Guide for Computer Vision (CS 319) Final Project:

Project Description: Our project leverages the power of Stable Diffusion, an advanced generative model, to create captivating and realistic images. We started with a pre-trained Stable Diffusion model and fine-tuned it using Dreambooth, incorporating images inspired by the iconic character Rapunzel. Below is a short description of each component of the project.

FinetuneDreamBooth.ipynb: The script completes the training process of DreamBooth. Considering the fact that the model is complex and our local computers do not have GPUs, we trained the model and opened this file by Google Colab.

Code: This folder contains essential scripts and codes.

- **train.py:** The code responsible for fine-tuning the Stable Diffusion model using DreamBooth. This script is pivotal in enhancing the model's capacity to generate images aligned with the Rapunzel-inspired theme. It ensures that the model learns and adapts to the training dataset.
- **model:** The model folder serves as a repository for all components related to the saved model after training.
- **inference.py:** This script loads the saved model and enables the generation of new images based on provided test texts. The use of a fixed seed ensures reproducibility, allowing collaborators to recreate results consistently.

Data: This folder includes three parts.

- **training_data:** Contains images featuring Rapunzel, serving as the foundation for training the Stable Diffusion model.
- **class_data:** A dynamic collection of images generated by the model to align with the characteristics of the trained class.
- **testing_data:** Includes test texts in yaml file for model testing during the inference phase.

Result: This folder stores the generated images produced by the saved model using test texts from the testing_data section. These images serve as a visual representation of the model's performance during validation.