Bill Watson

Data Scientist · Software Engineer

(+1) 516-543-1131 | nextbillyonair@gmail.com | nextbillyonair.github.io | nextbillyonair | nextbillyonair

Education

The Johns Hopkins University — *Baltimore, MD*

Master of Science in Engineering, Computer Science — *Aug. 2018 - May 2019*

Bachelor of Science, Computer Science — *Aug. 2014 - May 2018*

- Minor in Applied Mathematics and Statistics

Work Experience

S&P Global — *New York City, NY*

Data Scientist — *Jun. 2019 - Present*

- **Supply Chain Network**
  - Developed the ability to identify higher order risks for supply chain disruption, increasing transparency in a company's network for analysts.
  - Aggregated several data sources to provide empirical measures and weights on supplier importance in a real-time dashboard.
- **LIBOR Exposure and Risk Mitigation**
  - Full stack web application to review language in indentured documents for LIBOR exposure during phaseout transition.
  - Creation of LIBOR fallback classification and language extraction algorithms to provide interpretability and insight to stakeholders.
  - Mitigates oversight risk from manual efforts, with total LIBOR exposure estimated in the hundreds of trillions of dollars.
- **Table Extraction From Image Documents**
  - Proof of concept pipeline to identify, extract, and organize tabular content in unstructured financial documents.
  - Leveraged image segmentation, optical character recognition, and sequence to sequence modeling to improve extraction results over internal methods.
- **Criteria Validation Search & Citation Recommendation**
  - Created a full stack web application to allow multi-phrase search on a corpus of ratings criteria.
  - Improved the efficiency of validating proposed framework changes by providing insights to relevant and lateral connections between current criteria.
  - Utilized a transformer-based graph model to recommend citations for newly drafted criteria, reducing inconsistencies and liability from missing citations.
- **Data Science Reading Group**
  - Introduction to Differentiable Probabilistic Models (July 25th, 2019)
  - Introduction to Random Numbers, Sampling, and MCMC Methods (August 22nd, 2019)
  - Seq2Seq in Action: Column Segmentation (January 23rd & 30th, 2020)

Publications

Financial Table Extraction in Image Documents — *ICAIF 2020*

William Watson *and Bo Liu* — *Published*

Directed Criteria Citation Recommendation & Ranking Through Link Prediction — *ICAIF 2020: Extended Abstract*

William Watson *and Lawrence Yong* — *Published*

Modeling Color Terminology Across Thousands of Languages — *EMNLP-IJCNLP 2019*

Arya D. McCarthy, Winston Wu, Aaron Mueller, ***William Watson****, and David Yarowsky* — *Published*

Personal Projects

Differentiable Probabilistic Models

- Modular library to leverage modern techniques such as differentiable programing to statistical modeling.
- Programmed a simple interface to interact, learn, and visualize various distributions, transforms, and loss functions.
- Implemented advanced concepts such as Variational Inference methods (ELBO), MCMC sampling techniques, and adversarial losses.
- Reinterpreted classical algorithms such as regression, classification, clustering, and factorization as probabilistic models.

Skills & Abilities

Programming — Python, LaTeX, JavaScript

Frameworks, Libraries, and Tools — PyTorch, React, Flask, OpenCV, Elasticsearch