|  |
| --- |
| **Jonathan D. Groth, PhD** |

jon.groth.phd@[gmail](http://jon.groth.phd@gmail.com).com| (337) 581-2789 | | Princeton, NJ

https://github.com/<https://github.com/jgroth1> [| https://www.linkedin.com/in/jonathan-groth-phd/](https://www.linkedin.com/in/jonathan-groth-phd/)

**Summary**

* Data Scientist with 12 years’ experience creating solutions with data analysis. With expertise in experimental design, predictive analytics, time series forecasting, and natural language processing
* Excellent communication skills with the ability to present complex ideas to experts and novices. Developed during my 2 years teaching, 5 years teacher’s assistant, 10 conference presentations, 4 peer reviewed publications.
* Strong understanding statistical, machine learning, and AI techniques; developing analysis for A/B testing, classification, scoring, segmentation, time series analysis, forecasting, document clustering, document classification, key word extraction, topic modeling, and document similarity.

**Recent Projects**

Document job search (NLP):

The aim of this project was to redesign the way the user searches for a job. Instead of title search I matched jobs based on resume search using doc2vec and document similarity.

* Scraped Glassdoor.com for job listings using python’s splinter
* Used word2vec and doc2vec representations clustering and classification to determine the most similar job postings to input resume, CV, or LinkedIn profile.
* Plan to add ensemble method for document similarity and sentiment analysis using deep learning for skills ranking.

Time series forecasting:

The aim of this project was to forecast the price of commodity coffee over a 1-year period.

* Modeled the price using L1 regularized regression and ARIMA with weather and market metrics as predictors.
* Found that the price and predictors were highly seasonal so modeled forecast of predictors using weighted seasonality.
* Found with ideal data found an average error of $7.62 for a price on the order of $200.

**Experience**

Data Science Bootcamp – Rutgers University **05/18 – 10/25**

* Data acquisition and full stack development.
* Data visualization and storytelling using python and java script.
* Data collection and parsing using SQL, APIs, and web scraping.
* Developed machine learning systems for natural language processing, forecasting, and market segmentation.

Data Analyst / Research Scientist – Melior Discovery  **08/17 – 10/17**

* Developed new experimental models to be used in new marketable pharmaceutical testing models.
* Developed algorithms and software for the analysis of data. Resulting in new analysis procedures that enable new experimental paradigms to be brought to market.
* Performed feature selection and feature engineering extracting more robust predictors for analysis.

Instructor - New Jersey Institute of Technology  **09/15 - 05/16**

* Communicated complex concepts to a novice audience in a 25 to 32 person lecture and lab course on introduction to engineering concepts and analysis for biomedical engineering and electrophysiology.
* Problem solved student’s lab activities in programing and electronic design then guided them in developing solutions.
* Managed 25 to 32 students in an instrumentation lab course directing students in how to design, prototype, and validate the instrumentation and troubleshoot problems.

Graduate Researcher – New Jersey Institute of Technology **09/08 - 05/15**

* Developed and programmed Machine learning systems for brain computer interfaces using both primary motor cortex and cerebellar signals to forecast arm movements.
* Worked with Team members to improve experimental design and increase integration between different experiments being done on the project.
* Developed data analysis techniques for the analysis of electrophysiological time series signals from the cerebellum and primary motor cortex resulting in two publications.

**Technical Experience**

**Programing Languages**: Python, Matlab, some JavaScript, and R **|** **SQL Platforms**: PostgresSQL, MySQL, and SQLite **|** **NoSQL Platforms**: MongoDB **|** **Markup Languages**: HTML, CSS, and Markdown **|** **Shell Languages**: Bash **|** **Data Wrangling and Data Analysis**: Numpy, Pandas, Matplotlib, Plotly, D3.js, Scipy, Scikit-Learn, NLTK, Gensim, SpaCy, Tensorflow, and Keras

**Education**

PhD biomedical engineering - New Jersey Institute of Technology 05/15

MS Biomedical engineering - New Jersey Institute of Technology 05/08

BS Biomedical engineering - Louisiana Tech University 02/05