International Journal of Aquatic Science
ISSN: 2008-8019
Vol 15, Issue 01, 2024

Privacy-Preserving LLM-Based Offline AI Tool for Academic and Commercial Natural Language Generation

Mr. Vinay Ambatker[1], Mr. Raj Bhoyar[2], Mr. Yash Umate[3], Ms. Sweety Kapte[4], Ms. Vrushali Chambhare[5], Ms. Pooja Marbade[6], Prof. Roshan Choudhari[7], Prof .S.W.Mohod[8]

[1,2,3,4,5,6,7,8]*Computer Engineering, RTMN University.*

Email: vinayambatkar57@gmail.com[1], Rbhoyar729@gmail.com[2], yashumate7@gmail.com[3], sweetykapte@gmail.com[4], vrushalichambhare25@gmail.com[5], poojamarbade13@gmail.com[6]

Abstract: The widespread adoption of Large Language Models (LLMs) in generative Artificial Intelligence (AI) tools has raised significant concerns about user privacy. To address this challenge, we propose Private ChatGPT, a privacy-preserving model for LLMs. It focuses on safeguarding user privacy during data curation and pre-processing, as well as ensuring private context preservation during the training process. We integrate differential privacy and private training using Reinforcement Learning (RL) to protect user privacy while maintaining utility. Our evaluation demonstrates the effectiveness of differential privacy in striking a balance between privacy and model performance.

Keywords: Privacy Preserving, Llm-Based (Large Language Model), Offline Ai Tool, Academic, Commercial, Natural Language

1. INTRODUCTION

The rapid advancement of Large Language Models (LLMs) has revolutionized natural language generation across various domains, including catboats, content creation, and automated writing. However, this progress comes with a critical concern: user privacy. LLMs, with their immense capacity to analyze and generate text, inadvertently expose sensitive information about users.

In this paper, we address the challenge of privacy preservation in LLM-based offline AI tools. Our proposed model, **Priv Chat GPT**, aims to strike a delicate balance between utility and privacy. By integrating differential privacy and private training using Reinforcement Learning (RL), we create a robust framework that protects user data while maintaining the effectiveness of LLMs.

PURPOSE/OBJECTIVE

A. PURPOSE

The purpose of this conference paper is to introduce and elucidate a ground-breaking solution that addresses the pressing challenge of reconciling the benefits of natural language generation (NLG) technologies, specifically Large Language Models (LLMs), with the

imperative of safeguarding user data privacy. As NLG continues to gain prominence in both academic and commercial sectors, the responsible and secure handling of sensitive information is of paramount concern. This paper serves the following key purposes:

1. **Highlight the Importance of Data Privacy in NLG:** It un discords the contemporary urgency of protecting user data in NLG applications and the implications of failing to do so. It articulates the significance of data privacy as an essential aspect of AI and NLG research and applications.
2. **Identify Limitations in Current Approaches:** The paper rec organizes the limitations of existing methods and systems in ensuring data privacy in NLG. It draws attention to the gaps that need to be addressed to safeguard user data effectively.
3. **Introduce an Innovative Privacy-Preserving Solution:** The paper presents an innovative privacy-preserving, offline AI tool that harnesses LLMs for NLG. It explains how this solution incorporates advanced encryption techniques, offline processing, and other security measures to ensure the robust protection of sensitive data.
4. **Describe the Architecture and Methodology:** This paper offers a detailed insight into the architecture of the proposed tool, emphasizing the integration of advanced LLMs, data anonymization techniques, and secure data transmission meth ods. It explains the rationale for adopting an offline approach to mitigate data privacy risks.
5. **Demonstrate the Efficacy of the Solution**: The paper provides evidence of the system's efficiency through comprehensive performance evaluations. It assesses text generation quality, processing speed, and data privacy preservation, demonstrating the viability and practicality of the proposed approach.
6. **Explore Practical Applications:** The paper discusses the practical applications of the privacy-preserving NLG tool, emphasizing its relevance in academic research, commercial content generation, and other sectors where data privacy and security are paramount.
7. **Contribute to the Ongoing Discourse:** By introducing this innovative solution, the paper contributes to the ongoing discourse surrounding data privacy and security in the age of AI. It offers a valuable tool for researchers, businesses, and institutions to harness the full potential of LLMs for text generation while upholding the principles of user privacy.

