challenging colonial ideas and reviving Indian heritage.
23	23 environmental damage?	Indigo, with forced cultivation practices hurting Indian farmers.
24	24 What was the role of the printing press during the Indian independence movement?	It facilitated the spread of nationalist ideas through newspapers and pamphlets.
25	25 Who were the Dalits, and how did British rule impact their social status?	The "untouchables" of the caste system. British policies like separate electorates initially offered some opportunities for political participation.
26	26 What was the significance of the Simon Commission?	A British attempt at constitutional reform, ultimately rejected by the INC for not offering enough self-governance.

Figure 3.2: Dataset-2 Sample

Figure 3.2 features the Indian History dataset, curated for answering history-related questions. This dataset was compiled by scraping content from various websites that cover Indian history during the British colonial period. Stored in CSV format on an Excel sheet, it comprises 1000 rows and 3 columns. Each row likely corresponds to a unique historical question,

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

with the columns containing the question, its answer, and possibly additional context. This dataset is particularly valuable for educational and research purposes focused on India's colonial history, as it aids in developing and testing models designed to enhance automated question-answering systems in the historical domain. The three different columns are

- ID
- Question
- Answer

The dataset's two key features are utilized to develop question and response functionalities. Each row is arranged with a question in one column and its corresponding answer in the adjacent column. This straightforward yet effective format simplifies the pairing of questions with answers, thereby facilitating the development and testing of automated question-answering systems.

### 3.2 Web Scrapping

```
</tbody>
</table></div>, <div class="widget-area footer-widgets-3 footer-widget-area"><section class="widget widget_
ass="widget-wrap"><h4 class="widget-title widgettitle">Quick Resources</h4>
<div class="textwidget"><table>
<tr><td><a href="https://www.learnbse.in/english-grammar/">English Grammar</a></td>
<td><a href="https://www.learnbse.in/hindi-grammar/">Hindi Grammar</a></td></tr>
<tr><td><a href="https://www.learnbse.in/textbook-solutions/">Textbook Solutions</a></td>
<td><a href="https://www.learnbse.in/maths-ncert-solutions/">Maths NCERT Solutions</a></td></tr>
<tr><td><a href="https://www.learnbse.in/science-ncert-solutions/">Science NCERT Solutions</a></td>
<td><a href="https://www.learnbse.in/sst-ncert-solutions/">Social Science NCERT Solutions</a></td></tr>
<tr><td><a href="https://www.learnbse.in/english-ncert-solutions/">English Solutions</a></td>
<td><a href="https://www.learnbse.in/hindi-ncert-solutions/">Hindi NCERT Solutions</a></td></tr>
<tr><td> <a href="https://www.learnbse.in/ncert-exemplar-problems-solutions/">NCERT Exemplar Problems</a></td>
<td><a href="https://www.learnbse.in/engineering-entrance-exams/">Engineering Entrance Exams</a></td></tr>
<tr><td><a href="https://www.facebook.com/ncertsolutionsbooks" rel="nofollow">Like us on Facebook</a></td>
<td><a href="https://twitter.com/learnbse" rel="nofollow">Follow us on Twitter</a></td></tr>
<tr><td><a href="https://www.youtube.com/user/cbsepapers/videos" rel="nofollow">Watch Youtube Videos</a></td>
<td><a href="https://play.google.com/store/apps/details?id=in.learnbse.ncertsolutions&hl=en_US" rel="n
pp</a></td>
</tr>
</table></div>
</div></section>
</div>, <div class="widget-wrap"><h4 class="widget-title widgettitle">Quick Resources</h4>
<div class="textwidget"><table>
<tr><td><a href="https://www.learnbse.in/english-grammar/">English Grammar</a></td>
```

Figure 3.3: Web Scrapper Output

Web scraping, an automated technique, efficiently retrieves extensive data from online sources. This data, often in unstructured HTML on websites, requires conversion to structured formats like spreadsheets or databases for various applications. Central to web scraping is its

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

adept use of advanced automation, enabling rapid extraction of vast data sets. The process involves two components: a crawler and a scraper. The crawler, an AI-driven algorithm, explores the web for specific content by following links. Acting as a digital explorer, it systematically navigates webpages to locate desired information.

Conversely, the scraper serves as a specialized instrument tailored for extracting data from specific websites. It operates by adhering to predefined instructions to gather relevant data elements based on set criteria. Utilizing diverse techniques, the scraper navigates through the HTML structure of web pages to extract essential data fields like text, images, or links. The amalgamation of the scraper and crawler enhances the efficiency of the web scraping process, enabling the extraction of substantial data volumes from various online platforms. This technique furnishes organizations, researchers, and analysts with valuable insights, trends, and information sourced from the vast realm of the internet, facilitated by astute automation and systematic data retrieval. Its applications range from trend prediction and competitive analysis to market research.

Web Scrapers can extract all the data on particular sites or the specific data that a user wants. So, when a web scraper needs to scrape a site, first the URLs are provided. Then it loads all the HTML code for those sites and a more advanced scraper might even extract all the CSS and Javascript elements as well. Then the scraper obtains the required data from this HTML code and outputs this data in the format specified by the user. Mostly, this is in the form of an Excel spreadsheet or a CSV file, but the data can also be saved in other formats, such as a JSON file. A variety of libraries are created specifically for Web Scraping. Scrapy is a very popular open-source web crawling framework that is written in Python. It is ideal for web scraping as well as extracting data using APIs. BeautifulSoup is another Python library that is highly suitable for Web Scraping. It creates a parse tree that can be used to extract data from HTML on a website. BeautifulSoup also has multiple features for navigation, searching, and modifying these parse trees.

### 3.3 Proposed Framework

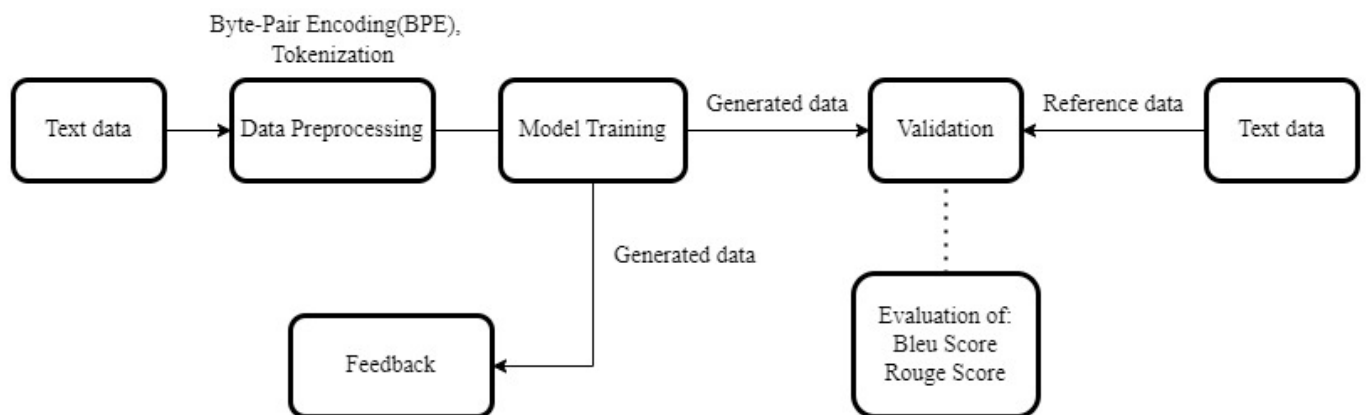


Figure 3.4: Proposed Framework

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

The process depicted in Figure 3.4 outlines how a student's query from a text file is analyzed, and personalized feedback is provided. Initially, the text data undergoes preprocessing to ensure its suitability for analysis. This preprocessed data is then utilized to train the deep learning model, sourced from platforms like Kaggle. Employing the Generative Pre-trained Transformer 2 (GPT-2), an advanced transformer model renowned for its ability to comprehend and generate coherent text, is integral to this approach. The GPT-2 model is trained to identify patterns within the input data, enabling it to generate feedback tailored to the student's query.

