

# Sree Bhattacharyya

📍 Kolkata, India

✉ sree.bhattacharyya2@gmail.com

☎ 9330248102

🔄 sreebhattacharyya

in Sree Bhattacharyya



## Education

---

2023/08

State College, PA

### Ph.D. in Informatics

*Pennsylvania State University*

Incoming PhD Student at Penn State - School of Information Sciences and Technology (IST). Graduate Research Assistant at the Wang Group [🔗](#).

2018 – 2022

Howrah, India

### Bachelor of Technology in Computer Science and Technology

*Indian Institute of Engineering Science and Technology, Shibpur (formerly: Bengal Engineering College)* [🔗](#)

- CGPA: 9.86/10, Class Rank: 2/73, University Rank: 3
- Transcripts [🔗](#)
- Graduated: 16 December, 2022

## Publications

---

2022/09

### Link Prediction for Social Networks using Representation Learning and Heuristic-based Features [🔗](#)

Samarth Khanna\*, Sree Bhattacharyya\*, Sudipto Ghosh, Kushagra Agarwal, Asit Kumar Das

Preprint, currently under review at MAISoN Workshop - IJCAI 2023.  
(\*equal contributors)

2022/05

### Towards Bengali WordNet Enrichment using Knowledge Graph Completion Techniques [🔗](#)

Sree Bhattacharyya, Abhik Jana

Proceedings of the Workshop on Resources and Technologies for Indigenous, Endangered and Lesser-resourced Languages in Eurasia (EURALI) within the 13th Language Resources and Evaluation Conference (LREC, 2022)

## Work Experience

---

2022/10 – 2022/12

### Research Collaborator

- Independently collaborated with graduate students at UMD College Park on the **SemEval 2023** task of **MultiCoNER II**.
- Approaching **low-context complex NER** through synthetic text generation using **diverse constrained NLG**.

2022/07 – present

Bangalore, India

### Software Engineer

*Microsoft*

- Part of **Azure Storage**, working on the Hadoop Azure Blob Filesystem Driver, contributing to the **open-source Apache Hadoop** repository.
- Responsible for bringing feature improvements within the driver, and maintaining the continuous integration setup.
- Collaborating with teams beyond Azure Storage and Microsoft to align on feature requirements.

2021/10 – 2022/05

Howrah, India

### Undergraduate Researcher (Thesis)

*Indian Institute of Engineering Science and Technology, Shibpur*

Work done for Bachelor Thesis under Prof. Asit Kumar Das, in two different parts:

**Classification of Alzheimer's Disease Progression using GCNs on functional MRIs**

- Independently found the research problem, designed methodology to apply GCNs on graph of regions of interest in the brain obtained from fMRIs
- Worked on pre-processing fMRIs using tools like fMRIPrep and NiLearn
- Further effort was paused due to lack of resources in the department

#### Study of Approaches for Link Prediction in Social Networks

- Compared **structural, neighborhood-based, Graph Attention and GNN-based** node and edge representation methods with featurization for edges using **Graph Heuristics**
- Used the edge representations with **ensemble ML models** and **custom neural network** architectures for predicting links on **4 social network datasets** (soc-Epinions, musae-Github, musae-Twitch, wiki-Vote)
- Finally, proposed a **combination of learned representations and heuristics** as the best feature set for the link prediction task
- Demonstrated **improvements** across most datasets and predictive modeling techniques by using this new feature set
- Publication linked above.

2021/06 – 2021/10

Remote

#### Research Intern

*University of Hamburg*

- Worked at the Language Technology group as a **DAAD-WISE fellow**, under the guidance of Prof. Dr. Chris Biemann and Dr. Abhik Jana.
- Used knowledge graph completion techniques (including DistMult, ComplEx, HolE) for **enrichment of an existing WordNet** in Bengali, a low-resource language.
- Also experimented with using a **distributional thesaurus** for the same task.
- Presented publication (linked above) on intermediate results as a poster at the EURALI Workshop at LREC 2022. Poster link [☞](#), presentation link [☞](#).

