Michael Kleyn

michaelkleyn@gmail.com | 206-817-9515 | mkleyn.com

PROFESSIONAL SKILLS

Programming Languages & Frameworks: Python, Go, Java, Typescript, Next.js
Libraries, Data Science & Miscellaneous Technologies: Python Libraries *(eg. scikit-learn, numpy, pandas, matplotlib),* Git, A/B Testing, Data Science Pipeline (cleansing, wrangling, visualization, modeling, interpretation), Statistics, Experimental Design

PROJECTS

Scalable Multiplayer Tic-Tac-Toe Game | Go, gRPC, Docker, Kubernetes, Terraform, Ansible

- Developed a scalable multiplayer Tic-Tac-Toe game server using Go and gRPC, containerized with Docker and deployed on Kubernetes.
- Automated infrastructure provisioning and Linux server hardening with Terraform and Ansible; authored comprehensive technical documentation.

Serverless URL Shortener | Go, Terraform, AWS Lambda, DynamoDB

- Developed a serverless url-shortening service with AWS Lambda, DynamoDB, and API Gateway, leveraging Terraform for fully automated infrastructure deployment.
- Implemented unique id generation with collision handling, DynamoDB for scalable url storage, and api testing to ensure robust, user-friendly functionality.

NLP Classifier | Python, matplotlib, scikit-learn

- Created a suite of ML tools (Logistic/Lasso/Ridge Regression, Random Forest, Naive-Bayes) to attempt a classification problem of classifying tweets as being possibly written by IRA agents, using bag-of-words and n-grams.
- Wrangled tweet data from ~3 million tweets with 27 features. Using NLP this was made into a sparse term-document matrix of 671,000 observations, and 883,000 unique features.

EXPERIENCE

KBK Management — Remote
IT Infrastructure & Automation Lead — 2023-Current

- Led company migration to Microsoft 365 services, implementing robust security protocols to ensure compliance with financial industry regulations for handling sensitive client information.
- Engineered an automated ETL pipeline with Azure Vision and Microsoft Graph API to implement OCR and NLP techniques to extract and normalize financial data from receipts stored in SharePoint, reducing the need for manual data entry.
- Engineered and implemented a scalable Azure-based data analysis solution for identifying revenue discrepancies from multiple data sources using Azure Data Factory, Databricks, and PowerBI to automate data ingestion, extraction, transformation, and visualization.

Tao He's Laboratory (in affiliation with UCSF) — San Francisco,CA
Research Assistant — 2022-2023

- Developed a classification suite using Python, sklearn, and TensorFlow for classification of cancer diagnoses.
- Encoded n-grams from ~18,000 TCR sequence data points to feed into multiple clustering algorithms.
- Researched and tested TCR classification models (CNN, Transformers) from multiple academic papers.

San Francisco State University — San Francisco, CA
Computer Science Tutor and Advisor — 2022 - 2023

- Holding weekly tutoring sessions which are accessible for ~300 entry level Computer Science students.
- Recruited to aid and advise the CS department restructuring of "CSC 210", including on relaying important Java based fundamentals missed by students. Attended weekly meetings to discuss and track student learning outcomes and KPIs.

EDUCATION

San Francisco State University, School of Engineering and Applied Science — 2021-2024
B.S. in Applied Mathematics - Computer Science Focus | CSC210 Tutor
Relevant Coursework: Data Structures, Statistical Learning, Data Mining, Probability & Statistics, Software Dev