Sunil Sahu

Bio Data

Flat No 20, 55 Hathersage Road
Manchester, UK
✆ (+44) 7404316373
✉ sunil.sahu@manchester.ac.uk
🖰 https://sunilitggu.github.io/

Research Interests

Natural Language Processing, Machine Learning, Deep Learning.

Education

2013-2017 **PhD (Computer Science and Engineering)**, *Department of Computer Science and Engineering, Indian Institute of Technology Guwahati, Assam, India*.

2011–2013 **MTech (Computer Science)**, *School of Computer and Information Science, University of Hyderabad, Hyderabad, Andhra Pradesh, India*.

2005–2009 **BE (Computer Science and Engineering)**, *Department of Computer Science and Engineering, Guru Ghasidash Vishwavidyalaya Bilaspur, Chhattisgarh, India*.

Professional Experiences

Oct-2017 to Present **Postdoctoral Researcher**, *National Centre for Text Mining*, **The University of Manchester**, Manchester, UK, `Project: Cross-sentence relation extraction using deep neural network` .

Aug-2013 to Sep-2017 **Graduate Student Researcher**, *Machine Learning and Data Mining*, **Indian Institute of Technology**, Guwahati, Assam, India, `My research focuses on information extraction from unstructured text` .

July-Oct.2016 **Research Intern**, *Graph and Text Analytics Group*, **Xerox Research Center India**, Bangalore, Karnataka, India, `Project : Phrase embedding for biomedical text using deep learning techniques` .

Feb-Apr.2016 **Research Intern**, *Text Mining and Analitics Group*, **GVKBioscience Private Limited**, Hyderabad, Telangana, India, `Project : Relation extraction from clinical texts using domain invariant convolutional neural network`.

June-July.2008 **Engineering Intern**, *Software Engineering Team*, **Orbit Technology**, Hyderabad, Andhra Pradesh, India, .

June-July.2007 **Engineering Intern**, *Information Technology Department*, **Bharat Aluminium Company Ltd. (BALCO)**, Korba, Chhattisgarh, India, .

Teaching Experiences

2017 **Deep Learning Techniques and its Applications**, *Course Instructor*, NIT Raipur, India, A short term course.

2016 **Intelligent Systems and Interfaces**, *Teaching Assistant*, IIT Guwahati.

2013 **Natural Language Processing**, *Teaching Assistant*, University of Hyderabad.

2012 **Data Mining**, *Teaching Assistant*, University of Hyderabad.

Technical Skills

Python, C, C++, Java, tensorflow, theano, scikit-learn, NLTK, CoreNLP, OpenNLP, MySQL, BANNER, LingPipe, UMLS, Metamap.

Awards and Honors

2 Sep 2018 **Our team achieved 3rd rank in n2c2 shared task 2018 track 2.**

6-12 Aug 2016 **Selected for the grant to attend ACL-2016 Conference from Google Research, Microsoft Research India and ACM-India/IARCS .**

2013-17 **MHRD fellowship to do PhD.**

2011-13 **MHRD fellowship to do MTech.**

Recent Publications

2018 **Drug drug interaction extraction from biomedical text using long short-term memory network**, *Sunil Kumar Sahu, Ashish Anand*, **Journal of Biomedical Informatics 86, 15-24**.

2018 **What matters in a transferable neural network model for relation classification in the biomedical domain?**, *Sunil Kumar Sahu, Ashish Anand*, **Artificial Intelligence in Medicine 87, 60-66**.

2017 **Learning local and global contexts using a convolutional recurrent network model for relation classification in biomedical text**, *Desh Raj, Sunil Kumar Sahu, Ashish Anand*, **CoNLL-2017**, Vancouver, Canada.

2017 **Biomedical Event Trigger Identification Using Recurrent Neural Network**, *Rahul PVS, Sunil Kumar Sahu, Ashish Anand*, **BioNLP at ACL-2017**, Vancouver, Canada.

