Sunil Kumar Sahu

Website: https://sunilitggu.github.io
Email: sunilitggu@gmail.com
LinkedIn: sunil-sahu-202b133b
GitHub: github.com/sunilitggu/

EDUCATION

IIT Guwahati Guwahati, India
Ph.D. in NLP, Advisor: Dr. Ashish Anand 2013–2017
  − Thesis: "Neural architectures for NER and RE in biomedical and clinical texts"

University of Hyderabad Hyderabad, India
M.Tech. in CSE, GPA: 8.03/10.00 2011–2013

Guru Ghasidas University Bilaspur, India
B.E. in CSE, GPA: 6.5/10.00 2005–2009

EXPERIENCE

Inception Institute of Artificial Intelligence Abu Dhabi, UAE
Research Associate March 2019–Currently

University of Manchester Manchester, UK
Post Doctoral Fellow Sep 2017–March 2019

Xerox Research Centre Bangalore, India
Research Intern Summer 2016

GVK Bioscience Hyderabad, India
Research Intern Winter 2016

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)

SKILLS

- **NLP:** NLTK, CoreNLP, OpenNLP, BANNER, LingPipe, UMLS, Metamap
- **ML:** Tensorflow, Pytorch, Fairseq, Scikit-learn
- **Database:** MySQL

LANGUAGES

- **Python:** Advance Level
- **Java:** Intermediate Level
- **C++:** Intermediate Level

Publications

[1] 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 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, ser. SIGIR '20, Virtual Event, China: Association for Computing Machinery, 2020, pp. 1685–1688, ISBN: 9781450380164.

[2] B. Chiu, S. K. Sahu, D. Thomas, N. Sengupta, and M. Mahdy, "Autoencoding keyword correlation graph for document clustering", in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020)*, Online: Association for Computational Linguistics, Jul. 2020, pp. 3974–3981.

[3] S. K. Sahu and A. Anand, "Unified neural architecture for drug, disease, and clinical entity recognition", in *Deep Learning Techniques for Biomedical and Health Informatics*, B. Agarwal, V. E. Balas, L. C. Jain, R. C. Poonia, and Manisha, Eds., Academic Press, 2020, pp. 1–19, ISBN: 978-0-12-819061-6.

[4] S. K. Sahu, D. Thomas, B. Chiu, N. Sengupta, and M. Mahdy, "Relation extraction with self-determined graph convolutional network", in *Proceedings of the 29th ACM International Conference on Information and Knowledge Management*, ser. CIKM '20, Association for Computing Machinery, 2020.

[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*, vol. 27, no. 1, pp. 39–46, Aug. 2019, ISSN: 1527-974X.

[6] 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 57th Annual Meeting of the Association for Computational Linguistics (ACL 2019)*, Florence, Italy: Association for Computational Linguistics, Jul. 2019, pp. 4309–4316.

[7] S. K. Sahu and A. Anand, "Drug-drug interaction extraction from biomedical texts using long short-term memory network", *Journal of Biomedical Informatics*, vol. 86, pp. 15–24, 2018, ISSN: 1532-0464.

[8] 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*, vol. 87, pp. 60–66, 2018, ISSN: 0933-3657.

[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 14th International Conference on Natural Language Processing (ICON 2017)*, Kolkata, India: NLP Association of India, Dec. 2017, pp. 273–282.

[10] D. Raj, S. 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 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, Vancouver, Canada: Association for Computational Linguistics, Aug. 2017, pp. 311–321.

[11] R. V S S Patchigolla, S. Sahu, and A. Anand, "Biomedical event trigger identification using bidirectional recurrent neural network based models", in *16th Workshop on Biomedical Natural Language Processing (BioNLP 2017)*, Vancouver, Canada, Association for Computational Linguistics, Aug. 2017, pp. 316–321.

[12] S. Sahu and A. Anand, "Recurrent neural network models for disease name recognition using domain invariant features", in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL 2016)*, Berlin, Germany: Association for Computational Linguistics, Aug. 2016, pp. 2216–2225.

[13] S. Sahu, A. Anand, K. Oruganty, and M. Gattu, "Relation extraction from clinical texts using domain invariant convolutional neural network", in *Proceedings of the 15th Workshop on Biomedical Natural Language Processing (BioNLP 2016)*, Berlin, Germany: Association for Computational Linguistics, Aug. 2016, pp. 206–215.

[14] M. TH, S. Sahu, and A. Anand, "Evaluating distributed word representations for capturing semantics of biomedical concepts", in *Proceedings of 14th Workshop on Biomedical Natural Language Processing (BioNLP 2015)*, Beijing, China: Association for Computational Linguistics, Jul. 2015, pp. 158–163.

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

- *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

Extracurricular Activities

- Program Committee
  AAAI-2021, ACL-2020, EMNLP-2020, AAAI-2020, AACL-2020, NAACL-2019, ACL-2019, EMNLP-2019, EMNLP-2018

- Reviewer
  Journal of Biomedical Informatics, IEEE/ACM Transactions on Computational Biology and Bioinformatics

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