

Benjamin J. Hammack

Chicago, IL

Cell: (630)746-1685 | Email: bjhammack@protonmail.com

Website: <https://benjaminhammack.com>

PROFESSIONAL OVERVIEW

- Extensive work leading, coordinating, and executing complex machine learning and data analytics projects
- Proficient in the design, development, and testing of predictive models (transformers, neural networks, logistic/linear regression models, support vector machines, etc.) using multiple languages and data systems
- Strong experience in multiple high-level programming languages, including, but not limited to: Python, R, and C++
- Strong background in the cleaning, transformation, and analysis of messy, un-coded, and raw data
- Extensive experience at all levels of the ETL process
- Strong understanding and experience working with RDMS including MS SQL Server, MySQL, PostgreSQL
- Experience working on and with cloud technologies, including: AWS (S3, EC2, etc.), Hadoop (Hive, HDFS, etc.), Google Cloud Services (BigQuery, Google Buckets, etc.), and Microsoft Azure
- Experience developing and maintaining Spark drivers and executors for scalable predictive modeling pipelines
- Highly proficient at picking up new skills, tools, and systems

EDUCATION

Lewis University, Romeoville, IL
Master of Science, Data Science

July 2020
GPA: 3.70

Ball State University, Muncie, IN
Bachelor of Science, Computer Science

May 2017

SKILLS

- **Languages:** Python, R, C#, C++, JavaScript, SQL, Bash
- **Databases:** SQL Server, MySQL, MongoDB, IBM DB2, PostgreSQL
- **O/S:** Windows, Linux, OSX, VMWare, Virtual Box
- **Cloud Computing:** Amazon Web Services, RDS, Hadoop, Apache Spark, Apache Hive, Google Cloud Platform, Microsoft Azure
- **Data Tools/Libraries:** Keras, TensorFlow, PyTorch, Hugging Face, Optuna, MatLab, NumPy, Pandas, Scikit-learn, SciPy, Matplotlib, ggplot2, Tableau, PowerBI, PySpark

SKILL CALLOUTS

- Transformer architecture
- Feature engineering
- Model deployment
- Data storytelling
- ETL pipeline development
- Stakeholder and client communication

EXPERIENCE

Nielsen

Chicago, IL

Lead Data Scientist

August 2021 – Present

As part of a team of talented data scientists and mathematicians, I built, evaluated, and maintained machine learning models. Developed advanced data pipelines that connected our models to our robust, cloud, data infrastructure. Lead projects focused on executing the above. Presented team research and products to leadership and stakeholders.

- Utilized Python, Spark, AWS, and Apache Airflow to develop a full-stack machine learning model to ingest data from S3, perform cleaning and analysis, model the data, output the results throughout our cloud infrastructure, and generate automated reports for quick analysis of results
- Performed statistical analysis on a variety of models, presenting the results to the head of data science, with suggestions for model improvements

- Established best practices within the data science team for workflow, maintaining codebases, and development best practices
- Provided mentorship to junior data scientists, providing both project support and guidance on data science problems in general
- Lead inter-team projects aimed to develop cutting edge products for Nielsen's clients
- Worked with external auditing companies to communicate our modeling processes, provide requested datasets, and useful analytics to complete our annual accreditations

Volkswagen Credit

August 2020 – August 2021

Libertyville, IL

Data Strategy Scientist

Working on a team of data scientists, analysts, and data governance experts, I analyzed, cleaned and distributed data. I established best practices for data usage and data consumption. I worked with business users to create robust, automated data pipelines. I provided analysis, support, data engineering, and validation for machine learning modeling efforts. I helped transition Volkswagen from an on-premises data company to an agile, cloud-based company.

- Converted large amounts of deprecated SAS programs to up-to-date and automated Python scripts for continued analysis and reporting by the leadership team and business units
- Utilized Python and Spark to convert type 2 database tables into transactional tables to be consumed by the leadership team and machine learning models
- Profiled critical database tables for business processes. Identified danger areas in the tables (eg. errors, out-of-date data, missing data, etc.) and designed approaches for resolving these issues
- Developed new database tables, based on business users' needs, to be used for analysis, reporting, and machine learning
- Reviewed old data processes in Business Objects, Excel, Tableau, and SAS, documenting their purposes and methods, designing new, improved methods to achieve better results and developing these new methods to be sent to production
- Assisted in the transition of Volkswagen to cloud-based company in AWS. Developed best practices for cloud processes, transitioned data into Redshift, assisted in the preparation of Sage Maker for machine learning, maintained EC2 instances for multiple data pipelines, and validated data to ensure accuracy
- Helped in the development of data governance strategies, mapping out areas of need, creating a "source of truth" for developing data, and working with business users to ensure all data processes was compliant

FarmLogix

October 2018 – July 2020

Chicago, IL

Data Services Manager

Working in a fast-paced environment as Data Visualization Lead and then Data Services Manager, created data visualizations for internal and client-side use, developed and maintained ETL processes, and led data projects for clients.

- Used Python (Pandas, SciPy, etc.) to clean and load large datasets of food transaction invoices from public school districts, major food distributors, and local farms
- Developed visualizations using PowerBI and Tableau and presented them internally and to clients
- Maintained and updated legacy pipelines using AWS, GCP, MySQL, and Apache Airflow
- Developed and maintained new ETL pipelines using AWS, GCP, and MySQL; creating a more streamlined process and reducing the cost to maintain FarmLogix's pipelines

- Led a project coordinating with seven of the largest school districts in the United States to receive, clean, combine, upload, and analyze their transaction and nutrient data to the cloud for storage and analysis. This project was a key factor in FarmLogix winning a new contract with the USFA
- Developed the ETL process for the reception and uploading of transaction data for one of the largest food distribution companies in the United States. This included: accessing new transaction JSON files on an FTP server, developing a process for cleaning and standardizing this data using Python, developing a script to separate new data from old data, and writing a Python script to upload the new transactions to a MySQL database

Perficient, Inc.

May 2017 – May 2018

Chicago, IL

Cloud Developer

CLIENT: Blue Shield of California

Joined a Perficient team in San Francisco to develop and improve systems that facilitate Blue Shield of California members quickly and easily matching with doctors. This was accomplished mainly through the use of Java, SQL/ESQL.

- Developed SQL scripts to transfer, copy, delete, and create data
- Used SQL/ESQL to develop extract files to gather both user and doctor information and ensure the two get matched correctly
- Wrote SQL scripts to test the veracity of fellow Developers' code and to verify the integrity of the data
- Developed UNIX scripts to backup database tables to ensure data integrity throughout development

CLIENT: Network Real Estate

Joined a Perficient team remotely and developed data management tools using Java, Spring Boot, and HTML for Network Real Estate to better serve them and their customers' needs.

- Developed backend Java code that incorporated Spring Boot to properly handle and interpret data
- Developed frontend HTML pages for user interaction that utilized available Spring Boot tools to read user input and display necessary backend data to them
- Used the MVC framework to develop clean and seamless database to frontend code to make future alterations to NRE's data easy and convenient

National Express

May 2016 – August 2016

Warrenville, IL

IT Intern

During a summer internship program, worked on various projects developing SharePoint intranet sites with C#, ASP.NET, JavaScript, Microsoft SQL Server, and HTML.

- Developed an interactive, intranet, news feed with SharePoint Lists and JavaScript
- Created and managed an SQL Server database to store and sort submitted internal audits
- Developed software to create, edit, search, comment on, and delete internal audits using C#/ASP.NET and SharePoint
- Used Microsoft PowerShell to deploy and manage intranet SharePoint sites