Evaluation of the accuracy and relevance of the model-generated feedback is conducted by comparing it against a reference response using metrics such as the BLEU Score and ROUGE Score. This systematic approach ensures that students receive individualized and pertinent responses to their queries, enhancing their overall educational experience. Leveraging state-of-the-art deep learning techniques like GPT-2 results in more precise and beneficial feedback for learners, thus improving their educational journey.

### 3.4 Technique Used

Text generation tasks have benefited from the utilization of deep learning models like Recurrent Neural Networks (RNNs) and Long Short Term Memory (LSTMs) Networks. However, their effectiveness is hindered by memory constraints when processing large datasets. To overcome this limitation, larger models become imperative. Transformers represent a novel approach to natural language processing (NLP), offering superior capability in deciphering intricate word relationships over extended contexts, resulting in unprecedented semantic accuracy. Unlike RNNs and LSTMs, transformers leverage parallel processing via self-attention mechanisms, eliminating sequential computation bottlenecks and enhancing scalability. This innovative architecture has propelled advancements in question answering, text summarization, and machine translation, facilitating the generation of more coherent outputs and deeper contextual understanding. With their efficient handling of extensive datasets, transformers have become the cornerstone of contemporary NLP, fundamentally reshaping our interaction with textual data.

Transformer models utilize a sophisticated two-step process to generate and analyze text. Initially, input phrases are transformed into vectors through a technique known as self-attention. This enables the model to identify important word associations and capture subtle contextual details crucial for understanding the input text. Following encoding, the data is decoded using the same self-attention mechanism to convert the encoded data back into output tokens. By applying self-attention during both encoding and decoding, the model ensures the contextual coherence of the input text is maintained throughout the production process, ensuring coherence and relevance in the output sequence.

Through the utilization of self-attention, transformer models dynamically prioritize different segments of the input text, facilitating the recognition of intricate relationships and generating reliable outcomes. This dynamic focusing mechanism allows the model to adapt

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

tively adjust its attention to various parts of the input text, enhancing its ability to capture complex linguistic structures and produce contextually appropriate outputs. Overall, the efficiency and adaptability of transformer designs rely on this innovative approach of utilizing self-attention for both encoding and decoding text data.

In transformer models, the attention mechanism plays a vital role in aiding comprehension of word relationships across sentences. It identifies key words in the input to generate output tokens during decoding. This dynamic allocation of significance helps the model understand linguistic complexities and long dependencies. By focusing on relevant elements, transformers better grasp contextual information, resulting in accurate outputs. The adaptability of this attention mechanism contributes to the success of transformer-based architectures in various natural language processing tasks. The different types of transformers are BERT, RoBERTa, DistilBERT, GPT.

### 3.4.1 GPT-2

OpenAI's GPT-2, the second iteration in its core GPT model series, is renowned for its extensive language capabilities. OpenAI researchers utilized web crawling to compile the substantial 40GB WebText dataset, which serves as the model's training data. This dataset enables GPT-2 to discern intricate linguistic patterns and generate text resembling human writing. Through a blend of supervised and unsupervised learning methods during training, GPT-2 learns from both labeled and unlabeled data. Post-training, the model can generate coherent text responses upon request, showcasing its applicability in various natural language processing tasks. However, discussions surrounding ethical considerations and responsible deployment arise due to GPT-2's adeptness at mimicking human-like writing, highlighting the need for cautious use of such advanced language models in real-world contexts.

With the capability to generate text akin to human writing, this model emerges as a potent tool for tasks within natural language processing professions. Integral to the Transformer neural network architecture renowned for its efficacy across diverse natural language processing domains, this model boasts an impressive 1.5 billion trainable parameters. This expansive parameter count enables it to grasp and adapt to a broad spectrum of language scenarios, showcasing its versatility and utility in tasks like text generation and language translation. Its adeptness at replicating human-authored content underscores its significance across various contexts, solidifying its status as a foundational technique in the realm of computational linguistics. Among its many functions are summarization, language translation, and text completion. GPT-2's architecture is based on the Transformer architecture, which was introduced by Vaswani et al. in the paper "Attention is All You Need."