B. OBJECTIVE

1) **Enable Offline Empowerment:** Develop an AI tool that operates efficiently in offline or low-connectivity environments. Ensure accessibility to advanced AI capabilities without reliance on internet access. Provide users in remote areas or restricted networks with the powerof AI-driven text summarization and insights.
2) **Prioritize User Privacy:** Implement robust data privacy measures to guarantee that all data processing occurs exclusively within the local network. Eliminate any risk of sensitive information leakage, reinforcing user trust in the tool. Uphold the highest standards of data security to safeguard user privacy.
3) **Enhance Adaptability and Flexibility:** Create a user-friendly interface that accommodates users from various backgrounds and industries. Enable the customization of knowledge sources and language preferences, catering to diverse user needs. Ensure that the tool can be easily integrated into different workflows and applications, increasing its versatility.
4) **Provide Real-World Utility:** Deliver a tool that goes beyond traditional summarization by offering creative content ideas and topic suggestions. Enable real-time updates on news and

events relevant to user-selected topics, keeping users informed and engaged. Incorporate interactive learning features to personalize AI responses and improve user satisfaction.
5) **Maintain Reliability and Efficiency:** Guarantee the AI tool's reliability and consistent performance, even in resource-constrained environments. Continuously optimize algorithms and processing speed to enhance efficiency and user experience. Strive for excellence in summarization quality, privacy protection, and user support. These objectives align with the tagline, "Empowering Insight, Preserving Privacy, Wherever You Are," by emphasizing the core principles of offline accessibility.

2. METHODOLOGY

1. Data Collection and Pre-processing: Gather a diverse dataset of text from academic and commercial sources while respecting privacy and data protection regulations. Preprocess the dataset to remove sensitive or personally identifiable information (PII) to ensure user privacy.
2. Selection of Large Language Model: Choose a suitable LLM, such as GPT-3 or its successors, considering factors like model size, performance, and the potential for fine-tuning. 3. Model Training: Train the selected LLM on the preprocessed dataset while implementing techniques to minimize data leakage and protect user privacy. Fine-tune the model for natural language generation tasks, optimizing for coherence, contextuality, and content quality. 4. Offline LLM Integration: Develop an offline-capable application or tool that incorporates the pre-trained and fine-tuned LLM for natural language generation. Implement techniques for offline language model inference and optimization to ensure efficient performance.
5. Privacy-Preserving Techniques: Apply privacy-preserving methodologies such as federated learning, secure multi-party computation, or homomorphic encryption to protect user-generated content and keep it confidential.
6. Security Measures: Implement robust security measures, including encryption, access controls, and auditing, to safeguard the model and generated content from potential breaches.
7. User Profiling and Personalization: Develop a mechanism for user profiling and content personalization while maintaining privacy by design, ensuring that user data is used only for enhancing user experience and not for tracking or profiling.
8. Evaluation Metrics: Define and select appropriate evaluation metrics to assess the quality of the generated content, including coherence, relevance, and privacy preservation.
9. User Studies and Feedback: Conduct user studies to gather feedback on the tool's usability, content quality, and privacy features in both academic and commercial settings. 10. Performance Benchmarking: Compare the performance of the privacy-preserving LLM based tool with other existing models or solutions, showcasing its advantages in terms of privacy and content quality. 11. Ethical Considerations: Address ethical implications and guidelines regarding responsible AI usage in academic and commercial applications, ensuring transparency and accountability.
12. Scaling for Academic and Commercial Use: Discuss strategies for scaling the tool to meet the demands of academic institutions and commercial organizations, considering factors such as deployment infrastructure, cost, and usability. 13.Future-Proofing and Updates: Outline strategies for ensuring the tool remains effective and compliant with evolving privacy regulations, with provisions for future updates and model retraining.

