

Dr. Sanjay Chatterji

sanjaychatter@gmail.com

Samsung R&D Institute India, Bangalore (3 Years)

Mailing Address

B-401, Sai Krupa Elite, Fern City Road
Doddanekundi, Marathahalli Post,
Bangalore, Karnataka, India - 560037
Phone: +91-9740122889, +91-9647342393

Permanent Address

10/5 B, Botanical Garden Lane,
Near Barafcall, Shibpur, Howrah,
West Bengal, India - 711103
Phone: +91-9433122824

ACADEMIC AND INDUSTRIAL EXPERIENCE

As **Technical Lead** in the Web Services, Education and in the S-voice groups for 3 years (February 2014 till date) in Samsung R&D Institute India Bangalore (SRIB).

As **Senior Project Officer** in *Indian Language Machine Translation Phase II* project Sponsored by Deity, MCIT, Govt of India for 2 year (March 2012 to February 2014) in IIT Kharagpur.

As **Senior Project Officer** in *The Bangla Treebank* project Sponsored by TDC-IL, CIIL, Mysore for 2 years 8 months (July 2009 to Feb 2012) in IIT Kharagpur.

As **Senior Project Officer** in *Indian Language Machine Translation* project Sponsored by Deity, MCIT, Govt of India for 2 years 4 months (March, 2007 to June 2009) in IIT Kharagpur.

As **Lecturer** in Dept. of CSE, Haldia Institute of Technology (WBUT), Haldia, Midnapur (East), West Bengal for 6 months (Sept 2005 to Feb 2006) and in Dept. of CSE, KNS Institute Of Technology (VTU), Bangalore, Karnataka, for 1 year (March 2006 to Feb 2007).

RESEARCH STATEMENT

In my Ph.D. I have worked on Indian Language Machine Translation domain specifically syntactic statistical and transfer based approaches to build Bengali-Hindi machine translation systems based on the resources that are available. We have built Morphological analyzer, POS tagger, Parser(dependency), etc. using the Bengali resources. We have targeted to correct the noise in the output of Machine Translation systems for Bengali Hindi language pair. Now in Samsung, I have started working on semantic analysis. I have worked on aspect based sentiment analysis, inter concept higher order predicate relation extraction, representation and application in search domain, DNN based classification of intents, etc. Overall, I have experience and expertise in application of Machine Learning algorithms for any NLP tasks.

PROJECTS

Semantic Analysis

Project Undertaken: Samsung R&D Institute India, Bangalore

Works: *Aspect Based Sentiment Analysis, Higher Order Knowledge Representation, Intent and Agent Classification using DNN*

Role: Technical Lead

Feb 2014 – December 2016

The Bangla Treebank (TBT), Indian Institute of Technology, Kharagpur

Project Undertaken: Central Institute of Indian Languages (CIIL-Mysore)

PI: Prof.Sudeshna Sarkar, CSE, IITKGP & Prof.Jayshree Chakraborty, HSS, IITKGP

Role: Senior Project Officer

July 2009 – Feb 2012

Indian Language Machine Translation (ILMT), Indian Institute of Technology, Kharagpur

Project Undertaken: Ministry of Communication and Information Technology (MCIT)

PI: Prof.Sudeshna Sarkar, CSE, IITKGP & Prof.Anupam Basu, CSE, IITKGP

Role: Senior Project Officer

March 2007 – June 2009

Extraction of Digital Straight Line Segments (DSSs) from Digital Curves, Bengal Engineering and Science University, Shibpur
M.E. Final Year Project
Supervisor: Prof. Partha Bhowmick
Feb 2004 – Jun 2005

SUBJECTS HANDLED **HIT(WBUT):** Database Management System, Op. System, Data Structure, C, C++
KNSIT(VTU): Principles of Programming Languages, Op. System, C
As TA in IIT KGP: Machine Learning, AI, NLP, Intelligent Systems, Cognitive Science

SKILLS Perl, C, C++, Java, Python
Automata, O.S., Algorithm, Natural Language Processing, Machine Learning, A.I.

EDUCATION **Indian Institute of Technology**, Kharagpur, India
Degree: Ph.D.
At: Department of Computer Science and Engineering
Year: 2015

Bengal Engineering & Science University, Shibpur, Howrah, India
Degree: Master of Engineering
At: Department of Computer Science and Technology
Year: 2005

Haldia Institute of Technology, Haldia, India
Degree: Bachelor of Technology
At: Department of Computer Science and Engineering
Year: 2003

West Bengal Council of Higher Secondary Education, West Bengal, India
Degree: Higher Secondary(XII)
At: Irhpala K.M. Institution, Irhpala, Paschim Medinipur
Year: 1998

West Bengal Board of Secondary Education, West Bengal, India
Degree: Madhyamik(X)
At: Irhpala K.M. Institution, Irhpala, Paschim Medinipur
Year: 1996

RESEARCH INTERESTS **Natural Language Processing:** Machine Translation, Bangla Treebank
Information Retrieval: Indexing and Searching, Lucene
Machine Learning: Parser, POS tagger, Chunker, NER, Semantics

PH.D. **Thesis Title:** Improving the Performance of a Bengali to Hindi Machine Translation System Through Systematic Processing of Errors.

