Lizi Chen

www.lizic.com

110 First St. Apt 31 E, Jersey City, NJ 07302
+1 631-561-8141
lizi.chen@nyu.edu
https://github.com/lizichen

Objective: Look for Full-Time Position in Machine Learning Engineer.

Education

Dec. 2018 **New York University**, *Courant Institute of Mathematical Sciences*, *New York, NY*.
M.S. in **Computer Science**

- Data Analytics Researcher at Department of Environmental Science, NYU. *('16 '17)*
- NYU SPIKE Fellowship, NYU AINexus Lab. *(Spring '17)*
- Grader for Undergraduate-level *(Fall '16)* and Graduate-level *(Spring '17)* Algorithms Courses.

Dec. 2013 **Stony Brook University**, *Stony Brook, NY*.
B.S. in **Computer Science**

- Teaching Assistant on: Discrete Mathematics(CSE215) and Java Programming(CSE114).

Work Experience

Jan. 2014 - Mar. 2016 **Goldman Sachs**, *Full-time, Capital Data and Infrastructure, Analyst*, *New York, NY*.

- Took ownership on SEC **Stress Test** projects that overlooked firmwide transactions, analyzed firm capital quantitative structures and risks models, and saved millions USD from Federal Reserve.
- Built **data streaming + analytics interface** front-to-end among 20+ teams and deployed **predictive models** in proprietary Java **ORMs** on multiple million-line-code **Java distributed systems**.
- Designed and sustained multiple **Bi-Temporal relational databases** with 870,000+ object classes and initiated **Big Data** reform with **applied Machine Learning** practices to improve risk analytics.
- Proactively proposed and implemented web services that automated the process of convert **legacy code and data** into use-able ones, while keeping the integrity of previous versions.

Programming Skills

Proficient: **Java**, **Python**(NumPy, MatplotLib, Pandas, Scikit-Learn), **C**(Unix), **SQL**(MySQL, Hive, Impala)

Intermediate: JavaScript, Scala(Spark), C#, MATLAB, Lua, Bash

Skills: Hadoop, Spark, HBase, Tableau, Hibernate, Maven, Nginx, AWS EC2, S3, DynamoDB, Tensorflow, PyTorch, ElasticSearch, MongoDB, Anaconda, Jupyter Notebook, D3.js

Recent Academic Projects

Spring 2018 **Semi-Supervised Learning on Noise Detection, a SoNYC project**.

- Devised **Active Learning** algorithm to build sound classifier on 1 million **unlabeled** audio clips.
- Created Feature abstraction pipeline to convert audio into **VGG-ish** vectors on **Spark Servers**.
- Used Least Confident and Smallest Margin sampling strategies results in F_1-score 0.93.
- Experimented with supervised, semi-supervised and unsupervised algorithms (Label Propagation, Uncertainty Sampling, K-means, PCA, t-SNE), and compared their effectiveness.

Spring 2018 **Audible Social Media**.

- Programmed a web service with **Python Flask**, **Google Speech API**, a **CNN** model in **Tensorflow**, to categorizes social media and converts text feed to steaming-able customized audio clips.
- Deployed services on **EC2** machines that periodically store users' timeline to **DynamoDB** and **S3**.

Fall 2017 **Knowledge Graph Project**.

- Investigated **Neural Network** techniques (Bi-RNN/GRU/LSTM) to extract **pairwise relations** from NYT10 and Wikipedia Biography Datasets .
- Outperformed traditional **Deep Learning** method in joint predictions by using hybrid Linear and Fully-Connected **Conditional Random Field** on global relations prediction.

Fall 2017 **Stock Price Prediction from Full-Text Financial News**.

- Deployed **web crawlers** periodically fetch full-page unstructured news data, and use Machine Learning algorithms to predict mentioned stock future prices with **Heat Map Visualization** on **Tableau**.
- Extracted stock entities and attributes from collected news, and mapped the transformed **word embeddings** to its stock prices; which were ingested from US exchanges via **Spark Streaming**.

More internship experience and project information, available at **www.lizic.com**