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*
- Strong focus on Machine Learning

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
- Praised for exceptional efficiency, adherence to deadlines, organizational skills, and competence as a staff member

PROJECTS

Research on Coordinate Descent Algorithms | *Python, Sci-Kit*   Jan. 2024 - Mar. 2024
- Authored comprehensive review paper on coordinate descent algorithms, analyzing various methods
- Conducted thorough research into existing coordinate descent algorithms, gathering insights and best practices to inform the development of the enhanced algorithm
- Synthesized selected algorithms for enhanced efficacy; achieved 40% loss reduction with max gradient approach and backtracking line search
- Demonstrated strong ML engineering skills through algorithm optimization and empirical research validation

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%
- Successfully configured GPU acceleration, achieving a 200% reduction in model training time
- Resolved mismatches between Transformers and TensorFlow, demonstrating adeptness in resolving technical issues

NFL Combine vs Career Length | *XGBoost, Sci-Kit, BeautifulSoup*   April 2023 - June 2023
- Mastered BeautifulSoup for advanced web scraping, creating a unique 8k-entry dataset, published on Kaggle.
- Addressed and cleaned 25% of the dataset containing null values, enhancing data quality and reliability
- Conducted EDA with Seaborn and Matplotlib; identified outliers, uncovered correlations for model development
- Evaluated multiple classifiers, regressors from XGBoost, Sci-Kit Learn, focusing on performance analysis
- Executed hyperparameter tuning, 5-fold cross validation, and grid-searching, achieving a peak accuracy of 70%

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