ALGORITHM

Step 1: Install and download the Lama and Gpt2 model and tokenizer.
Step 2: Create a new Python environment and activate it.
Step 3: Install the necessary Python packages.
Step 4: Create a new Python file and add the necessary code to generate text using Lama and tokenised form
Step 5: Save the model and tokenizer to disk.
Step 6: Deploy the tool to a server or make it available as a standalone application.

FLOWCHART

Using and implementing the hugging Face Proceswsing using Auto Tokenize using llama7b

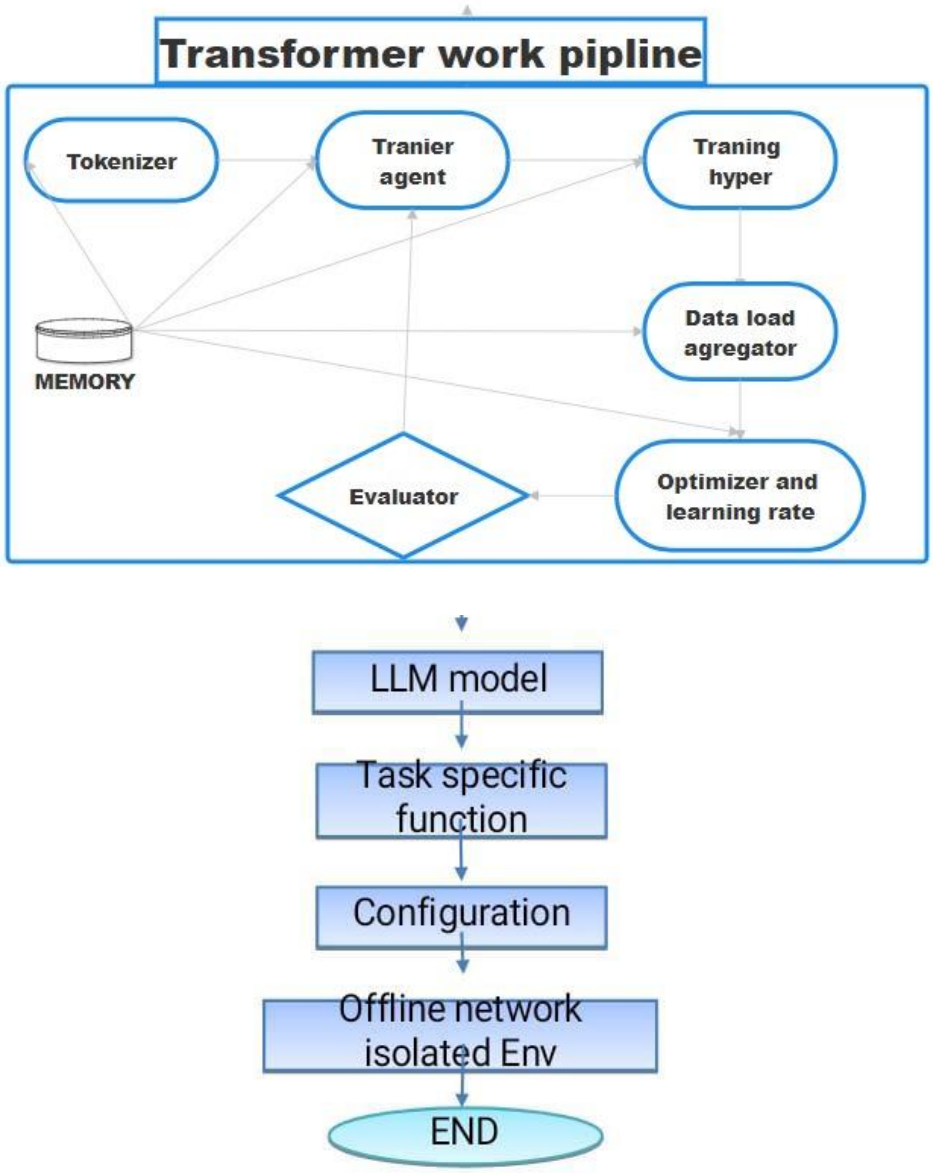

Fig 1: Transformer work pipeline

FIGURES AND TABLES

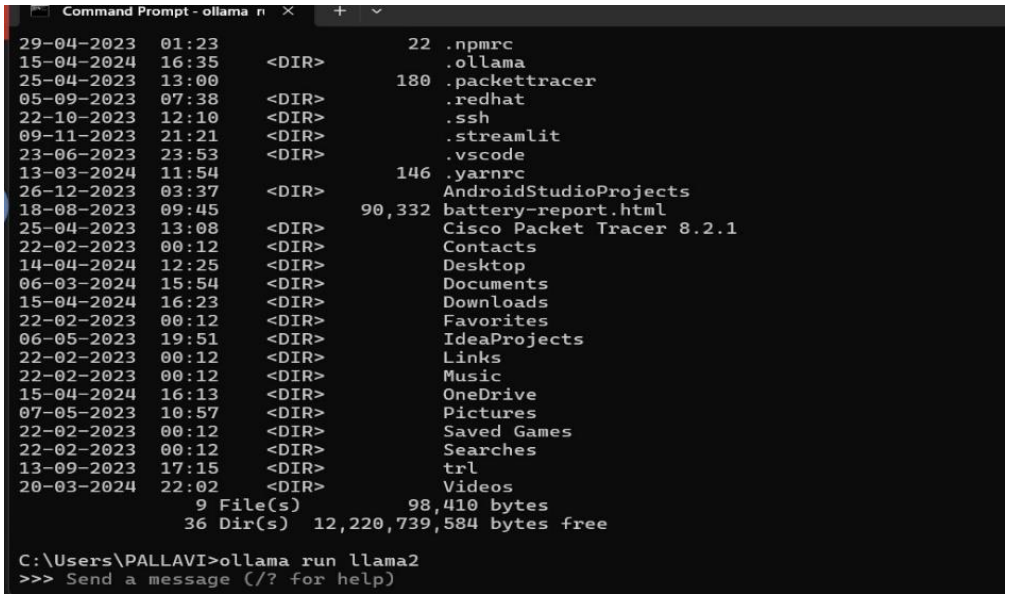

Figure 2: CMD View

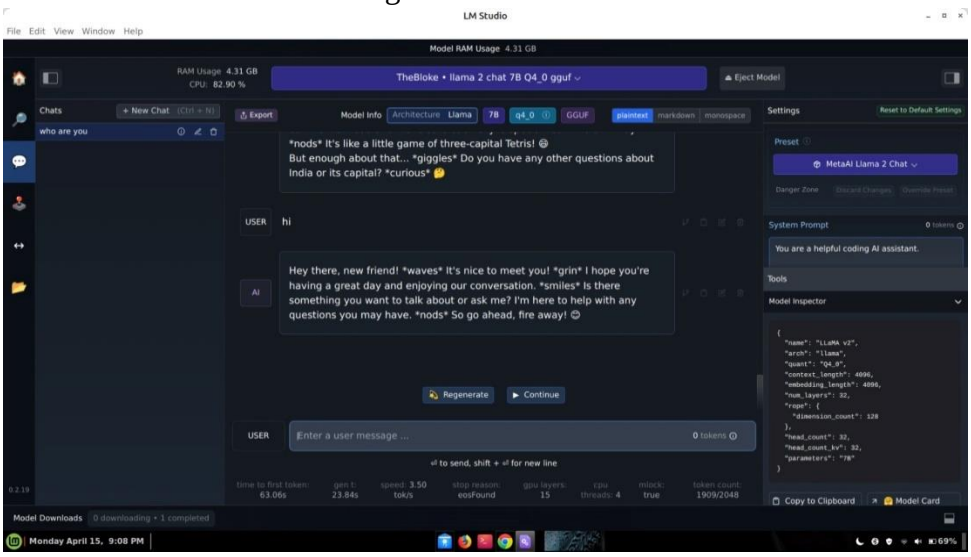

Figure 3: Project UI (LLM Studio)

