Aidan Draper

Data Engineer

adraper19@gmail.com | adraper2 | aidan-draper

Education

University of Notre Dame *Notre Dame, IN*

Graduate Coursework in Computer Science, Advised by Dr. Adam Czajka, GPA: 3.7/4.0 *Jun 2019 - June 2021*

Elon University *Elon, NC*

BS in Computer Science, BS in Statistics, GPA: 3.6/4.0 *Aug 2015 - May 2019*

Skills

Languages Python, JavaScript, React, SQL, Bash, C++, C, R

Libraries Airflow, Django, React Native, Tensorflow, PyTorch, CV2, Pandas, NumPy

Tools Git + GH Actions, PostgreSQL, MySQL Workbench, SQLite,JIRA, Notion, Tableau, Power BI, Advanced Excel/Sheets + Macros

Experience

Pathstream *New York, NY*

Data Engineer / Learning Experience Designer *Dec 2021 - Present*

- Building Airflow ETL pipelines in Python for data ingestion from sources, such as TikTok and Brightspace. I have learned Django to assist in building out the data warehouse for our data scientists. This codebase is managed and reviewed by our team of just three engineers.
- Write and review technical content for the data analytics courses offered. This role allows me to actively think about how I am presenting information to students, and in documentation I provide to other engineers.

Teaching Assistant - Data Analytics (Contract) *June 2021 - Dec 2021*

- The TA role included answering technical questions in office hours and in emails about topics in Google Sheets, Tableau, SQL, Statistics, and Python courses. I have improved upon syntax in the Python course and made language changes in the Statistics courses I reviewed.

University of Notre Dame *Notre Dame, IN*

Research Assistant *Jun 2019 - June 2021*

- Investigated the use of multiview scene rendering for distortion correction modelling on 3D fingerprint images for improved minutiae map matching accuracy with 2D fingerprints. Used high throughput computing to parallelize code across 20 GPUs.
- Studied the relationship between neural networks (NN) cortices of the brain by probing how activation maps can steer emotion through visual stimuli from NN-produced (Inceptionv3) images.

Elon Center for Organizational Analytics *Elon, NC*

Data Science Intern *Feb 2018 - May 2019*

- Lead a team in investigating a large IoT dataset in order to build a predictive model to measure machine efficiency.
- Used python, R, and Tableau to manipulate data, create a classification algorithm, and visualize the data story within a dashboard.

University of Notre Dame *Notre Dame, IN*

Summer Research Assistant *Jun - Aug 2018*

- Built a classifier to measure grass species abundance using remote sensing and Landsat satellite image processing in R.
- Reconstructed a Bayesian hierarchical model that used a Poisson point process and parallelized it to create an execution time speedup of 2x.

Projects

3D to 2D Unconstrained Fingerprint Flattening from Video

- Mimicking left-right camera consistency for stereo reconstruction from a single unconstrained camera to produce finger depth clouds for the purpose of flattening 3D contactless fingerprints to improve verification of contactless to contact-based matching.

Subliminal emotion-evoking signals in learning-based visual stimuli

- Searching for signals of specific emotional response from subjects when analyzing images produced at specific layers of networks to better understand network topology in relation to our knowledge of the human brain.

Echo State Network for Keystroke Dynamics User Identification

- Trained an Echo State Neural Network (ESN) to learn 10 users keystroke dynamics data recorded using a custom keylogger GUI.
- Achieved 91.43% accuracy and introduced a significantly faster model that just slightly under performs state-of-the-art metrics (-13% on 7 session of user data) on the Villani dataset.

Optimizing GPU Performance across Operating Systems

- Investigated optimizing energy usage, GPU temperature, and runtime across 3 tasks: DNN training, Scene Rendering and Cryptomining.
- Setup Nvidia profiles and compared settings (persistence mode, wddm/tcc, etc.) across the Windows 10 and Ubuntu operating systems.

aidandraper.com

- Taught myself enough of the React and JS to rebuild my portfolio site with reusable containers.
- Hosting on GitHub Pages and pay just $10 for the site yearly.

Survey on Image Quality Loss from Image Denoising

- Designed an RShiny app survey that asked participants to rate the appearance of poorly-exposed images that were ”denoised” using filters, such as Bilateral and Non-local Means filters, in comparison to properly exposed low-light photos.
- Collected data on 100 college students and performed an ANOVA and ANCOVA to analyze variance and covariance between participants’ filter opinions in comparison to standard image quality scores, like SSIM and PSNR.

Music Lyrics Study

- Pulled song data from the Spotify API and song lyrics from the Genius API and hosted the database in a MySQL relational database online.
- Implemented an LDA to model topics in song lyrics, ran a sentiment analysis using Vader, and visualized the results in Tableau.

Competitions

Manhattan College Business Analytics Competition

- Analyzed 10 million rows of NYC contract spending and allocation data with three teammates, built a logistic regression model to predict contract allocation across minority groups, created a data story, and then, presented a poster on our findings.
- Constructed a Glassdoor web-scraper overnight in order to collect employee review data on companies in NYC and ran a sentiment analysis on employees’ reviews of businesses. Placed **fourth out of 18 teams** and won the **Best Poster Award**.

HanesBrand Data Analytics Competition

- Collaborated with two teammates to implement three data mining algorithms, including a sentiment analysis, association rules, and entity matching on more than 1.8 million rows to give insight into consumer reviews for Champion products. Placed **first out of twenty teams**.