


Josephine Widjaja

✉ j4widjaj@uwaterloo.ca 🔗 josephine-widjaja.com  /in/josephine-w/

SKILLS

PROGRAMMING AND DATABASE:

Python, SQL, R, Javascript, C/C++, HTML/CSS

LIBRARIES/FRAMEWORKS:

Spark, React, Pandas, NLP, Numpy, Scikit-Learn, Matplotlib

OTHER SOFTWARE/TOOLS:

Power BI, Tableau, RedShift, Excel (Pivot Tables, V-Lookups, VBA), AWS, Jira, Airflow

EXPERIENCE

Data Engineer Intern

05/2024 – 08/2024

Meta

- Developed a robust measurement system to optimize video recommendation platform performance
- Created a new webpage for trigger surface intelligence using **React** components and **MySQL/Presto** Tables
- Implemented rigorous **data quality checks** to ensure data consistency and prevent data loss in pipelines

Data Engineer Intern

01/2024 – 04/2024

Tesla

- Spearheaded the migration of over **20 Spark jobs** across Tesla Energy products from legacy **Airflow** instance and **Jenkins** to a standardized **Airflow** environment
- Collaborated with cross-functional teams to develop comprehensive data tables to track fleets of approximately **5000+ Energy products**, facilitating efficient monitoring and management of assets.
- Addressed ad-hoc data requests from stakeholders, including **SQL query optimization**, handling complex query requirements, and contributing to the development of **ETL** jobs

Software Engineer Intern

05/2023 – 09/2023

Tatari

- Led the migration and development of an **ETL** pipeline from **Redshift** to **Databricks** using **Pyspark** and ensured seamless data processing for spend data calculations
- Participated in the successful implementation of the **Great Expectations** tool for comprehensive data quality checks
- Compiled distinct solutions for **Opsgenie alerts**, resulting in a 10% increase in efficiency for weekly on-call efforts

Data Science Intern

09/2022 – 12/2022

KPMG

- Developed an **NLP Model** using **spacy** text classification, **HuggingFace question-answering**, **camelot** table extraction, and **Python regex** to process PDF reports for business stakeholders.
- Drove integration efforts to enhance the compliance report tracking system by incorporating new features into the **Power BI Dashboard** for internal business stakeholders.
- Provided continuous technical support for **Power BI** and **Excel** needs of stakeholders.

Business Technology Analyst Intern

01/2022 – 04/2022

Bank of Montreal

- Spearheaded the creation of a dynamic **Power BI dashboard**, facilitating data visualization and automating fiscal reports, resulting in enhanced decision-making and time-saving for stakeholders.
- Engineered a highly efficient **Excel** data reconciliation tool using **VBA** and **Pivot Tables**, reducing manual effort and minimizing errors in financial data analysis.

Machine Learning Engineer Intern

05/2021 – 08/2021

Assent Compliance

- Built, trained, and tested pre-trained **Natural Language Processing** models through **AWS Sagemaker** to determine ideal models to process reports for internal stakeholders
- Contributed to the development of models with tasks like Named-Entity Recognition (NER) and Image Classification

EDUCATION

Candidate for Bachelor's of Computer Science

09/2020 – 05/2025

University of Waterloo