

Sunil Kumar Sahu

Website: <https://sunilitggu.github.io>

Email: sunilitggu@gmail.com

LinkedIn: [sunil-sahu-202b133b](https://www.linkedin.com/in/sunil-sahu-202b133b)

GitHub: github.com/sunilitggu/

MobileNo: +971 585347249

EDUCATION

IIT Guwahati Ph.D., Computer Science and Engineering	Guwahati, India 2013–2017
University of Hyderabad M.Tech., Computer Science, GPA: 8.03/10.00	Hyderabad, India 2011–2013
Guru Ghasidas University B.E., Computer Science and Engineering, GPA: 6.5/10.00	Bilaspur, India 2005–2009

RESEARCH INTEREST

General Areas: Natural Language Processing, Clinical Data Mining, Machine Learning

Specific Areas: Information Extraction, Machine Translation, Question Answering

EXPERIENCE

IIAI, Group 42 Applied Scientist	Abu Dhabi, UAE March 2019–Currently
NaCTeM, University of Manchester Postdoctoral Research Associate	Manchester, UK Sep 2017–March 2019
Xerox Research Centre Research Intern	Bangalore, India July–October 2016
GVK Bioscience Research Intern	Hyderabad, India Feb–May 2016

SCHOLARSHIPS AND AWARDS

- 5th rank in WMT-2020 shared task Similar Language Translation 2020
- 3rd rank in N2C2 shared task track 2 Medical Event Extraction 2018
- Travel grant to attend ACL-2016 from Google and Microsoft 2016
- MHRD fellowship to do PhD 2013–2017
- MHRD fellowship to do MTech 2011–2013
- State fellowship to do BE 2005–2009

TALKS

- *AI for NLP*, Marhaba Induction Program for G42 employees and students of UAE, G42 Abu Dhabi, UAE 2021
- *Deep Learning Techniques and its Applications*, TEQIP-III sponsored 5-day course, NIT Raipur 2017

- *Introduction of TensorFlow*, Intelligence System Course, IIT Guwahati 2017
- *Neural Probabilistic Language Model*, ML Group, IIT Guwahati 2015
- *Recent Trends in Machine Learning*, Research Conclave-2015, IIT Guwahati 2015

TEACHING

- **Teaching Assistant** at IIT Guwahati Spring 2016
Intelligent Systems and Interfaces (CS565)
- **Teaching Assistant** at University of Hyderabad Spring 2013
Natural Language Processing (AI482)
- **Teaching Assistant** at University of Hyderabad Spring 2012
Data Mining (CS422)

PUBLICATIONS

- [1] S. K. Sahu, D. Thomas, B. Chiu, N. Sengupta, and M. Mahdy, "Relation extraction with self-determined graph convolutional network", in *Proceedings of the CIKM*, 2020.
- [2] B. Chiu, S. K. Sahu, N. Sengupta, D. Thomas, and M. Mahdy, "Attending to inter-sentential features in neural text classification", in *Proceedings of the SIGIR*, 2020.
- [3] B. Chiu, S. K. Sahu, D. Thomas, N. Sengupta, and M. Mahdy, "Autoencoding keyword correlation graph for document clustering", in *Proceedings of the ACL*, 2020.
- [4] S. K. Sahu, F. Christopoulou, M. Miwa, and S. Ananiadou, "Inter-sentence relation extraction with document-level graph convolutional neural network", in *Proceedings of the ACL*, 2019.
- [5] F. Christopoulou, T. T. Tran, S. K. Sahu, M. Miwa, and S. Ananiadou, "Adverse drug events and medication relation extraction in electronic health records with ensemble deep learning methods", *Journal of the American Medical Informatics Association*, 2019.
- [6] S. K. Sahu and A. Anand, "Drug-drug interaction extraction from biomedical texts using long short-term memory network", *Journal of Biomedical Informatics*, 2018.
- [7] S. K. Sahu and A. Anand, "What matters in a transferable neural network model for relation classification in the biomedical domain?", *Artificial Intelligence in Medicine*, 2018.
- [8] D. Raj, S. K. Sahu, and A. Anand, "Learning local and global contexts using a convolutional recurrent network model for relation classification in biomedical text", in *Proceedings of the CoNLL*, 2017.
- [9] K. Chawla, S. K. Sahu, and A. Anand, "Investigating how well contextual features are captured by bi-directional recurrent neural network models", in *Proceedings of the ICON*, 2017.
- [10] R. V S S Patchigolla, S. K. Sahu, and A. Anand, "Biomedical event trigger identification using bidirectional recurrent neural network based models", in *Proceedings of the BioNLP*, 2017.
- [11] S. K. Sahu and A. Anand, "Recurrent neural network models for disease name recognition using domain invariant features", in *Proceedings of the ACL*, 2016.
- [12] S. K. Sahu, A. Anand, K. Oruganty, and M. Gattu, "Relation extraction from clinical texts using domain invariant convolutional neural network", in *Proceedings of the BioNLP*, 2016.
- [13] M. TH, S. K. Sahu, and A. Anand, "Evaluating distributed word representations for capturing semantics of biomedical concepts", in *Proceedings of BioNLP*, 2015.

EXTRACURRICULAR ACTIVITIES

- Program Committee
AAAI-2022, EMNLP-2021, ACL-2021, NAACL-2021, AACL-2021, ACL-2020, EMNLP-2020, AACL-2020, NAACL-2019, ACL-2019, EMNLP-2019, EMNLP-2018
- Reviewer
Journal of Biomedical Informatics, IEEE/ACM Transactions on Computational Biology and Bioinformatics

SKILLS

- **NLP:** NLTK, CoreNLP, OpenNLP, BANNER, LingPipe, UMLS, Metamap
- **ML:** Pytorch, Tensorflow, Fairseq, Transformers, Scikit-learn
- **Database:** MySQL
- **Languages:** Python, Java, C
- **Deployment Tools:** Docker, Flask API

REFERENCES

- Prof. Sophia Ananiadou, *Professor* School of Computer Science, University of Manchester, United Kingdom email id: sophia.ananiadou@manchester.ac.uk
- Dr. Ashish Anand, *Associate Professor*, Department of Computer Science, Indian Institute of Technology Guwahati, India, email id: anand.ashish@iitg.ernet.in