**PARTHO DAS** email: parthodas6176@gmail.com
daspartho.github.io GitHub: @daspartho

## RESEARCH & WORK EXPERIENCE

**UC Berkeley SPAR,** *Student researcher* Feb. 2023 — Present

- Working on RLLF (RL from language feedback) project under mentorship of Juan Rocamonde for UC Berkeley's SPAR research program.
- Worked on implementing photorealistic image augmentor using stable diffusion models for gym environment observation images.

**AI Safety @ UCLA,** *Collaborator* Feb. 2023 — Present

- Collaborating with AI Safety @ UCLA's alignment research team on mechanistic anomaly detection project.
- Worked on training the poisoned models, evaluating the impact of poisoining, creating activation dataset, training a binary classifier for anomaly detection and evaluating it.

## PERSONAL PROJECTS

**Prompt Extend,** *github / demo* Nov. 2022

- Text-Generation model to help with prompt engineering by generating suitable style cues for Stable Diffusion prompts, resulting in better image generations.
- Processed the diffusiondb dataset and trained a new tokenizer and a GPT-2 model on the dataset of stable diffusion prompts for generating style cues.
- Uploaded the model on HuggingFace Hub gaining 30k downloads. Received 2x $1000 grant from algovera.ai for the project.

**MagicMix,** *github / demo* Dec. 2022

- Implementation of MagicMix: Semantic Mixing with Diffusion Models paper. This technique allows for mixing two different concepts in a semantic manner to create a new concept using Diffusion Models.
- Implemented the paper in PyTorch using components from the diffusers library and successfully reproduced results from the paper.
- Added the implementation as a community pipeline to the diffusers library. Also deployed the implementation demo on HuggingFace Spaces.

**Predict Subreddit,** *github / demo* Oct. 2022

- Multi-class text classification model to predict the subreddit of a post based on its title.
- Wrote python scripts to scrape posts from top subreddits, cleaned and processed the collected data for training.
- Fine-Tuned DistilBERT model on the collected dataset of post title pairs from top the 250 subreddits using huggingface transformers library.

## OPEN SOURCE CONTRIBUTIONS

- huggingface/transformers
- huggingface/diffusers

## SKILLS

**Languages:** Python, HTML, CSS, JavaScript, SQL.
**Libraries**: PyTorch, HuggingFace, FastAI, NumPy, Pandas, Matplotlib, Ploty, Gradio, Streamlit, Flask, Selenium.
**Tools:** Git, GitHub, Jupyter, VS Code, Bash, Linux, AWS, TensorBoard, CI/CD.

## EDUCATION

**APS Nehru Road, Lucknow** Graduation Year: 2022

Achieved third place in a national level coding competition and secured the runner-up position in a state level coding competition representing my high school.