

Albert Wei

(203) 529-7086 | New York, NY | alwei@umich.edu | weialbert.github.io

Education

University of Michigan

MSE Industrial & Operations Engineering – Operations Research

Ann Arbor, MI

BSE Industrial & Operations Engineering

Experience

Palantir Technologies | Data Engineer

2025 — present

New York, NY

- Migrated petabyte-scale production datasets to iceberg table format, optimizing storage costs by 40+% and enabling real-time analytics capabilities for enterprise stakeholders

Capital One | Senior Data Analyst

2021 — 2025

McLean, VA

- Refactored legacy retail bank overdraft risk model monitoring pipeline, combining 8 cloud data sources and optimizing runtime from 4+ hours to 20 minutes through parallelization and Apache Spark compute optimization
- Owned US Card complaints data pipeline (Snowflake SQL and Python), migrated data source to S3 enabling real time updates, reduced dataset query time from 300 to 1 seconds, and oversaw change management for 100+ analysts onto production system serving SVP+ leadership
- Designed and implemented intelligent real-time recoveries analytics infrastructure using Airflow, Apache Hamilton, and Snowflake supporting \$10M+ core business decision-making
- Led federal Community Reinvestment Act (CRA) compliance data strategy coordinating cross-functional teams and delivering regulatory submissions to US Treasury OCC
- Built central data quality framework using Python and Apache Airflow, eliminating 20+ organization-wide adhoc solutions and enabling advanced statistical trend analysis capabilities with scalable infrastructure supporting 70+ production datasets

University of Michigan – Transportation Research Institute | Model Developer & First Author

2018 — 2021

Ann Arbor, MI

- Automated landmark and feature extraction from 1,000+ CT images with Machine Learning radial basis functions and cubic spline interpolation
- Published peer-reviewed first-author paper on parametric head geometry modeling using PCA and multivariate linear regression in MATLAB

Technical Skills

- **Programming & Core Tools:** C++, Rust, Python (pandas, numpy, scikit-learn, TensorFlow, PyTorch), SQL (Snowflake, PostgreSQL, NoSQL), R, CI & CD (Github Actions, Jenkins)
- **Data Engineering & Infrastructure:** Apache Spark, Airflow, Kafka, Hamilton, AWS (S3, Glue, Redshift, EMR, SageMaker, Lambda, Athena), Azure (Databricks, Blob Storage, Data Factory, Synapse), GCP (BigQuery, Cloud Storage, Dataflow, Dataproc), Docker
- **Analytics & Visualization:** Jupyter Notebooks, Tableau, Power BI, Statistical Regression Analysis, Time Series (ARIMA, Prophet), A/B Testing
- **Machine Learning & Optimization:** Predictive Modeling, Natural Language Processing, Anomaly Detection & Variance Testing, Linear & Nonlinear Optimization, Operations Research