2017 **Investigating how well contextual features are captured by Bi-directional Recurrent Neural Network Models**, *Kushal Kumar, Sunil Kumar Sahu, Ashish Anand*, **ICON-2017**, Kolkata, India.

2016 **Relation extraction from clinical texts using domain invariant convolutional neural network**, *Sunil Kumar Sahu, Ashish Anand, Krishna Dev Orungati, Mahanandeeshwar Gattu*, **BioNLP at ACL-2016**, Berlin, Germany.

2016 **Recurrent neural network models for disease name recognition using domain invariant features**, *Sunil Kumar Sahu, Ashish Anand*, **ACL-2016**, Berlin, Germany.

2016 **Predicting online doctor ratings from user reviews using convolutional neural networks**, *Ranti Dev Sharma, Samarth Tripathi, Sunil Kumar Sahu, Sudhanshu Mittal and Ashish Anand*, **ICMLC-2016**, Osaka, Japan.

2015 **Evaluation of distributed word representation for capturing semantics of biomedical text**, *Muneeb Th, Sunil Kumar Sahu, Ashish Anand*, **BioNLP at ACL-2015**, Beijing, China.

Unpublished Works

2018 **Siamese neural networks with random forest for detecting duplicate question pairs**, *Ameya Godbole, Aman Dalmia, Sunil Kumar Sahu*, https://arxiv.org/abs/1801.07288.

2017 **Unified neural architecture for drug, disease and clinical entity recognition**, *Sunil Kumar Sahu, Ashish Anand*, https://arxiv.org/abs/1708.03447.

2014 **Named entity recognition in tweets using conditional random field (CRF) and domain specific features.**

2014 **Object recognition in images through deep convolutional neural network (CNN) .**

2013 **Digit recognition system in speech signal using hidden markov model (HMM) .**

2013 **Language identification system for small (ISCII coded) text corpus using different linear non-linear classifier**.

Conference/Workshop/Summer School Participations:

2016 **ACL-2016**, *Humboldt University, Berlin, Germany*, August 7-12, Thanks to Google Research India, Microsoft Research India and ACM-India/IARCS for entire traveling and other expenses support.

2015 **ACL-2015**, *China National Convention Center Beijing, China* , July 30 to Aug 1, Thanks to IIT Guwahati for entire traveling and other expenses support.

2015 **Summer School on Machine Learning**, *Microsoft Research and IISc Bangalore (Karnataka)*, 15 June to 26 June, Thanks to Microsoft Research India for entire traveling and other expenses support.

2015 **Information and Communication Technology in Healthcare: Challenges and Promises**, *IIT Guwahati (Assam)*, 25 March.

2014 **National Workshop on GPU Programming and Application**, *IIT Guwahati (Assam)*, 12 Sep to 14 Sep.

2008 **Information Security : Practice and Challenges**, *ITGGU Bilaspur (CG)*, 9 Feb to 10 Feb.

Certified Courses

2014 **Machine Learning Course**, *by Prof. Andrew Ng, Coursera*, Sep to Dec 2014, Course Link .

Reviewer

Conferences:, *EMNLP-2018, NAACL-2019*.

Journals:, *Journal of Biomedical Informatics, IEEE/ACM Transactions on Computational Biology and Bioinformatics*.

Research Talks:

2017 **Intelligence System Course**, *Indian Institute Technology Guwahati*, **Introduction of TensorFlow**.

2015 **ML Group**, *Indian Institute Technology Guwahati*, **Neural Probabilistic Language Model**.

2015 **Techniche-2015**, *Indian Institute Technology Guwahati*, **Evaluating distributed word representations for capturing semantics of biomedical concepts**.

2015 **Research Conclave-2015**, *Indian Institute Technology Guwahati*, **Recent Trends in Machine Learning**.

Languages

First **Hindi**.

Second **English**.

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.