

SOMNATH BASU ROY CHOWDHURY

Residential Information

Carrboro, NC
Chapel Hill 27516, US.

Electronic Information

Web: <https://brcsomnath.github.io>
Email: somnath@cs.unc.edu

Research Interests

Representation Learning for text, Fairness in NLP and Information-Theoretic evaluation metrics.

Education

University of North Carolina (UNC) at Chapel Hill

2021 - present

Ph.D. in Computer Science

Advisor: Prof. [Snigdha Chaturvedi](#)

Indian Institute of Technology (IIT) - Kharagpur, India

2013 - 2018

B.Tech. & M.Tech. (Dual Degree Hons.) in Electrical Engineering

Rank - 3/31

Minor in Computer Science & Engineering (CSE)

Micro-specialization in Embedded Wireless Systems (EWS)

Advisor: Prof. [Jayanta Mukhopadhyay](#)

Publications

Google Scholar: [xGbyrIUAAAAJ](#). (Note: Most computer science researchers publish more in conferences than journals. Conferences are considered equally or more prestigious than journals, all conferences below including short papers have an acceptance rate < 25%.)

Conference papers:

- [9] **Adversarial Scrubbing of Demographic Information for Text Classification**
S. Basu Roy Chowdhury, S. Ghosh, Y. Li, J. Oliva, S. Srivastava and S. Chaturvedi
Accepted (Long, main), EMNLP 2021 [[pdf](#)]
- [8] **Is Everything in Order? A Simple Way to Order Sentences**
S. Basu Roy Chowdhury*, F. Brahman* and S. Chaturvedi
Accepted (Long, main), EMNLP 2021 [[pdf](#)]
- [7] **Learning beyond datasets: Knowledge Graph Augmented neural networks for NLP**
K.M. Annervaz*, S. Basu Roy Chowdhury* and A. Dukkipati.
Proceedings of NAACL-HLT, 2018. (Oral) [[pdf](#)]

Other refereed publications:

- [6] **Towards efficient data normalization using grammar-based parsing**
S. Basu Roy Chowdhury, A. Mishra and A. Banerjee
Proceedings of Microsoft Journal of Applied Research (MSJAR), 2020 [[pdf](#)]
- [5] **Instance-based Inductive Deep Transfer Learning by Cross-Dataset Querying with LSH**
S. Basu Roy Chowdhury, K.M. Annervaz and A. Dukkipati.
Proceedings of DeepLo Workshop, EMNLP, 2019. [[pdf](#)]
- [4] **AdGAP: Advanced Global Average Pooling**
A. Ghosh, B. Bhattacharya* and S. Basu Roy Chowdhury*
Proceedings of AAAI, Student Abstract and Poster Program, 2018. [[pdf](#)]
- [3] **Training Autoencoders in Sparse Domain**
B. Bhattacharya, A. Ghosh* and S. Basu Roy Chowdhury*
Proceedings of AAAI, Student Abstract and Poster Program, 2018. [[pdf](#)]

- [2] **Handwriting Profiling using Generative Adversarial Networks**
A. Ghosh*, B. Bhattacharya* and S. Basu Roy Chowdhury*
Proceedings of AAAI, Student Abstract and Poster Program, 2017. [pdf]
- [1] **SAD-GAN: Synthetic Autonomous Driving using Generative Adversarial Networks**
A. Ghosh*, B. Bhattacharya* and S. Basu Roy Chowdhury*
Proceedings of DLAI Workshop, NeurIPS, 2016. [pdf]

(* = equal contribution)

Theses:

- [1] **Lung Cancer segmentation in 3D CBCT Images**
Advisor: Prof. [Jayanta Mukhopadhyay](#)
M.Tech. Thesis, IIT Kharagpur [pdf]

Research and Work Experience

UNC Chapel Hill, Computer Science Dept. (2021 - present)
Graduate Research Assistant, [UNC-NLP](#)
Advisor: Prof. [Snigdha Chaturvedi](#)

Microsoft Corporation (July 2019 - Dec 2020)
Data Scientist II, Semi-Structured Information Retrieval team (Bing)

- Developed a context-free grammar framework leveraging *Earley parsing* and improved it to handle right-recursive scenarios enhancing down-stream schema-matching performance by $\sim 400 - 1200\%$
- Part of the team to setup a *semi-structured knowledge graph* from scratch and shipped 100M+ entity panes on Bing from unstructured data. Leveraged *entropy-based signals* from user-queries to perform *conflation* and ship over 200K entity panes for people segment

Apple Inc. (June 2018 - July 2019)
Software Engineer, Apple Security Engineering
Complex event processing for access management.

Accenture Technologies Lab (ATL) (Summer 2017)
Research Intern, Topic: Knowledge graph enhanced NLP
Advisor: [K. M. Annervaz](#) (ATL), Prof. [Ambedkar Dukkipati](#), (IISc)

Statistics and Machine Learning Group, IISc (May 2017 - April 2018)
Research Intern, Topic: Low Resource NLP
Advisor: Prof. [Ambedkar Dukkipati](#)

Visual Information Processing Lab - IIT Kharagpur (2016 - 2018)
Undergraduate Researcher, Topic: Lung Cancer Segmentation in 3D CBCT
Advisor: Prof. [Jayanta Mukhopadhyay](#)

National Brain Research Centre - Gurgaon, India (Summer 2015)
IAS Summer Fellow, Topic: Image Processing in MRI scans
Advisor: Prof. [Pravat Mandal](#)

Professional Service/ Teaching

- Program Committee Member/ Reviewer**
 - 2nd workshop on [Insights from Negative Results in NLP](#), EMNLP 2021
 - 3rd workshop on [Narrative Understanding Workshop](#), NAACL 2021

- *Teaching Assistant*
 (EE21101) Signals and Networks Lab, IIT Kharagpur 2018
 (EE19001) Electrical Technology Lab, IIT Kharagpur 2017

Talks

- ‘Probabilistically grounded word embeddings’, Bing paper reading forum, *Microsoft* (Nov 2019)
- ‘Secure Machine Learning’, *Apple India Research Forum* (April 2019)

Academic Achievements

- *All India Rank 7* (EECS) - Nationwide Education & Scholarship Test (NEST - II). 2017
- *ACM-ICPC Asia-Chennai Regionals* - Rank 33/1100 2015
- *IAS Summer Research Fellowship* - National Brain Research Center 2015
- *Department Change*, IIT Kharagpur - Rank 3/52 department students 2014
- *INSPIRE Scholarship* - Top 1 percentile in ISC (Grade 12). 2013
- *Top ranks* - JEE (Mains, Advanced): 99.9+ percentile, Top 1% in ISCE & ISC 2013

Miscellaneous Activities

- *Winner* of the WaterAid Marathon Hyderabad - *5km* category 2019
- *Best Innovation Award* - Apple EMEA Hackathon 2018
- *Special Mention in Technology GC* - Graduating batch LBS Hall. 2018
- *Silver-medal* winning Data Analytics team - Inter-hostel GC 2017
- *Gold-medal* winning Hardware Modeling team- Inter-hostel GC 2015
- *5th* year distinction in *Fine Arts* and Painting - All India Art and Crafts Society 2010