Chapin-I-1125A, 700 Health Science Drive.
Stony Brook, NY, 11790
✆ *(+1) 3158793236*
✉ *adoni1203@gmail.com*
✎ *adoni.github.io/~xfsun*

Xiaofei Sun

Education

2017–Present **Ph.D. of Computer Science**, *Stony Brook University*, USA.
- GPA: 3.77

2015–2017 **Master of Science**, *Harbin Institute of Technology*, China.
- Advisor: Prof. Ting Liu. Research Center for Social Computing and Information Retrieval.
- GPA: 84.18% (**Top 2/193**)

2013.9–2014.1 **Exchange Study**, *National Taiwan University*, Taiwan.
- Selected as one of two students in HIT supported by Chinese Scholarship Council.
- Studied graduate-level courses on computer science.
- GPA: 4.08/5

2011–2015 **Bachelor of Science**, *Harbin Institute of Technology*, China.
- Computer Science and Technology, graduated from the **Honors Program**.
- Thesis: User Profiling on Sina Weibo, **Outstanding Thesis Award (2%).**
- GPA 88.51% (**Top 1/5**)

Experience

2017.9–Present **Teaching Assistant**, *Course: CSE 519 Data Science*, Prof. Steven Skiena.

2017.3–2017.7 **Intern**, *Weizoom Media Company*, Beijing, China.
- Worked on project *Smart Choice*, an algorithm to recommend products to sell for a text-rich account.
- Build a recommendation for Weizoom to improve their sale and update their product pools.

2016.5–2016.10 **Executive Committee Member**, *Social Media Processing Technique Evaluation (SMP CUP 2016)*, Jiangxi, China.
- In charge of collecting data from Weibo. Dataset is released to the public. [Dataset]
- There were 197 teams from China, America, Canada and Singapore, among which 63 teams submitted final results.

2016 Spring **Teaching Assistant**, *Course: Software Design and Practice I: Data Structure*, Prof. Yang Liu.

2015.5–2015.9 **Summer Intern**, *HTC Research & Innovation Center*, Beijing, China.
- Mentor: Wayne Xin Zhao
- Worked on project *User Profiling on E-commerce Websites from Online Review*.

2015.5–2015.9 **Summer Intern**, *Printed Media Center of National Language Monitor and Research*, Beijing Language and Culture University, China.
- Main developer of the Semantic Cloud Platform, which detects the semantics vicissitude of Chinese.

Research Projects

2016 **Content Enhanced Network Embedding (CENE)**, [Code].
- A novel approaches for learning network embeddings that could jointly leverage structure and content information in a social network.
- Applications on node classification (attributes: gender, profession, location) showed that our proposed approach significantly improved the F1 score up to 7% when comparing to existing network embedding models.

2015 **Semantic Cloud Platform (SCP)**, [Code][Demo] .
- Supported by Beijing Language and Culture University. The main purpose was to detect semantic vicissitude.
- Using word2vec to learn word embeddings in each year and tooking the nearest neighbors as the possible semantics, which was clustered by Affinity Propagation a shared embedding space. The final clustering results were regarded as the semantics of each word.

2015 **User Profiling on E-commerce Websites from Online Review**.
- Using weakly-supervised learning method for demographic attribute prediction based on demographic mentions.
- Demographic mentions were phrases which revealed the demographic information of the purchaser. For example, if someone said: "I bought my wife iPhone 6s, and it worked well". Then the consumer is probable to be a married male.

2014–2015 **User Profiling on Weibo**.
- Part of my bachelor dissertation.
- Predicting user attributes (e.g. gender, age and location) in Sina Weibo using a semi-supervised representation model to combined different user representation approaches.

2014 **Weibo Spider**, .
- A distributed crawler to collect data on Weibo.
- Collected data was used in *SMP CUP 2016* and *User Profiling on Weibo* mentioned above.

Research Interests

Distributed representation learning for Social/Information Networks.

Deep Learning for Natural Language Processing.

Data mining on Social Networks.

Honors & Awards

2016 National Scholarship for Graduate.

2015 Best Paper Award, *Chinese National Conference on Social Media Processing 2015*, (4th author).

2015 Outstanding Undergraduate Graduation Thesis Award, *Harbin Institute of Technology*.

2015 Second Prize of National Level Innovation Project for Undergraduate, **Project Leader**.

2014 Microsoft Research Asia Young Fellow Scholarship Award.

2012 Reginal Outstanding Prize of the Android Application Challenge in Northeast China, **Team Leader**.

2012 Regional First Prize of the Undergraduate Mathematical Contest in Modeling.

2011 Second Prize of Freshman Innovation Project, **Project Leader**.

2011-2016 First Prize of People's Scholarship.

Publications

[1] **Xiaofei Sun**, Jiang Guo, Xiao Ding, Ting Liu. A General Framework for Content-enhanced Network Representation Learning[J]. arXiv preprint arXiv:1610.02906, 2016..

[2] Yang Li, Jing Jiang, Ting Liu, **Xiaofei Sun**. Personalized Microtopic Recommendation with Rich Information[C]. Chinese National Conference on Social Media Processing. Springer, 2015: 1-14.

[3] **Xiaofei Sun**, Xin Zhao, Ting Liu. Research of the Influence of User Representation Methods for User Attributes Classification on Weibo[C]. Chinese National Conference on Social Media Processing. 2015. (in Chinese).

[4] **Xiaofei Sun**, Xiao Ding, Ting Liu. Gender Identification on Social Media[C]. Chinese National Conference on Social Media Processing. Springer, 2014: 99-107.

Miscellaneous

Programming	**Python, C/C++, Java, LaTeX**, Ruby, Go, Matlab.
Skills	DyNet, PyTorch, Scikit-Learn, Pandas, Web Crawler, Website Building .