Yexi Jiang

Resume

📞 *305-458-8476*
✉ *yexijiang@gmail.com*
GitHub: `https://github.com/yxjiang,`
Homepage: `http://yxjiang.github.io/,`
LinkedIn: `https://www.linkedin.com/in/yxjiang`

Immigration Status

Permanent Resident of USA

Summary

A passionate result driven practitioner 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

2019-Now **Tech Lead Manager**, ***Uber Technologies Inc.***, San Francisco, CA, USA.
- Machine Learning Platform, Advanced Technologies Group
  Support the ML application track under the Machine Learning Platform pillar, including:
  - Support the inference service used in model evaluation for the use cases of perception, prediction, motion planning, simulation, and more for the autonomous vehicles.
  - Support the ML-as-a-Service efforts, connecting the ML tasks such as dataset generation, dataset management, model training, model management, and various model evaluation stages into a holistic ecosystem.

2014-2019 **Research Scientist**, ***Facebook Inc.***, Menlo Park, CA, USA.
- Tech Lead of Feed and Story Integrity team
  - Building the ML-based system to detect, measure, and enforce the violating (e.g. regulated goods, bullying) and low-quality (e.g. clickbait, ads farm) contents appeared in stories.
  - Leverage ML to tackle fake engagement on Facebook contents, designed and implemented the system for measurement, detection, and enforcement. Successfully reduced xx% of active accounts manipulated for fake engagement and reduced the impressions of the benefited contents by xxxM/day.
  - For both projects, worked with XFN and partner teams on problem formulation, road-mapping, system designing, and hold regular 1 on 1 with the junior engineers to make sure they are not blocked.
- Tech Lead of User Growth (People You May Know) team
  - Joined the team and then tech led the friend recommendation (PYMK) team after the 2nd year. The improvements on the service has increased Facebook's Monthly Active People (MAP) by xM/year through various optimization efforts, including model selection, model serving, data logging, training data preparing, feature engineering, and candidate retrieval.
  - Led the efforts of system stability improvement. Standardized and simplified the on-call/debugging guideline, and consolidated the monitoring of the system.
  - Major improvement of the friend recommendation system backend, increased the effectiveness of feature logging by 20+ Times.
- Intern of News Feed Ranking team
  - Designed and implemented the pipeline as well as the related analytic tools for personalized News Feed ranking at Internet scale, increased the AUC of the prediction by 0.x%.

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

- ○ Participated in the design and implementation of a prototype system called *Cloud Service Marketplace*, an RDF knowledge graph based interactive system that facilitates 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 by 80+%.

- ○ 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.

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.

2008-2009	**Data Mining Engineer**, *Mobey Inc.*, Chengdu, China.

- ○ The only data dedicated engineer in a yelp-like startup, working on web page crawling, Chinese word segmentation, and web page classification.

Open Source Projects

Open Source Community	**Apache Software Foundation**.

- ○ **PMC Member** of Apache Hama: A BSP 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**]
- ○ **Committer** of Apache Mahout: A machine learning library for Hadoop.
  - − Implemented the multilayer neural network model trained with stochastic gradient descent.

▬▬ 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 with 700+ citations, 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://goo.gl/4dRHrb`)

- ○ 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