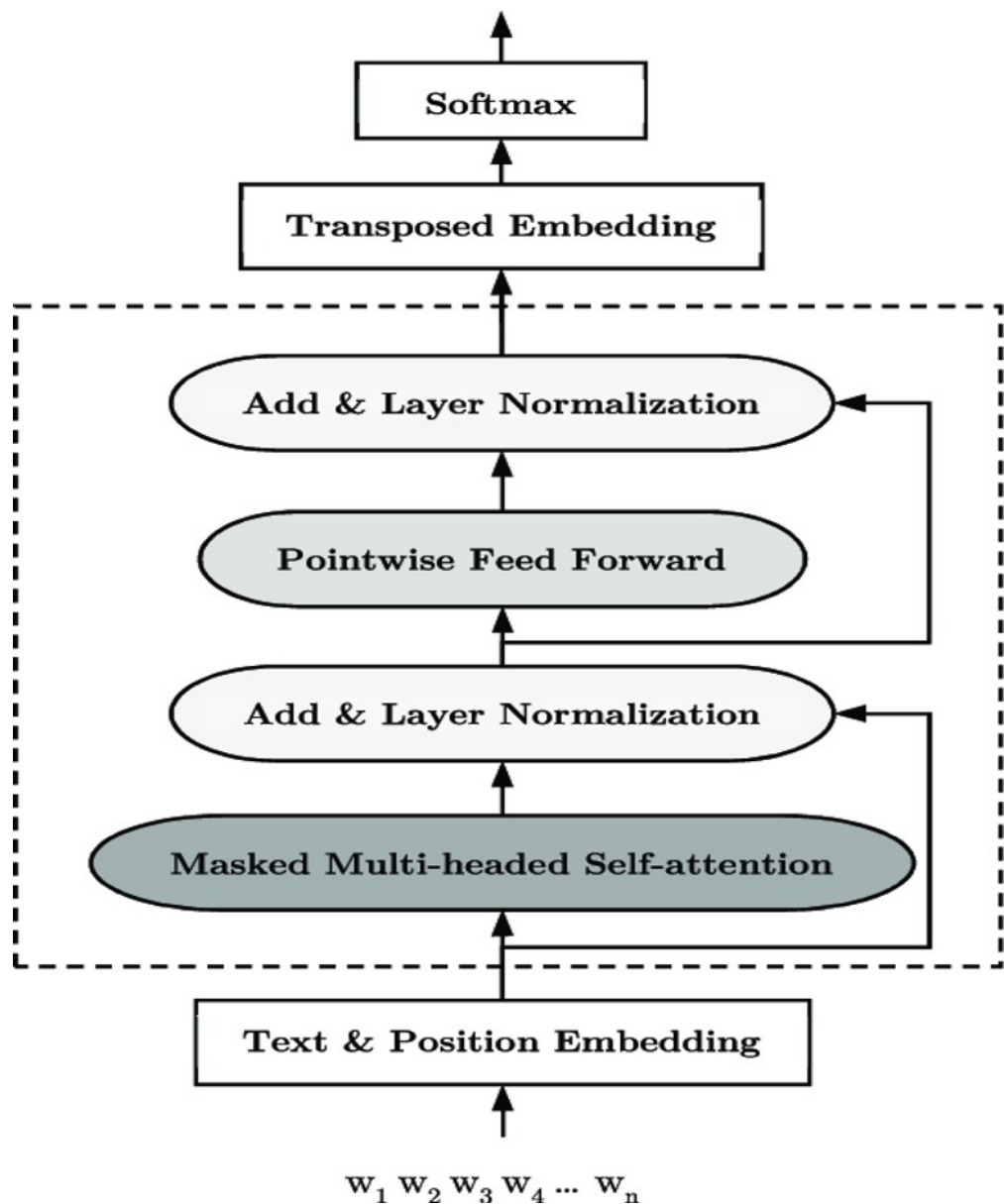


Figure 3.5: GPT-2 Architecture

The key components of GPT-2's architecture are:

- GPT-2 solely utilizes the decoder component, unlike the original Transformer model which incorporated both encoder and decoder elements for tasks like machine transla-

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

tion. This decision was guided by the model's focus on autoregressive language modeling tasks, where predicting the next token in a sequence depends on preceding tokens. By exclusively emphasizing the decoder, GPT-2 is adept at generating coherent and contextually appropriate text responses, rendering it suitable for various natural language processing applications.

- GPT-2 comprises multiple layers following the Transformer architecture. Each layer includes two key components: position-wise feedforward networks and a multi-head self-attention mechanism. Position-wise feedforward networks facilitate independent processing of information within each sequence position, while the multi-head self-attention mechanism allows the model to assess the importance of different words in the input sequence.
- In each layer of GPT-2, there is a position-wise feedforward network comprising fully connected layers. This structure allows the model to autonomously handle information within each position of the sequence. By transmitting input through interconnected neurons, the position-wise feedforward network aids in transforming information, enhancing the model's ability to understand text and generate coherent responses.
- In GPT-2, following each sub-layer, such as the feedforward network or the self-attention layer, two additional components are present: layer normalization and a residual connection. Layer normalization aids in stabilizing the outputs of each sub-layer to ensure consistent training performance. Meanwhile, the residual connection addresses the vanishing gradient problem by allowing the model to maintain information from the initial input, thereby promoting smoother gradient flow during training.

**Self-Attention Mechanism:** The self-attention mechanism is a crucial element found in many neural network architectures, particularly those involved in natural language processing tasks. This mechanism plays a pivotal role in the processing of input sequences before they are fed into the feedforward neural network. Through the self-attention mechanism, each position within the input sequence gains the ability to evaluate and prioritize different elements of the sequence, capturing intricate connections and dependencies between words. This capability is essential for the model to grasp the nuanced context embedded within the input sequence. By allowing each word to assess the significance and relevance of other words in the sequence, the self-attention mechanism facilitates a deeper understanding of how words interact with each other.

Consequently, the model can generate outputs that are more accurate and contextually appropriate, as it can recognize subtle semantic connections and contextual cues. In real-world applications, the self-attention mechanism significantly enhances performance across various natural language processing tasks. Its ability to capture complex dependencies and interrelationships within input sequences proves beneficial for tasks such as text generation, language translation, sentiment analysis, and document summarization. Therefore, the self-attention mechanism stands as a critical component of modern neural network architectures for natural language processing, as it improves the model's ability to produce outputs that

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

are coherent and contextually relevant.

Self-attention tracks the route of each token in the segment, with the Query, Key, and Value vectors being crucial. These vectors are vital for assessing the significance of a word and assembling its representations within the context.

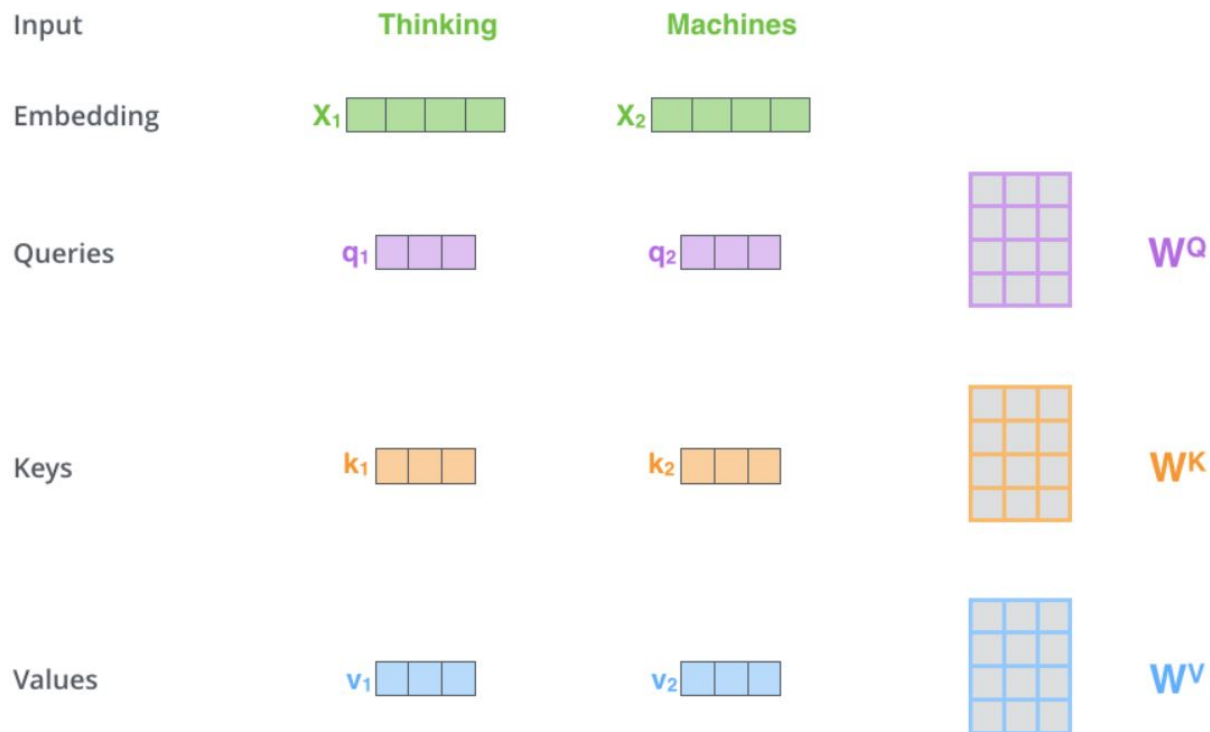


Figure 3.6: query,keys and values

- **Query:** The current word's representation, known as the query, is compared to each word's key within the sequence to assess its relevance. This process helps determine the significance of each word relative to the current one. For precise comparisons within the model, only the token's query under evaluation is taken into account during processing.
- **Key:** Key vectors act as identifiers for individual words within a segment, playing a vital role in locating and aligning relevant words during searches. By comparing the query with these vectors, the model can facilitate tasks such as information retrieval and semantic comprehension, identifying the most contextually significant words within the dataset.
- **Value:** Value vectors depict the real word representations within the context. After assessing the importance of each word, these values are combined to accurately represent the current term. In this process, words are assigned numerical values based on

