**Vivekanand Education Society's Institute of Technology**

****

**Department of Computer Engineering**

**Group No: 08**

**Date: 02-08-2024**

**Project Synopsis (2024-25) - Sem V**

Social Stories Generator: An LLM-based learning tool for specially-abled children

Dr. Sujata Khedkar

Associate Professor, Computer Engineering Department

Sarang Pavanaskar Akshat Mahajan Tanmay Maity Mohit Vaidya

V.E.S.I.T V.E.S.I.T V.E.S.I.T V.E.S.I.T

2022.sarang.pavanaskar 2022.akshat.mahajan 2022.tanmay.maity 2022.mohit.vaidya

@ves.ac.in @ves.ac.in @ves.ac.in @ves.ac.in

**Abstract:**

We propose a new method for creating social stories for specially-abled children using artificial intelligence technology, specifically Large Language Models (LLMs). Our approach involves two main steps: generating personalized stories and adding multimedia elements. We will use LLMs to create the text of the stories and incorporate audio, images, and videos to make them more appealing and effective in teaching social norms and behaviors to children who need more personalized education. This will be integrated with a website as the user interface where users can interact. By automating the story creation process, we aim to address the challenge of producing social stories that meet each child’s unique requirements. This approach would not only support the learning and development of specially-abled children but also demonstrate the potential of AI in personalized education, providing a valuable tool for parents, educators, and caregivers to create impactful social stories with ease.

**Introduction:**

This project is planned to innovate how educational narratives are created and delivered to teach social norms and appropriate behaviors to specially-abled children. These stories, which can address specific situations or challenges, are a valuable tool for parents, educators, and caregivers. With advancements in artificial intelligence, particularly Large Language Models (LLMs), there is an opportunity to automate the creation of these stories, making them more engaging and accessible. This project will use LLMs and generative AI technologies to produce personalized stories in various formats, including text, audio, images, and videos, providing a valuable learning experience for these children.

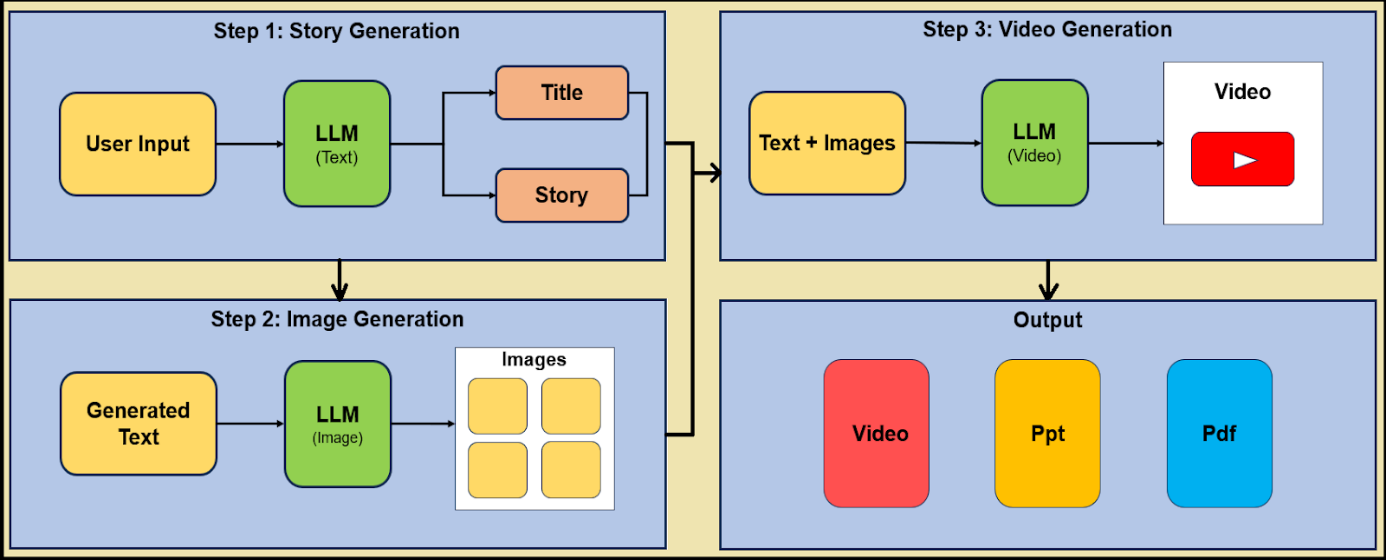
**Problem Statement:**

Creating Social stories for specially-abled children is often challenging and time-consuming. These stories need to be customized to address specific situations and learning needs. Traditional methods usually involve manual creation and rely heavily on text, which may not be engaging or accessible to all children.

