Peng Xu

Curriculum Vitae

Education

2016-present **M.Sc. in Computing Science**, *University of Alberta*, Edmonton, *GPA: 4.0*.

2011-2015 **B.Eng. in Electronic Information Engineering**, *Beijing University of Posts and Telecommunications*, Beijing, *Major: 89.62/100, Overall: 87.06/100*.

Research Experience

2016.5 - present **Research Assistant, Database Group**, *University of Alberta*, Edmonton, *Supervised by Denilson Barbosa, focus on knowledge harvesting from the web, link and relation prediction on knowledge base, question answering based on knowledge base*.

2014.9 - 2015.7 **Research Assistant, Knowledge Engineering Group**, *Tsinghua University*, Beijing, *Supervised by Jie Tang, focus on machine learning, data mining and information retrieval. More specifically, name disambiguation in large scale database and cross-domain knowledge linking*.

2014.4 - 2015.2 **Research Assistant, High Performance Computing Laboratory**, *Tsinghua University*, Beijing, *Supervised by Zhihu Du, focus on parallel computing via GPU hardware and distributed systems. More specifically, efficient partical-mesh spreading on GPUs*.

Publications

2015 Xiangyu Guo, Xing Liu, **Peng Xu**, Zhihu Du, Edmond Chow, ***Efficient Particle-Mesh Spreading on GPUs***,In Proceedings of the Fifteenth International Conference On Computational Science (ICCS'15)

Project Experience

2016 **Inferring Infobox Template Class Mappings From Wikipedia and WikiData (under DBpedia)**, *Python*, Google Summer of Code 2016

First, exploit the Wikipedia pages (instances) already mapped with a DBpedia class and their cross language links to the pages that the template to be mapped to infer new mappings. With this approach, we obtain 456 high-quality mappings for Chinese as output of the project. Then, I use tensor factorization and graph embeddings to give link predictions on DBpedia. All codes and documents are maintained in this ***repo***.

(780)6558396 • pxu4@ualberta.ca

2014-2015 **Arnetminer II**, *Python, Scala, Shell*, Research Project
Arnetminer is a website that offers comprehensive search and mining services for academic community. I has worked on the second version of it, and been responsible for designing and implementing API of the back-end server and handle the data in the databases, including experts recommendation, conference ranking, affiliations merge, linking patents and so forth.

2014-2015 **Crawler for Google Scholar**, *Python*, Research Project
Develop a web crawler to crawl the whole co-author network with all their papers on Google Scholar through bfs. In order to avoid blockage of Google, use proxies via proper strategy. In addition, maintain the crawled data and match them to the corresponding databases.

2014 **Efficient Particle-Mesh Spreading on GPUs**,*C, CUDA*, Research Project
Study various approaches for particle-mesh spreading on GPUs. A central concern is the use of atomic operations. We are also concerned with the case where spreading is performed multiple times using the same particle configuration, which opens the possiblity of preprocessing to accelerate the overall computation time. Experimental tests show which algorithm are best under different circumstances. Our work is accepted as full paper in proceedings of the Fifteenth International Conference On Computational Science (ICCS'15)

Teaching

2016 Fall **Teaching Assistant**, *CMPUT466/551 Machine Learning*, University of Alberta.

MOOC Certificates

- Verified Certificates: Functional Programming Principles in Scala (EPFL, 100%), Functional Program Design in Scala (EPFL, 100%), Parallel Programming (EPFL, 95%), Neural Networks for Machine Learning (UofTornoto)
- Statements of Accomplishment **with Distinction**: Machine Learning (Stanford, 100%), Game Theory (Stanford, 92.9%), Mining Massive Datasets (Stanford, 93.5%), Introduction to Big Data with Apache Spark (UCB, 100%), Scalable Machine Learning (UCB, 97%)
- Statements of Accomplishment: Statistical Learning (Stanford, 82.4%)

Skills and Qualifications

Languages Python, C/C++, Scala, CUDA, Matlab(Octave), Bash, R

OS Microsoft Windows, Mac OS, Linux

Tools LaTeX, MongoDB, Tensorflow, Keras, Numpy, Panads, Scikit-learn, XGBoost, Scrapy

Awards and Honors

- **Kaggle: Home Depot Product Search Relevance** Top 10%
- 2013 Bronze Medalist in **ACM-ICPC Hunan Invitational Programming Contest**
- 2013 Gold Medalist in **Programming Contest of BUPT**
- 2013-14 Third Grade Scholarship for Undergraduates
- 2012-13 Second Grade Scholarship for Undergraduates
- 2011-12 Third Grade Scholarship for Undergraduates

(780)6558396 • pxu4@ualberta.ca