Ph.D. Supervisors Prof. Sudeshna Sarkar, CSE, IIT Kharagpur
&
Prof. Anupam Basu, CSE, IIT Kharagpur

REFEREED CONFERENCE PROCEEDINGS AND JOURNALS ————— 2016 —————
Sanjay Chatterji; Sudeshna Sarkar; Anupam Basu, *Experiences in Building Transfer Based Bengali to Hindi Machine Translation System*, Communicated to Journal of Multimedia Tools and Applications (SCI Indexed), Springer

Sanjay Chatterji; Nitish Varshney; Parnab Kumar Chanda; Vibhor Mittal; Bhavi Bhagwan Jagwani, *Extracting and Representing Higher Order Predicate Relations between Concepts*, 21st International Conference on Applications of Natural Language to Information Systems (NLDB 2016), University of Salford, MediaCityUK Campus, Salford, UK, June 2016, pp 175–186

Sanjay Chatterji; Nitish Varshney; Ranjan Kumar Rahul, *AspectFrameNet: A FrameNet Extension for Analysis of Sentiments around Product Aspects*, Journal of Supercomputing (SCI Indexed), DOI: 10.1007/s11227-016-1808-6, 2016, Springer

————— 2015 —————

Sanjay Chatterji; Ranjan Kumar Rahul; Anoop Arora, *A Hybrid Approach for Identifying Sentiments around Aspects*, 2nd IEEE International Conference on Recent Trends in Information Systems (ReTIS 2015), Kolkata, India, September, 2015

————— 2014 —————

Pragati Dhang; **Sanjay Chatterji**; Tanaya Mukherjee Sarkar; Sudeshna Sarkar; Jayashree Chakraborty; Anupam Basu, *Syntactico-semantic subject (karta) in Bangla*, International Journal of Bengal Studies (IJBS 13-14), 4(5), International Society of Bengal Studies, Viswakos Parisad, Print ISSN 2277-7717, pp 434–439

Sanjay Chatterji; G. S. Sreedhara; Maunendra Sankar Desarkar, *An Efficient Tool for Syntactic Processing of English Query Text*, MIKE 2014, Cork, Ireland, December, 2014, pp 278–287

Sanjay Chatterji; Subrangshu Sengupta; Bagadhi Gopal Rao; Debarghya Banerjee, *A Tool for Converting Different Data Representation Formats*, MIKE 2014, Cork, Ireland, December, 2014, pp 288–297

Sanjay Chatterji; Tanaya Mukherjee Sarkar; Pragati Dhang; Samhita Deb; Sudeshna Sarkar; Jayshree Chakraborty; Anupam Basu, *A Dependency Annotation Scheme For Bangla Treebank*, Language Resources and Evaluation (SCI Indexed), 48(3), pp 443–477, 2014, Springer Netherlands, DOI 10.1007/s10579-014-9266-3, Print ISSN 1574-020X, Online ISSN 1574-0218.

————— 2013 —————

Sanjay Chatterji; Pragati Dhang; Tanaya Mukherjee Sarkar; Jayshree Chakraborty; Sudeshna Sarkar; Anupam Basu, *Syntactico-Semantic Subject (Karta) in Bangla Sentences*, 3rd International Congress of Bengal Studies, Kolkata, India, November, 2013.

————— 2012 —————

Sanjay Chatterji; Diptesh Chatterjee; Sudeshna Sarkar, *An Efficient Technique for De-Noising Sentences using Monolingual Corpus and Synonym Dictionary* in Proceedings of COLING 2012: Demonstration Papers, Mumbai, India, December, 2012, pp 59–66

Sanjay Chatterji; Nabanita Datta; Arnab Dhar; Biswanath Barik; Sudeshna Sarkar; Anupam Basu, *Repairing Bengali Verb Chunks for Improved Bengali to Hindi Machine Translation* in Proceedings of the 10th Workshop on Asian Language Resources, Mumbai, India, December, 2012, pp 65–74

Sanjay Chatterji; Arnab Dhar; Sudeshna Sarkar; Anupam Basu, *Translations of Ambiguous Hindi Pronouns to Possible Bengali Pronouns* in Proceedings of the 10th Workshop on Asian Language Resources, Mumbai, India, December, 2012, pp 125–134

