2024 Nittany AI Challenge

Submission ID: 12

TEAM INFORMATION

1. Project Team Name:

SPECTRUM STORIES

2. Team Members:
List each team member including Name, PSU Email, Campus, and College
(ex: Jane Doe, jxd123@psu.edu, Shenango, Engineering)

Savitha Kolar (svk3@psu.edu) and Frank Long (fvl5170@psu.edu), World Campus, MPS in AI.

3. Primary Contact/Team Lead:

Savitha Kolar, svk3@psu.edu, 814-753-1930

PROJECT INFORMATION

4. Problem / Opportunity Statement
Documentation should include an overview of the problem and the method used to address that problem. This should demonstrate that you clearly understand the problem or opportunity you are addressing with AI. Response is limited to 300 words.

In the USA alone, an estimated 15-20 percent of the world's population exhibits some form of neurodivergence. Neurodiversity describes the variation in the human experience of the world, in school, at work, and through social relationships. By not making books and educational material accessible and inclusive, we are essentially ignoring a large part of the population and leaving them with no means to advance their educational skills and be independent, functioning members of society.
By developing inclusive reading material and illustrations, our hope is to bring alive the magic of classics (to start with), and ensure educational materials are more inclusive in nature.

5. MVP Use Case
Provide a sample use case for the tool. Describe how someone will use the MVP functionality you intend to build and the benefits or impact the MVP will provide. Response is limited to 300 words.

As a Minimum Viable Product, we chose the story of Goldilocks and the Three Bears as a use case.
For our prototype, we have identified 10 potential steps in generating a neurodivergent-friendly illustration and story.
The user would first select a story, in this example, it would be Goldilocks and the Three Bears.
They would copy the text of the story and input it into our app.
They would then select the neurodivergent sub-category, for instance, Autism Spectrum Disorder, Dyslexia, etc.
The app will use Natural Language Processing to summarize, tokenize, and use Named-entity Recognition on the story text. For example, some entities could be Goldilocks, Three Bears, etc.

The app will use these tokens and entities that were extracted using NLP in the previous process as prompts into the Stable Diffusion API script. There will also be negative prompts to ensure cleaner images.

The app will use custom LoRAs which are Low Rank Adaptations that allow us to use low-rank adaptation technology to quickly fine-tune diffusion models. The LoRA training model makes it easier to train Stable Diffusion on different concepts, such as characters or a specific style. The response of the stable Diffusion API script will output the generated images. The generated images will be combined with the text to produce a neurodivergent-friendly storybook.

6. Data Availability

Detail the data sources leveraged within the prototype as well as the data sources necessary if this project moved to MVP. If available, please detail the location and availability of the data sources and/or the plan for collecting the necessary data. Remember that while we can provide some assistance with finding data sources, finding and gaining access to those sources is the team's responsibility. Response is limited to 300 words.

The dataset we plan to use is from within Stable Diffusion, which is built off a 5 Billion image public dataset.

We plan to use stories that are open source and freely available from the Library of congress.

We are also planning to use public LoRAs from Civit AI.

7. Technology

Provide a technical description of the approach the team used to achieve its proposed goal, including the ways in which the selected AI platforms are used within the prototype and how the team anticipates using those and other services in the MVP phase. Specifically, the documentation should include a list of the components of the selected AI platforms that are leveraged in the prototype, any additional components that may be leveraged in the development of the MVP, and additional services that may be necessary for continued development. Response is limited to 350 words.

The technology we plan on using to create Spectrum Stories will involve Natural Language Processing (especially the parts involving summarization, tokenization, and Named Entity Recognition) using Python as the programming language.

We will use Stable Diffusion and its API script along with custom LoRAs to generate text-to-images.

8. Prototype Video Overview

All teams submitting a prototype for review are required to submit video demonstrations of their working prototypes. The videos must:

- *be <u>no more</u> than 5 minutes in length.*
- *explain the intent, goals, and potential impact of the solution.*
- *demonstrate the basic, working functionality of the prototype.*
- *be available through a YouTube link accessible for viewing by the Challenge reviewers.*

The production value of the videos will not be factored into the review, but they must clearly and accurately represent the prototype functionality. To help, Media Commons at Penn State provides free One Button Studio options throughout the Commonwealth.

**See Video Provided in the Folder