Aney Kanji

aneykanji@gmail.com | (254)-239-4972 | linkedin.com/in/aney-kanji/ | github.com/aneykanj12 | aneykanji12.vercel.app/

Education

Texas A&M University – College Station, TX — Class of 2027 | Third-Year Student (Junior)
B.S. in Computer Science (Honors) and B.S. in Statistics — GPA: 3.84 | Graduate Coursework GPA: 4.00

Relevant Coursework: Program Design & Concepts · Data Structures & Algorithms · Computer Organization · Programming Languages · Data Science · Software Engineering · Computer Systems · Probability

Graduate Coursework: Statistical Research · Statistical Computing · Algorithms · Data Mining & Analysis

Experience

Machine Learning Research Intern — July 2025 – Present
Cincinnati Children's Hospital Medical Center — *Remote*

- Developed and researched AI and machine learning models for use in bioinformatics and improving medical infrastructure.
- Worked on natural language processing techniques for medical LLM integration with microscopy workflows along with constructing RAG architecture. Worked in a team of 5 to deploy the LLM at the Hospital for treatment reinforcement.
- My impact has helped improve the model's accuracy and classification of malaria infections within red and white blood cells.

Data Science & Analyst Intern — September – Present
BroadStreet Institute — *Remote*

- Researched reduction in maternal and infant mortality rates as well as safer pregnancy. Contributed to a cross-functional project with a team of 30+ interns to clean, preprocess, and visualize datasets.
- Reduced the complexity of medical data reports by constructing SQL queries in a Postgres database and developing data visualizations using Seaborn, Matplotlib, Pandas, and NumPy, as well as Power BI.
- My impact has allowed stakeholders to track pregnancy outcomes in real time through data portfolios.

Undergraduate Teaching Assistant — January 2025 – Present
Texas A&M University — *College Station, Tx*

- Assisted professors in teaching a Data Science (fall semester) course and a Discrete Structures (spring semester) course to 200 students.
- Worked in a team setting along with other teaching assistants to deliver a structured and well-manner learning environment for students who want to learn industry-standard tools and technologies in various fields.
- My impact has allowed for improved productivity and course deliverables such as grading and instruction for students.

Mathematics Research Lab Assistant — August 2024 - December 2024
Texas A&M University — *College Station, Tx*

- Developed a custom greedy algorithm in permutation with the Zeckendorf Decomposition algorithm and Chung-Graham theorem to study summands within a sum of even-indexed Fibonacci numbers. Paper has been publish in the PUMP article [Link]
- My impact has allowed our team to compute big data using a custom Python script I wrote to compute values using the greedy algorithm.

Projects

VeriCare – Python, LangChain, MongoDB, Heroku, Django, ReactJS, TailwindCSS, CUDA (2025) [Documentation]

- Worked as team lead to develop a chatbot to reduce complexity in medical insurance policies. The chatbot is uses custom **pipelines** and **tokenizers** for **text-simplification** and **web scraping** using **HuggingFace Transformers**, **PyTorch**, and **LangChain**.
- Used **MongoDB** to store medical information and previous chats and built the backend with **Python** and **Django**. Accelerated **PyTorch** inference using **Nvidia CUDA** and integrated the frontend using **ReactJS**, **TailwindCSS**, and **JavaScript**.
- Containerized the project using **Docker** and deployed using **Heroku**.

SafeSpeak – Go (Golang), SQL, Postgres, LangChain, DiscordJS, FastAPI (2025)

- Built a toxicity and foul language discord bot to moderate channel chats using a machine learning **pipeline** with **LangChain**, **Python** and **HuggingFace Transformers**. Strict backend project using **DiscordJS** and **FastAPI** and **Go** for the primary language.
- Designed using Postgres to store chats, **SQL** to query the data, and **Go** for the bot's output to the channel.

Blood Cell Counter – Python, TensorFlow, Keras, Docker, Neural Networks (2025) [Documentation]

- Developed a machine learning model pipeline to predict red and white blood cell counts from medical metrics (age, platelet count, hemoglobin, MCV, MCH, etc.) using **Random Forests**, **Neural Networks** and **Linear Regression**.
- Preprocessed 500+ cases using **NumPy** and **Pandas**; trained multi-layer neural network with **ReLU** and **Tanh** activations; used **Random Forest** model to reduce overfitting. Achieved accuracy of 99.94% and performed statistical validation using **SHAP**.

Publications

Fixed-Term Decompositions Using Even-Indexed Fibonacci Numbers –**The PUMP Journal of Undergraduate Research** (August 2025) [Link]

Skills

Languages: Python · Java · C/C++ · JavaScript · TypeScript · R · SQL · Haskell
AI/ML: Scikit-Learn · PyTorch · TensorFlow · Keras · PyTorch · LangChain · NLTK · LlamaIndex · HuggingFace Transformers
Libraries & Frameworks: NumPy · Matplotlib · Pandas · Seaborn · ReactJS · TailwindCSS · Bootstrap · Flask · FastAPI · NodeJS · Streamlit
Tools: Git · GitHub · Docker · VS Code · Linux/Unix (Bash & ZSH) · MS Suite · Postgres · MongoDB · Vercel · AWS · Heroku