Daiwik Pal

(774) 541-0162 | daiwikpal@gatech.edu| linkedin.com/in/daiwik-pal | github.com/daiwikpal

Education

Georgia Institute of Technology | Atlanta, GA *August 2022 – May 2025*
Bachelor of Science in Computer Science; GPA: 3.83; CS GPA: 4.00 Faculty Honors
Concentration: Intelligence & Information/Internetworks
Coursework: Data Structures and Algorithms, Computer Systems and Networking, Statistics, Linear Algebra

Worcester Polytechnic Institute | Worcester, MA *August 2021 – May 2022*
High School Dual Enrollment Classes with Massachusetts Academy of Math and Science; GPA 4.00 Presidents List

Skills

Languages: Java, Python, JavaScript, C/C++, R, SQL
Technologies: Next.js, Express + React.js, Node.js, MongoDB, Flask, Android, Java Spring Boot, Tensorflow, Google Cloud Platform, OpenCV

Experience

Deloitte *June 2024– August 2024*
Incoming Software Engineering Intern Boston, MA

Peak Technologies *May 2023– August 2023*
Machine Learning & Software Engineering Intern Greater Boston Area, MA

- Developed backend architecture to store and perform video analytics on warehouse videos using Java Spring Boot and Python
- Designed and developed an end-to-end ML Application using Flask, JavaScript, and Python by deploying an object detection model using the Tensorflow.js library
- Implemented a shift to client-side computing for increased security by allowing clients to demo and analyze key model metrics about Peak's package detection models on sensitive images locally
- Developed a pipeline using Node.js and Flask to automate model execution on datasets (88000+ images) and with a process to monitor and evaluate the performance, drift, and accuracy of deployed models
- Reduced pipeline latency by 200ms by refactoring redundant code and minimizing costly session storage calls

Siena Analytics *June – August 2021, March -August 2022*
Data Analytics & Software Engineering Intern Greater Boston Area, MA

- Created data-pipeline to perform image pre-processing on 40,000+ barcode images before feeding into decoding model and improving barcode decoding rates by 25% using Python and OpenCV
- Migrated local running ML models to Google Cloud Vertex AI by deploying docker files to train custom AI Models
- Increased model efficiency through researching hyper-params and implementing Google Brain Team's EfficientDet-D0 Object Detection model on Vertex AI

Organizations

DataScience@GT – BioASQ Synergy Competition *January 2024 – Present*
Research Assistant Atlanta, GA

- Developed and fine-tuned large language model architecture to answer yes/no, factoid, list, and summary questions in biomedical domain, leveraging context from PubMed database and literature to address real-life information needs in biomedical research
- Used Spacey and PubMed API, to extract key words from query and retrieve relevant literature to answer query
- Generated embedding vectors through SentenceTransformer package to find text snippets on retrieved literature
- Retrieved and analyzed responses from mistral_7B, llama2_7B, and chatGPT3.5 LLMs using Textsynth API

GT Web Dev Club & Bits of Good – Georgia Tech Hack4Impact Chapter *September 2022 – Present*
Project Manager, Software Developer, & Bootcamp Instructor Atlanta, GA

- Created real-world webapps for clients such as local non-profits and organizations in Atlanta region
- Developed CRUD webapps using React+Express.js, Next.js, MongoDB, and REST APIs in an Agile workflow
- Punchshot Pickleball (Client Project) – A social media application for Pickleball players, organizations, and tournaments
- GT Video Streaming Platform – Education platform to upload lectures, view lectures, and take collaborative notes with peers
- Architected a scalable backend infrastructure leveraging GCP Cloud Run, Pub/Sub, and Cloud Storage to deploy a custom Node.js server that automatically transcribes uploaded videos to lower file size using FFMPEG tool for latency free streaming
- Implemented authentication features using "bcrypt" encryption library and JWT Tokens
- Crafted and taught lessons to 140 new members on topics such as JavaScript, APIs, middleware, and backend architecture

Big Data Big Impact Club *September 2022 – December 2023*
Project Lead & Data Scientist Atlanta, GA

- Created a SteamLit application using a Random Forest Regression model to predict housing prices in Georgia without 3rd-party bias
- Cleaned and enhanced Georgia housing data by curating relevant and unique features such crime rate, job rate, education
- Developed algorithms such as decision trees and random forest regression, achieving a mean absolute error of ~2000
- Demonstrated leadership skills by guiding the team's analytical approach and fostering a collaborative environment that welcomed constructive disagreements and facilitated conflict resolution through decision matrices

Projects

Predicting COVID-19 Transmission Risk | Python, Pandas, StreamLit, Data Analytics, Modeling

- Conducted an independent research project to develop a probabilistic model to calculate COVID-19 transmission risk
- Developed a webapp that allows users to compare the transmission risk of contracting COVID-19 between two venues
- Used Python, Pandas, and the Streamlit Framework for backend and Streamlit Framework for frontend