their relevance, which are then aggregated to offer a comprehensive understanding of the word within the model's structure.

**Feedforward Neural Network (FFN):** After the self-attention mechanism's output, the input undergoes processing by a feedforward neural network (FFN), crucial for executing non-linear transformations and identifying complex patterns within the data. Typically, an activation function, often a Rectified Linear Unit (ReLU), is positioned between two linear transformations in the FFN structure. Initially, the input is transformed into a higher-dimensional space to introduce non-linearity via the ReLU activation function. Subsequently, the data is projected back to its original dimensionality through the second linear transformation, enabling the model to learn and encode intricate associations present in the input data. In neural network architectures, the FFN assumes a vital role by enabling the identification of potential non-linear patterns and relationships within the input. Through the introduction of non-linearity via activation functions like ReLU, the FFN facilitates the learning of more intricate data representations. Additionally, the FFN's multiple layers enable the extraction of hierarchical features, capturing increasingly abstract representations of the input data.

Furthermore, the FFN's capability to map the input data to a higher-dimensional space and then project it back to the original dimensionality improves the separability and discriminability of the data. This enhances the model's performance in classification or prediction tasks by aiding in the differentiation between various classes or categories within the data. In essence, the FFN plays a crucial role in neural network architectures by enabling non-linear transformations that allow the model to capture and express complex relationships within the input data. The model's ability to learn and extract relevant features from the data is facilitated by the FFN's structural design and sequential processing, ultimately enhancing its performance across a range of machine learning and artificial intelligence tasks.

### 3.5 PEFT (Parameter Efficient Fine Tuning)

As models grow larger and more complex, fully fine-tuning them on consumer hardware becomes increasingly challenging. This challenge is further intensified by the substantial storage demands associated with optimizing models separately for each downstream task, given that these optimized models retain the same size as the original pre-trained versions. For example, a GPT-2 model can expand to around 30 GB after being fine-tuned on a specific custom dataset. Managing numerous models for various tasks thus becomes a significant issue, driving up costs and complicating deployment logistics due to these high storage requirements.

Parameter-Efficient Fine-Tuning (PEFT) methods have been devised to address these challenges. By focusing only on a small subset of new parameters, PEFT significantly reduces the number of parameters needing fine-tuning while leaving the majority of the pre-

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

trained model's parameters unchanged. This approach lowers the required computing power and storage, and it also mitigates the risk of catastrophic forgetting—when a model loses its pre-trained knowledge after being fully optimized for a new task. By preserving most of the pre-trained parameters, PEFT ensures that the original model's broad capabilities remain intact. This enhances the efficiency and effectiveness of customizing large language models for specific tasks without the high costs associated with full fine-tuning. By utilizing only a small fraction of the total parameters, Parameter-Efficient Fine-Tuning (PEFT) has demonstrated excellent performance in various applications, such as text generation and image classification. This efficiency is achieved by merging the finely tuned weights with the original pre-trained weights, thus avoiding the need to fully retrain or store the model for each new task.

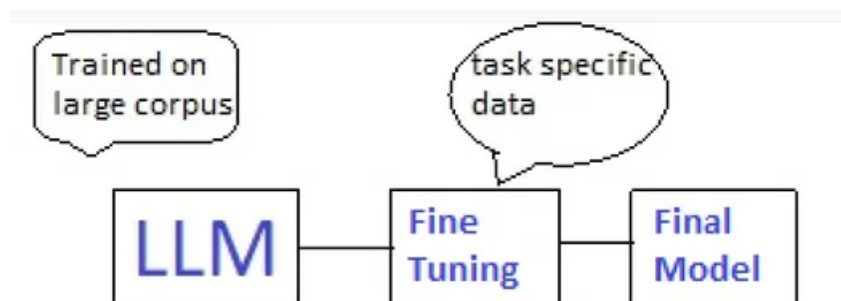


Figure 3.7: PEFT

The PEFT library simplifies this process by incorporating several popular techniques, including Low-Rank Adaptation (LoRA), Prefix Tuning, AdaLoRA, Prompt Tuning, MultiTask Prompt Tuning, and LoHa, into frameworks like Transformers and Accelerate. This integration allows users to easily access advanced large language models (LLMs) and benefit from efficient, scalable fine-tuning capabilities. LoRA is a prominent PEFT method known for reducing the parameter count of the GPT-2 model. It achieves this by implementing low-rank adaptations to the pre-trained model, significantly cutting down the number of trainable parameters. This approach modifies only a small subset of parameters while leaving the majority untouched. By selectively fine-tuning in this manner, LoRA minimizes the risk of catastrophic forgetting, a common issue where models lose previously learned information during full fine-tuning for new tasks. Additionally, this method lowers both computational overhead and storage requirements.

One example of how LoRA is strategically used in the GPT-2 design is how important it is to improving the model's scalability and efficiency. LoRA facilitates the process of adapting to different downstream demands by minimizing the amount of parameters that require fine-tuning. This increases the practicality and reduces the resource consumption of the fine-tuning process, making GPT-2 models more accessible to a wider variety of users and applications. In addition, LoRA's incorporation into the PEFT library, along with other

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

cutting-edge methods, highlights the library’s dedication to offering state-of-the-art model fine-tuning options. This guarantees that practitioners can create and implement highly effective and scalable LLMs that are customized to match the various requirements of contemporary AI applications by utilizing the most recent developments in PEFT.

### 3.5.1 LoRA (Low-Rank Adaptation of Large Language Models)

Due to its straightforward training approach, LoRA (Low-Rank Adaptation of Large Language Models) has gained popularity within the natural language processing field. LoRA incorporates trainable rank decomposition matrices into each layer of the Transformer architecture, effectively reducing the number of trainable parameters while maintaining model performance. This method involves freezing the majority of the original Large Language Model (LLM) weights, leading to improved memory efficiency and accelerated training. Consequently, the resulting models are notably smaller, typically a few hundred megabytes in size. This compactness facilitates easier sharing among researchers and practitioners and simplifies storage. LoRA’s ability to balance model performance with resource efficiency makes it an appealing option for developing more accessible and practical natural language processing solutions. The inclusion of LoRA during training marks a significant step forward in optimizing large language models. By strategically positioning rank decomposition matrices, LoRA achieves a delicate equilibrium between reducing parameters and preserving model performance, resulting in accelerated training and alleviating resource limitations. Furthermore, the ensuing reduction in model size creates fresh opportunities for diverse applications, particularly in environments with limited resources.

