Daniel Lee
Software Engineer

www.danleedata.com/portfolio
www.github.com/Dan-S-Lee
daniellee94@uchicago.edu
949-310-1784

EXPERIENCE

Software Engineer | Enterprise ML | Capital One

Feb 2022 - Present, Chicago, IL

- Develop applications for sensitive data detection efforts using NLP and machine learning techniques to mitigate potential security risks and protect customer information.
- Designed and developed a scalable ETL pipeline in AWS that moves 50 million data points to the enterprise data lake daily
- Built and maintained data processing pipeline for card embossing and anti-money laundering analytics
- Implemented sensitive data masking API calls and used multi-threading to decrease execution time by 90%

Associate | Consumer ABS | S&P Global

Sep 2017 - Feb 2022, New York, NY

- Developed and updated modules to model and project structured finance transactions
- Analyzed millions of loans to identify risky account segments
- Developed, deployed, and maintained automation applications that save hundreds of hours a month in manual input
- Performed statistical analysis to identify impacts of defaults, prepayments, and market assumptions

RELEVANT PROJECTS

Chat Disentanglement and Harassment

Sep 2022 - Present

- Developed a machine learning model that extracts conversations and harassment from short message chat logs
- Successfully utilized cutting-edge NLP techniques and Python to create a unique framework to disentangle chat data
- Employed TensorFlow to develop and validate the model and achieved a 96% accuracy score

COVID-19 NYC Dashboard

Mar 2020 - Present

- Designed and developed full-stack interactive web dashboard that presents zip code breakdown of NYC COVID-19 cases using NYC Department of Health and US Census
- Conducted hypothesis testing and ran regressions to identify key demographic trends that affected infection rates
- Created map and graph visualizations using Python, Dash, HTML, and Javascript

LANGUAGES

Python

Java

SQL

Javascript

TOOLS

PySpark

AWS

Docker

React

Scikit Learn

TensorFlow

SKILLS

Data visualization

Data modeling

Machine learning

Model validation

EDUCATION

University of Chicago
M.S. Applied Data Science
2021-2023

Columbia University
Non-Degree Computer Science
2021-2021

New York University
B.A. Economics and Mathematics
2013-2017

COURSEWORK

Graduate

Statistical Analysis

Big Data Platforms

Natural Language Processing

Data Mining

Neural Networks