

RESEARCH AREAS

Natural language generation, with specific interests in style transfer, data-to-text generation, document/story-level generation tasks and low-resource generation.

EDUCATION

- **Language Technologies Institute, CMU** Pittsburgh, Pennsylvania
PhD in Language Technologies; **GPA:** 3.97/4.33 Sep. 2016 – Ongoing
 - **Advisor:** Eduard Hovy
 - Key Courses: *Grammars & Lexicons, Machine Translation, Language Grounding For Vision & Control, Neural Networks For NLP, Algorithms for NLP, Deep Reinforcement Learning, Structured Prediction for NLP*
- **Indian Institute of Technology, Madras (IIT-M)** Chennai, India
B.Tech/M.Tech. in Computer Science and Engineering; **CGPA:** 9.27/10 Aug. 2011 – May. 2016
 - **Advisor:** Balaraman Ravindran
 - Key Electives: *Machine Learning, NLP, Reinforcement Learning, Memory Based Reasoning, Searching & Indexing, Knowledge Representation, Distributed Algorithms, Ontologies, Graph Theory*
 - Core Courses: *Networks, Databases, Operating Systems, Automata Theory, Compilers, Algorithms*
 - **GRE:** 337/340 (**Quant:** 170/170 **Verbal:** 167/170 **Analytical Writing:** 5.5/6)
 - **TOEFL:** 118/120 (**Reading:** 30/30 **Writing:** 30/30 **Listening:** 30/30 **Speaking:** 28/30)

PUBLICATIONS

- BERTing RAMS: What and How Much does BERT Already Know About Event Arguments? – A Study on the RAMS Dataset [[ArXiv](#)]
Authors: Varun Gangal, Eduard Hovy
Accepted for **BlackBoxNLP 2020 Workshop, EMNLP 2020**
- GenAug: Data Augmentation for Finetuning Text Generators [[ArXiv](#)]
Authors: Steven Y. Feng, Varun Gangal*, Dongyeop Kang, Teruko Mitamura, Eduard Hovy*
Accepted for **Deep Learning Inside Out Workshop, EMNLP 2020**
- SCDE: Sentence Cloze Dataset with High Quality Distractors From Examinations [[PDF](#)] [[DATA](#)] [[CODE](#)]
Authors: Xiang Kong, Varun Gangal*, Eduard Hovy*
Accepted for **ACL 2020, Seattle**
- Likelihood Ratios and Generative Classifiers For Unsupervised OOD Detection In Task-Based Dialog [[PDF](#)] [[DATA](#)]
Authors: Varun Gangal, Abhinav Arora, Arash Einolghozati, Sonal Gupta
Accepted for **AAAI 2020, New York City**
- (Male, Bachelor) and (Female, Ph.D) have different connotations: Parallely Annotated Stylistic Language Dataset with Multiple Personas [[PDF](#)] [[CODE+DATA](#)] [[PRESENTATION](#)]
Authors: Dongyeop Kang, Varun Gangal, Eduard Hovy
Accepted for **EMNLP 2019, Hong Kong**
- Learning to Generate Move-by-Move Commentary for Chess Games [[PDF](#)] [[POSTER](#)]
Authors: Harsh Jhamtani, Varun Gangal*, Eduard Hovy, Graham Neubig, Taylor Berg-Kirkpatrick*
Accepted for **ACL 2018, Melbourne**
- Charmanteau: Character Embedding Models For Portmanteau Creation [[ARXIV](#)][[DEMO](#)][[CODE](#)][[DATA](#)]
Authors: Varun Gangal, Harsh Jhamtani*, Graham Neubig, Eduard Hovy, Eric Nyberg*
Accepted for **EMNLP 2017, Copenhagen**
- Detecting and Explaining Causes From Text For a Time Series Event [[ARXIV](#)]
Authors: Dongyeop Kang, Varun Gangal, Ang Lu, Zheng Chen, Eduard Hovy
Accepted for **EMNLP 2017, Copenhagen**
- Shakespearizing Modern Language Using Copy-Enriched Sequence-to-Sequence Models [[ARXIV](#)][[CODE](#)]
Authors: Harsh Jhamtani, Varun Gangal*, Eduard Hovy, Eric Nyberg*
Accepted for **EMNLP 2017 Workshop on Stylistic Variation, Copenhagen**

- HEMI: Hyperedge Majority Influence Maximization [PDF]
Authors: Varun Gangal, Balaraman Ravindran, Ramasuri Narayanam
Accepted for **The Second IJCAI Workshop on Social Influence Analysis (SocInf 2016), New York**
- Trust And Distrust Across Coalitions: Shapley Value Centrality Measures For Signed Networks [PDF]
Authors: Varun Gangal, Abhishek Narwekar, Balaraman Ravindran, Ramasuri Narayanam
Accepted for **AAAI Student Abstract 2016**
- Trust And Distrust Across Coalitions: Shapley Value Centrality Measures For Signed Networks [PDF]
Authors: Varun Gangal, Abhishek Narwekar, Balaraman Ravindran, Ramasuri Narayanam
Accepted for **NIPS 2015 Workshop on Networks In the Social And Information Sciences**

REVIEWING EXPERIENCE

- **ACL 2019, Generation Track & Machine Learning Track**
- **EMNLP 2019, Generation Track**
- **AAAI 2020, Main Track, Natural Language Processing**
- **ACL 2020, Generation Track**
- **COLING 2020, Generation Track**
- **HAMLETS 2020 WS, NEURIPS 2020**

PROFESSIONAL EXPERIENCE

- **Unsupervised OOD Detection For Task Based Dialog** Facebook Conversational AI, Menlo Park, CA
Mentors : [Sonal Gupta](#), [Arash Einolghozati](#), Abhinav Arora May – August 2019
 - Task-based dialog systems on deployment often get user inputs which aren't actually intents pertaining to any domain, such as rhetorical remarks, subjective questions and ill-specified search queries.
 - If not filtered, these inputs can wreak havoc on downstream components like slot detection. Furthermore, it is infeasible to curate training data for these "OOD" inputs. Hence, we need unsupervised approaches to detect these at test-time jointly with intent classification.
 - We explore **likelihood ratio** with a **background** likelihood as an alternative to plain likelihood. We find this to consistently improve OOD detection for multiple types of likelihood functions.
 - We propose learning a **generative classifier** and computing a marginal likelihood (ratio) for OOD detection. This outperforms approaches based on simple likelihood as well as discriminative classifiers.
 - **Accepted at AAAI 2020**
- **Centrality and Influence in Unconventional Social Networks** IBM Research, India
Mentor : [Ramasuri Narayanam](#) May – July 2015
 - Worked on centrality measures for signed networks, multiplex networks and hypergraphs.
- **Tongue Twister Generation** Snap Research, Venice, CA
Mentor : [William Brendel](#) May – August 2018
 - Worked on learning natural language models for generating tongue twisters - sentences which are semantically meaningful as well as phonetically unusual.
- **Building Ad Customer Profile Using ML (CUPRUM)** Bing Ads Team, Microsoft India
Mentor : [Prashant Rajoria](#) May – July 2014
 - Worked on building models of various aspects of Bing advertisers such as spamminess, malware and serveability using supervised machine learning methods, based on features from their ad pages.

TEACHING EXPERIENCE

- **Teaching Assistant - Neural Networks For NLP, Spring 2018** CMU
- **Teaching Assistant - Grammars & Lexicons, Fall 2017** CMU
- **Teaching Assistant - Machine Learning MOOC, Spring 2016** NPTEL & IIT Madras
- **Teaching Assistant - Reinforcement Learning, Spring 2016** IIT Madras