

Alec B. Hon

www.linkedin.com/in/alecbhon · github.com/abhon
(714)929-9772 · alecbhon@gmail.com · New York, NY

EDUCATION

- New York University, Masters of Science, Data Science** Sep 2019 - May 2021
- **Coursework:** Machine Learning, Deep Learning, Search and Discovery, Tools and Techniques for Machine Learning, Big Data, Fundamental Algorithms, Mathematical Tools for Data Science, Probability and Statistics, Linear Algebra
- Columbia University, Bachelor of Arts, Chemistry,** Sep 2014 - May 2018
- Dean's List

EXPERIENCE

- New York University Center for Data Science** Current
Graduate Teaching Assistant and Grader, DSGA-1004 Big Data New York City, NY
- Taught graduate students the fundamentals of Version Control(git) and Distributed Computation infrastructures.
 - Primary focus on teaching how to utilize SQL, Hadoop File Systems, MapReduce, Spark, and Dask, as well as how to request and utilize remote resources such as NYU's High Performance Computing Clusters.
- New York University Center for Data Science Summer Incubator** Jun 2020 - Aug 2020
Research Volunteer New York City, NY
- Worked with Prof. Leo Miolane and other student volunteers to implement AlphaGoZero in Pytorch for multiple board games. Created residual neural network architecture and designed game architecture to correspond with self-play pipeline.
 - Assisted other student groups in researching and implementing Deep Learning Techniques for Computer Vision.
- Lambert Lab, Columbia University Department of Chemistry** Jul 2016 - Sep 2016
Research Assistant New York City, NY
- Researched reaction methods to synthesize derivatives of research group's cyclopentamethylester chiral catalyst.
 - Quantitatively analyzed reaction results utilizing ^1H NMR, Thin-Layer Chromatography, Mass Spectroscopy, and further optimized reactions for desired products.

PROJECTS

- Variational Autoencoders and Hotel2Vec for Hotel Recommendation and Ranking** Fall 2020
- Worked with proprietary e-Commerce Data from RocketMiles to develop novel Deep Learning techniques to compare with their production recommender and ranker.
 - Created data preprocessing, training, and prediction pipeline, as well as network architecture. Despite not outperforming the current RocketMiles Ranker, model learned relevant rankings with extremely minimal data compared to other models.
- Wavelet Pooling Techniques in Convolutional Neural Networks** Spring 2020
- Researched different Discrete Wavelet Transforms as a pooling technique in CNNs on MNIST and CIFAR-10 with Pytorch.
 - Compared CNN's utilizing Wavelet pooling with CNN's utilizing standard max/mean pooling techniques.
- Company Size Prediction Utilizing Quarterly Financial Data** Fall 2019
- Combined and cleaned data from the Securities Exchange Committee with scraped data from Yahoo Finance to study stock data with quarterly company financial data.
 - Utilized a variety of machine learning models(SVM, Random Forest, etc.) to improve classification of company business size by 10%(AUC) from original baseline logistic regression model.

PUBLICATIONS

Methods for the Synthesis of Functionalized Pentacarboxycyclopentadienes Gheewala, C. D.; Radtke, A. M.; Hui, J.; Hon, A. B.; Lambert, T. H. *Org. Lett.* **2017**, 19, 4227-4230.

SKILLS

- | | |
|----------------------------|--|
| Computer Skills | Python (pandas, sklearn, pyspark, dask, XGBoost, Pytorch, Tensorflow)
Spark, Hadoop, MapReduce, bash, R, SQL(MySQL, SQLite), Tableau, git |
| Domain Knowledge | Data Science, Machine Learning, Deep Learning, Hypothesis Testing, Bandits |
| Professional Skills | Classical Cellist (10+ years of perf. exp. at venues such as Carnegie Hall, Lincoln Center) |