JANHI ONG

(202) 330-7326 | janhiong04@gmail.com | linkedin.com/in/janhiong/ | github.com/Janhi2004

EDUCATION

Drexel University - College of Computing and Informatics **Philadelphia, Pennsylvania**
Bachelor of Science in Data Science - **GPA: 3.95**/4.0 | Dean's List Graduating: June 2027

TECHNICAL SKILLS

- **Tools:** Python, SQL, Tableau, Power BI, Excel, R, HTML, Tux, JavaScript, HTML/CSS, Looker, Qlik, Angular, Agile, Azure, Java, Linux, C++, JSON, Cloud Computing, Salesforce, Jira, Hadoop, Business Intelligence
- **Libraries/Frameworks:** A/B Testing, Pandas, Matplotlib, NumPy, SciKit-Learn, dplyr, ggplot2, Tidyverse, TensorFlow, jQuery, React.js, Apache Spark, Pytorch, Keras, BigQuery, Snowflake

WORK EXPERIENCES

Data Science Analyst Intern March 2025 - Present
Panasonic Remote - New Jersey
- **Parsed and normalized raw invoice data** from unstructured CSV files by merging fragmented key-value pairs using Python (Pandas), enabling accurate extraction of financial fields for downstream processing
- Maintained and analyzed a centralized **Lake Database** containing customer and tax data by using **PySpark** on **Microsoft Fabric** to join and transform 15+ raw tables, enabling the company to **loyal customers**
- **Developed interactive Power BI dashboards** to visualize trends in customer behavior and tax transactions, enabling stakeholders to make data-driven decisions and monitor KPIs in real-time

Research Assistant February 2024 - May 2024
Drexel University Gettysburg, PA
- Enhanced data coherence using advanced Excel functions (Power Query, VLOOKUP, Pivot Table), developing long-term usability and improving attention to detail skills and critical thinking
- Conducted **data mining** and **data cleaning** using R, ensuring that over 100,000 student records were properly formatted, improving problem-solving skills, organizational skills and decision making skills

Data Science Analyst Intern July 2024 - December 2024
Chase Cost Management (CCM) New York, NY
- Performed **exploratory data analysis**, **data visualization,** and **data modeling** for 8+ financial datasets using **MySQL** and **Tableau** to report findings to Sales team, enhancing revenue forecasting accuracy for 2024
- Automated **data warehouse** with **Apache Airflow** and **cloud technologies** and collaborated with vice president and Technology team to streamline file matching, reducing manual work by 35%
- Performed project management using **Jira, Agile, and DevOps** methodologies while collaborating with cross-functional department and customers, ensuring customer service transformation and product management

Data Science Analyst Intern May 2024 - July 2024
FPT Software Corporation Vietnam
- Applied **hypothesis testing** and **business analytics** to decide key factors influencing customer behavior while shopping online using Python and Google Analytics, expected to result in an increase of $3.5M/year
- Created **metrics** using BI tools **Power BI**, demonstrate **data storytelling** on E-commerce Dashboard and delivered product analytics and data models to **Marketing teams**, improving **verbal** and **written communication skills**
- Optimized **ETL processes** using **Google Cloud** and **MySQL**, providing technical support and reducing time spent on manual data entry by 25% and improving the accuracy of user behaviors database

PROJECTS

Facial Recognition with Supervised Learning - Machine Learning (github.com/Facial-Recognition**)** October 2024
- Applied **regression models** and deep learning to distinguish individuals, achieving 82.9% cross-validation accuracy
- Implemented a **CI/CD pipeline** in **GitLab** to automate testing and deploying data processing application
- Applied **regression models** and deep learning to distinguish individuals, achieving 82.9% cross-validation accuracy

A/B Testing on Udacity Free Trial Screen - A/B Testing (github.com/Udacity-Free-Trial-Screen**)** August 2024
- Implemented **A/B test** to optimize Udacity's free trial enrollment, improving key metrics and product features
- **Analyzed large datasets** using **advanced analytics** and **predictive models** with **Python** and **Jupyter Notebook**, visualized control and experimental group performance, and conducted sanity checks on key metrics
- Applied **regression models** and deep learning to distinguish individuals, achieving 82.9% cross-validation accuracy