

# SHUBHANSHU MISHRA

mishra@shubhanshu.com · shubhanshu.com

## EDUCATION

### University of Illinois at Urbana-Champaign

Doctor of Philosophy (Ph.D.) Library and Information Science

Aug 2013 - May 2020

### Indian Institute of Technology Kharagpur

Bachelors and Masters in Science Mathematics and Computing

Jul 2007 - May 2012

## EXPERIENCE

### Twitter, Inc.

Machine Learning Researcher, Content Understanding Research

Aug 2019 - Present

Intern, Content Understanding & Applied Deep Learning

Jul 2018 - Sep 2018

### Citrix, Bangalore

Software Engineer, NetScaler INFRA Team

Jul 2012 - Jul 2013

### Barclays Capital, Singapore

Global Technology Analyst, Commodities

May 2011 - Jul 2011

### Global Venture Lab, Finland

Lead Web Developer

Dec 2009 - Jan 2010

## ACADEMIC HONORS & AWARDS

Impact Recognition Award - CSCW

Oct 2021

Best Poster Award - UIUC Student Poster Session

Mar 2020

Best student paper award - ASIST SIGMET Workshop

Nov 2018

Graduate Teacher Certificate

May 2018

University of Illinois GIS Day Runner-up (Research Quality)

Nov 2017

Kishore Vaigyanik Protsahan Yojana Scholar

2007-2012

3rd rank in Regional Mathematics Olympiad, Uttar Pradesh, India

Dec 2006

## TEACHING

Tutorial presenter, ACM Hypertext

Sep 2019

*Tutorial on hands on advanced machine learning for information extraction from tweets tasks, data, and open source tools.*

Co-instructor - Network Analysis

Spring 2018

*Teaching network analysis applications on social media data using python.*

Teaching Assistant - Network Analysis,

Summer 2017

*As part of the Information Science and Engineering Summer Program organized for Chinese students, sponsored by Global Education and Training (GET)*

Teaching Assistant - Foundations of Information Processing

Spring 2017

Co-instructor - Data Mining Applications

Fall 2016

***Listed in Teachers Ranked as Excellent By Their Students!***

## PUBLICATIONS

- [1] Vivek Kulkarni, Shubhanshu Mishra, and Aria Haghighi. “{LMSOC}: An Approach for Socially Sensitive Pretraining”. In: *Findings of the Association for Computational Linguistics: EMNLP 2021*. Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 2967–2975.
- [2] Shubhanshu Mishra and Aria Haghighi. “Improved Multilingual Language Model Pretraining for Social Media Text via Translation Pair Prediction”. In: *Proceedings of the Seventh Workshop on Noisy User-generated Text (W-NUT 2021)*. Online: Association for Computational Linguistics, Nov. 2021, pp. 381–388.
- [3] Sudhanshu Mishra, Shivangi Prasad, and Shubhanshu Mishra. “Exploring Multi-Task Multi-Lingual Learning of Transformer Models for Hate Speech and Offensive Speech Identification in Social Media”. In: *SN Computer Science* 2.2 (Apr. 2021), p. 72. ISSN: 2662-995X. DOI: 10.1007/s42979-021-00455-5.
- [4] Kyra Yee, Uthaipon Tantipongpipat, and Shubhanshu Mishra. “Image Cropping on Twitter: Fairness Metrics, their Limitations, and the Importance of Representation, Design, and Agency”. In: *Proceedings of the ACM on Human-Computer Interaction* 5.CSCW2 (Oct. 2021), pp. 1–24. ISSN: 2573-0142. DOI: 10.1145/3479594.
- [5] Kanyao Han et al. “WikiCSSH: Extracting Computer Science Subject Headings from Wikipedia”. In: *Workshop on Scientific Knowledge Graphs (SKG 2020)*. 2020.

