

Jeffrey Hsieh

(551) 775 - 2150 | ch35@njit.edu | [linkedin.com/in/jeffrey-hsieh-chialin/](https://www.linkedin.com/in/jeffrey-hsieh-chialin/) | jeffrey-hsieh-portfolio.vercel.app/

EXPERIENCE

Bank of America

Newark, NJ

Data Analyst

Jan 2022 – May 2022

- Achieved 93.6% R^2 value of predicting customer impact from weather info using XGBoost and Random Forest Regressor
- Enhanced predictive model dynamically by incorporating user search data, real-time information, and previous prediction results, leading to a 25% reduction in incorrect predictions and a notable improvement in customer satisfaction
- Developed an interactive and real-time dashboard to visualize predictions and key metrics by using Python and HTML
- Implemented an ETL pipeline to automate datasets processing, handling 8K+ rows, improving 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

New Jersey Institute of Technology

Newark, NJ

Teaching Assistant

Jan 2022 – Dec 2022

- Coordinated study events to enhance students' Python programming skills and understanding of deep learning coursework
- Guided 10+ students through an accelerated onboarding process, leading to a 20% improvement in the final scores

PROJECTS

Amazon Food Reviews Sentiment Analysis

Mar 2023 – Jun 2023

- Detected sentiment patterns and trends by analyzing the sentiment polarity and intensity across customer segments, enabling targeted actions to improve customer satisfaction and loyalty
- Unveiled valuable insights into user experiences and potential areas for improvement by analyzing the significant contrast between the given rating and the sentiment expressed in the comments
- Conducted sentiment analysis on 560K+ rows text data using two NLP techniques including VADER and RoBERTa
- Enhanced data interpretation by utilizing Seaborn Pairplot to visually present and compare the results of the analysis

SpaceX Landing Success Predictor

Feb 2023 – Mar 2023

- Engineered predictive models with an impressive 88% accuracy for Falcon 9 first-stage landings, employing a diverse range of advanced algorithms including Decision Tree, Support Vector Machine, KNN, and Regression
- Collected and cleaned historical launch records by conducting data collection and pre-processing, 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
- Constructed 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 multiple Convolutional Neural Networks (CNN) using TensorFlow and Keras, achieving exceptional recognition accuracy of 99.6% in identifying handwritten digit images, ranking in the top 10% among over 2K+ teams
- Employed Cross validation and RandomSearch techniques to optimize parameters, resulting in a 20% increase in accuracy compared to the previous model

Baseball Data Management Optimization

Sep 2022 – Dec 2022

- Created a high-performance database system for managing a vast dataset of baseball games, consisting of 450k+ rows, resulting in a 30% enhancement in querying efficiency
- Devised a customized script aligned with the database structure and query demands, resulting in a 25% reduction in average wait time and a 15% increase in accuracy and speed for performance data retrieval
- Crafted a user-friendly interface that simplifies the process of accessing performance records, allowing effortless selection of individual players and games

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

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