Caleb Tucker Dame

calebdame@gmail.com
www.calebdame.com
(707) 305-8453

Experience

Data Engineer – M Science LLC Jun. 2021 – Present
Murray, UT

- Design ETL frameworks and maintain 15+ pipelines that are run daily to export 3.5 Tb of custom data transforms to Financial Analyst teams using Databricks, Airflow, and Snowflake and worked with teams to perform quality assurance on existing pipeline logic and test the purity of their data.
- Train and parametrize clustering algorithms to regularly and automatically identify and flag potential noise and anomalies, increasing data purity by 10-15% on average. Changes are automatically visualized in Tableau dashboards to explore potential noise and anomalies to help explain changes to analysts and management.

Summer Data Analyst – Goldman Sachs Consumer Investment Management Jun. 2020 – Aug. 2020
Salt Lake City, UT

- Leveraged existing databases across divisions using SQL and Alteryx to design a data pipeline that dynamically generates 30+ regularly updated SEC-compliant reports.
- Designed scripts to find and extract necessary data from existing reports to automatically store in spreadsheet format, saving 60+ hours monthly for each team member.

Machine Learning Researcher – Brigham Young University, Economics Apr. 2020 – Jul. 2020
Provo, UT

- Trained dozens of ML models (XGBOOST, Neural Network, Naïve Bayes, etc.) in Python and managed a large dataset in SQL Server to link census records across time and match genealogical profiles in FamilySearch and Ancestry databases.
- Improved accuracy by 80% in linking census profiles by implementing advanced feature engineering informed by graph theory and increased efficiency by implementing new rules-based approaches for selecting candidate matches when multiple classifications cannot coexist.

Education

Brigham Young University, Mathematics Department Scholarship Honoree April 2021
B.S. Applied and Computational Mathematics, Secondary Major Economics; 3.85 GPA
Dean's List Honoree (2018, 2020)

RELEVANT COURSEWORK

- Deep Learning
- Thompson Sampling
- Monte Carlo Simulation
- Bayesian Analysis and Statistics
- A/B Testing
- Analysis of Variance
- Computational Linear Algebra
- Fourier Analysis
- Econometrics and Causal Analysis
- Convex Optimization
- Machine Learning for Forecasting

Languages and Competencies

Python Packages

- Numpy, Scipy, Matplotlib, Scikit-Learn, Keras, Tensorflow, Pytorch, StatsModels, Pandas, PyMC3, PyQT5

Other Languages, Softwares, and Frameworks

- SQL, Tableau, Pyspark, Airflow, Snowflake, Databricks, AWS, C++, STATA, Git, Agile, and Unix Shell

Projects

Battleship: Neural Network Players v.s. Thompson Sampling Players

- Built various battleship solvers by leveraging Neural Networks, conducting Thompson Sampling on random board configurations, and Deep Reinforcement Learning to determine optimal gameplay strategies.

LSTM Mozart Music Generator

- Trained an LSTM Network to predict upcoming notes in Mozart Sonatas with accuracy of 50% when simultaneously predicting the correct next note's frequency, length, and offset. The Music21 package in Python was used to retrieve notes, chords, lengths, and note offsets from Mozart Sonatas formatted in .mid files.