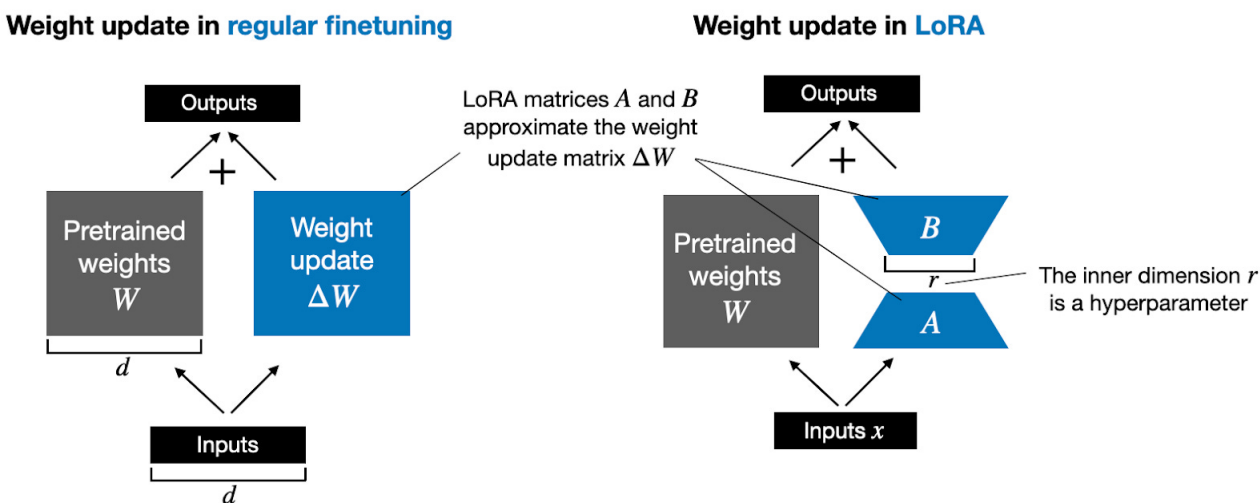


Figure 3.8: LoRA Model

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

Overall, LoRA epitomizes the ongoing quest for scalability and efficiency in artificial intelligence, offering promising avenues for further exploration and enhancement. Its potential to drive advancements in natural language processing and other domains is underscored by its ability to streamline training processes while upholding model quality. Hence, LoRA represents a notable milestone in the pursuit of more efficient and scalable AI solutions, poised to spark further innovations in the field.

Table 3.1: Comparison of size before and after applying LoRA

Technique	Model Size
Before Applying LoRA	30 GB (GigaBytes)
After Applying LoRA	124 MB (MegaBytes)

Table 3.1 provides a comparison of sizes before and after implementing LoRA. This technique enhances training efficiency and reduces the hardware requirements by up to threefold, particularly evident when utilizing adaptive optimizers. With LoRA, the need for calculating gradients or maintaining optimizer states for most parameters is eliminated. Instead, the focus is primarily on optimizing the injected, significantly smaller low-rank matrices. This streamlined approach not only accelerates the training process but also minimizes the computational burden, making it more accessible for a broader range of hardware configurations and lowering the barrier to entry for utilizing large language models in various applications.

## Chapter 4

# Experimental Analysis and Results

### 4.1 Experiments on data pre-processing task

This section outlines the data pre-processing procedures required to prepare the text in the training dataset for model input. Initially, the texts varied significantly in length, so they were resized to a standard size to ensure consistency and compatibility with the model. After resizing, tokenization is the next essential pre-processing step. Tokenization involves breaking the text into smaller, more manageable chunks known as tokens, which can be letters, words, or subwords. This step is crucial for converting raw text into a format that the model can process. For this task, the GPT-2 Tokenizer from the Transformers library is utilized. This tokenizer is specifically designed for use with the powerful transformer-based language model, GPT-2.

GPT-2 uses Byte Pair Encoding (BPE) for tokenization. BPE begins with the text's individual characters and iteratively merges the most frequent character pairs to form new tokens. This process continues until the entire text is tokenized into a sequence of these newly formed tokens. The advantage of BPE lies in its ability to handle rare or unknown words effectively by breaking them down into more common subword units. This approach ensures that the model can manage a wide range of vocabulary, including rare or unique terms, thereby enhancing its comprehension and generation accuracy. Overall, these pre-processing steps are crucial for optimizing text data, enabling the model to understand and generate relevant responses based on the input data.

### 4.2 Experimental Setup

#### 4.2.1 Environment Used

The experiments conducted with Large Language Model (LLM), specifically Generative Pre-trained Transformer 2 (GPT-2), were performed using Google Colab, a cloud-based Python Integrated Development Environment (IDE). The environment utilized a GPU with specifications including 1x Tesla K80, compute 3.7, comprising 2496 CUDA cores, and 12GB GDDR5 VRAM. This setup facilitated efficient processing and training of the models on large datasets. Google Colab's access to GPU resources provided computational power necessary for training deep learning models like GPT-2, enabling researchers and developers to

explore and experiment with complex natural language processing tasks effectively.

Web scraping was conducted using Microsoft's Visual Studio, a source-code editor compatible with Windows, Linux, macOS, and web browsers. The process was executed on a high-performance computer featuring an Intel Core i7 8th generation processor, a 256 GB SSD, 16 GB of RAM, and an Nvidia GeForce GTX 1060 graphics card. The robust infrastructure ensured smooth and efficient execution of the web scraping operations. Visual Studio's versatility and ability to work across different platforms made it a suitable choice for performing the scraping process effectively on various operating systems and configurations.

### 4.3 Data Augmentation

By employing data augmentation techniques, the dataset underwent expansion with the primary aim of augmenting both its volume and diversity. Various methods within data augmentation were utilized to generate additional data points from existing ones, thus artificially inflating the dataset's size. These techniques involve making minor adjustments to existing data or generating entirely new instances using deep learning models. Such supplementation of the dataset provides machine learning models with access to a larger and more diverse training set, potentially enhancing their performance and outcomes. Augmentation approaches are indispensable for furnishing training datasets with a broader and more comprehensive range of cases, thereby enhancing the model's ability to generalize and make accurate predictions on unfamiliar inputs.

Through data augmentation, the model's robustness to real-world data is bolstered, and the risk of overfitting is mitigated by exposing it to a wider array of scenarios and variations. Ultimately, data augmentation serves as a valuable strategy for improving the quality of training datasets and enhancing the effectiveness of machine learning models across various applications and fields. It contributes to overall model performance and generalization capabilities by providing models with a larger and more diverse set of examples to learn from. With the advancement of machine learning, the importance of data augmentation in ensuring model quality and reliability is expected to grow, making it an indispensable tool in the machine learning toolkit.

