AI-Powered Story Generation Chatbot

Presented by TEAM ALPHA NLP

Origin of the creative idea

The development of creative chatbots is driven by the need for more natural, engaging, and efficient interactions in a digital world. By blending advanced AI technologies with creative design, chatbots not only meet these needs but also open up new avenues for innovation and user engagement.

Project vision and mission

Our goal here was to develop a fully functional and efficient chatbot that helps us generate stories with its multiple parameters ex: length, genre, plots etc.

01. Evaluate top LLM models like Llama 2, gpt 2, gpt 3.5 turbo, Llama 3 based on there efficiency and Data and pick one suitable for our project.

02. Develop a Interacticve UI and integrate it the backend code.

03. Delivering a Fully Functional and to the mark chatbot that performs the best what it is intended to do.

LLM MODELS

01
GPT 2.0

02
Llama 2

03
GPT 3.5 turbo

04
Llama 3 - 70b

Llama 3

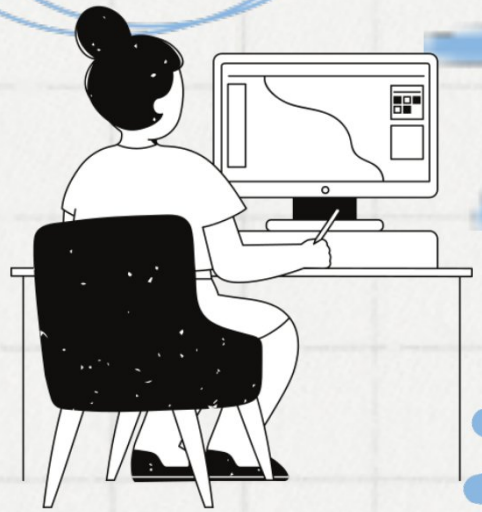
LLaMA 3 stands out as a powerful, efficient, and versatile language model, making it an excellent choice over GPT-2, GPT-3.5 Turbo, and LLaMA 2 for anyone looking to leverage cutting-edge AI for superior natural language understanding and generation.

Meta Llama 3 Instruct model performance

	Meta Llama 3 8B	Gemma 7B - It Measured	Mistral 7B Instruct Measured
MMLU 5-shot	**68.4**	53.3	58.4
GPQA 0-shot	**34.2**	21.4	26.3
HumanEval 0-shot	**62.2**	30.5	36.6
GSM-8K 8-shot, CoT	**79.6**	30.6	39.9
MATH 4-shot, CoT	**30.0**	12.2	11.0

	Meta Llama 3 70B	Gemini Pro 1.5 Published	Claude 3 Sonnet Published
MMLU 5-shot	**82.0**	81.9	79.0
GPQA 0-shot	39.5	**41.5** CoT	38.5 CoT
HumanEval 0-shot	**81.7**	71.9	73.0
GSM-8K 8-shot, CoT	**93.0**	91.7 11-shot	92.3 0-shot
MATH 4-shot, CoT	50.4	**58.5** Minerva prompt	40.5

LangChain & GROQ-API



To accelerate development and empower our project with large language model capabilities, we opted for the LangChain framework and the GROQ API. This innovative approach bypassed the traditional method of downloading a bulky pre-trained model. Instead, the GROQ API provided seamless access to the LLM, saving on development time and disk space. This also simplified deployment and holds promise for future scalability, as the GROQ API can potentially integrate with cutting-edge LLMs that might not be readily downloadable.

Enchaced Performance

01.

Fine Tuning:

process of taking a pre-trained language model and further training it on a specific dataset to adapt it to a particular task or domain, improving its performance and accuracy for that specific use case.

