Joshua Chen

408-505-5987 | joshua.ycc.chen@gmail.com | linkedin.com/in/joshua-ycc-chen | github.com/Zaanis

Education

UC San Diego — San Diego, CA
Masters of Science in Business Analytics — *Aug. 2023 – June 2024*

UC San Diego — San Diego, CA
Bachelors of Science in Cognitive Science, Specialization in Machine Learning — *Sep. 2021 – June 2023*

- Magna Cum Laude, GPA: 3.966, Major GPA: 4.0

Experience

Machine Learning Engineer — March. 2024 - Present
AlphaTrAI — *San Diego, CA*

- Developed a LLM to analyze sentiment from earnings call transcripts, enhancing market prediction accuracy
- Engineered a system to integrate sentiment analysis with 900K rows of historical stock price data and technical indicators, utilizing Snowflake and SQL for efficient ETL processes
- Designed an LSTM model to forecast stock prices using sentiment and historical data, incorporating a custom loss function for enhanced prediction accuracy
- Conducted in-depth data analysis and validation, ensuring the robustness and reliability of predictive models for financial market forecasting

Teaching Assistant - COGS108 Data Science in Practice — Sep. 2023 - Present
UC San Diego — *San Diego, CA*

- Developed a Python script to automate assignment grading, enhancing efficiency in the evaluation process
- Provided one-on-one mentorship to students, contributing to a 10-15% improvement in their grades
- Published a tutorial video instructing students on setting up SSH keys for GitHub integration with DataHub and local Jupyter Notebooks, facilitating efficient project management and coding workflow

Projects

Restaurant Type Prediction Using Text Reviews | *NLTK, Gensim, Sci-Kit* — Jan. 2024 - Mar. 2024

- Secured 1st in a Kaggle competition by developing a classification model to predict restaurant types
- Applied NLP preprocessing techniques and enhanced TFIDF features for improved model performance
- Boosted model performance by applying training data labels to test predictions, identified via targeted EDA

Comparative Analysis of NLP Techniques | *TensorFlow, Transformers, NLTK* — Jan. 2024 - Mar. 2024

- Authored an ACM format paper focusing on TFIDF, Word2Vec, Doc2Vec, and DistilBERT for text classification
- Developed a DistilBERT model, enhancing news content categorization accuracy by over 5%
- Resolved mismatches between Transformers and TensorFlow, demonstrating adeptness in resolving technical issues

Experimentation in Custom Neural Network Development | *Python* — Jan. 2024 - Mar. 2024

- Engineered a custom neural network including softmax regression and backpropagation, achieving notable improvements in model accuracy with detailed accuracy plotting.
- Optimized network with momentum and regularization techniques; momentum enhanced performance by 6%, while L1 regularization outperformed L2, improving accuracy by 1%.
- Conducted extensive experiments with various activation functions, finding ReLU to be the most effective, and managed configurations using YAML, ensuring efficient and organized testing.

Generation of Code using CharRNN | *Python, PyTorch* — April 2023 - June 2023

- Developed a CharRNN using GRU architectures, trained on a dataset of over 100,000 lines of Python code
- Optimized the model with CUDA, achieving a 250% reduction in training time
- Conducted hyperparameter experimentation with over 50 combinations, achieving 20% reduction in validation loss

Technical Skills

Languages: Python, SQL (PostgreSQL, T-SQL, MySQL), R, Java
Softwares: Snowflake, Excel, Microsoft Windows, Tableau, Tableau Prep Builder, PowerBI, JMP Pro
Developer Tools: Git, Docker, VS Code, RStudio, Jupyter Notebook
Libraries: Sci - Kit Learn, Transformers, BeautifulSoup, TensorFlow, PyTorch, XGBoost, NLTK, Gensim