To enhance the accuracy and efficiency of machine learning models, access to extensive and diverse datasets is crucial. However, the process of gathering and annotating data can be both time-consuming and expensive. Consequently, many businesses turn to data augmentation strategies to supplement their datasets and alleviate these operational challenges, thereby reducing costs. These modifications enhance the model's robustness and its ability to handle real-world scenarios by introducing variations into the data. Synonym replacement is a widely employed data augmentation technique for this purpose, enriching the dataset's vocabulary and enhancing the model's understanding of subtle linguistic nuances through the substitution of terms with synonyms.

The implementation of data augmentation techniques significantly enhances machine learning models' overall performance and adaptability across various scenarios. By exposing models to a broader range of examples, data augmentation aids in improving their gener-

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

alization capabilities and accuracy in predicting unseen data. Furthermore, it reduces the expenses associated with data labeling and collection, making machine learning applications more accessible and cost-effective for businesses. Overall, the utilization of data augmentation techniques, including synonym replacement, is imperative for enhancing the effectiveness and efficiency of machine learning models. These methods enable models to manage complex and diverse datasets more efficiently by introducing modifications and expanding their vocabulary, ultimately resulting in improved performance and cost reduction in machine learning applications.

### 4.3.1 Synonym replacement

The objective of synonym replacement is to enhance the diversity of text data by substituting terms with suitable synonyms. Utilizing resources like WordNet, a comprehensive lexical database, simplifies the process of identifying relevant synonyms. By introducing subtle linguistic variations into the dataset, this augmentation method enriches it and strengthens the model's resilience to diverse linguistic patterns. This approach broadens the dataset's lexicon by incorporating synonyms, potentially enhancing the model's understanding of subtle language nuances. Moreover, it exposes the model to a wider range of language expressions, thereby improving its ability to handle various inputs and addressing concerns related to overfitting. In conclusion, synonym replacement is a valuable technique in data augmentation, contributing significantly to the overall robustness and effectiveness of natural language processing models across tasks such as text generation, classification, and sentiment analysis.

id	question	new_question
1	What was the first European trading company to establish a foothold in India?	What was the first european deal company to establish a foothold in India?
2	What key Battle helped solidify British influence in India?	What key battle helped solidify brits influence in India?
3	Who was a key figure in Indian social reform movements during British rule?	who was a key figure in Indian mixer reform movements during British rule?
4	What was the Sepoy Mutiny of 1857?	What was the Sepoy mutiny of 1857?
5	What policy aimed to assimilate Indians into British culture?	What policy aimed to assimilate american indian language into british people culture?
6	What was the significance of the Indian National Congress (INC)?	What was the meaning of the Indian subject Congress (INC)?
7	Who were two prominent leaders of the Indian independence movement?	Who were two big leadership of the Indian independence movement?
8	What was the impact of the Rowlatt Act of 1919?	What was the bear upon of the Rowlatt play of 1919?
9	What was the significance of the Government of India Act of 1935?	What was the implication of the politics of India Act of 1935?
10	What was the Quit India Movement?	What was the stop bharat Movement?
11	What role did World War II play in the Indian independence movement?	What role did World war II maneuver in the Indian independence movement?
12	What were the key figures behind the partition of India in 1947?	What were the key figures derriere the partition off of India in 1947?
13	What were the consequences of the partition of India?	What were the import of the partition off of India?
14	What economic impact did British rule have on India?	What economical impact did British prevail have on India?
15	What social and cultural changes occurred in India during British rule?	What societal and ethnic changes occurred in India during British rule?
16	How did the British justify their rule in India?	How did the british people justify their regulation in India?
17	What are some lasting legacies of British rule in India?	What are some endure legacies of brits rule in India?

Figure 4.1: Sample of synonym replacement on question

# STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

Sl.No	answer	new_answer
1	The British East India Company, arriving in 1600 with permission from Mughal Emperor Akbar.	The British East India Company, arriving in 1600 with license from Mughal emperor Akbar.
2	The Battle of Plassey (1757), where a smaller British force strategically defeated the Nawab of Bengal.	The Battle of Plassey (1757), where a smaller British force strategically disappointed the nabob of Bengal.
3	Raja Ram Mohan Roy, advocating against Sati (widow burning) and founding the Brahmo Samaj.	rajah jam Mohan Roy, advocating against Sati (widow burning) and founding the Brahmo Samaj.
4	A large-scale rebellion by Indian soldiers against the British East India Company, ultimately leading to British Crown rule.	A large-scale rising by Indian soldiers against the brits East India Company, ultimately leading to brits Crown rule.
5	The English Education Act of 1835, promoting English language education and Western ideas.	The side education Act of 1835, promoting side language education and Western ideas.
6	A major political party formed in 1885, spearheading the Indian independence movement.	ampere major political company formed in 1885, spearheading the Indian independence movement.
7	Mahatma Gandhi, advocating for non-violent resistance, and Bhagat Singh, known for his revolutionary actions.	mahatma Gandhi, preach for non-violent resistance, and Bhagat Singh, known for his revolutionary actions.
8	It granted the British broad emergency powers, leading to protests like the Jallianwala Bagh Massacre.	it granted the British broad emergency powers, leading to protests alike the Jallianwala Bagh Massacre.
9	It introduced limited self-government in India, but also solidified the separation of British India and Princely States.	It introduced limited self-government in India, but also solidified the legal separation of British india and Princely States.
10	A nationwide campaign launched by Gandhi in 1942, demanding an immediate British withdrawal from India.	A nationwide campaign launched by Gandhi in 1942, need an immediate british people withdrawal from India.
11	It weakened British power and boosted Indian nationalism, with movements like the Azad Hind Fauj.	It weakened British power and hike Indian nationalism, with campaign like the Azad Hind Fauj.
12	Muhammad Ali Jinnah, leader of the Muslim League, and Jawaharlal Nehru, representing the INC.	muhammad cassius clay Jinnah, leader of the Muslim League, and Jawaharlal Nehru, representing the INC.
13	Mass displacement and violence along religious lines, creating the independent nations of India and Pakistan.	Mass shift and violence along religious lines, creating the independent nations of india and Pakistan.
14	Exploitation of resources for British benefit, while neglecting Indian infrastructure and development.	victimization of resource for British benefit, while neglecting Indian infrastructure and development.
15	A rise in Western education, changes in dress and customs, and the emergence of new social movements.	A rise in western sandwich education, changes in dress and customs, and the emergence of new mixer movements.
16	Often through the concept of the "White Man's Burden," claiming to bring civilization and progress to India.	often through the concept of the "White Man's Burden," claiming to bring civilization and build up to India.
17	The English language, a parliamentary system of government, and a complex social and religious landscape.	The English language, a parliamentary organization of government, and a complex mixer and religious landscape.
18	With a critical eye, acknowledging both modernization efforts and the exploitation of India's resources and people.	With a critical eye, acknowledging both modernisation feat and the exploitation of India's resources and people.

