ANNA NYULUND

19409 Melwas Way, Pflugerville, TX 78660 | 512-927-7184 | anyulund@gmail.com | GitHub: anyulund

EXECUTIVE SUMMARY

Experienced product manager and data scientist with 7 years of experience in building teams and products. Deep experience in machine learning, natural language processing, predictive modeling and data engineering.

WORK/ CONTRACTING EXPERIENCE

Northrop Grumman **Austin, TX**
Principal Data Scientist (Contract) *2020*
- Facing C-Suit (engineering management and centers of excellence) and communicating requirements to the Data Science team
- Guiding the data science team in project development according to C-suite requirements
- Doing yearly projections for project development and doing appropriate hiring
- Transitioning current Data Science team to the next level: from self-serving BI Tools and Predictive Models to Autonomous Human-Machine Interaction

Oleum Scientia **Austin, TX**
Data Science Product Manager *2018 to 2020*
- Customer facing Data Scientist acting as a liaison between engineers, geologists and data scientists
- Met with different clients and translated their vision and strategy into a road-map and prioritized features of highest impact and fastest time-to-market
- Defined metrics and KPIs to measure the success of the features of team builds
- Developed an expert level understanding of each client's data to drive solutions that result in increased user management and business optimization
- Worked closely with stakeholders, data science, engineering, and QA teams to enable efficient, collaborative, and high-quality product development process
- Anticipated potential issues, necessary points of integration, and needs beyond the basic product requirements
- Removed roadblocks for the data science team quickly, addressing all questions and needs for on-time delivery
- Developed Oil and Gas lease and deed parsing algorithms and profile optimization analytics strategies resulting in savings/ profit over 100 million dollars

RigUp **Austin, TX**
Sr. Data Analyst *2018 to 2018*
- Built Deep Averaging Network performing polarity semantic analysis in contractor reviews
- Debiased word embeddings using Scikit learn library cosine simularity functions (vector angle adjustment)
- Implemented debiased word embeddings instead of GloVe
- Semantic analysis combined with ranking algorithm served as a base algorithm for the recommendation engine resulting in profit of at least a million dollars in the first year of operation
- Transitioned the company from local data warehouse into AWS in two months resulting in smoother and faster running algorithms
- Mapped the whole database using ERDs
- Performed resume parsing using custom-built NER engine sitting on top of Python Spacy package
- Interviewed and selected junior analysts and machine learning engineers for the Data Science team

DataJudo **Austin, TX**
Sr. Data Scientist *2017 to 2019*
- Royal Dutch Shell Oil Company:
  Successfully completed optimization of a Petroleum Producing Assets Portfolio: Advanced Computer Model Development (Python) resulting on profit over 300 million dollars over 5 years.
  Simulated distribution of reserves and a set of expected production profiles using Monte-Carlo Analysis.
- Chevron Corporation:
  Developed petroleum price forecasts based on Sequential Gaussian Simulation.
  Calculated after-tax cash flows, estimated performance indicators for each realization, thus yielding Distribution of return for each project.
  Estimated covariance between return distributions of individual projects and compiled them into Portfolios.

Merk (Contract) **Austin, TX**
Sr. Data Architect *2018 to 2018*
- Developed a model that predicted risk of nonadherence for Januvia patient cohort resulting in savings of billions of dollars

Sensoleak (Contract) **Austin, TX**
Sr. Data Scientist *2017*
- Advised the CEO on setting up operations in the USA and building a local Data Science team, helped to select potential hires, conducted the interviews and mentored junior Data Scientists and Engineers
- Managed the workload, scheduled the assignments and deadlines based on the expertise of the team members and conducted the interdisciplinary business meetings between the clients and the team members

- Developed a time-series pipeline detection model

USAA **San Antonio, TX**
Data Architect *2017*
- Developed a behavior based ML model to identify fraud and FICO accuracy testing algorithm using Kolmogorov-Smirnov statistics, Area Under the Curve, Receiver Operating Charachetristic and Population Stability Index.
- Built Convolutional Neural Network prototype, which predicted the number of transactions based on the customer's qualifications.
- Improved Exploratory Data Analysis by implementing the Apriori Algorithm to find Association Rules in customer behavior.