Some children require a different approach to learning; they need to be taught every single etiquette, behavior, and value from scratch. This makes it essential to have highly specialized educational content. Additionally, these stories often lack multimedia elements, such as audio, images, and videos which are important for capturing the attention of children with different learning preferences. The absence of these elements can limit the effectiveness of the stories in teaching important social norms and behaviors. Furthermore, it can be difficult for parents and educators to create and update these stories regularly, as it requires a lot of time and effort.

There is a need for a more efficient solution that can quickly produce engaging, multimedia-rich educational stories, making them more appealing to each child's unique needs. This project aims to address these challenges by using advanced AI technologies like LLMs to enhance the process of creating Social Stories.

**Block Diagram:**

****

**Methodology:**

1. User Input: Users provide a text prompt of the story they want to generate.

2. Text Generation: A fine-tuned LLM generates the story text based on the input.

3. Audio Generation: Text-to-speech technology converts the generated text into audio.

4. Image Generation: Text-to-image models create relevant illustrations.

5. Video Generation: A video synthesis system combines text, audio, and images into a coherent video with captions.

6. Output: The final product is a fully generated multimedia Social Story.

**Hardware and Software Requirements:**

Hardware Requirements:

High-end computer with the following configuration:

* + Processor: Intel Core i7 (7th Generation or later)
  + RAM: 16GB or more
  + Storage: 512GB or higher
  + Graphics Card: Dedicated GPU with at least 4GB VRAM (for multimedia processing)

Software Requirements:

1. Programming Language: Python for scripting and developing core functionalities
2. Libraries and Frameworks:
   * LangChain: For building and managing language models and text generation
   * OpenCV: For video processing and image manipulation
   * Pillow: For image processing tasks
   * Llama 3: For advanced language model capabilities and customization
   * LLM with vision: For integrating visual understanding and generation capabilities
   * NLTK or SpaCy: For natural language processing tasks
3. Development Tools:
   * IDE/Code Editor: PyCharm, VSCode, or any preferred Python IDE
   * Version Control: Git and GitHub for version control and collaboration
   * Virtual Environment: Anaconda for managing dependencies

**Proposed Evaluation Measures:**

**Usability Evaluation:**

* Ease of Use: The website will be evaluated on how user-friendly it is for special children, their parents, and educators.

**Functional Parameters:**

* Precision of Prompts:The accuracy and relevance of the generated prompts will be evaluated. This includes assessing how well the prompts align with the intended learning objectives and the specificity of the information provided.
* Accuracy:The accuracy of the content, including factual correctness, will be measured.

**Non-Functional Parameters:**

* Readability: The clarity and simplicity of the language used in the stories will be evaluated to ensure it is easily understandable by the target audience.
* Effectiveness: The effectiveness of the stories in achieving educational goals includes evaluating how well the stories engage children and facilitate learning.
* Understandability: The ease with which users can comprehend the content, including the use of simple language and clear visuals, will be evaluated.
* Predictability: The consistency and predictability of the website's behavior will be assessed. This includes ensuring that the system's responses are reliable and that similar inputs yield consistent outputs.

**Conclusion:**

This project aims to introduce an innovative approach to creating social stories for specially-abled children using artificial intelligence. By integrating Large Language Models (LLMs) with multimedia elements like audio, images, and videos, we aim to produce engaging and personalized educational content. This method addresses the challenges of traditional story creation, which can be slow and often lacks multimedia features like audio, images, and videos. Our solution seeks to make educational stories more accessible and effective, supporting the learning needs of every child while demonstrating the potential of AI in personalized education. The outcome will be a valuable tool for parents, educators, and caregivers to create impactful social stories efficiently for special children.

**References:**

1. Alabdulrahman, M., Khayyat, R., Almowallad, K., & Alharz, Z. (2024). Sarid: Arabic storyteller using a fine-tuned LLM and text-to-image generation. 2024 16th International Conference on Computer and Automation Engineering (ICCAE).
2. Gozalo-Brizuela, R., & Garrido-Merchan, E. C. (2023). A survey of generative AI applications.
3. Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., Liu, P., Nie, J.-Y., & Wen, J.-R. (2023). A survey of large language models.
4. Alabdulrahman, M., Khayyat, R., Almowallad, K., & Alharz, Z. (2024). Sarid: Arabic storyteller using a fine-tuned LLM and text-to-image generation. 2024 16th International Conference on Computer and Automation Engineering (ICCAE).
5. Zhu, H., Qin, S., Su, M., Lin, C., Li, A., & Gao, J. (2023). Harnessing large vision and language models in agriculture: A review.
6. Alabdulkarim, A., Li, S., & Peng, X. (2023). Automatic story generation: Challenges and attempts.
7. Ding, R., Han, S., & Zhang, D. (2023). InsightPilot: An LLM-empowered automated data exploration system. Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023).
8. Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). LLaMA: Open and Efficient Foundation Language Models.
9. Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2019). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension.
10. Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., Liu, P., Nie, J.-Y., & Wen, J.-R. (2023). A Survey of Large Language Models.