Hsiang Yu Huang (Anna)

Tel: +1 617-319-5044 | Mail: huanganna1004@gmail.com | GitHub: github.com/annaandmandy

EDUCATION

Boston University Boston, MA
Master of Science in Data Science (Cumulative GPA: 3.7/4.0) *Dec. 2025*
Relevant Courses: Deep Learning, Graduate Databases, Data Engineering, Time Series

National Taiwan University of Science and Technology Taiwan
BBA in Industrial Management and Bachelor Degree Program of Finance, minor in Computer Science (Cumulative GPA: 3.85/4.3) *June 2023*
Relevant Courses: Algorithms, Machine Learning, Data Analytics, Statistics

RESEARCH EXPERIENCE

- **Research Assistant**: Machine learning for sales forecast in graphic card manufacturing
  Supervised by Prof. Chia-Yu Hsu from NTUST Artificial Intelligence and Decision Analysis Lab *Nov. 2023 - Jul. 2024*
  - Designed and implemented a machine learning pipeline to forecast sales performance, optimizing inventory management and procurement strategies for a graphics card manufacturer.
  - Decomposed sales data into fixed trend and residual components, leveraging ARIMA for trend analysis and XGBoost for residual prediction to enhance forecasting accuracy.
  - Developed a conditional rolling window model to dynamically update predictions, ensuring adaptability to real-time data changes.
  - Achieved a significant improvement in model performance, increasing the R^2 metric from 8.3% (ARIMA alone) to 73.4% using the combined ARIMA-XGBoost approach.
- **Research Assistant**: The Deployment and Shelf Analysis Recommendations for Smart Vending Machines
  Supervised by Prof. Shi-Woei Lin from NTUST Decision Analysis and Applied Statistics Lab *Apr. 2023 - Sep. 2023*
  - Developed data-driven recommendations for optimizing on-shelf product in smart vending machines based on district and area.
  - Utilized K-means clustering to categorize items by sales performance, incorporating metrics such as mean sales, coefficient of variation, revenue, and unit price.
  - Grouped clusters by unit price levels and constructed a classification tree to identify patterns and relationships in product performance.
  - Delivered actionable insights on high-performing items and provided tailored recommendations for product placement, improving inventory efficiency and profitability.
- **Research Project**: Modeling power system frequency regulation reserve trading volume based on ensemble learning
  Supervised by Prof. Shi-Woei Lin from NTUST Decision Analysis and Applied Statistics Lab *Sep 2021 - May. 2022*
  - Analyzed electricity demand patterns in ancillary market across different time periods and holidays to forecast overall demand trends.
  - Categorized electricity demand into distinct levels and developed a forecasting model using SARIMAX and Backpropagation Neural Networks (BPN) in Python.
  - Achieved a prediction accuracy of 69.44% for electricity demand levels, demonstrating the effectiveness of ensemble learning approach.

PROJECTS

Citale citaleco.com *Jan 2025 - Present*
BU Spark! Launch Lab - Back End Developer *Boston, MA*
- Designed and developed a social networking platform focused on Boston event recommendations and user networking, enhancing community engagement.
- Implemented core features for information sharing and social interaction, ensuring seamless user experience.
- Integrated Google Maps API to provide location-based map views, enabling users to visualize event locations.
- Optimized database performance using SQL queries, enabling fast and accurate search functionality tailored to user needs.

Equity in Federal Budget Earmarking Processes *Sep. 2024 - Dec. 2024*
Course: Tools for Data Science *Boston, MA*
- Conducted data mining on PDF files from the Senate Committee on Appropriations website, extracting Massachusetts funding request data using Python (Camelot library). Developed an automated workflow for data cleaning and exporting.
- Utilized Census API keys to fetch decennial race and income data from census.gov, enabling demographic analysis.
- Analyzed equity trends and funding allocations from Senator Markey's office, identifying patterns and disparities in resource distribution.
- Created visualizations to illustrate the relationship between race, income, and earmark funding, presenting to stakeholders.
- Built an interactive dashboard using Looker Studio to showcase insights and facilitate data exploration.

SKILLS

- **Computer Languages**: Python, SQL, R, C++, JavaScript, CSS, HTML, Git, LINUX, React
- **Familiar Techniques**:
  DL & ML: PyTorch, TensorFlow, Scikit-Learn, Pandas
  Database & Cloud Computing Platform: Azure, Supabase
  Visualize & Interface Design: Looker Studio, Photoshop, Indesign