

# Xian Lai

Tel. : 917 593 3051

E-mail : [XianLaaai@gmail.com](mailto:XianLaaai@gmail.com)

Webpage : <https://xianlai.github.io/>

Address:

311 W. 55th st.

New York, NY, 10019

---

## **Education:**

- **Fordham University** (Sept, 2016 - present)  
Master of Science in Data Analytics, Computer Information Science, GSAS.  
Graduate in May. 2018. GPA: 3.973/4.000
- **Columbia University** (May, 2012 - May, 2013)  
Master of Science in Advanced Architectural Design, GSAPP.

## **Experience & Related Works:**

**Personal Projects:** <https://xianlai.github.io/>

- Learn features from tweets using Spark Streaming and visualize the results in real-time.  
GitHub: [https://github.com/Xianlai/streaming\\_tweet\\_feature\\_learning](https://github.com/Xianlai/streaming_tweet_feature_learning)
- Clustering and visualize New York rental apartments with online rental posting data.  
GitHub: [https://github.com/Xianlai/Manhattan\\_rental\\_apartment\\_clustering](https://github.com/Xianlai/Manhattan_rental_apartment_clustering)
- Classify online news' popularity with various models and compare the performance.  
GitHub: [https://github.com/Xianlai/online\\_news\\_popularity\\_classification](https://github.com/Xianlai/online_news_popularity_classification)
- Implement and visualize tree searching algorithms including BFS, DFS, A\* search etc.  
GitHub: <https://github.com/Xianlai/Tree-Search-and-Visualization>

## **Previous Working Experience:**

- Architectural Designer in RUR Architecture D.P.C, New York (2013 - 2016).

## **Knowledge & Skills:**

### **- Programming Languages and Applications:**

Python (pandas, numpy, sklearn, tensorflow, pyspark, matplotlib, Bokeh, etc.)

Matlab, R, SQL, bash.

QGIS, Tableau, Weka, Spark, Adobe Creative Suite 6

### **- Discrete Math, Statistics, Linear Algebra and Calculus.**

Able to solve problems using mathematics knowledge including sets, propositional logic and first order logic, vectors and matrices manipulation and decomposition, various optimization methods, probability and distributions, hypothesis testing.

### **- Database Systems and Distributed Computing Frameworks:**

# Xian Lai

Tel. : 917 593 3051

E-mail : [XianLaaai@gmail.com](mailto:XianLaaai@gmail.com)

Webpage : <https://xianlai.github.io/>

Address:

311 W. 55th st.

New York, NY, 10019

---

Understand how database works and able to use SQL to update or query data. Able to use distributed computing frameworks like Apache Spark to analyze large datasets or analyze streaming data in real-time.

## **- Unsupervised Learning Algorithms:**

Able to implement or apply models of linear matrix decomposition(PCA, ICA etc.), non-linear manifold learning(LLE, t-SNE, etc.) and graphical probability model(BN, MRF) to learn the informative or latent variables of dataset and reduce the dimensionality. Able to implement or apply clustering methods to learn internal structure of dataset.

## **- Supervised Learning Algorithms:**

Able to implement or apply learning algorithms like generalized linear models, SVM, tree models, neural networks etc. Able to evaluate the results using appropriate measuring or combine the decisions from various models.

## **- Searching, Planning Algorithms, Logic Knowledge Base Encoding and Inference :**

Able to implement or apply tree/graph search on problems with or without heuristic information or come up with an admissible heuristic. Able to use classic or graph planning to find optimal sequence of actions to achieve goal state. Able to encode knowledge base of logic agent and making inference.

## **- Dynamic Bayesian Network, HMM, Kalman Filter:**

Able to learn the structure or parameters of general time-dependent model like dynamic Bayesian networks or special cases like HMM, Kalman filter from temporal or spatio datasets. Able to perform probabilistic inference using forward-backward algorithms. Able to perform likelihood weighting sampling or MCMC to approximate the desired probabilities.

## **- General Data Structures and Algorithms.**