Arnab Dhar; **Sanjay Chatterji**; Sudeshna Sarkar; Anupam Basu, *A Hybrid Dependency Parser for Bangla* in Proceedings of the 10th Workshop on Asian Language Resources, Mumbai, India, December, 2012, pp 55–64

Sanjay Chatterji; Arnab Dhar; Sudeshna Sarkar; Anupam Basu, *A Three Stage Hybrid Parser for Hindi* in Proceedings of the Workshop on Machine Translation and Parsing in Indian Languages, Mumbai, India, December, 2012, pp 155–162

————— 2011 —————

Sanjay Chatterji; Praveen Sonare; Sudeshna Sarkar; Anupam Basu, *Lattice Based Lexical Transfer in Bengali Hindi Machine Translation Framework* in Proceedings of 9th International Conference On Natural Language Processing (ICON 2011), Chennai, India, December, 2011, pp 53–60

Sanjay Chatterji; Arnab Dhar; Biswanath Barik; Moumita PK; Sudeshna Sarkar; Anupam Basu, *Anaphora Resolution for Bengali, Hindi, and Tamil Using Random Tree Algorithm in Weka* in Proceedings of ICON2011 NLP Tools Contest: Anaphora Resolution in Indian Languages, Chennai, India, December, 2011, pp 7–12

————— 2010 —————

Sanjay Chatterji; Sudeshna Sarkar; Anupam Basu, *Implementation of a Bengali Morphological Generator* in Proceedings of 8th International Conference On Natural Language Processing (ICON 2010), Kharagpur, India, December, 2010, pp 300–305

————— 2009 —————

Sanjay Chatterji; Devshri Roy; Sudeshna Sarkar; Anupam Basu, *A Hybrid Approach for Bengali to Hindi Machine Translation* in Proceedings of 7th International Conference On Natural Language Processing (ICON-2009), Hyderabad, India, December, 2009, pp 83–91

Sanjay Chatterji; Praveen Sonare; Sudeshna Sarkar; Devshri Roy, *Grammar Driven Rules for Hybrid Bengali Dependency Parsing* in Proceedings of ICON09 NLP Tools Contest: Indian Language Dependency Parsing, Hyderabad, India, December, 2009, pp 37–41

Sanjay Chatterji; Tanaya Mukherjee Sarkar; Sudeshna Sarkar; Jayshree Chakraborty, *Karak Relations in Bengali* in Proceedings of 31st All-India Conference of Linguists (AICL 2009), Hyderabad, India, December, 2009, pp 33–36

————— 2008 —————

Sujan Kumar Saha; **Sanjay Chatterji**; Sandipan Dandapat; Sudeshna Sarkar; Pabitra Mitra, *A Hybrid Approach for Named Entity Recognition in Indian Languages*. in Proceedings of IJCNLP-08 Workshop on Named Entity Recognition for South and South East Asian Languages, Hyderabad, India, January, 2008, pp 17–24

AWARDS AND
ACHIEVEMENTS

Jury Award for a demo in Nipun, 2015 in Samsung R&D Institute India - Bangalore
First Prize in NLP Tool Contest, ICON 2011
First in NERSSEAL Workshop, IJCNLP 2008
Qualified in GATE (CS) in 2003 with 97.98 Percentile
MHRD Scholarship during M.E. from August 2003 to July 2005

EXTRA
CURRICULAR
ACTIVITIES

Organized Summer School on NLP and Text Mining, CSE, IIT Kharagpur, June, 2009.
Organized ICON 2010, IIT Kharagpur, December, 2010.
Committee member of Durga Puja for Howrah Youth Academy.
Committee member of VSRC Hall of Residence, IIT Kharagpur
Captain of VSRC Volleyball team for two consecutive years in IITKGP.

REFERENCES

Prof. Sudeshna Sarkar (Dept of CSE, IIT Kharagpur)(Ph.D. Supervisor)
Prof. Anupam Basu (Dept of CSE, IIT Kharagpur)(Ph.D. Supervisor)
Prof. Partha Bhowmick (Dept of CSE, IIT Kharagpur)(M.Tech. Supervisor)
Prof. Pushpak Bhattacharya (Director, IIT Patna)(Project Coordinator)
Dr. Monojit Choudhury (Microsoft Research (India)) (Senior from IIT Kharagpur)

PERSONAL
DETAILS

Date of Birth: 14/01/1981
Marital Status: Married
Passport Details: M3009476, Kolkata, Date of Expire: 26.10.2024
PAN Details: AHAPC7150L, Jurisdiction:WBG/W/179/1

DATED

9th of December, 2016.

PLACE

Bangalore, India