

# Evan Ditter

Machine Learning Engineer | [evannditter@gmail.com](mailto:evannditter@gmail.com) | [evan@evanditter.com](mailto:evan@evanditter.com) | 763-367-0379  
[evanditter.com](https://www.linkedin.com/in/evanditter/) | in <https://www.linkedin.com/in/evanditter/> | <https://github.com/evanditter>

## SKILLS

**Programming Languages:** Intermediate/Advanced - Python | SQL | Familiar with - C/C++ | Java | HTML | CSS | JavaScript | TypeScript  
**Cloud Technologies:** Azure: ADF/Functions/Logic Apps/VMs/Key Vault/Blob/CosmosDB | AWS: EC2/Lambda/S3/Redshift/DynamoDB  
**Engineering Toolkit:** Git | Spark | Databricks | Snowflake | dbt (data build tool) | NoSQL | Jenkins | Flyway | Jupyter | Astronomer/Apache Airflow | Power BI | Kafka | Excel/Microsoft Suite | Windows/Mac OS/Linux  
**Concepts and Design:** OOP | Agile | Software Development/Testing | Problem Solving | Big Data Engineering | Data Analytics | ETL | Data Architecture | Machine Learning | Descriptive/Predictive Analytics | CI/CD | Artificial Intelligence

## EXPERIENCE

### JP Morgan Chase

May 2024 - present

Machine Learning Engineer

New York, NY

- Trained & fine-tuned Computer Vision Machine Learning models to increase our ability to read financial documents.
  - Trained / Fine-tuned a Vision Encoder Decoder which performed Optical Character Recognition (OCR) which lead to an increase in accuracy from the previous model by 2 percent (93 to 95%) and allowed for a 6 percent increase in straight through processing for checks. This equates to thousands of more transactions a day being processed directly.
  - Fine-tuned multiple image segmentation models to recognize segments of financial documents.
- Designed & created the dataset for training and testing ML models.
  - Created an API to grab production data and any associated annotations/labels to train or test AI/ML models.
- Developed an algorithm to expand the bounding boxes in our snippets of specific portions of our financial documents that were previously cutoff. This allowed an increase in straight through processing (STP) from 0 to 25% for a new set of documents.
- Implemented a workflow after recognizing a pattern in customer payments that increased STP on key fields by 2 to 10 percent.
  - Created collections in MongoDB representing key customer concepts that are populated through custom API calls and used as lookup tables for verifying attributes based on previous customer transactions.

### Inspire11

June 2022 - April 2024

Data Engineer

Charlotte, NC and Minneapolis, MN

- Transformed a Pharmacy Benefit Manager's (PBM) data landscape from relying solely on 3rd parties for data, to a performant Azure Cloud in-house solution allowing for previously unachievable same day insights and customer claim resolutions.
  - Designed / Implemented a generic JSON parser with data from 60+ Google Cloud Storage Buckets (GCS) every 10 minutes with dynamically changing schemas and built a Databricks Medallion Architecture. This led to increased visibility of data and allowed the PBM to provide details on claims / inquiries from their patients in hours to days faster than before.
  - Created APIs in an Azure Functions to send digital insurance cards to patients through Azure Blob Storage. These APIs allowed the PBM to track insurance cards that were sent and reduced the time patients had to wait for their cards.
- Designed, Migrated, and Implemented a Cloud Modern Data Platform for a Logistics Company:
  - Migrated 4 on-prem databases with over 300 tables to Azure SQL DBs while creating a new Data Architecture to enhance BI Capabilities through increased performance with a near real time refresh rate. Utilized / Configured the following tools: Azure Data Factory, BLOB, Logic Apps, VMs, Azure SQL DBs, Private Endpoints, Virtual Networks, and ExpressRoute.
- Developed recruiting coding challenges in Python and SQL. Created a new data engineering onboarding process at Inspire11 with a bootcamp that performs data processing in Databricks notebooks which then flows into Snowflake where dbt is utilized.

### West Monroe Partners

March 2021 - June 2022

Experienced Data Engineer

March 2022 - June 2022 | Minneapolis, MN

- Automated a prospecting campaign and migrated data for a Tax Software Company:
  - Setup an EC2 instance (AWS) to run Apache Airflow to orchestrate a customer-prospecting campaign which included optimizing python scripts, creating DAGs with sensors to Snowflake tables and Airflow's metadata, SNS email alerting, and setup of the Airflow instance using a Postgres database. Migrated a database using Flyway which involved enhancements on existing views / creating new views in Snowflake. Created Continuous Integration and Continuous Deployment pipelines using Gitlab Runners to automate deployment.

Data Engineer

March 2021 - March 2022 | Minneapolis, MN

- Designed, implemented, and tested a Modern Data Platform for a mobile gaming company:
  - Created 271 metrics leveraging Databricks (PySpark / Spark SQL) and Snowflake using mobile users' data with a streaming architecture from phone to Data Lake (Databricks Lakehouse - Delta Lake) to Data Warehouse (Snowflake) to reports in Tableau. This event driven pipeline had an SLA of under 30 minutes from the time of a user click to reporting insights.
  - Wrote UAT test cases in SQL and facilitated the UAT process leading to higher trust in the data and a quicker release cycle.

## EDUCATION

### Bachelor of Science in Engineering - Computer Science

Sep 2016 - May 2020

University of Minnesota - Twin Cities

Minneapolis, MN

- Teacher Assistant for Algorithms and Data Structures - UMNTC Computer Science Department for 2 years
- Carlson MIS Case Team - Competed at local and International Case Competitions with an emphasis on Technology Solutions.

## NOTABLE PROJECTS

see - [evanditter.com](https://www.evanditter.com) | **NBA Data Engineering Project in Databricks** | **Twitter Prediction** | **Spark Covid Analysis**

2020-2025

**NoSQL ScyllaDB Stock Predictions Based on NYC 311 requests / Weather Data** | **NFL QB Success Prediction**

2021-2025

## CERTIFICATIONS

Databricks Associate Data Engineer

November 2023