

Sunil Sahu

Bio Data

Room No S5 Brahmaputra Hostel
IIT Guwahati, Assam, India

☎ (+44) 7404316373

✉ sunil.sahu@manchester.ac.uk

📄 <https://sunilitggu.github.io/>



Research Interests

Natural Language Processing, Machine Learning, Deep Learning.

Education

- 2013-2017 **PhD**, *Department of Computer Science and Engineering, Indian Institute of Technology Guwahati, Assam, India*, Thesis Submitted.
- 2011-2013 **M Tech (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 Vishvavidyalay Bilaspur, Chhattisgarh, India*.

Professional Experiences

- Oct-2017 to Present **Research Associate**, *National Centre for Text Mining, The University of Manchester, Manchester, UK*, Project: Metabolic reaction extraction from biomedical text using deep neural network.
- July-Oct.2016 **Research Internship**, *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 Internship**, *Text Mining and Analytics Group, GVKBioscience Private Limited, Hyderabad, Telangana, India*, Project : Relation extraction from clinical texts using domain invariant convolutional neural network.

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*.
- 2015 **Operating System and Lab**, *Teaching Assistant, IIT Guwahati*.
- 2014 **Programming Language**, *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.

Scholarships/Grants

- 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 M Tech.**

Recent Works (Published and Submitted)

- Under Review **Unified Neural Architecture for Drug, Disease and Clinical Entity Recognition**, *Sunil Kumar Sahu, Ashish Anand*, **AI in Medicine Journal**, <https://arxiv.org/abs/1708.03447>.
- Minor Revision **What matters in a transferable neural network model for relation classification in the biomedical domain?**, *Sunil Kumar Sahu, Ashish Anand*, **AI in Medicine Journal**, <https://arxiv.org/abs/1708.03446>.
- 2017 **Investigating how well contextual features are captured by Bi-directional Recurrent Neural Network Models**, *Kushal Kumar, Sunil Kumar Sahu, Ashish Anand*, **ICON-2017**, <https://arxiv.org/abs/1709.00659>.
- 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**, <http://aclweb.org/anthology/K/K17/K17-1032.pdf>.
- 2017 **Biomedical Event Trigger Identification Using Recurrent Neural Network**, *Rahul PVS, Sunil Kumar Sahu, Ashish Anand*, **BioNLP at ACL-2017**, <http://www.aclweb.org/anthology/W/W17/W17-2340.pdf>.
- Minor Revision **Drug Drug Interaction Extraction from Biomedical Text Using Long Short Term Memory Network**, *Sunil Kumar Sahu, Ashish Anand*, **Journal of Biomedical Informatics**, <https://arxiv.org/abs/1701.08303>.
- 2016 **Relation extraction from clinical texts using domain invariant convolutional neural network**, *Sunil Kumar Sahu, Ashish Anand, Krishna Dev Orungati, Mahanandeeswar Gattu*, **BioNLP at ACL-2016**, Berlin, Germany, <https://aclweb.org/anthology/W/W16/W16-2928.pdf>.
- 2016 **Recurrent neural network models for disease name recognition using domain invariant features**, *Sunil Kumar Sahu, Ashish Anand*, **ACL-2016**, Berlin, Germany, <http://aclweb.org/anthology/P/P16/P16-1209.pdf>.
- 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, <http://www.ijmlc.org/vol6/590-L017.pdf>.
- 2015 **Evaluation of distributed word representation for capturing semantics of biomedical text**, *Muneeb Th, Sunil Kumar Sahu, Ashish Anand*, **BioNLP at ACL-2015**, <https://aclweb.org/anthology/W/W15/W15-3820.pdf>.

Unpublished Works

- 2016 **Disease synonyms extraction from biomedical corpora using different Word Embedding techniques.**
- 2015 **Biomedical entity recognition in PubMed articles through character level word embedding and Recurrent Neural Network (RNN) .**
- 2014 **Named entity recognition in tweets through Conditional Random Field (CRF) .**
- 2014 **Object recognition in images through deep Convolutional Neural Network (CNN) .**
- 2014 **Parts of speech tagging in English language text through Hidden Markov Model (HMM) .**

- 2013 **Digit recognition system in speech signal by Hidden Markov Model (HMM)** .
- 2013 **Spam mail detection through Support Vector Machine (SVM)** .
- 2013 **Digit recognition in hand written images through Multi Layer Perceptron (MLP)**.
- 2013 **Language identification system for small (ISCII coded) text corpus through 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.

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**.

Personal

- Full Name **Sunil Kumar Sahu**.
- Other Interest **Athletics,Swimming**.

References

- Dr. Ashish Anand**, *Assistant Professor, Indian Institute Technology Guwahati*, email id: anand.ashish@iitg.ernet.in.
- Dr. Manjira Sinha**, *Research Scientist, Conduent Labs India*, email id: manjira.sinha@conduent.com.
- Dr. Niranjana Uma Nares**, *Applied Scientist, Microsoft Seattle*, email id: Niranjana.Uma@microsoft.com.