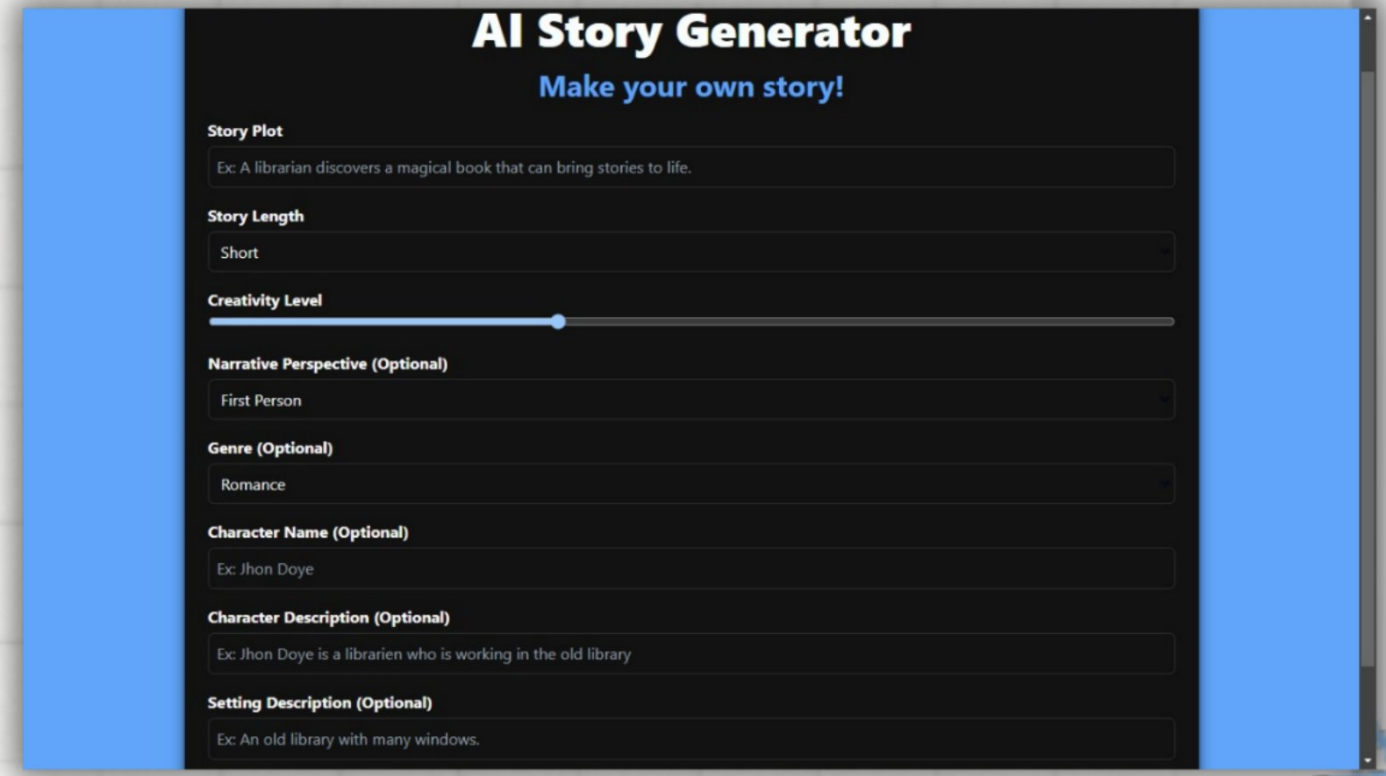
02.

Prompt Engineering:

Technique of designing and optimizing the input prompts given to a language model to guide it in generating the most relevant and accurate responses for a desired outcome.