Parsley Energy **Austin, TX**
Spotfire Engineer *2016 to 2017*
- TDeveloped Data Analytics Lifecycle and Project Request Process for Spotfire
- Provided support for the development of production and reservor databases for Investments
- Used time series modeling to develop decline curve analysis and to set up a system of alerts to identify problematic wells
- Developed a pump failure prediction model which resulted in savings over 100 million dollars
- Developed Arps Curves code for production forecasting
- Developed Tableau and Spotfire dashboards for Lease Operating Expenses. The automation saved over 10,000 hours.
- Developed advanced code and automated dashboard for the Business Intelligence department

Devon Energy **Oklahoma City, OK**
Data Scientist *2016 to 2016*
- Guided Integrated Reservoir Characterisation team in decision making process for acquisitions and divestment (multi-million dollar decisions)
- Developed ML learning algorithm which determined intermittent sedimentation. Algorithm resulted in savings of over 1000s of hours.
- Developed prediction model for corrosion in completions casing which prevented well abandonment. Each well cost ranges from 5 to 15 million dollars.The company had over 300 wells.
- Developed dashboards for Drilling, Production, Geology, Environmental Health and Protection and communicated all insights to C-Suite on a monthly basis.

Schlumberger Information Systems **Anchorage, AK**
Data Scientist *2014 to 2015*
- Developed a complex reservoir simulation analytics solution using advanced mathematical models and is being used today by different companies including Pattern Computer.
- The solution development helped facilitate a sale of the Schlumberger oilwell services in the amount of 200 million dollars
- The development resulted in a large academic accomplishment

Gulf Interstate Engineering **Houston, TX**
Mechanical Engineer *2012*
- Developed a complex analytical model to identify pressure flow in a pipeline for the client (Pacific Oil and Gas)
- Mapped the pipeline using GIS technology
- The solution was a part of new government regulations and resulted in savings of hundreds of millions of dollars and hundreds of human lives

Halliburton **Houston, TX**
Directional Driller *2011*
- Was in charge of over 40 staff on a drilling rig. Guided the team to conduct safe drilling operations.
- Was the first female Directional Driller at Halliburton.
- Developed projections to the well and interacted daily with clients (operating companies)

EDUCATION

- MS, Engineering, University of Alaska Fairbanks, 2015
- BS, Engineering, University of Texas at Austin, 2010

CERTIFICATIONS/ TRAINING/ CLASSES

- SAS: Data Science Learning Path, 2016
- Natural Language Processing (post-graduate work), University of Texas at Austin, 2018
- Data Structures, University of Washington, 2019
- Post-Graduate Program in Machine Learning and Artificial Intelligence, University of Texas at Austin, 2020
- Computer Vision I,II,II Open CV, 2019 -2020

TECHNICAL SKILLS

- **Natural Language Processing:** Semantic Analysis (DAN, LSTMs, encoder-decoder models, BERT(prepackaged), parsing of unstructured and semi-structured data (NER (custom CRF and Spacy), Regex, Topic Modeling (Gensim))
- **Computer Vision:** Open CV, Teseract (OCR)
- **Data Processing:** NLTK, Spacy, OpenCV, SKLearn, Pandas, Numpy, SQL
- **Data Visualization:** Tableau, Spotfire, SAS VA, Seaborn, Matplotlib, Bokeh, Plotly

- **Data Engineering:** SQL, PySpark, PostgreSQL, Docker, ERD construction
- **Machine Learning:** Neural Nets (CNN, RNN, LSTM, DAN), Linear Regression, Logistic Regression, Auto-Regressive Moving Average, Isotonic Regression, Random Forest, Decision Trees, Principal Component Analysis, Design of Experiments, K-means clustering, Monte-Carlo Analysis
- **Big Data:** Hive, Hadoop HDFS, Kafka (KSQL)
- **Python libraries used:** Tensorflow, PyTorch, SKLearn, Spacy, NLTK, SparkML, SciPy, Beautiful Soup, Scrapy, XGBoost, SweetViz
- **Data Science Project Management:** SEMMA, KDD, CRISP-DM, Agile