

ANISH THITE

anishthite@gmail.com · (203) · 909 · 9118 · anishthite.github.io

Objective: Seeking Spring or Summer '21 positions in Software Engineering or Machine Learning.

EDUCATION

Georgia Institute of Technology, Atlanta GA

Expected May 2022

M.S. in Computer Science (Intelligence and Devices threads) B.S. expected May 2021

Overall GPA: 3.74/4.00 Major GPA: 4.00/4.00

EXPERIENCE

The Agency (AI/ML Club)

January 2020 - Present

External Relations Officer (President)

Atlanta, GA

- Sys admin for club server, lecture on NLP topics, coordinate guest lecturers, and advise project groups

Georgia Tech Research Institute

January 2019 - Present

Machine Learning Intern

Atlanta, GA

- **Using genetic programming to evolve machine learning models** with the GTRI Evolutionary Auto-ML Framework
- Expanding the library to perform neural network architecture search (using **Python, MySQL, Scikit-learn, Keras, Tensorflow**)
- **Out-performed Google Auto-ML** on the Wikidetox toxic comment classification dataset

NLP and Computational Social Science Lab:

November 2019 - Present

Student Researcher

Atlanta, GA

- Working under Dr. Diyi Yang on keyword-based **neural humor generation** and classification using **transformers**
- Further study includes analyzing effects of humor on mental health and the COVID pandemic.

PROJECTS

EleutherAI GPT-3 Replication

July 2020 - Present

- Contributing to open-source project to **replicate and open-source GPT-3 weights**
- Working on **HTML + PDF text extraction** and filtering to make the 400GB text dataset for the model
- Efficiently scraped + filtered entire Hacker News for language model training

GT Hacker Home

June 2020

- Created + deployed **reddit-like forum** with **Svelte.js** Frontend and **MySQL + Rust** (Rocket framework) back-end for GT students to discuss side projects
- **Try it:** <https://gthackerhome.github.io/>

AILight: Autonomous Highlighting Web App

July 2019 - March 2020

- Used **BERT** (Deep Learning model) for **extractive summarization** and **automatically highlight** the important elements of a given document using **OCR** libraries
- Built a **REST API** and **React** frontend to interact with model, deployed to **GitHub Pages**
- **Try it:** <https://anishthite.github.io/ailight>

ATHENA (A Talkative Hip Extremely Nit-picky Assistant)

April 2019 - January 2020

- Fine-tuned **GPT-2** on conversations in order to create a **conversational agent**, intend to create a personal assistant
- Set up **Linux** server in garage and deployed **Flask** back-end to it (using **Supervisord + gunicorn**)
- **Try it:** <https://pal-ai.github.io/chat>

SKILLS

Languages	Python, Java, Rust, C, Scala, Bash, JS,
Tools/OS	Git, AWS, GCP, Firebase, Docker, Linux, Windows, SQL, Slurm, CUDA
Frameworks	Flask, Django, React, PyTorch, TensorFlow, Numpy, Pandas