# Yexi Jiang

*Resume*

*34849 Blackstone Way*
*Fremont, 94555, CA*
*305-458-8476*
*yexijiang@gmail.com*
*GitHub: https://github.com/yxjiang,*
*Homepage: http://yxjiang.github.io/,*
*LinkedIn: https://www.linkedin.com/in/yxjiang*


## Summary


A passionate applied researcher and software developer with hands-on skills on data mining/machine learning algorithm development and research. Strong eagerness to work with talented and energetic people on promising projects.


## Experience

### Industrial Experience

05/2015 **Research Scientist**, ***Facebook Inc.***, Menlo Park, CA, USA.

- Tech lead of News Feed High Quality Engagement team.
  - Initiated and framed the problem of fake engagement detection and mitigation.
  - Worked with XFN and partner team on problem formulation, measurement, opportunity sizing, and road-mapping.
  - Drove the designing and implementation of fake engagement detection system and mitigation system.
  - Leveraged data mining and machine learning techniques to detect accounts conducting 10M+/day of fake engagement, took actions on 100k+ of accounts directly, and reduced the impressions of contents benefit from fake engagement by 500M+/day.
  - Had 1 on 1 with the junior engineers to make sure they executed well.
- Tech lead of Facebook's friend recommendation system PYMK since the second year in team.
  - Worked with XFN on idea initiation, opportunity sizing, and project prioritization.
  - Incrementally increased Facebook's Monthly Active People (MAP) by 5+ millions through various friend recommendation optimization efforts, including model selection, model serving, data logging, training data preparing, feature engineering, and candidate retrieval.
  - Drove efforts of system stability improvement from the aspects of oncall/debugging procedure standardization, and alert system improvement.
  - Collaborated with partner growth teams on joint projects, such as friend request conversion optimization and single user multiple accounts detection.

Summer of 2014 **Intern**, ***News Feed Ranking team, Facebook Inc.***, Menlo Park, CA, USA.

- Interned at news feed ranking team. Designed and implemented the pipeline as well as the related analytics tools for personalized News Feed ranking at Internet scale, making the recommendation more accurate for active users in Facebook.

Summer of 2011, 2012, 2013 **Intern**, ***Service Management team, IBM T.J. Watson Research Center***, New York, NY, USA.

- Participated in the design and implementation of an interactive service retrieval system called *Cloud Services Marketplace* to facilitate the cloud services acquisition. This prototype evolved into IBM 2013 Global Technology Outlook project "Scalable Services Ecosystem".
- Designed and implemented a time-series prediction algorithm to help the cloud capacity planning and reduce the VM fulfillment time of IBM's Smart Cloud Enterprise from the perspective of resource prediction.
- Participated in the design and implementation of a tool that leverages data mining techniques to improve the efficiency for both the customers and service providers during server configuration.

03/2009-07/2009 **Intern**, ***Data Intelligence and Tools team, Microsoft Research***, Beijing, China.

- Designed and implemented a distributed frequent sequential mining algorithm to discover the users' popular operation sequences from the log of Office software to help the UI designers ameliorate the layout of the next version of Office.

### Open Source Projects

Apache Software Foundation **PMC Member**, ***Apache Hama***: *a BSP computing framework for distributed computing.*

- Working on the distributed machine learning package, including the design and implementation of distributed multilayer perceptron, linear regression, logistic regression, and auto-encoder. [**GitHub Link**]

## Publication

- Program committee member and reviewer of the top academic conferences/journals in the area of data science, such as: *ACM Conference on Knowledge Discovery and Data Mining (**SIGKDD**)*, *International Conference on Data Mining (**ICDM**)*, *ACM Conference on Recommender Systems (**RecSys**)*, *IEEE Transactions on Knowledge and Data Engineering (**TKDE**)*, *ACM Transactions on Knowledge Discovery from Data (**TKDD**)*, *Knowledge and Information System (**KAIS**)*, *Information Sciences*, *IEEE Transactions on Cybernetics*, *Data and Knowledge Engineering (**DKE**)*, etc.
- Published 20+ papers at the top journals/conferences in the area of data science, such as: *ACM Conference on Knowledge Discovery and Data Mining (**SIGKDD**)*, *International Conference on Data Mining (**ICDM**)*, *ACM Conference on Recommender Systems (**RecSys**)*, *ACM Conference on Information Retrieval (**SIGIR**)*, *ACM Conference on Information and Knowledge Management (**CIKM**)*, *IEEE Transactions on Knowledge and Data Engineering (**TKDE**)*, *ACM Transactions on Intelligent Systems and Technology (**TIST**)*, etc. (Details at: `https://scholar.google.com/citations?user=ojlglBsAAAAJ&hl=en`)
- Published 3 patents during the internships at IBM,

## Education

2015 **Ph.D. in Computer Science**, *Florida International University*, Miami, FL,USA.
Research Area: Temporal Data Mining, Distributed Data Mining

2010 **B.S and M.S. in Computer Science**, *Sichuan University*, Chengdu, Sichuan,China.
Research Area: Data Mining