Chia-Lin Hsieh

+886 989587647 | jeffrey.hsieh14@gmail.com | linkedin.com/in/jeffrey-hsieh-chialin/ | jeffrey-hsieh-portfolio.vercel.app/

Education

New Jersey Institute of Technology Newark, NJ
M.S. in Data Science - Computational Track GPA: 3.75/4.0 Jan 2021 – Dec 2022
Coursework: Machine Learning, Deep Learning, Data Analytics with R Program, Enterprise Database Management
Certificate (Coursera): IBM Data Science Professional Certificate, Google Data Analytics Professional Certificate

Fu Jen Catholic University New Taipei City, Taiwan
B.B.A. in Statistics and Information Science Sep 2015 – Jul 2019
Coursework: Mathematical Statistics, Data Mining, Regression Analysis, Time Series Analysis, Multivariate Analysis

Work Experience

Rain Spring Technology Co. Ltd Taipei, Taiwan
AI Engineer | Data Scientist Aug 2023 – Present

- Deployed and optimized generative AI models to enhance workflows, improve performance, and reduce operational costs through AI-driven automation.
- Applied LangChain and Retrieval-Augmented Generation (RAG) techniques to fine-tune large language models (LLMs) for business-focused knowledge management.
- Integrated advanced AI models for image generation, video synthesis, and voice cloning, fine-tuning hyperparameters to optimize quality and assess model performance for business impact.
- Developed AI-driven automation and decision-support systems to improve efficiency, optimize resources, and refine sales strategies through intelligent algorithms.
- Built efficient ETL data pipelines and conducted exploratory data analysis (EDA) to provide real-time, actionable insights for improving marketing strategies and customer personalization.

Bank of America Newark, NJ
Data Analyst Jan 2022 – May 2022

- Created a prediction model to determine the scale of service down times as a result of weather impacts through XGBoost and Random Forest regressors
- Improved customer satisfaction by reducing incorrect model predictions by 25% through feature engineering and incorporation of user search data and real-time data
- Developed an interactive, real-time dashboard and Flask web application a to visualize predictions and key metrics through Python and HTML
- Implemented an ETL pipeline to automate data processing and to improve data quality and consistency
- Conducted exploratory data analysis using Python, including examining correlations with the target variable, creating scatter plots, distribution plots, and utilizing heat maps to identify meaningful patterns and relationships within the data

Project Experience

Amazon Food Reviews Sentiment Analysis Mar 2023 – Jun 2023

- Detected sentiment patterns across customer segments through VADER and the RoBERTa transformer model to enable targeted marketing strategies to improve customer satisfaction and loyalty
- Analyzed low-scoring reviews with positive sentiments to unveil areas for user experience improvement and product changes
- Utilized scatterplots to compare the performance of NLP models and to improve data interpretability

SpaceX Landing Success Predictor Feb 2023 – Mar 2023

- Predicted the reusability of the Falcon 9 first-stage landers to determine launch costs through machine learning models such as decision trees, SVM, KNN, and regression to motivate
- Developed data ingestion and preprocessing pipelines for historical launch records from diverse data sources such as web scraping and API requests
- Analyzed and presented insights on the factors behind rocket launch successes to aid in bidding strategies for the most successful long-term items
- Constructed an interactive dashboard and map via Dash to display the success rate for the landings based on locations

Technical Skills

Programming Languages: Python, SQL, R, Java, JavaScript, HTML/CSS
Methodologies: Machine Learning, Deep Learning, Natural Language Processing; Statistical Inference, A/B Testing, Hypothesis testing, Probability, Simulations, Confidence Intervals, Correlation, Regression
Frameworks & Tools: Power BI, Tableau, Matplotlib, Tensorflow, Pytorch, Git, JIRA, Docker, MS Office, Jupyter, Pega
Data Engineering: ETL, MS SQL Server, IBM Db2, Azure, Hive, Snowflake, Spark, MongoDB, AWS