Aaron **Glover**

Data Scientist

1602 Elgin St Apt 14 Houston, TX 77004

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

Analytically minded by nature. Solving complex problems with the simplest solution possible is the desired outcome. Solid capabilities in modeling, statistical analysis, data munging, and developing data products.

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*

Skills

Programming	Python, SQL, Scala, R
Packages & Frameworks	Pandas, NumPy, scikit-learn, xgboost, Keras, Tensorflow, git
Platforms & Engines	Hadoop, Spark, Linux, Windows, Google Cloud, Azure, AWS
Databases	PostgreSQL, MongoDB, MySQL, Microsoft SQL Server, Oracle, neo4j

Experience

Sanchez Energy — *Houston, TX*

Data Scientist — *Apr 2017 - Current*

- Developed a model fitting solution using peak detection methods and Kalman filters for determining a well's spontaneous (SP) log curve. The end result is used for identifying potential fields to target.
- Developed a Markov Chain Monte Carlo solution for simulating a well's decline curve and ultimate recovery.
- Contributed to the development of a multi model prediction framework for predicting well production.
- Contributed heavily to the development of automated methods for feature engineering.
- Built and productionized a LSTM neural net for predicting a well's future flowback amount.
- Developed operational alerts utilizing realtime data. Alerts include a deterministic model for detecting tubing leaks.
- Built an OCR application that extracted data from PDF documents generated by the Texas government.
- Developed a recurrent neural net for classifying types of vendor emails with 99% accuracy.
- Analyzed, extracted, and quality checked numerous data sources and data stores required for data science projects.
- Developed a multitude of data pipelines and workflows from a variety of sources (relational, non-relational, SCADA systems, etc).

Occidental Petroleum — *Houston, TX*

Data Analytics — *Jan 2012 - Mar 2017*

- Developed and productionized a Monte Carlo simulation to determine the optimal number of workover rigs a field needs.
- Built a logistic regression model for predicting the likelihood of IT tickets breaching their SLAs; achieved 87% accuracy.
- Implemented the first graph database (using neo4j) at Oxy for visualizing relationships between key data points. Visualization of complex data flows simplified the understanding of how data moves throughout the organization's systems.