Andrew Chang

github.com/itsmandrew | linkedin.com/in/itsmandrew | theachang1688@gmail.com | Los Angeles, CA

EDUCATION

University of California, Irvine — Irvine, CA
BS in Computer Science, Minor in Statistics & Informatics, Dean's List (7x) — *September 2020 – December 2024*

EXPERIENCE

Data Engineer Intern — September 2023 – Present
NBA - San Antonio Spurs — *San Antonio, TX*
- Implemented an extract, transform, and load pipeline using Apache Kafka and Airflow to ingest data from Dynamix CRM, Archtix API, Oracle Eloqua, and other sources into our Azure SQL Data Warehouse, resulting in a 20% reduction in data redundancy, 15% decrease in storage costs, and 25% faster read times for analysis
- Collaborated with the analytics team to pull data from our Azure SQL Data Warehouse to train a recurrent neural network to show company executives using Spark a real-time streaming dashboard on public sentiment about the Spurs
- Migrated the entire ETL library from OneDrive to a GitHub repository automating pull requests, leading to a 40% improvement in version control efficiency and reducing code collaboration conflicts by 30%.

Undergraduate Researcher — June 2023 – August 2023
UC Irvine Texera — *Irvine, CA*
- Collaborated with a team to optimize our relational database, implementing Spark's distributed clusters which reduced processing times by 45% and improved large-scale data analytics efficiency by 30%
- Utilized hash, range, and bucketing partitioning techniques in Spark, achieving a 60% reduction in query latency and boosting data processing efficiency by 40% for extensive datasets
- Devised comprehensive unit tests using the PyTest testing framework to validate the performance of PySpark SQL queries leading to a test coverage of 95% and reduction of data inconsistencies

Programming Instructor — January 2022 – November 2022
WhizKids Computer Center — *San Jose, CA*
- Lectured classes of 3-5 students in topics including Python, Java, front-end web development, PyGame, competitive programming, and machine learning by creating in-lecture challenge activities through Github
- Motivated student interest in computer science through personally developed take home projects such as weather forecasting, quiz/studying, and type racing based web-apps shown by a returning student rate of 90%

Instructional Assistant — December 2021 – June 2023
UC Irvine Donald Bren ICS — *Irvine, CA*
- Assigned exercises on Github and graded weekly homework assignments for UC Irvine's ICS 46: Data Structure Implementation and Analysis & ICS 32A: Python Programming and Libraries (Accelerated)
- Oversaw open lab hours three times weekly to help debug course projects and implement data structure concepts such as DAG graphs, hash maps, and AVL trees resulting in a 95% positive anonymous student-tutor feedback

PROJECTS

Instagram Lite | *React, Express, Node.js, MongoDB*
- Engineered a social media platform capable of storing 50,000+ users and posts dynamically in MongoDB with responsive user interaction for mobile and desktop using React/Redux
- Implemented token-based authentication using JSON Web Tokens (JWT) upon successful user logins and user requests with the server and created RESTful API endpoints using Express to handle CRUD operations

Search Engine | *Python, Flask, HTML/CSS*
- Restructured an inverted index algorithm to effectively store 30,000+ documents locally and incorporated a TF-IDF/cosine similarity algorithm to efficiently retrieve information with a query response time less than 150ms
- Increased the rate of web-crawling the UCI directories by over 150% via cache server by implementing multithreaded web-crawlers and scraped web pages using nltk and BeautifulSoup

Minesweeper Solver | *Python*
- Coordinated and distributed work among a team of four to modularize the program into 4 different deterministic layers and 2 probabilistic layers, resulting in our algorithm being in the 98th percentile for algorithmic speed
- Utilized heuristic algorithms and the Monte Carlo method to deduce the best possible solution for a board, resulting in 85%, 80%, and 34% solve rates for beginner, intermediate, and expert boards respectively

TECHNICAL SKILLS

Languages: Python, Java, Go, C/C++, Scala, SQL (Postgres), JavaScript, HTML/CSS
Developer Tools: Git, Docker, Jenkins, Postman, Linux, Unix
Frameworks: Node.js, flask, pandas, Keras, PyTorch, scikit, TensorFlow, Apache Spark
Current Tools: ETLs, SQL (Postgres, Redshift, MySQL), NoSQL (MongoDB), Airflow, Microsoft Azure