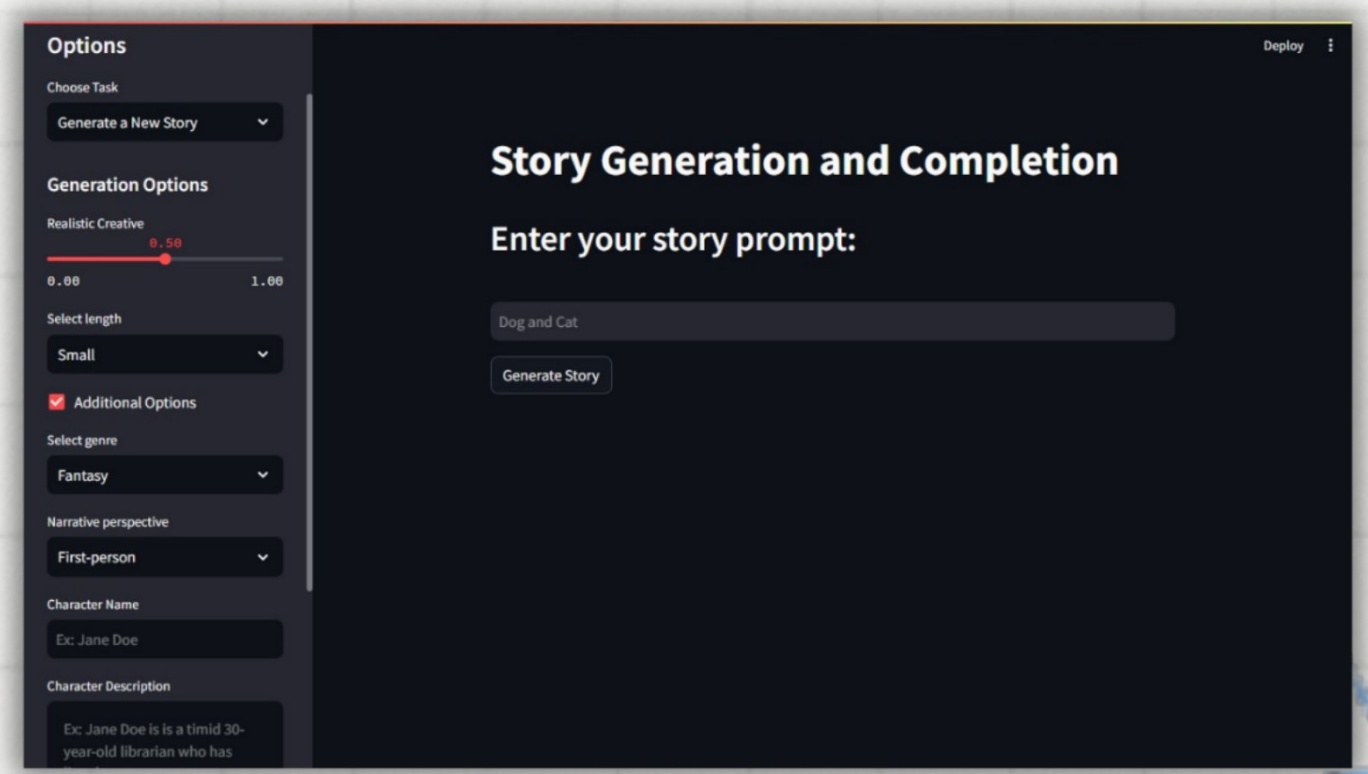
Typescript/React UI

focuses on building scalable, component-based user interfaces using statically typed JavaScript (TypeScript) and a powerful UI library (React), suitable for complex web applications requiring robust architecture and state management.



Streamlit UI

- tailored for data-driven applications in Python, emphasizing simplicity and interactivity without extensive front-end development knowledge.



Fast-Api

FastAPI is our backend framework, linking our TypeScript/React frontend to the Python backend where our model is stored. It handles requests from our frontend, ensuring smooth data flow. TypeScript in React keeps our code safe and efficient, while FastAPI manages data processing using Python's tools like Pydantic for reliability. This setup makes our web app fast, scalable, and easy to maintain for real-time updates and complex data tasks.

Final reflections and future steps

To make our storytelling chatbot more engaging and user-friendly, we plan to improve its understanding of natural language and personalize interactions based on user preferences. We'll add pictures and videos to enhance storytelling and create dynamic stories based on user input. Making the chatbot available on different platforms and learning from user feedback will ensure it evolves to meet users' needs. These steps aim to enhance the overall experience and make the chatbot more appealing and interactive

Thank you very much!

Team Alpha NLP