Peter Jourgensen

1520 N Hayworth Ave. #17, Los Angeles, CA. 90046
pjourgensen@gmail.com | (925) 818-0552

Education

University of California, Los Angeles, Los Angeles, CA. *expected 12/19*
Masters of Science in *Computer Science*

- GPA: **3.7/4.0**
- Relevant Coursework: Machine Learning, Artificial Intelligence, Algorithms, Applied Probability

Northwestern University, Evanston, IL. *2015*
Bachelors of Science in *Applied Mathematics* | Minor in *Economics* | Certificate in *Financial Economics*

- Major GPA: **3.8/4.0**, *Cum Laude*
- Relevant Coursework: Dynamic Systems, Econometrics, Numerical Analysis, Statistics, Optimization

Skills

- **Programming:** Python, R, MATLAB, SQL
- **Machine Learning:** SKLearn, Feature Engineering, Normalization, Cross-Validation, Tensorflow, NLP
- **Visualization:** MatPlotLib, Seaborn, Choropleth Maps

Quantitative Experience

Valkyrie Trading Llc, Chicago, IL.
Trader *2017-2018*
Junior Trader *2015-2017*

- Led a team of 3 in the collaborative execution and risk management of fixed income future options; achieved a net profit of >$1,000,000 during my tenure
- Identified a growth opportunity in calendar-spread trading, developed a risk pricing model, syndicated with partners, and led implementation with developers
- Analyzed time-series data of 15 crude oil futures expirations spanning 2 years to develop a model for assessing covariance of futures as a function of their time to expiration
- Trained 3 Junior Traders; empowered each to independently manage their own portfolios

Data Science Project Portfolio

Acute Leukemia Classification *2019*
Computational Genomics - "Identifying Significant Genes via Ensembled Neural Networks"

- Constructed a Densely Connected Neural Network framework via Tensorflow to be used as a base model
- Applied bootstrap aggregation, whereby 4,680 individual neural networks were trained on a subset of the training samples and a subset of genes
- Noticed most models exclusively predicted the majority class, so I subtracted the corresponding accuracy from the model's accuracy and passed the result through a ReLu function to weight each model
- Scored the individual genes by iterating through each model and taking the average weight of the models that the gene was present in; assumed higher scoring genes were stronger predictors of cancer class
- Analyzed cross-validation of KNearestNeighbors algorithm and performed a grid search over the number of highest scoring genes and the number of neighbors to find an optimal model; achieved 97.2% accuracy

Loan Data - Lending Club *2019*
Kaggle Exploratory Dataset - "Predictive modeling for loan default"

- Assessed 30 features of 900,000 samples for missing values, outliers, and categorical encodings
- Ran a random forest baseline model before undersampling the majority class to test model performance at varying ratios of class balance; best model achieved 84% accuracy with 39% correct default predictions
- Applied Synthetic Minority Oversampling Technique (SMOTE) and tested models at the same varying ratios of class balance; best model achieved 88% accuracy with 14% correct default predictions
- Reapplied undersampling before training a densely connected neural network constructed with Tensorflow; achieved 75% accuracy with 45% correct default predictions
- Concluded that final model choice would be dependent on cost of lending to someone who will default versus opportunity cost of not lending to someone who will not default