Taha Al-Nufaili

College Park, MD ♦ talnufai@terpmail.umd.edu ♦ (301)-814-8069 ♦ github.com/ta0a2000t ♦ linkedin.com/in/taha-al-nufaili

EDUCATION

University of Maryland, College Park
B.S in Computer Science & Minor in Statistics *Expected: May 2024*
GPA: 3.7 / 4.0
Coursework: Advanced Algorithms, Intro to ML, Advanced Data Structures, Intro to Data Science, Discrete Structures, Computer Vision, Systems Architecture, Web Application Development, Game Programming, Object-Oriented Programming
Full-ride KGSP scholarship, awarded yearly to 100 students.

SKILLS

Coding: Python, Swift, JavaScript, Java, C#, C, Ruby
Data/ML: Keras, PyTorch, Pandas, Sklearn
Cloud Databases: Firebase, MongoDB
Front-end: SwiftUI, React
Tests: 841/850 on CodeSignal's General Coding Assessment

WORK EXPERIENCE

Application Developer Intern (mascot-vis.github.io) *College Park, MD | Sep - Now 2023*
Human-Data Interaction Group
- Working as part of a development team on Mascot.js, a data visualization API, using JavaScript and React.
- Implementing high-level abstractions and flexible rendering options in SVG and WebGL.

Research Intern (https://shorturl.at/fuN39) *Thuwal, SA | June - Aug 2023*
KAUST AI Initiative
- Led the design and implementation of a standardized benchmarking pipeline for multi-modal Large Language Models (LLMs), creating an object-oriented structure to encapsulate any model for seamless integration.
- Utilized PyTorch to modify existing LLMs for batched inference, optimizing for A100 GPUs in the IBEX remote cluster.
- Evaluated LLMs on object and action recognition tasks, using metrics calculated against image & video datasets.

Machine Learning Intern (https://shorturl.at/kqEX7) *College Park, MD | Feb - May 2023*
NASA Harvest
- Employed Google Earth Engine API to split labeled 1-year image composites into 3-week intervals, overlayed coordinates, and exported imagery to Google Drive.
- Selected features post-preprocessing using both Random Forest Ranking and custom-designed Separability Metrics.
- Utilized Random Upsampling to balance data, then built a 1D CNN model in TensorFlow and evaluated using F-1 scores.
- Presented findings on stage at the Global Sustainable Development Congress 2023.

Software Engineer Intern *Lanham, MD | Jun - Aug 2022*
Science Systems and Applications, Inc. (SSAI)
- Collaborated with NASA scientists to develop cloud-detection neural networks using Keras, based on 18 features.
- Achieved 6X faster processing by applying multi-threading and managed a 5-million-observation dataset with Xarray.
- Used Matplotlib & Bokeh for visualization and DVC for optimization; presented findings to 50+ people at SSAI.

PROJECTS

CasuaFolio (https://terpconnect.umd.edu/~talnufai/) *Aug - Now 2023*
- Built & published an open source React portfolio template, to present a well-rounded version of an individual.
- Allowed one-file customization without requiring coding skills, making it accessible and customizable.
- Wrote comprehensive setup and personalization guides.
- Designed for static website hosting, providing a lightweight and easy-to-host solution.

Social IOS App *Aug - Now 2023*
- Incorporated Firebase for login authentication & real-time data syncing between local storage and Cloud Firestore.
- Structured the app using three core classes: User, Community, and Event, and designed object-oriented relationships.
- Implemented the MVVM design pattern in SwiftUI & Swift for a scalable and maintainable codebase.

3D Cellular Automata (https://shorturl.at/astLU) *May 2023*
- Extended Conway's Game of Life to 3D in Unity; demonstrated patterns with the (4555) rule.
- Increased efficiency by only checking neighbors of alive cells after each generation.

MIPS Processor Simulator *Nov 2021*
- Built a cycle-accurate pipelined C program that reads and compiles MIPS assembly instructions.
- Implemented forwarding between stages in the 5-stage pipelined processor to reduce stalls.

Data Science Tutorial (https://ta0a2000t.github.io/) *Dec 2021*
- Showcased the steps needed to tidy tables into easy-to-manipulate data frames using Pandas.
- Demonstrated skewed data handling techniques, like logarithmic representation and data trimming, using Sklearn.