Ze (Zac) Tan

zac.ztan@gmail.com | 425-470-3019 | github.com/zacztan | Seattle, WA | US Permanent Resident

Summary   Data science enthusiast with experience in backend development and data engineering. Seeking to build scalable systems and robust data pipelines that enhance business performance.

Experience

Software Engineer (Backend / Data) | **AdTech Startup** | Seattle, WA | 2020 - Present
- Built API endpoints in Flask and implemented functionality to support various prediction logic.
- Improved Python SDK usability and functionality for campaign management tasks.
- Created internal tools that simplified data science and client success workflow.
- Optimized regex-based keyword filtering, reducing cloud compute costs by ~40%.
- Automated data preparation and report generation tasks for analysis and metrics dashboards.
- Worked on customer usecase definition, ETL / ELT data integration and pipeline configuration.
- Built helper functions library that simplify data pipelines creation in the AWS ecosystem.

Python Developer (Data Science) | **Freelancer** | San Francisco, CA | 2017 - 2019
- Developed a multithreaded web scraper to collect over 750,000 points of interest.
- Created database tables and relationships using an Object Relational Mapper (ORM).
- Implemented native full-text search and resolved concurrency issues in PostgreSQL.
- Worked on data preparation (cleaning, wrangling, aggregation) and feature engineering.
- Worked on model selection, training, validation and hyperparameter optimization.
- Built machine learning pipelines for supervised / unsupervised learning (classification, regression, clustering, dimensionality reduction, visualization) and neural networks.
- Worked on natural language processing tasks (text cleaning, tokenization, lemmatization, word cloud, bag of words, tf-idf, word embedding and topic modeling)

Chip Verification Engineer | **Broadcom Corporation** | Irvine, CA | 2007 - 2014
- Developed and managed VB software implementation from pre-silicon to production.
- Analyzed chip functionality and performance data across various operating conditions to ensure compliance with datasheet specifications.
- Performed code reviews to optimize execution time, contributing to cost reduction initiatives.
- Developed standalone VBA tools to minimize errors in software release.

Education

University of California, Los Angeles, CA
- M.S. Engineering - Signal Processing and Communications

Texas A&M University, College Station, TX
- B.S. Computer Engineering (Minor in Mathematics)

Skills   **Database:** PostgreSQL, SQLAlchemy, SQL
Data Mining / Engineering: Requests, BeautifulSoup, Selenium, Airflow, dbt, Docker, Flask
Data Science: Pandas, NumPy, SciPy, Matplotlib, Seaborn
Machine Learning: Scikit-Learn, Keras, TensorFlow
Cloud: AWS (DynamoDB, ElastiCache, Kinesis, Lambda, Redshift, RDS, SES, MWAA, Athena)