Chris Yuhao Liu

University of California, Santa Cruz
1156 High Street
Santa Cruz, CA 95064
yliu298@ucsc.edu
https://chrisliu298.io/

Research Interests

I am currently working on algorithms that can improve the sample complexity of neural networks. I am also interested in the generalization ability of neural networks and their interpretability.

Education

University of California, Santa Cruz September 2017 - June 2021
Bachelor of Science in Computer Science, GPA: 3.75/4.0

Research Experience

Student Researcher, JLab at UC Santa Cruz April 2020 - present

- Worked on a research project that aims to make learning algorithms learn at the fast rate (advised by Professor Jeffrey Flanigan)
- Presented weekly research progress and participated in research paper reading, presentation, and discussion

Work Experience

Tutor, Baskin School of Engineering at UC Santa Cruz October 2020 - December 2020

- Graded student work, including written homework and programming assignments
- Held office hours to respond to student's questions about course materials and assignments

Projects

Fine-Tuning GPT-2 to Generate Research Paper Abstracts

- Fine-tuned a pre-trained GPT-2 (774M) model using all research paper titles and abstracts under cs.AI, cs.LG, cs.CL, and cs.CV on arXiv and built a machine learning paper abstract generator
- Ranked 1st in the text generation competition of CSE142 Machine Learning at UC Santa Cruz
- Open-sourced the model, which is hosted by Hugging Face at `https://huggingface.co/chrisliu298/arxiv_ai_gpt2`

TAPT: Text Augmentation Using Pre-Trained Transformers With Reinforcement Learning

- Fine-tuned a distilled RoBERTa model as a text classifier on the IMDb Large Movie Review Dataset
- Fine-tuned a GPT-2 (345M) model as a text generator with the RoBERTa classifier using proximal policy optimization (PPO)
- Built a text generation pipeline to generate augmented text data and a text classification pipeline to verify the quality of the generated text
- Led the group members to conduct experiments and analysis

Sentiment Analysis With Transformers

- Fine-tuned a RoBERTa (355M) model using the IMDb dataset
- Ranked 1st in the sentiment analysis competition of CSE142 Machine Learning at UC Santa Cruz

Honors and Awards

Dean's Honors, University of California, Santa Cruz

- The top 15% of the academic group
- 7 times out of 12 quarters

Relevant Coursework

Undergraduate: Calculus, Vector Calculus, Linear Algebra, Probability and Statistics, Applied Discrete Mathematics, Machine Learning, Introduction to Natural Language Processing, Applied Machine Learning

Graduate: Natural Language Processing, Advanced Topics in Natural Language Processing, Seminar in Natural Language Processing

Skills

Programming languages (ranked by proficiency): Python, C, MATLAB, Java

Frameworks: scikit-learn, auto-sklearn, NumPy, Pandas, matplotlib, NLTK, PyTorch, PyTorch Lightning, TensorFlow, Keras with TensorFlow backend, HuggingFace Transformers

Miscellaneous: Git, Unix, Jupyter Notebook, Google Colab, Google Cloud Platform, LaTeX, Markdown, HTML&CSS