Figure 4.2: Sample of synonym replacement on answer

This approach leverages synonyms from WordNet, a comprehensive lexical database, to subtly yet effectively enhance the dataset. The goal is to improve the Insurance Kaggle dataset, providing the machine learning model with a more complete and representative training set. By incorporating synonyms, the dataset is enriched with a broader range of linguistic terms, potentially boosting model performance. The method involves systematically replacing certain terms in the dataset with their synonyms. WordNet's extensive database of English terms, including definitions, synonyms, antonyms, and hyponyms, makes it an ideal tool for this task. This synonym replacement process expands the dataset's vocabulary, preventing the model from overfitting to a narrow set of terms and enhancing its ability to generalize across various inputs.

Table 4.1 illustrates a comparison of the dataset before and after synonym integration. This comparison highlights the increased linguistic diversity and expanded vocabulary. Introducing synonyms enriches the dataset, exposing the model to a wider array of term variations. This is particularly valuable in the context of insurance-related documents, where terminology can vary significantly. The objective is to enhance the model's language comprehension rather than merely adding more words to the dataset. Synonym replacement enables the model to recognize different ways of expressing the same concept, improving its adaptability and ability to process diverse linguistic inputs. For example, terms like "claim" and "assertion" or "premium" and "cost" can be used interchangeably in different contexts. By learning these variations, the model can more accurately process and analyze real-world data, resulting in more reliable predictions and insights. Ultimately, integrating synonyms from WordNet significantly enhances the quality and robustness of the training data. This leads to machine learning models that are more effective and accurate in handling insurance-related tasks, improving their practical utility and performance in real-world applications.

Table 4.1: Comparison on dataset-1 before and after data augmentation

Model	Before Augmentation	After Augmentation
<b>ROUGE Score</b>	0.188	0.692
<b>BLEU Score</b>	1.483 e-78	0.498

## 4.4 Results

Evaluation metrics like BLEU Score and ROUGE Score are widely utilized in appraising text generation models. The primary objective of model evaluation is to determine the extent to which the model produces the desired output. This evaluation involves comparing the machine-generated text with either reference text or an intended result predefined by the model. Hence, BLEU Score and ROUGE Score are pivotal in this assessment process. While ROUGE Score evaluates the quality of machine-generated summaries against human-authored summaries, BLEU Score measures the similarity between machine-generated text and reference texts. These metrics furnish researchers and developers with crucial insights into their model's performance, enabling them to identify strengths and weaknesses in their text generation algorithms and improve overall efficacy.

### 4.4.1 BLEU Score

BLEU (BiLingual Evaluation Understudy) is a metric employed to automatically evaluate machine-generated text. It provides a BLEU score, ranging from zero to one, assessing the similarity between machine-generated text and a set of high-quality reference texts. This precision-centric measure calculates the overlap of n-grams shared between the generated and referenced texts. A BLEU score of 1 indicates a perfect match with the reference translations, reflecting excellent quality, while a score of 0 signifies no overlap, indicating poor quality in the machine-generated output.

BLEU serves as a crucial resource for researchers and developers by enabling the unbiased evaluation of machine-generated text quality compared to reference texts. It allows for thorough analysis of diverse natural language processing algorithms and models. Its precision-oriented approach enables detailed scrutiny of linguistic accuracy and quality in machine-generated output, offering valuable insights for enhancing future text generation systems. Ultimately, BLEU plays a pivotal role in guiding the refinement and advancement of natural language processing technologies by providing a standardized framework for evaluating the strengths and weaknesses of text generation models.

$$BLEU = (\text{BLEU} = \text{AVG}(\text{precision across different range of n-gram sizes})) \quad (4.1)$$

### 4.4.2 ROUGE Score

Rouge and BLEU metrics differ mainly in their focus on distinct aspects of text assessment. Rouge prioritizes recall over precision, contrasting with BLEU's emphasis on precision. Rouge was specifically designed to evaluate machine-generated summaries against human-authored reference summaries. It accomplishes this by examining the overlap of n-grams, which are word sequences shared by both the machine-generated and reference texts. Trigrams, bigrams, and unigrams are commonly employed n-grams for evaluation purposes. The Rouge score quantifies the similarity between the two texts by comparing the recall of n-grams present in the machine-generated text to those in the reference text.

This comprehensive analysis enables researchers and developers to identify areas where machine-generated summaries may deviate from human-written ones, offering valuable insights for refining summarization algorithms and enhancing the quality of machine-generated summaries in various natural language processing applications. Rouge serves as a crucial metric for evaluating the effectiveness of machine-generated summaries in conveying the main ideas of source material, placing a specific emphasis on recall. This comprehensive evaluation aids in identifying areas where machine-generated summaries differ from those created by humans, offering valuable insights for academics and developers to refine their summarization algorithms.

In contrast, BLEU focuses primarily on the precision of exact word matches between reference texts and machine-generated texts. When used together, these evaluation metrics provide comprehensive tools for assessing and improving the quality of machine-generated summaries across various natural language processing applications. By leveraging both Rouge and BLEU, developers can enhance the precision and efficiency of summarization systems, gaining a better understanding of the strengths and weaknesses of their summarization models.

The formula for ROUGE score is as follows:

$$\text{Rouge-1 (Recall)} = \frac{\text{Unigram Matches (b/w generated and reference)}}{\text{Unigram in reference text}} \quad (4.2)$$

$$\text{Rouge-1 (Precision)} = \frac{\text{Unigram Matches (b/w generated and reference)}}{\text{Unigram in generated text}} \quad (4.3)$$

$$\text{Rouge-1 (F-1)} = \frac{2 \times \text{Precision} \times \text{Recall}}{(\text{Precision} + \text{Recall})} \quad (4.4)$$

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

The final results obtained from the model is shown in table 4.2, figure 4.3 and figure 4.4 is obtained by training the GPT-2 model on insurance dataset:

Table 4.2: Evaluation metrics 1

<b>ROUGE-1 Score (Recall)</b>	0.583
<b>ROUGE-1 Score (Precision)</b>	0.614
<b>ROUGE-1 Score (F-1)</b>	0.687
<b>BLEU Score</b>	0.498

```
[Q] Can Veterans Get Life Insurance?  
[A] The Veterans Administration can get life insurance through their state  
insurance commissioner, but they are only allowed to use it for military purposes.
```

Figure 4.3: Sample Output 1

```
[Q] Is Car Insurance Prepaid?  
[A] Yes, automobile insurance is typically paid in advance.  
Normally no less than thirty days at a time.
```

Figure 4.4: Sample Output 2

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

The final results obtained from the model are presented in Table 4.3, Table 4.4, and Table 4.5. Additionally, Figure 4.5 and Figure 4.6 illustrate key findings and visualizations derived from the data. These results were achieved by training the GPT-2 model on a dataset focused on Indian history.

Table 4.3: Evaluation metrics 2

<b>ROUGE-1 Score (Recall)</b>	0.659
<b>ROUGE-1 Score (Precision)</b>	0.719
<b>ROUGE-1 Score (F-1)</b>	0.687
<b>BLEU Score</b>	0.452

Table 4.4: Evaluation metrics 3

<b>ROUGE-2 Score (Recall)</b>	0.571
<b>ROUGE-2 Score (Precision)</b>	0.571
<b>ROUGE-2 Score (F-1)</b>	0.571
<b>BLEU Score</b>	0.452

Table 4.5: Evaluation metrics 4

<b>ROUGE-3 Score (Recall)</b>	0.333
<b>ROUGE-3 Score (Precision)</b>	0.333
<b>ROUGE-3 Score (F-1)</b>	0.333
<b>BLEU Score</b>	0.452

[Q] What was the Sepoy Mutiny of 1857?  
[A] A large-scale rebellion by Indian soldiers against the British East India Company, ultimately leading to British Crown rule.

