David Trakhtenberg

847-331-8165 | davidtrakh@gmail.com | 567 6th St, Brooklyn, NY 11215

EXPERIENCE

Inspire11   **Chicago, IL/Remote**
Senior Consultant, Data Engineering   July 2022 - Present
Consultant, Data Engineering   August 2020 - July 2022

- Modeled ten tables to create star schema which ingested 20M rows of Qualtrics survey data
- Translated five Alteryx workflows into 27 SQL (Snowflake) and two Python scripts
- Collaborated on development of a Spark architecture on top of Kubernetes which required parsing through active GitHub pull requests to develop novel solutions
- Consulted on the following clients: science research, government services nonprofit, trading firm

Clarity Insights   **Chicago, IL**
Senior Associate Consultant   December 2018 - March 2020
Associate Consultant   June 2017 - December 2018

- Developed PySpark code that met source to target requirements and created ETL pipelines, unit tested ETL code, and strengthened platform's common functions in code base
- Analyzed various industry's datasets by cleaning, transforming, and modeling (i.e. linear models)
- Utilized the following technologies: Hadoop ecosystem to manage file system, Microsoft Azure Cloud and TFS to build and release code, and SQL to validate data
- Documented best practices relating to useful Hadoop/Git commands, Spark error handling, a creative Git branching strategy, effective Python functions, and general knowledge transfer
- Conducted several stakeholder interviews to identify data consumption needs of the client and proposed appropriate solutions to sign new statements of work
- Consulted on the following clients: insurance, healthcare insurance, financial services

Zillow/NYU Capstone   **New York, New York**
Graduate Research Scientist   September 2021 - December 2021

- Implemented a novel algorithm (LayoutLMv2) to perform error analysis in the document understanding space
- Trained three modal baselines—BERT for text, ResNeXt-FPN for visual, and LayoutLMv2 for layout with the downstream task of document classification
- Learned various practical deep learning skills—working with Hugging Face, tuning deep learning models (using WandB), and communicating/presenting on complex data science subject matter

---

EDUCATION

New York University   **New York, New York**
Data Science, M.S.   September 2020 - December 2022
Center for Data Science   GPA: 3.67

University of Illinois Urbana-Champaign   **Champaign, Illinois**
Aerospace Engineering, B.S.   August 2012 - December 2016
James Scholar Honors Program   GPA: 3.65

SKILLS

Technologies: Git, ETL, Hadoop, Apache Spark, AWS, Microsoft Azure, UNIX, Linux, Tableau, Talend, Alteryx, several databases (Oracle, MongoDB, Snowflake etc.), Jira, Docker

Languages: Python (3 Years exp.), SQL (4 Years exp.), R (1 Year exp.)

NYU coursework: Machine Learning, Probabilistic Time Series Analysis, Text As Data, Natural Language Processing, Linear Algebra, Big Data, Responsible Data Science, Deep Learning, Fundamental Algorithms