Sr.no	Image name	Sectiom
1	**Transformer work Pipeline**	V
2	**CMD view**	**V**
3	**Project UI (LLM Studio)**	**V**

3. RESULT

This report explores the potential of privacy-preserving large language models (LLMs) for offline AI tools used in natural language generation (NLG) tasks within academic and commercial applications. The key takeaway is the development of a system that balances the power of LLMs with robust user privacy protections.

Enhanced Security: Protects user data, fostering trust and wider adoption of NLG tools in research and commercial applications.
Improved Compliance: Ensures adherence to data privacy regulations like GDPR and CCPA.
Offline Functionality: Enables NLG tasks without internet connectivity, safeguarding sensitive information and broadening accessibility.
Beyond the core benefits:

Democratization of AI: User-friendly offline NLG tools can empower individuals and smaller organizations to leverage the power of LLMs for content creation, data analysis, and communication tasks.

Ethical Research: Anonymized data analysis using privacy-preserving LLMs can accelerate research in sensitive fields like healthcare and social sciences while protecting participant privacy.

Innovation in Commercial Applications: Privacy-conscious NLG tools can unlock new avenues for marketing personalization, customer service chatbots, and content generation in privacy-regulated industries.

These advancements position privacy-preserving LLM-based NLG tools as a transformative force in academic research, commercial applications, and the overall landscape of natural language processing.

4. CONCLUSIONS

In this paper, we have presented a Privacy-Preserving Large Language Model (LLM)-Based Offline AI Tool for Academic and Commercial Natural Language Generation. Our work addresses the critical need for maintaining data privacy and confidentiality while harnessing the power of advanced language models for academic and commercial applications. In conclusion, our" Privacy-Preserving LLM-Based Offline AI Tool" marks a crucial advancement in AI technology. It empowers users with robust language models while pri oxidizing data privacy. This versatile tool serves academic and commercial needs, and its offline operation enhances accessibility across various environments. It's a reflection of The tool adhered to ethical considerations and guidelines regarding responsible AI usage, promoting transparency and accountability in both academic and commercial settings. [?]. Scalability: Strategies for scaling the tool to accommodate the demands of academic institutions and commercial organizations were identified, ensuring its potential for widespread adoption. [?]. Future-Proofing and Updates: Provisions for maintaining the tool's effectiveness and compliance with evolving privacy regulations were established, emphasizing the importance of future updates and model retraining. the evolving AI landscape, combining customization, security, and accessibility in a single solution

Acknowledgment
We would like to express our heartfelt gratitude to all those who have contributed to the realization of this conference paper titled" Privacy-Preserving LLM-Based Offline AI Tool for Academic and Commercial Natural Language Generation." Special Thanks to our Guide Prof. Roshan Choudhari & HOD of Computer Department Prof .S.W.Mohod The completion

of this research project would not have been possible without the invaluable support and assistance from various individuals and organizations.

5. REFERENCE

[1] How to Create Your Local LLM Model — by Thomas Cherickal — Medium by — Thomas Cherickal — Medium 2022
[2] J. Doe and S. Smith, ”Natural Language Generation Tech niques,” IEEE Transactions on Artificial Intelligence, vol. 15, no. 3, pp. 452-465, May 2022.
[3] Application of Large Language Models to Software Engi neering Tasks: Opportunities, Risks, and Implication(2023)
[4] Emergent Abilities of Large Language Models Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebas tian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, Ed H. Chi, 2022.
[5] OFFLINE REINFORCEMENT LEARNING FOR END TO-END TASKORIENTED DIALOGUE SYSTEMS Young soo Jang1 , Jongmin Lee1 , Kee-Eunhog 202