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. Extensive R&D experience at the world's leading companies and open source community like Facebook, IBM, Microsoft, and Apache Software Foundation. More than 20 research paper published at recognized international conferences and journals. Strong eagerness to work with talented and energetic people on promising projects.

Experience

Industrial Experience

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

- Tech lead of Facebook's friend recommendation machine learning system.
- Increased Facebook's Monthly Active People (MAP) by 3.7M through various friend recommendation optimization, including the integration of neural network and large scale sparse logistic regression and the improvement of training data quality.
- Refactored the friend recommendation backend, improved friend recommendation feature logging efficiency by 30x and increased the training data collection volume by 25x.

Summer of 2014 **Intern**, ***Facebook Inc.***, Menlo Park, CA, USA.

- 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**, ***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**, ***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 on top of Hadoop*.

- 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**]

Academic

08/2010-04/2015 **Research Assistant**, *Knowledge Discovery Research Group, FIU*, Miami, FL, USA.

- Designing and implementing a distributed monitoring system to conduct the continuous monitoring over the computer clusters. [**GitHub Link**]
- Participating in the design and implementation of FIU-Miner: An integrated and user-friendly data mining system under distributed environment. [**Web Page Link**]

Programming Skills

Languages	Java (proficient), Python (proficient), C/C++ (familiar), Scala (prior experience)
Frameworks & Libries	Apache Hive, Apache Hama, Apache Hadoop, ActiveMQ, Storm

Patents

1. *Cloud Provisioning Accelerator*. With Rong Chang, Mihwa Choi, Meir Laker, Chang-Shing Perng, Hidayatullah Shaikh, Edward So, Tao Tao. **US 20130139152**.
2. *Interactive Acquisition of Remote Services*. With Rahul Akolkar, Thomas Chefalas, Jim Laredo, Chang-Shing Perng, Anca Sailer, Frank Schaffa, Alla Segal, Ignacio Silva-Lepe, Tao Tao, Yang Zhou. **US 20140337010**.
3. *Complex Service Network Ranking and Clustering*. With Rahul Akolkar, Thomas Chefalas, Jim Laredo, Chang-Shing Perng, Anca Sailer, Frank Schaffa, Alla Segal, Ignacio Silva-Lepe, Tao Tao, Yang Zhou. **Filed**.

Research

- 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`)
- 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.

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