

YITONG WANG

91 Lupari, Irvine, CA 9268

(781)-786-1766 • yw652@nyu.edu • www.linkedin.com/in/yitong-wang-a485a158

SKILLS AND INTERESTS

- **Programming:** Python, Java, Scala, Node.js, React.js
- **Skills:** Docker, Hadoop, SQL, Couchbase, CouchDB, ElasticSearch, Kibana, d3, Tableau, L^AT_EX
- **Machine Learning:** Scikit-learn, Pandas, Scipy, Numpy, NLTK, Theano, Tensorflow, PyMC3, XG Boost, Spark
- **Open Source Project:**
 - Python Library for Instant Visualization of Twitter Data (Prototype: dahliallc.github.io)
 - Mongoose-inspired Couchbase ODM: <https://github.com/yw652/lounge-extend-index/tree/master>

EXPERIENCE

Hart, Inc (June 2016 - Present)

Data Scientist – Platform

- Collaborated in building **menrva**, a machine learning platform for training and serving models.
- Built an **MRI Diagnostics Tool** to detect early signs of tumor, which allows automated segmentation, labeling, and training of images.
- Containerized models for production deployment using Docker and Kubernetes containers
- Trained various machine learning and probabilistic model to draw diagnostic prediction for patients based on their **Electronic Medical Records**.
- Incorporated software engineering principles into pushing data science projects toward a faster deployment cycle, including implementing extensive unit-testing, regulating development cycles and processes.
- Maintained code integrity as the owner of critical projects, and responsible for rapid troubleshooting of reported issues in production.

HookLogic, Inc (June 2015 - Sep 2015)

Data Science Intern

- Managed various databases and automated acquisitions of newly incoming data from AWS Redshift to local.
- Analyzed and configured the effects of machine learning components in advertisement delivery performances (RPM, CPC) for each active algorithms.
- Designed automated algorithmic configuration system for constructing algorithms and assigning shares of traffic, and A/B tested the performances of each machine learning algorithm components.

New China Life (May 2014 - Aug 2014)

Actuarial Science Intern

- Analyzed life insurance clauses and practiced risk management for the existing products using VBA
- Prepared and assisted in constructing the third generation life table for Chinese population
- Updated life table database with newly collected data on mortality rates
- Performed detailed analysis of liability documents to ensure correct information on risk modeling

PROJECTS

VQA: Experimenting methods for Visual Question Answering (Sep 2015 - Dec 2015)

- Applied word-embedding on questions and answers of MSCOCO (Microsoft Common Objects in Contexts) dataset in preparation for later VQA tasks.
- Extracted features for each accompanying image and processed them into vectors.
- Explored various Long Short Term Memory neural networks in deep learning for training to achieve the state of the art outcome.

Dahlia: Python Library for Instant Visualization Dashboard of Twitter Data (Sep 2015 - Dec 2015)

- Constructed data pipeline for managing and reducing the data into proper formats for graphical purposes.
- Designed the functions and interactions of each graph contained in the dashboard, based on d3 in JavaScript.
- Hosted all visualizations in JavaScript environment for better user experiences and interactive presentations.

Topic modeling in categorizing website contents in Machine Learning (Jan 2015 - May 2015)

- Extracted real-time website visiting data from multiple publishers using web-scraping.
- Performed supervised learning using raw categorization results from third party data science company.
- Employed both parametric and non-parametric models in training processes for outcome optimization.

EDUCATION

New York University — M.S in Data Science (May 2016)

Brandeis University — B.A in Economics, Mathematics (May 2014)

University of Pennsylvania — Corporate Finance (Summer 2012)