# Aney Kanji

aneykanji@gmail.com | 13522 Pyrite Dr. Austin, TX | (254)-239-4972 | LinkedIn | GitHub

## 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

## Experience

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

- Aided a professor in running a Data Science course covering topics such as **Git**, **Docker**, **Databases**, and **Machine Learning** (Fall 2025).
- Worked in a team of Teaching Assistants to provide instructional support for a professor in running two discrete math courses for computer science and engineering majors ensuring smooth management of the classroom (Spring 2025).
- Delegated roles of grading homework, assignments, and exams as well as constructing new material for promoting growth in knowledge of the material.
- Held office hours multiple times during the week to foster assistance to students in need of help regarding anything revolving the class and its content.
- My impact has allowed students to leave with a better understanding of the material, as well as encouraging a healthy and fostered environment both inside and outside the classroom for the students, my peers and professor.

Peer Teacher/Student Assistant — August 2024- May 2025
*Texas A&M University* — *College Station, Tx*

- Reinforced learning for students by aiding professors and teaching assistants with assistance with student matters regarding content from courses, grading coursework, holding office hours and weekend reviews, as well as monitoring exams.
- Helped a Teaching Assistant run a Program Design and Concepts instructional lab. The class was taught in **C++** and covered content from **Unix** and terminal commands, programming fundamentals, **Stack** and **Heap** Memory Concepts, Memory Allocation, **Object Orientated Programming**, linear Data Structures, and **Recursion**.
- Aided a professor with a team of peer teachers for an introduction to engineering computation course taught in **Python**. The course covered programming fundamentals, basic frameworks such as **Matplotlib** and **NumPy**, and file/computer systems.
- My impact has allowed professors and teaching assistants to run a more productive learning environment for students to facilitate learning at a high level.

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

- Worked in a team of three to build off the Fixed Term Zeckendorf Decomposition Algorithm to test new sets of data and theorems including the Chung-Graham Theorem using non-negative integers.
- Developed programing model to comprehend pattern recognition of Greedy Algorithms using Chung-Graham Theorem which allowed for an observation of alternating Fibonacci constants in all sets of Fibonacci sequences.
- Impact led to the discovery of developing a new dataset of alternating terms in the Fibonacci Sequence
- Built skills in collaborative thinking, problem solving, and writing mastery. **(LaTeX, Python, Fibonacci)**

## Projects

Document Analysis Assistant – Python, LangChain, HuggingFace Transformers, PyTorch, CUDA (2025) [Documentation]

- Engineered a shell-based assistant to analyze and summarize user documents using **BART** and **RoBERTa LLMs**.
- Leveraged **HuggingFace Transformers** for **tokenization**, **embeddings**, and summarization **pipelines** with **LangChain chaining/invocation**.
- Accelerated **PyTorch** inference using **NVIDIA CUDA** to optimize large-scale document processing.
- Integrated a scalable architecture supporting multiple file types and extensible **NLP** workflows.

Credit Score Classifier – Python, Scikit-Learn, Docker (2025) [Documentation]

- Developed a credit score classification pipeline using **Random Forest**, **XGBoost**, and **EnsembleSV**; evaluated performance with AUC, ROC curves, and precision-recall analysis.
- Preprocessed the dataset and applied **k-fold cross-validation**; conducted in-depth statistical analysis using **Cook's Distance**, **Kernel Density Estimation**, and residual diagnostics.
- Containerized the project using **Docker** for reproducibility and versioned the pipeline with **Git**; utilized **scikit-learn** and supporting libraries for modeling and deployment.

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 99.94% validation accuracy with strong precision and recall; performed feature reduction and statistical validation using **SHAP**.

PDF Summarizer – Python, LangChain, Streamlit (2025) [Documentation]

- Designed a lightweight application for PDF uploads with real-time, context-aware summarization using LangChain and Gemini Flash LLM.
- Implemented a Streamlit interface with manual controls for document interaction and summarization.

## Publications

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

## Skills

**Programming Languages**: Python · Java · JavaScript · C/C++ · Go (Golang) · R · SQL · HTML · CSS · Haskell
**Libraries & Frameworks:** NumPy · Pandas · Matplotlib · BeautifulSoup · Seaborn · Scikit-Learn · TensorFlow · Keras · PyTorch · LangChain · LlamaIndex · HuggingFace Transformers · NLTK · Streamlit
**Tools:** Git · Docker · Jupyter · VS Code · Linux/Unix (Bash & zsh) · MS Office
**Relevant Coursework:** Program Design & Concepts · Discrete Structures · Data Structures & Algorithms · Computer Organization · Programming Languages · Computational Data Science · Software Engineering · Computer Systems · Probability
**Graduate Coursework:** Applied Statistics for Research · Statistical Computing · Algorithms · Data Mining & Analysis