Urvashi Khandelwal

CONTACT
INFORMATION

353 Jane Stanford Way
Gates 232
Stanford CA, 94305

urvashik@stanford.edu
https://nlp.stanford.edu/~urvashik/

EDUCATION

Stanford University, CA
Ph.D. Computer Science
Advisor: Dan Jurafsky

2015 - Present

University of Illinois Urbana-Champaign, IL
B.S. Computer Science, Minor in Mathematics
Advisor: Jiawei Han

2011 - 2015

RESEARCH
INTERESTS

Understanding and improving the generalization capabilities of neural language models and models for sequence generation; Non-parametric methods for improving out-of-distribution performance; Model interpretability and algorithmic fairness; Representation learning for better language understanding.

AWARDS

Microsoft Research Dissertation Grant
Winner

2020

CRA Outstanding Undergraduate Researchers Award
National Winner

2015

C.W. Gear Outstanding Undergraduate Award
Winner, University of Illinois Urbana-Champaign

2015

PUBLICATIONS

Nearest Neighbor Machine Translation.
Urvashi Khandelwal, Angela Fan, Dan Jurafsky, Luke Zettlemoyer and Mike Lewis.
International Conference on Learning Representations (ICLR), 2021.

With Little Power Comes Great Responsibility.
Dallas Card, Peter Henderson, **Urvashi Khandelwal**, Robin Jia, Kyle Mahowald and Dan Jurafsky.
Empirical Methods in Natural Language Processing (EMNLP), 2020.

Emergent Linguistic Structure in Artificial Neural Networks Trained by Self-Supervision.
Chris Manning, Kevin Clark, John Hewitt, **Urvashi Khandelwal** and Omer Levy.
Proceedings of the National Academy of Sciences (PNAS), 2020.

Generalization through Memorization: Nearest Neighbor Language Models.
Urvashi Khandelwal, Omer Levy, Dan Jurafsky, Luke Zettlemoyer and Mike Lewis.
International Conference on Learning Representations (ICLR), 2020.

What does BERT look at? An Analysis of BERT's Attention.
Kevin Clark, **Urvashi Khandelwal**, Omer Levy and Christopher D. Manning.
BlackboxNLP, 2019. **(Best Paper Award)**

BAM! Born-Again Multi-Task Networks for Natural Language Understanding.
Kevin Clark, Minh-Thang Luong, **Urvashi Khandelwal**, Christopher D. Manning

and Quoc V. Le.
Association for Computational Linguistics (ACL), 2019.

Sample Efficient Text Summarization Using a Single Pre-Trained Transformer.
Urvashi Khandelwal, Kevin Clark, Dan Jurafsky, Lukasz Kaiser.
ArXiv Preprint, 2019. Presented at WestCoast NLP, 2019.

Sharp Nearby, Fuzzy Far Away: How Neural Language Models Use Context.
Urvashi Khandelwal, He He, Peng Qi and Dan Jurafsky.
Association for Computational Linguistics (ACL), 2018.

ClusCite: Effective Citation Recommendation by Information Network Based Clustering.
Xiang Ren, Jialu Liu, Xiao Yu, **Urvashi Khandelwal**, Quanquan Gu, Lidan Wang and Jiawei Han.
International Conference on Knowledge Discovery and Data Mining (KDD), 2014.

Personalized Entity Recommendation in Heterogeneous Information Networks with Implicit User Feedback.
Xiao Yu, Xiang Ren, Yizhou Sun, Quanquan Gu, Bradley Sturt, **Urvashi Khandelwal**, Brandon Norick and Jiawei Han.
International Conference on Web Search and Data Mining, (WSDM), 2014.

HeteRec: Entity Recommendation in Heterogeneous Information Networks with Implicit User Feedback.
Xiao Yu, Xiang Ren, Yizhou Sun, Bradley Sturt, **Urvashi Khandelwal**, Quanquan Gu, Brandon Norick, and Jiawei Han.
International Conference on Recommender Systems (RecSys), 2013.

REPORTS   Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies.
Report submitted to the Administrative Conference of the United States (ACUS).
February 2020.
Served as the technical lead for two case studies:
- Informal Adjudication at the U.S. Patent and Trademark Office
  Daniel E. Ho, **Urvashi Khandelwal**, Alex Yu

- Formal Adjudication at the Social Security Administration
  Daniel E. Ho, Derin McLeod, **Urvashi Khandelwal**, Liza Starr, Emma Wang

INVITED TALKS   **Generalization through Memorization: Nearest Neighbor Language Models**
Berkeley NLP Seminar. November, 2019.

Sharp Nearby, Fuzzy Far Away: How Neural Language Models Use Context
Bay Area Research in NLP and ML Meetup. March, 2019.

Media Portrayals of AI
Stanford AI Lab - AI Salon. April, 2017.

