Chia-Yi Yen

+1(925)819-9608 • Seattle, WA • chia-yi.yen@utdallas.edu • https://www.linkedin.com/in/chiayi-yen/

SKILLS

Programming: Python, SQL, Hive SQL, Spark, R, SAS
Database: SQL, MySQL, MongoDB, Hive, RDBMS, Snowflake
Visualization: Tableau, PowerBI, Python
Other: Excel, ETL, Hadoop, Data Modeling, Machine Learning, Knime, Salesforce, VS Code
Languages: Chinese (Mandarin), Hokkien, Italian

PROFESSIONAL EXPERIENCE

PACCAR Financial | *Sales and Service Analytics Intern* — Washington, US **June 2024 - September 2024**

- Designed and developed ETL pipelines to extract data from Snowflake, process, and analyze in Python, exporting the results to Excel for reporting. Enhanced the reliability and accessibility of the pipelines, improving data processing efficiency by 10x.
- Initiated and developed UTC (Used Truck Center) Recommender using Python, identifying optimal UTCs for selling different type of truck. Significantly enhanced sales efficiency and increasing daily operations rates by 50%.
- Conducted inventory days forecasting for current stock using logistic regression and random forest models, providing the pricing and sales teams with valuable insights on how to adjust their pricing and business strategies.
- Developed and automated Tableau dashboards for real-time sales date visualization and trend identification, and created Salesforce reports to provide insights to senior management, enhancing sales performance with data-driven recommendations.

Taroko Door and Window Technologies Inc. | *Project Manager* — Kaohsiung, Taiwan **February 2020 - August 2021**

- Constructed and implemented ETL pipelines to aggregate data from diverse sources, fulfilling 90% of analysis requirements and enabling efficient export of findings to Excel for reporting.
- Conducted customer segmentation and competitor analysis using SQL and Excel, informing pricing and product positioning decisions for the company's first e-commerce product, successfully launched on Amazon during the pandemic.
- Orchestrated cross-functional meetings across four departments to establish a new process for developing products, improving sales productivity and efficiency metrics, resulting in a 50% faster project delivery.
- Conducted sales forecasting and created reports to present insights for executive leadership, while improving business processes by streamlining documentation and order flow charts, resulting in a 60% increase in efficiency.

PROJECTS

Sentiment Analysis of Online Food Reviews Using NLP, Applied Natural Language Processing **February 2024 - May 2024**

- Applied NLP techniques (tokenization, sentiment analysis, topic modeling, named entity recognition) using Python (NLTK, SpaCy) to analyze online customer reviews, uncover keywords, sentiment scores, and inform business decisions.

Convolutional Neural Network Fine-tuning, Applied Deep Learning **September 2023 - December 2023**

- Achieved an impressive test accuracy of 99.46% with a low loss of 0.0202 by using Keras Tuner to fine-tune hyperparameters for a Convolutional Neural Network (CNN) model, optimizing performance and enhancing predictive accuracy.

CO2, Methane, GDP, Population, and Temperature Relations, Applied Machine Learning **March 2023 - May 2023**

- Created and assessed various models, including Ridge, Lasso, Random Forest, Decision Tree, and K-Nearest Neighbors, to evaluate performance using k-fold cross-validation to determine the most effective solution.

Truck Data Analysis, Big Data **February 2023 - May 2023**

- Leveraged Spark and Hadoop to process extensive datasets, using clustering analysis to identify driving behaviors of truck drivers, and developed visualizations and dashboards with Tableau for enhanced data insights.

Classification- Decision Tree Analysis, Modeling for Business Analytics **September 2022 - December 2022**

- Developed decision tree models to predict customer behavior, using cross-validation for training and performance evaluation. Applied random forests and KNN to identify customer clusters, aiding targeted marketing strategies.

Movie Box Revenue Prediction, Modeling for Business Analytics **August 2022 - November 2022**

- Used Principal Component Analysis (PCA) to identify significant movie features and principal components, deploying them in predictive models to predict movie revenue, resulting in 88% accuracy.

CERTIFICATION

Graduate Certificate in Applied Machine Learning **December 2023**
Google Advanced Data Analytics Certificate **January 2024**

EDUCATION

The University of Texas at Dallas — Texas, US **August 2024**
MS, Information Technology and Management GPA 3.83
National Sun Yat-sen University — Kaohsiung, Taiwan **June 2018**
MS, International Relations GPA 3.99
Fu Jen Catholic University — New Taipei, Taiwan **June 2015**
BA, Italian Language and Culture (Exchange program: L'Università Ca' Foscari — Venice, Italy) GPA 3.10