- [6] Shubhanshu Mishra. “Information Extraction from Digital Social Trace Data with Applications to Social Media and Scholarly Communication Data”. In: *ACM SIGIR Forum* 54.1 (2020).
- [7] Shubhanshu Mishra. “Non-neural Structured Prediction for Event Detection from News in Indian Languages”. In: *Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation*. Ed. by P Mehta et al. Hyderabad, India: CEUR Workshop Proceedings, CEUR-WS.org, 2020.
- [8] Shubhanshu Mishra and Daniel Collier. “A Framework for Generating Annotated Social Media Corpora with Demographics, Stance, Civility, and Topicality”. In: *SSRN Electronic Journal* (2020). ISSN: 1556-5068. DOI: 10.2139/ssrn.3757554.
- [9] Shubhanshu Mishra and Sudhanshu Mishra. “Scubed at 3C task A - A simple baseline for citation context purpose classification”. In: *Proceedings of the 8th International Workshop on Mining Scientific Publications*. Wuhan, China: Association for Computational Linguistics, 2020, pp. 59–64.
- [10] Shubhanshu Mishra and Sudhanshu Mishra. “Scubed at 3C task B - A simple baseline for citation context influence classification”. In: *Proceedings of the 8th International Workshop on Mining Scientific Publications*. Wuhan, China: Association for Computational Linguistics, 2020, pp. 65–70.
- [11] Sudhanshu Mishra, Shivangi Prasad, and Shubhanshu Mishra. “Multilingual Joint Fine-tuning of Transformer models for identifying Trolling, Aggression and Cyberbullying at TRAC 2020”. In: *Proceedings of the Second Workshop on Trolling, Aggression and Cyberbullying*. Marseille, France: European Language Resources Association (ELRA), 2020, pp. 120–125.
- [12] Nikolaus Nova Parulian et al. “Effectiveness of the Execution and Prevention of Metric-Based Adversarial Attacks on Social Network Data †”. In: *Information* 11.6 (June 2020), p. 306. ISSN: 2078-2489. DOI: 10.3390/info11060306.
- [13] Mihai Valentin Avram et al. “Adversarial perturbations to manipulate the perception of power and influence in networks”. In: *2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. Vancouver, Canada: IEEE, 2019, pp. 986–993.
- [14] Daniel Collier et al. “Who is Most Likely to Oppose Federal Tuition-Free College Policies? Investigating Variable Interactions of Sentiments to America’s College Promise”. In: *SSRN Electronic Journal* (2019). ISSN: 1556-5068. DOI: 10.2139/ssrn.3423054.
- [15] Daniel A. Collier et al. “Americans ‘support’ the idea of tuition-free college: an exploration of sentiment and political identity signals otherwise”. In: *Journal of Further and Higher Education* 43.3 (Mar. 2019), pp. 347–362. ISSN: 0309-877X. DOI: 10.1080/0309877X.2017.1361516.
- [16] Shubhanshu Mishra. “Multi-dataset-multi-task Neural Sequence Tagging for Information Extraction from Tweets”. In: *Proceedings of the 30th ACM Conference on Hypertext and Social Media - HT ’19*. New York, New York, USA: ACM Press, 2019, pp. 283–284. ISBN: 9781450368858. DOI: 10.1145/3342220.3344929.
- [17] Shubhanshu Mishra and Jana Diesner. “Capturing Signals of Enthusiasm and Support Towards Social Issues from Twitter”. In: *Proceedings of the 5th International Workshop on Social Media World Sensors - SIdEWaYS’19*. New York, New York, USA: ACM Press, 2019, pp. 19–24. ISBN: 9781450369039. DOI: 10.1145/3345645.3351104.
- [18] Shubhanshu Mishra and Sudhanshu Mishra. “3Idiots at HASOC 2019: Fine-tuning Transformer Neural Networks for Hate Speech Identification in Indo-European Languages”. In: *Proceedings of the 11th annual meeting of the Forum for Information Retrieval Evaluation*. Kolkata, India, 2019, pp. 208–213.
- [19] Shubhanshu Mishra and Jana Diesner. “Detecting the Correlation between Sentiment and User-level as well as Text-Level Meta-data from Benchmark Corpora”. In: *Proceedings of the 29th on Hypertext and Social Media - HT ’18*. New York, New York, USA: ACM Press, 2018, pp. 2–10. ISBN: 9781450354271. DOI: 10.1145/3209542.3209562.
- [20] Shubhanshu Mishra et al. “Expertise as an aspect of author contributions”. In: *Metrics 2018: Workshop on Informetric and Scientometric Research (SIG/MET)*. Vancouver, 2018.
- [21] Shubhanshu Mishra et al. “Self-citation is the hallmark of productive authors, of any gender”. In: *PLoS ONE* 13.9 (Sept. 2018). Ed. by Niels O. Schiller, e0195773. ISSN: 19326203. DOI: 10.1371/journal.pone.0195773.
- [22] Aseel Addawood et al. “Developing an Information Source Lexicon”. In: *Prioritising Online Content workshop co-located at NIPS*. 2017.
- [23] Shubhanshu Mishra. “SCTG: Social Communications Temporal Graph – A novel approach to visualize temporal communication graphs from social data”. In: *UIUC Data Science Day*. Oct. 2017.

- [24] Shubhanshu Mishra and Jana Diesner. “Semi-supervised Named Entity Recognition in noisy-text”. In: *Proceedings of the 2nd Workshop on Noisy User-generated Text (WNUT)*. Osaka, Japan: The COLING 2016 Organizing Committee, 2016, pp. 203–212.
- [25] Shubhanshu Mishra and Vetle I. Torvik. “Quantifying Conceptual Novelty in the Biomedical Literature.” In: *D-Lib magazine : the magazine of the Digital Library Forum* 22.9-10 (Sept. 2016). ISSN: 1082-9873. DOI: 10.1045/september2016-mishra.
- [26] Shubhanshu Mishra et al. “Sentiment Analysis with Incremental Human-in-the-Loop Learning and Lexical Resource Customization”. In: *Proceedings of the 26th ACM Conference on Hypertext & Social Media - HT '15*. New York, New York, USA: ACM Press, 2015, pp. 323–325. ISBN: 9781450333955. DOI: 10.1145/2700171.2791022.
- [27] Shubhanshu Mishra et al. “Enthusiasm and support: alternative sentiment classification for social movements on social media”. In: *Proceedings of the 2014 ACM conference on Web science - WebSci '14*. Bloomington, Indiana, USA: ACM Press, June 2014, pp. 261–262. ISBN: 9781450326223. DOI: 10.1145/2615569.2615667.