Dan Huynh

226-606-9284 | dan.huynh@uwaterloo.ca | LinkedIn | GitHub | Portfolio

Overview

Languages: C/C++, Java/Scala, Python, TypeScript, GO, PHP, SQL, Swift, MATLAB, VHDL
Technologies: Docker, Azure, GCP, AWS, Node, Spring, React.js, Angular, Scikit-learn, Tensorflow, ROS2, OpenCV
5+ years of experience in designing mechanical equipment using SolidWorks and the Autodesk suite

Professional Experience

Data Scientist @ theScore Sept 2024 – Present

- Conducted an **EDA** using **pandas** on proprietary stolen base odds generated from an **Sklearn** logistic regression model, correcting systematic under-predictions for base runners and reducing consensus odds deviation by 56%.
- Deployed an automated **Python** data monitor for rapidly scaling MLB odds models to validate projection availability and sparsity across 27 tables in **BigQuery** and **Postgres**, integrating configuration-driven **SQL** and z-score / distributional divergence-based anomaly detection which trigger alerts via the **Slack API** and reports to **DataDog** for pipeline remediation within an **Argo Workflow**.
- Built an in-play odds validation **Argo Workflow** in **Python** to track real-time MLB odds, generate market suspension analysis, and to pair external to internal betting lines which enable real-time adjustments to market odds improving model accuracy and mitigating exposure.
- Currently enhancing **machine learning** infrastructure to improve pipeline robustness and optimize MLB projections used in market odds modeling.

Software Development Engineer in Test @ Vivid Seats Jan 2024 – Apr 2024

- Modified a **Selenium** grid interface to display repository-specific metadata, leading to a 100% increase in test identification across 71+ regression suites.
- Designed a **Stoplight**-documented test-data Backend for Frontend (BFF) with **Spring** and **OpenAPI**, exposing RESTful endpoints for **JPA** entity generation and management that insert regression-agnostic data for E2E checkout tests into 4+ Vault-authenticated **AWS Aurora** databases.
- Developed a distributed subscriber for the test-data BFF which enables on-demand and autonomous data cleanup, via client API calls and a proprietary cleanup micro-service, allowing for concurrent data management.
- Integrated **SonarQube** for static code analysis, leading to the development of 210+ unit and integration tests using **JUnit5**, **Mockito**, and **Spring** that achieved 96% code coverage for the test-data BFF.

Data Scientist @ PureFacts Financial Solutions May 2023 – Aug 2023

- Developed and tuned an **Scikit-learn** Bayesian optimized random forest regressor with a mean percentage error of 17.32% that forecasts client revenue movements, whilst providing interpretable explanations for model predictions using **SHAP**.
- Designed a dashboard using **Plotly Dash** that features dynamic visualizations of investor revenue, AUM, transactions, and customer trends over time for PureFacts clients, encouraging data-driven decision making.
- Led development of a **Flask + React** tool tailored to the PureFacts tech stack utilizing **OpenAI APIs** that empower non-technical personnel with accessible information and optimizes engineer labor time whilst maintaining data confidentiality.

Software Engineer @ Ford Motor Company Sept 2022 – Dec 2022

- Created components including a data-model agnostic autocomplete component using **React Typescript**, that queries 1000+ **Firestore** records for objects that fit a Regex string on one of 7+ record properties.
- Wrote asynchronous RESTful methods using **Axios** that reads/writes to 1000+ records in a **Firestore** database.

Projects

Perceptions Lead @ Watonomous — LiDar Object Detection | Github Sept 2023 – Sept 2024

- Developed a data loader for **OpenPCDet** to processes 32-beam, 4/5 feature Velodyne point clouds into **NumPy** arrays, optimized for **VoxelNeXt**, **TransFusion**, and **PV-RCNN** predictions.
- Wrapped **OpenPCDet** in a **ROS2-humble** node that processes a point cloud rosbag feed, publishing real-time bounding box predictions through the **Foxglove** WebSocket protocol for immediate data visualization.
- Modified **OpenPCDet** visualization utilities to render static **PV-RCNN** bounding-box predictions using **XVFB**, ensuring compatibility without reliance on a native **X-11 server**.
- Collaborated in the design and implementation of a end-to-end perceptions pipeline that tracks and associates objects in real time.

Education

University of Waterloo — GPA: 3.9/4.0 Waterloo, Canada
Honours B.ASc Candidate in Mechatronics Engineering (Option in Software Engineering) *2021 – 2026*