Jeffrey Hsieh

+1 (551) 292-0745 | ch35@njit.edu | linkedin.com/in/jeffrey-hsieh-chialin/ | jeffrey-hsieh-portfolio.vercel.app/

EDUCATION

New Jersey Institute of Technology GPA: 3.75/4.0 Newark, NJ
Master of Science in Data Science Jan 2021 – Dec 2022
Coursework: Machine Learning, Deep Learning, Data Analytics with R Program, Enterprise Database Management

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

EXPERIENCE

Bank of America Newark, NJ
Lead Developer Jan 2022 – May 2022

- Extracted 8K+ rows datasets using APIs and imported the data into a SQL server allowed for easy access
- Conducted data preprocessing techniques, including handling missing values and normalizing columns
- Implemented exploratory data analysis (EDA) using Python, including examining correlations with the target variable, creating scatter and distribution plots, and utilizing heatmaps to identify patterns in the data
- Trained and tested XGBoost and Random Forest Regressor models to predict the impact of customer from power outages, achieving a remarkable R-squared value of 93.6%
- Built an interactive and real-time application using Python and HTML to visualize the predictions and provide additional features for users to explore the data

New Jersey Institute of Technology Newark, NJ
Teaching Assistant Jan 2022 – Dec 2022

- Coordinated study events to improve students' Python programming skill and deep learning coursework study
- Assisted students in accelerating their onboarding process, and enhancing their skills in building deep learning network, resulting in a 20% improvement in their course scores

ACADEMIC PROJECT

Amazon Food Reviews Sentiment Analysis Mar 2023 – Apr 2023

- Conducted sentiment analysis on 560K+ rows text data using two NLP techniques including VADER and RoBERTa
- Visualized and compared the results of sentiment analysis using Seaborn pairplot to highlight the differences between the predictions made by the two techniques
- Extracted valuable user feedback on user experience issues by identifying contradicting ratings and comments, and saved them for further analysis to identify common product issues and extract key words

Space X Falcon 9 Landing Prediction Feb 2023 – Mar 2023

- Developed predictive models with 88% accuracy for Falcon 9 first-stage landings using various algorithms such as Decision Tree, Support Vector Machine, KNN, and Regression
- Collected and cleaned historical launch records from multiple sources by conducting comprehensive data collection and preprocessing, including web scraping and API requests
- Analyzed and presented insights on the success rate of rocket launches under different variables by utilizing data visualization techniques such as scatter plots, bar charts, and line charts
- Developed an interactive dashboard using Dash library to display the success rate for the landings in different locations, combined with a map visualization.

Digital Recognizer | Kaggle Competition Oct 2022 – Nov 2022

- Developed a Convolutional Neural Network (CNN) using TensorFlow and Keras to accurately recognize handwritten digit images with a high accuracy of 99.6%
- Optimized hyperparameter by using Cross validation and RandomSearch techniques to be increased 10% accuracy compared to previous model

Efficiently Managing Baseball Dataset Sep 2022 – Dec 2022

- Developed a comprehensive SQL database for a baseball dataset, implementing advanced queries and functions to extract insights from the data and enhance decision-making processes
- Demonstrated proficiency in database management by creating tables and views, designing a user-friendly interface, and optimizing the database for maximum efficiency and accuracy
- Leveraged transaction processing techniques to ensure data consistency and reliability, mitigating potential errors and improving the overall quality of the database

SKILLS

- **Programming Languages**: Python, SQL, R, Java, JavaScript, HTML/CSS
- **Frameworks & Tools**: Power BI, Tableau, Jupyter Notebook, Scikit-learn, Matplotlib, Tensorflow, Pytorch, Git, JIRA
- **Database & Data Warehouses**: MS SQL Server, IBM Db2, Azure, Hive, Snowflake, Spark, MongoDB