Neural Text Summarization
Stanford Data Science Initiative. October, 2016.

PRESS COVERAGE   Facebook's AI speeds up natural language processing without additional training.
VentureBeat, February 19, 2020.

Helpful Neighbors.
The Batch by Andrew Ng. January 29, 2019.

Stanford policy lab explores government use of artificial intelligence.
Stanford News Service, February 28, 2019.

Khandelwal receives CRA Outstanding Undergraduate Researcher award.
Department of Computer Science - CS@Illinois News, January 23, 2015.

SKILLS

Languages: Python, C++, C, MATLAB, LaTeX
Frameworks: PyTorch, Tensorflow

WORK
EXPERIENCE

Facebook
Research Intern - Facebook AI Research   Summer, 2020
Managers: Luke Zettlemoyer, Mike Lewis

Worked on k-nearest neighbor machine translation models (kNN-MT), an extension of
nearest neighbor language models to conditional generation tasks. Through a series
of carefully designed experiments, we show that retrieving nearest neighbors from a
datastore containing translation examples can improve model generalization without
additional training or data. It can also be used to adapt translation systems to new
domains by adding domain-specific data to the datastore, and to specialize multilingual
translation systems to specific language pairs by adding language-specific data to the
datastore, all without any added training costs.

Facebook
Research Intern - Facebook AI Research   Summer, 2019
Managers: Mike Lewis, Omer Levy

Worked on k-nearest neighbor language models (kNN-LM) which involved extending
pre-trained autoregressive LMs with a module to retrieve the nearest examples, in
the pre-trained LM's representation space, from any text collection without any fur-
ther training. This technique allows the model to generalize better using not only the
LM's training data, but also by scaling up to larger corpora and adapting to different
domains, by simply varying the text collection, without any additional training.

Google
Research Intern - Google Brain   Summer-Fall, 2018
Manager: Lukasz Kaiser

Worked on large pre-trained language models in the context of language generation and
multi-task learning. Demonstrated the sample efficiency of a pre-trained decoder-only
Transformer for abstractive text summarization by concatenating the source article
and the target summary and using the decoder to process both. The model achieved
surprising performance from fine-tuning on only 3,000 training examples. In addition,
collaborated on born-again multi-task learning for fine-tuning BERT on a number of
language understanding tasks (the GLUE benchmark). A single multi-task model is
fine-tuned via knowledge distillation from several single task models. This method
consistently improves over standard single-task and multi-task training.

Facebook
Software Engineering Intern - Data and Targeting, Ads   Summer, 2014

Designed, implemented and deployed an A/B testing framework for clustering similar
users with the goal of ads targeting. The aim of the project was to quantify, analyze
and track gains and losses introduced by the development and deployment of new clus-

ters by different teams within the organization. The project included two main components: a recurring automated data extraction, processing and metrics calculation pipeline implemented in Hive and a front end interface for launching and monitoring A/B experiments and compiling the results into automatically generated visualizations. The correctness and utility of the project was verified by studying the impact of the pipeline on real users in live experiments, which led to significant increases in revenue for the specific categories that made use of them.

Google

Software Engineering Intern - Gmail Backend   Summer, 2013

Developed a performance analysis framework for the Gmail backend pipeline. The framework involved the design, implementation and testing of an anonymizer module as well as its integration with the existing Bigtable setup, where user privacy was a key concern. The module, implemented in C++, was designed to meet timing and memory constraints.

Qualcomm

Engineering Intern - Qualcomm CDMA Technologies   Summer, 2012

Designed Production Level APIs for the Camera Sensor Driver team to debug Hardware Software Interface issues. The APIs penetrated the high level Software in C++ to the lower Hardware layer code in C, using registers and I2C Reads to retrieve the information from the Sensor.

TEACHING EXPERIENCE

CS124 - From Languages to Information
Head Teaching Assistant, Stanford   Winter 2019, 2020

CS225 - Data Structures
Teaching Assistant, UIUC   Spring 2013 - Fall 2014

CS173 - Discrete Mathematics
Teaching Assistant, UIUC   Fall 2013

ECE110 - Introduction to Electrical and Computer Engineering
Teaching Assistant, UIUC   Spring 2012

SERVICE

NeuralGen 2019 - Workshop on Methods for Optimizing and Evaluating Neural Language Generation
Co-Organizer   2019

Stanford Computer Science PhD Admissions
Committee Member   2018-2019

SAIL Undergraduate Mentorship Program
Mentor   2018-2019

Stanford Women in AI
Organizer   Fall 2017

Stanford AI4ALL
NLP Research Project Mentor   Summer, 2016

Reviewer
ACL, NAACL, EMNLP, NeuralGen (Meta-reviewer), DeepGen, ACL-Student Research Workshop, KDD, RecSys