

# TUHIN CHAKRABARTY

tuhinjubcse.github.io

Upper West Side, New York, NY 10027  
(+1)347-630-6405 ♦ tc2896@columbia.edu

## EDUCATION

---

### Columbia University, New York

PhD in Computer Science (Major in Natural Language Processing)  
Advisor : **Smaranda Muresan**

*Present*

GPA : 3.91

### Columbia University, New York

Master of Science (Major in Natural Language Processing)  
Department of Computer Science

*August 2017 - 2019*

GPA : 3.91

### Jadavpur University, Kolkata, India

Bachelor of Engineering in Computer Science

*July 2010 - June 2014*

GPA : 3.85

## PUBLICATIONS

---

### ARCHIVAL CONFERENCE PAPERS

---

#### 1. **R<sup>3</sup>: Reverse, Retrieve, and Rank for Sarcasm Generation using CommonSense Knowledge**

**Tuhin Chakrabarty**, Debanjan Ghosh and Smaranda Muresan and Nanyun Peng

Proceedings of 2020 Annual Conference of the Association for Computational Linguistics (Long)

#### 2. **DeSePtion: Dual Sequence Prediction and Adversarial Examples for Improved Fact-Checking**

Christopher Hidey, **Tuhin Chakrabarty**, Tariq Alhindi, Siddharth Varia ,Kriste Krstovski, Mona Diab and Smaranda Muresan

Proceedings of 2020 Annual Conference of the Association for Computational Linguistics (Long)

#### 3. **AMPERSAND: Argument Mining for PERSuAsive oNline Discussions**

**Tuhin Chakrabarty**, Christopher Hidey, Smaranda Muresan, Kathy McKeown and Alyssa Hwang  
Proceedings of 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong

*\* Accepted as a poster at Natural Language, Dialog and Speech (NDS) Symposium , New York Academy Of Sciences (Long)*

#### 4. **IMHO Fine Tuning Improves Claim Detection**

**Tuhin Chakrabarty**, Christopher Hidey, Kathy McKeown

Proceedings of 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics, (NAACL) (Short)

## WORKSHOP PAPERS

---

#### 1. **Discourse Relation Prediction: Revisiting Word Pairs with Convolutional Networks**

Siddharth Varia, Christopher Hidey, **Tuhin Chakrabarty**

Proceedings of the 20th Annual meeting of Special Interest Group on Discourse and Dialogue (SIG-DIAL), 2019 (Long)

## 2. Pay “Attention” to your Context when Classifying Abusive Language

**Tuhin Chakrabarty**, Kilol Gupta and Smaranda Muresan

Proceedings of 3rd Abusive Language Workshop, ACL 2019 (Long)

## 3. The Answer is Language Model Fine-tuning

**Tuhin Chakrabarty** and Smaranda Muresan

Proceedings of the 13th International Workshop on Semantic Evaluation, NAACL-HLT, Minneapolis, NAACL 2019 (Short)

## 4. Robust Document Retrieval and Individual Evidence Modeling for Fact Extraction and Verification

**Tuhin Chakrabarty**, Tariq Alihindi and Smaranda Muresan

Proceedings of 1st Workshop on Fact Extraction and Verification, EMNLP 2018 (Short)

## RESEARCH EXPERIENCE

---

### Research Assistant

*August 2017 - Present*

**Advisor: Smaranda Muresan**

**Columbia University**

- Figurative Language Generation
- Learning from less labeled data for improved Dialogue and Discourse Understanding
- Argumentation (Structures and Semantics)
- Adversarial Attacks in Veracity/ Fact checking models

### Research Assistant

*August 2019 - Present*

**Advisor: Nanyun Peng**

**Information Sciences Institute, University Of Southern California**

- Creative text generation.

## WORK EXPERIENCE

---

### Summer Intern PhD, IBM Research

*April 2020 - July 2020, YorkTown Heights, NY*

Generative models for Question Answering.

### Applied Scientist Intern, Amazon.com

*May 2018 - August 2018, Boston, MA*

Improving classification results of Alexa NLU utterances via diverse data generation using Weighted Finite State Transducers.

### Machine Learning Engineer, UBER

*Jan 2016 - August 2017, Hyderabad, India*

Improving transparency by preventing partner and credit card fraud in emerging markets using machine learning techniques.

### Software Engineer, Amazon.com

*July 2014 - Jan 2016, Hyderabad, India*

Backend Engineer in Customer Returns and Refunds team responsible for handling resolutions such as Return, Exchange, Replacement.

## RELEVANT COURSEWORK

---

Machine Learning, Natural Language Processing, Deep Learning, NLP in Context: Computational Models of Social Meaning, Analysis Of Algorithm, Fundamentals of Speech Recognition

## AWARDS AND FELLOWSHIPS

---

- Graduate Research Fellowship at Columbia University (Includes full tuition waiver) awarded to 3 out of 250 students.

- NAACL 2019 Student Volunteer Award.
- Google Travel Award for ACL 2019.
- Amazon Student Bursary Award for EMNLP 2019.

## PROFESSIONAL ACTIVITIES

---

- Program Committee at ACL 2020, EMNLP 2020 , AACL 2020, \*SEM 2020
- Program Committee at 1st,2nd and 3rd Workshop on Fact Extraction and Verification
- Program Committee at 4th Abusive Language Workshop.
- Program Committee at 6th Workshop on Noisy User-Generated Text (W-NUT)

## REFERENCES

---

- **Dr. Smaranda Muresan**  
Research Scientist, Data Science Institute  
Adjunct Associate Professor  
Columbia University, New York, NY 10027  
Email: sm761@columbia.edu
- **Dr. Kathy McKeown**  
Henry and Gertrude Rothschild Professor of Computer Science  
Columbia University, New York, NY 10027  
Email: kathy.cs@columbia.edu
- **Dr. Nanyun Peng**  
Research Assistant Professor, Department of Computer Science  
Computer Scientist, Information Sciences Institute  
University of Southern California  
Email: npeng@isi.edu
- **Dr. Kriste Krstovski**  
Associate Research Scientist, Data Science Institute  
Adjunct Assistant Professor, Columbia Business School  
Email: kriste.krstovski@columbia.edu