Son Tran

352-709-6136 | sm.tran@ufl.edu | linkedin.com/in/msontran | github.com/smtran8

Education

University of Florida August 2024 – May 2028
Bachelor of Science in Data Science – GPA: 3.73 *Gainesville, FL*

- Courses: Programming Fundamentals, Introduction to Computational Math, Programming with Data in R, Computational and Linear Algebra, Application of Discrete Structures, Database Design and Operational Business Intelligence (UCI), financial analysis (UIUC)

Skills Summary

Languages: Python, C/C++, SQL, R
Tools: Git, Google Cloud Platform, VS Code, Visual Studio, PyCharm, OpenAI API, WorldQuant BRAIN, pandas, NumPy, Matplotlib, PostgreSQL, Power BI, QlikView, FastAPI, Socrata API
Interests: Software Engineering, Quantitative Research, Data Science, Data Visualization, Natural Language Processing, Full-Stack Development, Mobile & Web Applications, Artificial Intelligence, Machine Learning, Mathematics, Backtesting

Experience

Data Engineer March 2025 – Present
University of Florida *Gainesville, FL*

- Collaborated cross-functionally with technical peers to engineer and deploy a **full-stack tutor reporting platform** for **E2S**, a nonprofit serving **200+ UWC students**, improving system scalability and maintainability
- Designed and automated an LLM-powered reporting pipeline with **OpenAI API**, generating **500+ structured session summaries** and progress **reports with 90% less manual effort**
- Developed a responsive HTML interface and public website using modern web frameworks, increasing tutor applications by **30%** and enhancing program accessibility

Event Productions Student Assistant January 2025 – Present
Reitz Union, University of Florida *Gainesville, FL*

- Delivered technical support for **1000+ events** at Reitz Union, coordinating logistics for furniture and equipment with **98%** setup accuracy across **50+** weekly setups
- Managed audio and lighting systems for **200+** events, ensuring optimal performance; conducted weekly inventories and cross-checked Event Management System documents for **100%** compliance with event specifications
- Provided rapid customer service via radio, resolving **95% of 300+ assistance requests** within 5 minutes, enhancing event operations efficiency

Projects

Alpha Signal Development | *Python, Pandas, NumPy, WorldQuant BRAIN* May 2025 – Present

- Developed and backtested 3 alpha signals (mean-reversion, earnings announcement, sector momentum) on **3,000+ US equities** over **5 years**, achieving **12.5% annualized returns** and **Sharpe 2.1** with industry neutralization
- Implemented signals using BRAIN's expression language after Python prototyping; optimized via **grid search**, reducing drawdown by **8%** and turnover by **40%** while **outperforming S&P 500 by 9%**

CivicOps-311: SLA Risk Prediction | *Python, SQL, PostgreSQL, FastAPI, Power BI* August 2025 – Present

- Collected, cleaned and analyzed **5,000+ 311** tickets; warehoused in **PostgreSQL** and visualized in **Power BI** for **ops stakeholders**; engineered 10+ features (rolling volume, backlog, SLA metrics)
- Trained and deployed a **Random Forest** for SLA-breach prediction model (**ROC-AUC 0.703**; **36%** recall on **4,714** closed tickets); served real-time scoring via **FastAPI** (3 endpoints) with Swagger docs
- Built ETL and analytics with **5 SQL KPI queries** (SLA attainment, P90 response) in a CI-ready repo (42 files); applied Postgres with **Docker** and documented deliverables; model accuracy **70%+**

Smart Schedule Planner | *Python, NumPy, Pandas, JSON, Git* April 2025 – Present

- Collaborated with University of Florida students to engineer a **predictive task planner**; leveraged custom linear regression models on task attributes (e.g., energy, time) to forecast success probability and grades with **92%** accuracy on simulated datasets of **500+ tasks**
- Implemented from-scratch linear regression (NumPy matrix inversion) with **sigmoid probabilistic outputs** and Pandas feature analysis; cut **prediction error** by **35%** and sped iteration cycles by **40%** via modular, clean code