Afshin Rahimi

afshinrahimi.github.io

Research Interests

My research interests fall within the fields of Natural Language Processing, Social Network Analysis and Machine Learning. I'm particularly interested in using large amounts of structured and unstructured language data to help machines understand natural language. My PhD research was focused on combining text and network information in a semi-supervised setting to learn user demographics and specifically Twitter user geolocation which resulted in state-of-the-art geolocation performance published in ACL, NAACL and EMNLP.

Education

2014–2018 **PhD in Computer Science**, *The University of Melbourne*, Melbourne, Australia.
Thesis: Geolocation of Social Media Users Exploiting Both Text and Network Information.

2010–2013 **MSc. Computational Linguistics**, *Sharif University of Technology*, Tehran, Iran.

2002–2006 **BSc. Computer Science**, *Sharif University of Technology*, Tehran, Iran.

Publications

Afshin Rahimi, Trevor Cohn, and Timothy Baldwin. 2018. Semi-supervised user geolocation via graph convolutional networks. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL 2018)*, Melbourne, Australia. Association for Computational Linguistics.

Afshin Rahimi, Trevor Cohn, and Timothy Baldwin. 2017b. Continuous representation of location for geolocation and lexical dialectology using mixture density networks. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP 2017)*.

Afshin Rahimi, Trevor Cohn, and Timothy Baldwin. 2017a. A neural model for user geolocation and lexical dialectology. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017)*.

Afshin Rahimi, Trevor Cohn, and Timothy Baldwin. 2016. pigeo: A python geotagging tool. *Proceedings of ACL-2016 System Demonstrations*, pages 127–132.

Afshin Rahimi, Trevor Cohn, and Timothy Baldwin. 2015. Twitter user geolocation using a unified text and network prediction model. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics — 7th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2015)*.

Afshin Rahimi, Duy Vu, Trevor Cohn, and Timothy Baldwin. 2015. Exploiting text and network context for geolocation of social media users. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics — Human Language Technologies (NAACL HLT 2015)*, Denver, USA.

Building 261, 203 Bouverie St,
Carlton Victoria, Australia (Permanent Residence)
✉ *afshinrahimi@gmail.com* • *afshinrahimi.github.io* • *Google Scholar*

Work Experience

2018- **Research Fellow in NLP**, The University of Melbourne, Melbourne, Australia.
I'm working on NLP for low-resource languages with Trevor Cohn.

2016–2017 **Internship (5 months)**, Telstra TELCO, Melbourne, Australia.
My internship was on customer intent prediction in interactive voice response (IVR) systems based on historical events (website visits, previous calls). My work involved working with technologies like GreenPlum and Theano to create learning models that can predict the intent of a customer (e.g. talking about billing, a problem with their Smart TV service or a failing modem) from their previous events (e.g. page visit, previous calls, previous bill) using classification, language modelling and association rules which improved routing of customers to agents resulting in customer satisfaction and cost reduction. The CTO of the company mentioned the results of my work in a general meeting and supported the investment of the company in the same line of work to improve customer experience.

2013–2014 **Researcher**, Linneaus University, Stockholm, Sweden.
I was a member of Sentiment Analysis team which focused on the study of stance/sentiment in big data real-time text streams like Twitter. My job involved working on Vector Space Models, Feature Extraction and Classification/Clustering of Stance/Sentiment in real time big data streams.

2011–2013 **Java/MapReduce Developer**, MITRC, Tehran, Iran.
I was responsible for a team of 4 developers working on the crawler component of a large scale web search engine. I could improve the crawler's performance from just 1 million fetched URLs to 1 billion at the end. The crawler had about 2 billion URLs where 1 billion ones were fetched and updated regularly. The data was about 30 TB. We used Hadoop's HDFS, MapReduce and HBase. We built our web crawler on top of Apache Nutch, an open source popular web crawler.

2006–2011 **Software Developer and System Administrator**, KECO, Tehran, Iran.
Main responsibilities include: Software Development (Web and Desktop) using .NET (C# and ASP) and MSSQL. Cisco Switching Administrator (50+ access switches and 4 distribution switches in a 2 tier network architecture, 500+ nodes, vlan, portsec, spanning-tree, redundancy, architecture, load balancing, high-availability), Windows Network Administrator (Active Directory, High Availability, Centralized Management), ERP Administrator (Supply Chain, Salary, HR), Microsoft SQL Server Administrator and Developer. I learned to work with technology under heavy pressure where a medium sized company (2k+ users) relies on your design and administration 24/7.

Academic Activities

Reviewing Reviewed for ACL, EMNLP, EACL and Transactions of Information Systems.

Teaching Teaching Assistant for Machine Learning Course, 2017, University of Melbourne and Computational Linguistics Course, 2012, Sharif University of Technology.

Workshop Collaborated on organizing W-NUT shared task for EMNLP 2016.

Technical skills

NLP Familiar with Core NLP tasks in Statistical NLP (e.g. tagging, Parsing, ASR, LM, MT)

Machine Learning	Familiar with optimization concepts, supervised (classification), unsupervised (e.g. autoencoders, dimensionality reduction, embeddings, clustering) and (graph-based) semi-supervised models. Hands-on experience with Scikit-Learn, Tensorflow and Theano. Experience with neural network models such as MLP, Convolutional, Seq2Seq, RBF and Mixture Density Networks (MDN)
SNA	Experienced in data collection (e.g. from Twitter), data curation to create datasets, graph construction, label propagation and random walks in the graph (e.g. LP, Modified Adsorption) and using both node features and network structure for inductive/transductive learning (e.g. using Graph Convolutional Network)
Collaboration	Git, SVN, JIRA, GreenHopper
Big Data	Experienced in Hadoop MapReduce, NoSQL and Hive, Familiar with Spark
Database	MS/My/Posgre SQL, GreenPlum, HBase, Lucene
Network	Cisco Switching Design and Maintenance
Programming	Experienced in Python, Java and .NET, Familiar with C++ and MATLAB
OS	Linux

Languages

English	Equal to IELTS 8 in all skills
Persian	Native

Honors and Awards

2017	My paper on geolocation was selected to be presented in the outstanding session in ACL 2017.
2015	Awarded Melbourne University's Travel Scholarship to attend NAACL 2015 conference.
2015	Awarded Google's Travel Scholarship to attend ACL 2015 conference.
2014	Awarded Australia's government scholarship (IPRS) for PhD studies in The University of Melbourne.
2013	Graduated as the top student of the M.Sc. program in Computational Linguistics from Sharif U. of Tech.
2002	Ranked 241st among more than 500k high school students in national entrance exam for universities in Iran.

References

Tim Baldwin tb@ldwin.net	My PhD supervisor at The University of Melbourne.
Trevor Cohn t.cohn@unimelb.edu.au	My PhD supervisor at The University of Melbourne.