

ANDRE L. CHEN

443-939-0597 ♦ andrechen@nyu.edu ♦ www.github.com/andrec279

EDUCATION

New York University
M.S., Data Science

Anticipated Graduation: May 2023
GPA: 4.0

University of Maryland, Honors College
B.S., Chemical Engineering

Graduated: May 2019
GPA: 3.84

Relevant Coursework: Machine Learning, Big Data, Probability and Statistics, Mathematical Statistics, Linear Algebra and Convex Optimization, Calculus III, NLP With Representation Learning, Responsible Data Science

Relevant Certifications: [AWS Developer Associate \(2021\)](#), [Harvard CS50x: Introduction to Computer Science](#)

KEY PROJECTS

Full Project Portfolio: <https://andrec279.github.io/>

- **Zillow Multi-Modal Knowledge Graph:** Collaborative research project with Zillow Applied Science using **graph neural networks (GraphSAGE)** and **multimodal data** to improve Zillow search
- **NLP Research Project:** Validated results in SimCSE paper (Gao, Yao, and Chen 2021) and adapted their sentence encoder for independent research in **neural sentiment classification** on Twitter data
- **Movie Recommender:** Machine learning (latent factor modeling) approach for building a personalized movie **recommender system** trained on ratings / content data on a compute cluster (**Hadoop**)
- **Contextify:** Web application hosted on **AWS** (<http://contextify.us-east-1.elasticbeanstalk.com/>) using logistic regression to categorize Spotify library into playlists based on feature data from **Spotify API**

EXPERIENCE

ASAPP, Inc.
Data Science Intern

New York, New York
June 2022 – August 2022

- Created **SQL** query scripts to extract 1M+ rows of customer service chat data from **S3** data repositories, as well as **Python** scripts to conduct exploratory analysis and feature engineering
- Developed **production machine learning model in Python** for customer service chat platform to monitor interactions and recommend automated workflows to the agent when appropriate, improving agent efficiency
- Prototyped and evaluated more sophisticated models trained on **AWS GPUs** for future development efforts in automated workflows, including **neural network language models** and **XGBoost**

Deloitte Consulting US
Analytics Consultant

Arlington, Virginia
September 2019 – September 2021

- Analyzed usage data and developed dashboards using **Tableau** and **SQL** to help clients understand user adoption of their new online mortgage data delivery platform
- Developed workflow automation tools in **Python** for client software testing, reducing testing time by ~95%
- Developed **Tableau** Dashboard and **Alteryx** pipeline with Deloitte financial data sources to help leadership understand and improve account-level financial performance across \$100M+ in active engagements
- Performed data exploration and POC dashboard development using **SQL** (Google BigQuery) and **Looker** to investigate project opportunities for newly acquired proprietary datasets

Tesla Motors
Data and Process Engineering Intern

Fremont, California
September 2017 – May 2018

- Designed and implemented new data pipeline and reporting system across Model S, X, and 3 manufacturing operations for efficient and accurate reporting of production data and trends

TECHNICAL SKILLS / INTERESTS

Analytical Tools and Frameworks: Pandas, PySpark, Dask, Numpy, Scikit-learn, PyTorch, Tensorboard, Matplotlib / Seaborn, Plotly-Dash, SHAP, Flask, Tableau, Google CloudSQL, Google BigQuery, AWS Athena

Programming Languages: Python, SQL, MATLAB, JavaScript, HTML, CSS

Interests: Cooking, Rock Climbing, Hiking, Soccer, Jazz Saxophone, Guitar