Figure 4.5: Sample Output 3

[Q] What was the first European trading company to establish a foothold in India?  
[A] The British East India Company, arriving in 1600 with license from Mughal Emperor Akbar.

Figure 4.6: Sample Output 4

The final results obtained from the model after applying LoRA are presented in Table 4.6, Table 4.7, and Table 4.8. Additionally, these results are visually depicted in Figure 4.7. These tables and figure collectively illustrate the model's performance and the impact of the LoRA application.

Table 4.6: Evaluation metrics 5

<b>ROUGE-1 Score (Recall)</b>	0.437
<b>ROUGE-1 Score (Precision)</b>	0.5
<b>ROUGE-1 Score (F-1)</b>	0.466
<b>BLEU Score</b>	0.311

Table 4.7: Evaluation metrics 6

<b>ROUGE-2 Score (Recall)</b>	0.2352
<b>ROUGE-2 Score (Precision)</b>	0.285
<b>ROUGE-2 Score (F-1)</b>	0.25806
<b>BLEU Score</b>	0.311

Table 4.8: Evaluation metrics 7

<b>ROUGE-3 Score (Recall)</b>	0.0625
<b>ROUGE-3 Score (Precision)</b>	0.07692
<b>ROUGE-3 Score (F-1)</b>	0.06896
<b>BLEU Score</b>	0.311

[Q]What was the impact of the Rowlatt Act of 1919?

[A]It granted broad emergency powers to the British, sparking protests and violence everybody the Indian revolutionary movement.

Figure 4.7: Sample Output 5

## Chapter 5

# Conclusion and Future scope

The primary aim is to develop a deep learning model capable of delivering personalized and immediate responses to students' questions, thereby enriching their educational journey. This model accommodates various forms of interaction, supporting both text and audio inputs to cater to individual preferences. Textual inputs undergo preprocessing to ensure their suitability for analysis. Utilizing the dataset obtained through web scraping, the model is trained to recognize patterns within the provided information. This training equips the deep learning model with the necessary skills and expertise to generate accurate and tailored responses to students' queries. This holistic approach aims to optimize the educational experience by harnessing advanced technology to offer students timely and personalized assistance.

The Generative Pre-trained Transformer 2 (GPT-2) is a widely recognized deep learning model chosen for its proficiency in understanding and generating coherent text. After training on the original dataset, GPT-2 analyzes patterns in the input data to generate personalized feedback for individual students. Following this, a validation process compares the model-generated text with the reference text to assess the relevance and accuracy of the feedback. Evaluation metrics, including ROUGE and BLEU scores, are then computed based on this comparison. Unfortunately, both BLEU and ROUGE scores yielded notably low results in the assessment. These outcomes indicate disparities between the generated feedback and the reference text, revealing potential deficiencies in the model's ability to provide precise and pertinent responses. Rectifying these issues is crucial to enhancing the dependability and efficacy of the feedback system in delivering valuable insights to students.

Data augmentation techniques, such as synonym replacement for dataset expansion, were employed to enhance the BLEU and ROUGE scores. By substituting synonyms for words, the dataset was expanded, providing the model with a broader range of examples. Furthermore, parameter-efficient fine-tuning methods like Low Rank Adaption were utilized to reduce the model's size. This reduction in size improves the model's computational efficiency and memory usage. Consequently, the model can produce results more efficiently, potentially maintaining or even improving the quality of generated responses while accelerating feedback delivery. These strategies synergize to enhance the efficiency and effectiveness of the feedback system, ultimately offering students more accurate and insightful information.

## **STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH**

---

Additionally, the model undergoes specialized training on an Indian history dataset, granting it in-depth knowledge within this domain. Its proficiency in addressing queries related to Indian history stems from this focused training. However, owing to its flexibility, the model can also undergo training with datasets spanning diverse disciplines. Through this process, it gains insights into fields such as science, literature, and geography. Consequently, users can pose questions on these varied subjects, receiving accurate and relevant responses from the model. Its adaptability positions the model as an effective tool for delivering precise information across a broad spectrum of topics.

# References

- [1] Karpagam, K., K. Madusudanan, and A. Saradha. "Deep learning approaches for answer selection in question answering system for conversation agents." *ICTACT Journal on Soft Computing*, 2020, 10 (2): 2040-2044 (2020).
- [2] Nuruzzaman, Mohammad, and Omar Khadeer Hussain. "IntelliBot: A Dialogue-based chatbot for the insurance industry." *Knowledge-Based Systems* 196 (2020): 105810.
- [3] Tan, Ming, et al. "Improved representation learning for question answer matching." *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2016.
- [4] Criscuolo, Marcelo, et al. "MilkQA: a dataset of consumer questions for the task of answer selection." *2017 Brazilian Conference on Intelligent Systems (BRACIS)*. IEEE, 2017.
- [5] Serban, Iulian, et al. "A hierarchical latent variable encoder-decoder model for generating dialogues." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 31. No. 1. 2017.
- [6] Grenander, Matt, et al. "Deep discourse analysis for generating personalized feedback in intelligent tutor systems." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. No. 17. 2021.
- [7] Song, Linfeng, et al. "CASA: Conversational aspect sentiment analysis for dialogue understanding." *Journal of Artificial Intelligence Research* 73 (2022): 511-533.
- [8] Ice, Philip, et al. "Using asynchronous audio feedback to enhance teaching presence and students' sense of community." *Journal of Asynchronous Learning Networks* 11.2 (2007): 3-25.
- [9] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).
- [10] Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).
- [11] Radford, Alec, et al. "Language models are unsupervised multitask learners." *OpenAI blog* 1.8 (2019): 9.
- [12] Vinyals, O. and Le, Q. A neural conversational model. *arXiv preprint arXiv:1506.05869*, 2015.

## STUDENT-CENTERED AUTOMATED FEEDBACK USING TRANSFORMER BASED APPROACH

---

- [13] Al-Rfou, R., Choe, D., Constant, N., Guo, M., and Jones, L. Character-level language modeling with deeper self-attention. arXiv preprint arXiv:1808.04444, 2018.
- [14] A. M. Dai and Q. V. Le. Semi-supervised sequence learning. In *Advances in Neural Information Processing Systems*, pages 3079–3087, 2015.
- [15] Oleksii Kuchaiev and Boris Ginsburg. Factorization tricks for LSTM networks. arXiv preprint arXiv:1703.10722, 2017.
- [16] Sepp Hochreiter, Yoshua Bengio, Paolo Frasconi, and Jürgen Schmidhuber. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies, 2001.
- [17] Sainbayar Sukhbaatar, arthur szlam, Jason Weston, and Rob Fergus. End-to-end memory networks. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 2440–2448. Curran Associates, Inc., 2015.