TENGXIAO SONG

Email: tengxiao@seas.upenn.edu; Mobile: 223-280-7571

EDUCATION

University of Pennsylvania — **Philadelphia, PA**
Master of Computer and Information Technology — Aug 2023 – May 2025

Franklin & Marshall College — **Lancaster, PA**
Bachelor of Arts, Major in Mathematics, GPA 3.84/4.00 — Aug 2018 – May 2022

- Awards: University Honors 2019-2022, John Kershner Scholar 2022, Dean's List 2018-2019

TECHNICAL SKILLS

Domains: Object-oriented Design and Programming, Software Development, Computer Systems, Computer Architecture, Assembly Programming, Machine Learning, Deep Learning and Neural Networks, Data Modeling and Analysis
Languages and Tools: C/C++, Python, JAVA, HTML, CSS, R, SQL, MATLAB, Git, Jupyter Notebook, VS Code, AWS
Frameworks/SDK: PyTorch, Scikit-Learn, Apache Spark, Apache Storm

SELECTED PROJECTS

Cardiovascular Disease Risk Prediction Kaggle Project (Python, Scikit-Learn)

- Applied machine learning algorithms over CDC's 2021 BRFSS dataset containing over 300,000 instances and 19 features.
- Conducted data wrangling using one-hot encoding, principal component analysis, and k-folds validations.
- Employed logistic regression, random forest and XGBoost for predictions with GridSearchCV for hyper-parameter tunning.
- Evaluated models' performance using confusion matrix and Area Under the Receiver Operating Characteristic curve. The ultimate model has a recall score of 0.93 and AUROC of 0.82.

Street View House Numbers Classification (Python, PyTorch)

- Developed a classification model for recognizing house numbers from the SVHN dataset, utilizing 73,257 training samples and 26,032 testing samples from Google Street View.
- Implemented a Multinomial Logistic Regression classifier as a baseline, laying the foundation for subsequent models.
- Leveraged Convolutional Neural Networks (CNNs) for image classification, optimizing key hyperparameters (learning rate, beta1, beta2, and number of fully connected layers). Achieved a final model Test Accuracy of 0.85.

LC4 Reverse Assembler (C, Assembly)

- Engineered a disassembler to reverse assemble PennSim-generated .OBJ files, translating hex instructions into LC4 assembly code following the Instruction Set Architecture (ISA).
- Implemented a linked list architecture for LC4's program and data memories, accommodating 29 distinct ISA instructions.
- Managed complexities in memory allocation, label association, content storage, assembly representation, and node traversal within the linked list structure.

Evil Hangman (Java, JUnit, JavaDoc)

- Developed a strategic variation of the classic Hangman with dynamic programming for maximize potential future moves.
- Created a comprehensive suite of JUnit tests to ensure code reliability, covering various scenarios and edge cases.
- Generated detailed documentation using JavaDoc, enhancing code maintainability and providing a reference for developers.

PROFESSIONAL EXPERIENCE

Meituan-Dianping — Shanghai, China
Case Analyst and Data Strategist Intern — May 2023 – Aug 2023

- Analyzed and allocated rationales for 900 unsatisfied order cases per week, leveraging AI software for comprehensive case analysis; provided actionable recommendations for compensations and penalties to the involved parties.
- Generated weekly departmental reports using SQL to extract raw data from the company's database. Calculated core metrics such as refund rates and complaint rates and analyzed reasons behind metric fluctuations while forecasting future trends.
- Formulated merchant agreements on after-sales services, conducting cross-platform market research and data analysis. Achieved a 25% reduction in company operational costs.
- Conducted audits of merchant activities, assessing operational risks for over 600 merchants weekly. Reduced their visibility in the recommendation system and improved their ratings.

Poizon — Shanghai, China
Data Analytics Intern — Jan 2023 – Apr 2023

- Leveraged SQL to extract monthly sales data for merchants and their corresponding product categories, developing return margin strategies to improve transaction efficiency by 30%.
- Proposed and built databases for luxury elements, facilitating machine learning and automated plagiarism detection software, leading to cost savings equivalent to 18 employees.
- Utilized the company's image recognition software to retrieve and analyze daily new product data, distinguishing the homogeneity of over 1000 new products and executing necessary actions for quality control and diversity maintenance.