Aaron **Glover**

Data Scientist, Machine Learning Engineer

281.221.5071 | aglove2189@gmail.com | aglove2189.github.io | aglove2189 | aglove2189 | @aglove2189

I am value driven, motivated to maximize productivity through simplicity. I take a liberal arts approach to data science, refining decision-making across any organizational level. I find the answers in the questions, clarifying the "why" and predicting "what's next".

Skills

Languages & Packages Python, SQL, pandas, numpy, scikit-learn, keras, pytorch, matplotlib, streamlit, shap, mlflow

Platforms Docker, Kubernetes, AWS, GCP, git

Experience

Enterprise Products — *Houston, TX*

Data Science Manager - Commercial — *Feb 2022 - Current*

- Focused on developing predictive models to assist with commodity trading and decision making. Since inception, we have productionized sixteen models that predict four different commodities utilizing thousands of features. Assets under management are in the multi-millions.

Lead Data Scientist — *May 2019 - Feb 2022*

- Tasked with developing predictive models to assist with commodity trading. The inception of the project started with a team of two and has now grown to eight.
- Developed a recurrent neural net for predicting heat exchanger failures. Model was the first deployed solution at Enterprise, implemented in the first 8 weeks of hire.
- Built a Monte Carlo simulation that optimized spare parts inventory, potential savings in the multi-millions.

Sanchez Energy — *Houston, TX*

Data Scientist / Machine Learning Engineer — *Apr 2017 - May 2019*

- Developed a model fitting solution for determining a well's spontaneous (SP) log curve using peak detection methods and Kalman filters. The end result is used for identifying potential oil field plays to target.
- Developed a Markov Chain Monte Carlo solution for simulating a well's decline curve and ultimate recovery. This augmented Engineering's decision making on how much a well will make over its lifetime.
- Optimized a geophysics simulator written in Python which decreased runtime by 6x and lines of code were reduced 10x.
- Contributed to the development of a multi model prediction framework for predicting well production. The solution was a significant improvement on the industry standard decline curve fitting.
- Implemented a real time alert for detecting tubing leaks which resulted in a cost savings in the six figures. An industry standard deterministic model was required by Engineering, the model was optimized by sampling the search space with a Tree-structured Parzen Estimator.

Occidental Petroleum — *Houston, TX*

Data Scientist — *Jan 2012 - Mar 2017*

- Developed a Monte Carlo simulation to determine the optimal number of workover rigs for a given field. Implemented in fields across Texas and California with a savings in the high six figures.
- Built a logistic regression model for predicting the likelihood of IT tickets breaching their SLAs.

Analytics Engineer

- Developed over 350 SSRS reports and Spotfire dashboards over the course of two years.
- Designed and developed multiple SSAS cubes for operational and well servicing data, query times were reduced 1,000X.
- Automated the delivery of partner reports utilizing SSRS and SQL which resulted in an 80% reduction in man hours.

Education

Texas A&M | M.S. in Analytics *Houston, TX*

Major GPA: 3.8 *May 2017*

- Thesis: Predicting the likelihood of ESP well failures utilizing survival analysis and gradient boosting.

University of North Texas | B.S. in Information Systems *Denton, TX*

Major GPA: 4.0 *Dec 2011*

Projects

- appias - A library for gluing together a few of the standard steps when exploring a dataset and building a model.
- Taleb but AI - GPT2 model trained on Nassim Taleb's quotes.
- cookiecutter_ds - Repository template for starting data science projects.
- Awair - Python library for viewing / downloading Awair data.