NealDawson-Elli

Chemical Engineer - Data Scientist - Machine Learning Engineer

(607) 857-9791 nealde@uw.edu nealde nealde.github.io scholar /neal-dawson-elli

Skills

Machine Learning Engineering
- Spark/Spark Streaming, Kafka, Redis, SQL, Python, Cython, Kubernetes, FastAPI
- ETL Pipelines, ML Training Automation, High-Performance APIs, Data Visualization / Dashboards

Data Science
- Python Software Design, Analysis Stack (NumPy, SciPy, Pandas), Keras, Tensorflow
- Machine Learning (Convolutional, LSTM Neural Networks, GANs, Bayesian Networks, GBMs)

Projects

2019 **What Can Electrochemistry Learn from Chess?**
- Li-Ion battery cycle life can be extended 2x by model-predictive control, but calibration is very difficult
- Architected neural-network-based multi-objective optimization framework using 200,000 time-series simulations to efficiently calibrate expensive nonlinear models using Keras and Tensorflow
- Reduced in-house Li-Ion model calibration time by 60% and improved fit by 30x over Genetic Algorithm

2018 **Ampere**
- Li-Ion models are useful for control and design, but are slow, complex, and have costly licenses
- Developed and distributed open-source Python package using Cython to wrap high-performance in-house models solved in C
- Implemented Sci-Kit Learn-Like API for model manipulation, outperformed industry standard by 100x

2018 **WYNS**
- Designed a Twitter sentiment analysis batch process for predicting climate change sentiment
- Trained Bi-Direction LSTM using term frequency and bag of words analyses on 6000 labeled tweets
- Deployed interactive map of tweets on Google Cloud Platform using Plotly Dash for visualization

Education

2019 **Ph.D. in Chemical Engineering** Univ. of Washington, Seattle, WA
- Option in Advanced Data Science

2015 **B.S. in Chemical Engineering** Rochester Inst. of Tech., Rochester, NY
- Minor in Musical Performance

Experience

Current **Data Engineer** PayScale, Seattle, WA
- Crafted high-performance ETL systems for Snowflake Snowpipe ingestion
- Developed and deployed multiple high-throughput MicroServices on AKS using Python and C#
- Productionized Data Science analytics and metrics, improving performance and reducing cost

2019 **Insight Data Engineering Fellow** Insight Data Science, Seattle, WA
- Deployed streaming ML pipeline to automatically detect duplicates using TF-IDF and cosine similarity
- Cut computation time by 85% and costs by 90% using custom distributed Cython functions
- Mirrored pipeline using both Spark Streaming and Kafka to compare performance metrics