Su WANG

305 E. 23rd Street STOP B5100
Austin, TX, 78712
(716) 578-5920
shrekwang@utexas.edu

EDUCATION

Ph.D. in Computational Linguistics (GPA 4.0) — 2015 - Present
Advisors: Katrin Erk (Linguistics); Greg Durrett (CS)
M.S. in Statistics (GPA 4.0) — 2016 - Present
University of Texas at Austin, Austin, TX

M.A. in General Linguistics — 2012 - 2015
Computer Science Minor (GPA 4.0)
University at Buffalo, Buffalo, NY

M.A. in Applied Linguistics — 2009 - 2012
Yunnan University, China

PUBLICATIONS

Peer-reviewed

[1] **Su Wang**, Eric Holgate, Greg Durrett and Katrin Erk. *Picking Apart Story Salads*. EMNLP 2018 (oral/poster: TBD).
[2] **Su Wang**, Greg Durrett and Katrin Erk. *Modeling Semantic Plausibility by Injecting World Knowledge*. NAACL 2018 (oral).
[3] **Su Wang**, Stephen Roller and Katrin Erk. *Distributional model on a diet: One-shot word learning from text only*. IJCNLP 2017 (oral).
[4] **Su Wang**, Elisa Ferracane and Raymond Mooney. *Leveraging discourse information efficiently for authorship attribution*. IJCNLP 2017 (oral).

Other work

[1] **Su Wang** and Joshua Levy. (2017). *Named Entity Recognition in Practice*. Journal of KUST. Vol2.
[2] **Su Wang**. (2017). *Non-Cooperative Behavior in Conversational Interaction*. Journal of KUST. Vol1.
[3] **Su Wang**. (2013). *The Representation of Topological Relations in Naxi Language – A Statistical Approach*. in Cha-Ma Ancient Trail Research Anthology, (Shiyuan Wang eds.). Vol3. Yunnan University Press: Kunming.
[4] **Su Wang**. (2013). *On a Descriptive Word Classification of "Proto-Unergative–Proto-Unaccusative" Verbs*. Journal of Yunnan Normal University. Vol3.

Talks

[1] *An Introduction to Topic Modeling*. NLP Day 2017, Austin TX.
(Link: http://globaldatageeks.org/nlpdaytx17/sessions#wang)
[2] *Exploring Question-Answering System: Named Entity Recognition & Sentence Similarity Measure in Practice*. DataDay Texas 2017, Austin, TX.
(Link: http://datadaytexas.com/sessions#wang).

ACADEMIC EXPERIENCE

Research Assistant 2017 - Present
EMPLOYER: Katrin Erk, PhD.,
Department of Linguistics, University of Texas at Austin.
PROJECT: Deep Natural Language Understanding
with Probabilistic Logic and Distributional Similarity (DEEPsem)
(NSF Grant No. 1523637)
RESEARCH FOCUS:
- Semantic space and representation.
- Inference from embeddings.
- Neural networks.

Research Assistant 2016 - 2017
EMPLOYER: Katrin Erk, PhD.,
Department of Linguistics, University of Texas at Austin.
PROJECT: Deep Exploration and Filtering of Text (DEFT)
RESEARCH FOCUS:
- Word Property Learning with Bayesian Hierarchical Models.
- Representation of Abstract Knowledge.

Teaching Assistant 2015 - 2016
EMPLOYER: Katrin Erk, PhD.,
Department of Linguistics, University of Texas at Austin.
COURSES:
- LIN353C Introduction to Computational Linguistics (Spring 2016)
- LIN350 Analyzing Linguistics Texts (Fall 2015)

RESPONSIBILITIES: Grading and answering programming questions;
occasional lectures.

Lab Researcher 2012 - 2014
LAB NAME: Semantic Typology Lab, University at Buffalo.
RESEARCH FOCUS:
- Quantitative methods in crosslinguistic study of semantic typology.
- Sino-Tibetan languages (Yi, Bai and Naxi).

Research Assistant 2009 - 2012
EMPLOYER: Liquan Yang, PhD. (Peking University),
Department of Linguistics, Yunnan University.
RESEARCH FOCUS:
- Statistical modeling in semantics (indigenous languages of Yunnan).
- Language Documentation (Yi, Bai and Naxi).

ACADEMIC AWARDS

Research Excellence Award in Natural Language Processing (NNSFC) 2018 - 2020

Professional Development Award 2017 - 2019
SPONSOR INSTITUTION: Department of Linguistics, UT Austin

Professional Development Scholarship 2012 - 2015
SPONSOR INSTITUTION: Hua Ruan Software
PROGRAM: Master in General Linguistics (CS Minor), University at Buffalo

Graduate Scholarship 2009 - 2012

SPONSOR INSTITUTION: Yunnan University
PROGRAM: Master in Applied Linguistics, Yunnan University

PROFESSIONAL EXPERIENCE

Software Engineer PhD Intern Summer 2018
EMPLOYER: Google Inc., Mountain View, CA
SUPERVISOR: Nancy Chang
CO-SUPERVISORS: Rahul Gupta, Jason Baldridge
PROJECTS: Efficient Paraphrase Generation with Structured Semantic Knowledge.

NLP Data Scientist Intern, Technical Consultant 2016 - Present
EMPLOYER: OJO Labs Inc., Austin, TX
SUPERVISOR: Joshua Levy, Chief Scientist.
(PhD. in Computer Science, University of North Carolina at Chapel Hill)
RESPONSIBILITIES: Consultation for linguistics-related NLP problems;
Software engineer; Development of Question-Answering system.
PROJECTS:

- Named Entity Recognition System for OJO Chatbot.
- Information Retrieval system for real estate search.

Technical Consultant 2012 - Present
Software Engineer 2009 - 2012
EMPLOYER: Hua Ruan Software, Yunnan, China
RESPONSIBILITIES: Software engineer; NLP solutions; Statistics consultation.
PROJECT: Identity Verification System.

Test Analyst, Instructor 2009 - 2011
EMPLOYER: Century Youth Education, Yunnan, China
RESPONSIBILITIES: Teaching test-taking skills IELTS and TOEFL.
AWARDS: Best Instructor (2010,2011); Best Average Scoring (2011).

TECHNICAL SKILLS

PROGRAMMING: Python, Java, C/C++, Prolog, R (Most Proficient in Python), working knowledge of HTML, CSS, JavaScript.
STATISTICS: Neural Network Architectures, Bayesian Inference; Hierarchical Models; Mathematical Statistics.

LANGUAGE SKILLS

MANDARIN CHINESE: Native.
ENGLISH: Fluent.
JAPANESE: Beginner.

TECHNICAL TUTORIALS

[1] *A Brief Tutorial on Machine Learning.*

[2] *Notes on Statistical Learning.*

[3] *Generative Adversarial Networks (GAN): A Gentle Introduction.*

[4] *Bayesian Statistical Methods: A Primer.*

[5] *EM-algorithm and Clustering: a Tutorial.*

[6] *Topic Modeling: A Complete Introductory Guide.*