2021/06 – 2021/07

Remote

#### Software Engineer Intern

*Microsoft (ADO Scanner Team, Digital Security and Risk Engineering, India)*

- Introduced **Incremental Scan feature** to Azure DevOps Security Scanner, reducing no. of resources scanned (for large organisations) from **~160k to 1000**.
- Explored **vulnerabilities in Azure DevOps**, recorded and demonstrated their end-to-end exploitation to DSR Redmond Leadership, including the Chief Information Security Officer at Microsoft.

2020/09 – 2021/03

Remote

#### Student Contributor

*CMS GEM Collaboration: CERN*

Part time developer in the **CMS experiment** (GEM Data Acquisition Team). **Designed a scalable database** for storing configuration parameters of the GEM detector components, and configurations history. Completed implementation for a single detector component (VFAT). Techstack used: PostgreSQL, C++ and CMAKE.

2020/06 – 2020/07

Remote

#### Mentee - Engage 2020

*Microsoft*

- Mentorship program by Microsoft. Created an **AI powered Tic-Tac-Toe** game webapp [☞](#) (GitHub Repo [☞](#)) using Minimax algorithm, completed research essay tasks on AI and ethics, algorithmic and coding tests and passed continuous problem-solving evaluations by mentor (software engineer from Microsoft).
- Ranked as the **7th** top sole contributor among **500+ participants** all over India, for the final project, and received a direct internship offer from Microsoft.

## Projects

---

2022/09

#### Microsoft Global Hackathon

- Worked on building **Adapt AI** - a visual question-answering system for the specific use case of aiding visually impaired users, extending upon Microsoft's Seeing AI.

- Created the proof-of-concept experimenting with **CNN-LSTMs** and **Multimodal Transformers**.



2021/07

#### Amazon Machine Learning Challenge

- In this 3-day long ML Hackathon, the task was to **classify 2.9 million product descriptions into ~9.9k "browse node IDs"** (categories).
- Tried different techniques including ML modelling, few-shot learning, pre-trained language models and a custom encoder-decoder architecture.
- Obtained best results with **Naive Bayes with incremental learning**.
- Was ranked us among the **top 2%**, with a **rank of 59** (Team Random\_State).

### Awards and Achievements

---

- Awarded the merit-based **Jordan-Rednor Graduate Fellowship** by IST, Penn State.
- Received the **Smt. Uma Rani Neogi Scholarship** (from Alumni Association, IESTS) in March 2022 for being the **woman student with the highest CGPA across all undergraduate disciplines**.
- Received the **DAAD Working Internships in Science and Engineering fellowship**  2021 for research internship in a public German Institution.
- Was awarded **IAS-INSANA Summer Research Fellowship (Indian Academy of Sciences' SRFP 2021)** for a research internship at Machine Intelligence Unit, ISI Kolkata (could not continue due to work as a DAAD fellow).
- Selected for **IIT Roorkee SPARK 2021** research fellowship (could not continue due to work as a DAAD fellow).
- Recipient of the "Bigyani Kanya Medha Britti" (development of young female "scientist") Scholarship from **Jagadish Bose National Science Talent Search (JBNSTS) Organization**  (2018-22 batch).
- Felicitated by the Chief Minister of West Bengal, India, for higher secondary (ISC) examination score - Percentage: 97.75% - All India Rank: 8 out of 80,880 candidates
- All India Rank 10 in secondary school examinations (ICSE 2016).

### Skills and Expertise

---

- **Languages:** Python, C++, C, SQL, Java
- **Tools:** Machine Learning: NLTK, Keras, Seaborn, Scikit-Learn, Pandas, Numpy, Matplotlib, Tensorflow; Git, CMAKE, MS Azure, Azure DevOps

### Extra-curricular Activities

---

- Member of **Graph Representation Learning Journal Club** - an independent group that discusses relevant research (papers, code etc.)
- Student Mentor for **Women In Technology Chapter of Google Developer Student Clubs**, IEST: responsible for closely mentoring 5 girls in the cohort and assisting them in career development, and participating in overall mentorship through internally organized hackathons, discussions, and lightning talks.
- Member in Les Thespians(dramatics society), Euphony (music society).
- Theatre actor in 'Sayak' (2009-15): part of the children's wing of this professional group, acted in over 5 productions, performing all over the state.
- Performer and trainee in classical and folk music vocals (since 2004): personally received training and performed in concerts, television programmes and the radio.