Samuel W. Fang

B.S. Computer Science, University of Michigan
May 2025 Expected Graduation | Senior (GPA: ~3.8)

906-231-2692 | 1514 Gilbert Ct., Ann Arbor, MI 48105 | swfang@umich.edu

I am a Computer Science student at the University of Michigan seeking a **full-time position** with a special interest in **machine learning/AI** and **software engineering.** I am hoping to be able to apply my knowledge and prior experience to be able to contribute to innovative ideas and projects, by working with logic and solving problems with simple, but elegant methods. I've worked on a large array of different projects pertaining to different skill sets, including writing code at the assembly level, writing OS code in C, designing machine learning models in Python, and implementing different algorithms and data structures in C++.

Skills

Programming Languages: C++, Python, C#, JavaScript (incl. React), C, Java, OCaml/Functional Programming, Verilog

Packages: NumPy, Pandas/Polars, PyTorch, Flask and Jinja, Tkinter

Additional Experience: SQL, HTML w/ CSS, API development with RestAPI, .CSV/.TSV, JSON, XML file formats

Dev Tools and IDEs: VSCode, Visual Studio, XCode, Unity, GitHub, Docker, Basic Linux Shell Scripting, HPC Systems, Microsoft Office

Work, Internship, and Class Experience

DCode Ovcharenko Research Group: NIH-NLM Summer Intern (June-Aug 2024)

Worked with Dr. Ivan Ovcharenko (Supervisor) and Dr. Di Huang (Mentor) to develop a pipeline and present a poster outlining a process of identifying Drug Discovery candidates through their association with gene regulatory elements using Machine Learning models. Involved the use of Python with PyTorch/NumPy/Pandas, fine-tuning models utilizing HPC clusters, using transformer models (DNABert), in-silico gene transformations, and general data management across clusters. More Info About the Group at https://dcode.org/wiki/

NCI Summer Intern with Dr. Umit Topaloglu and Dr. Robinette Renner (June-Aug 2023)

Internship involving understanding and applying various methods of machine learning, including Linear Regression, Support Vector Machine, and Neural Networks, along with data cleansing/preprocessing techniques and an understanding of utilizing NIH databases. Involved developing a project investigating the effects of data preprocessing (e.g. normalization) on model performance.

Dr. Jiang Research Lab: Michigan Technological University (May-Aug 2022)

Project with Dr. Jinfeng Jiang involving use of Python to sort and process ultrasound DICOM images (including use of machine learning for image recognition and Tkinter for GUI), along with use of R in machine learning. The goal was to predict cases of heart disease through analysis of thrombosis and heart disease through ultrasound images.

Personal Project: Visual Novel Framework in C#-Unity (Sept 2024- Ongoing)

Currently on-going personal project involving the design of a ground-up framework with functionality for parsing custom visual novel scripts through C#, and object oriented programming to code for the basic functionalities of a Visual Novel Game in Unity. Relevant Knowledge includes Planning and Designing Objects to Efficiently and Coherently Implement Necessary Features and Organization of Assets and Scripts for a Large-Scale Project.

Operating Systems (EECS 482: Jan-May 2024)

Worked on projects writing low-level OS code for multi-threaded processes and parallelism (mutexes/locks, CVs, semaphores), virtual-physical memory management (page table, page fault and eviction logic), and disk file management via networking.

Compilers (Jan-May 2024) : Built several stages of a compiler in OCaml for compiling and optimizing a subset of C functionality (called OatC) to Assembly Code through Lexing, Parsing, Type Checking, use of LLVM IR, and Control Flow Analysis/Optimization. Also included work on assembly execution and writing basic structure of an interpreter.

Instagram Clone using Dynamic Pages (Sept-Dec 2024): Project involving creating a Instagram-like website, including use of HTML, use of Python and Jinja to generate HTML, use of Python and Flask to implement server-side dynamic pages for login and posting functionality, and implementing RestAPIs in Python and React in JavaScript to implement a client-side dynamic index page for instantaneous "like" and comment functionality.

Implementing MapReduce over a Network (Sept-Dec 2024): Project involving implementing a distributed system of workers and a manager to be able to use parallel processing with the MapReduce framework to complete a user-submitted task. Required knowledge of socket communication, multi-threaded python programming, and understanding the MapReduce (Hadoop) framework.

RISC Assembler with Pipelining and Caching (Sept-Dec 2023): Designed an assembler for a RISC architecture called LC-2K which (1) took in code written in assembly and translated it into machine code and (2) executed machine code through a simulated processor. Adjustments made later on in the course included the modification of the processor to take advantage of pipelining as well as the introduction of a cache to expedite the retrieval of data.

Machine Learning Class Projects (Sept-Dec 2024):

Classifying Reviews using Support Vector Machine: Project involving using Python and scikit-learn to classify reviews as negative, neutral, or positive using their word content as features. Included use of hyperparameter tuning, data preprocessing (e.g. removing/altering words), weighting unequal/biased datasets, and understanding OvO/OvA multiclass classification.

Classifying Images using CNNs: Project involving using Python and PyTorch to classify images of two different landmarks. Involved implementing a custom CNN with convolutional, pooling, and fully connected layers, understanding earlystopping and transfer learning, and data augmentation.

Search Engine (Sept-Dec 2023): Project involving creating a Search Engine, using Information Retrieval (td-idf and link analysis) to sort and select articles based on a search query.

4-Function Bitwise Calculator (Jan-May 2023): Built and Coded a 4-Function Calculator for a basic Circuit Board in Verilog.

Education

B.S., Computer Science, University of Michigan–Ann Arbor Expected Graduation: April/May 2025 GPA: 3.776/4.00 (current)

Relevant Classes:

- Discrete Mathematics
- Foundations of Computer Science
- Data Structures and Algorithms
- Intro. Operating Systems
- Compilers
- Intro. Computer Security
- Web Systems
- Intro. Machine Learning
- Databases
- Machine Learning Research Experience
- Intro. Computer Organization

High School Diploma, Houghton High School, Michigan: Graduated 2021 GPA: 3.98/4.00