



Tathagata Dey
Computer Science & Engineering
Indian Institute of Technology Bombay

22M0765
M.Tech.
Gender: Male
DOB: 24/03/2000

Examination	University	Institute	Year	CPI / %
Post Graduation	IIT Bombay	IIT Bombay	2024	8.92
Graduation	MAKAUT	GCETTS	2022	9.13
Graduation Specialization: Computer Science and Engineering				

WORK EXPERIENCE & INTERNSHIPS

- **Artificial Intelligence Developer Intern | LG Soft India Pvt. Ltd.** (May'23-July'23)
 - Implemented a model for the problem statement of Table-to-text generation on Rotowire dataset.
 - Attained near state-of-the-art results with sequence-to-sequence model using Open Neural Machine Translation.
 - Gained expertise in linearizing tabular data, conducted extensive experiments on model parameters, refined data processing techniques, performed literature survey and developed a re-usable codebase.

M.TECH THESIS & SEMINAR

- **Table-to-text Generation with Subjectivity and Objectivity** (M.Tech Project | Guide: Prof. Pushpak Bhattacharyya) (June'23-Present)
Current work:
 - Scraped tables of different genres and their target subjective text from the web. Studied the pragmatics behind emotional inference generated from that data.
 - Synthetically generated data instances of different genres. Developed an interface to annotate the data instances.**Future work:**
 - A suitable baseline model has to be developed for the objective and required data for model training or fine-tuning or prompting has to be manually annotated. This dataset will be the first published data with this objective.
 - Human evaluation has to be performed to determine efficiency as it will also be first-of-this-kind work.
- **Text Generation** (M.Tech Seminar | Guide: Prof. Pushpak Bhattacharyya) (Jan'23-May'23)
 - Explored text generation, focusing on movie scripts, studying objectives and techniques in-depth.
 - Evaluated leading models, examined applications in text generation, enhancing understanding of the landscape.

RESEARCH & DEVELOPMENT PROJECT

- **Natural Language Generation from RDF Triples** (R&D Work for course CS 691 | Guide: Prof. Pushpak Bhattacharyya) (Jan'23-May'23)
 - Developed a state-of-the-art model to generate text sentences from RDF triples with coverage and coherence.
 - Sequence-to-sequence model, such as T5-base with pipeline approach has been used to convert a knowledge graph into sentences. Results are generated on the WebNLG corpus with ablation studies and evaluations.

COURSE PROJECTS

- **Argument Generation with Prompting** (CS 772 | Prof. Pushpak Bhattacharyya) (Jan'23-April'23)
 - Implemented a state-of-the-art paper and duplicated the results on argument generation using T5 model.
 - Employed multi-task training, cumulative loss objective to predict arguments on CMV dataset arguments.
- **Picture-based Word Sense Disambiguation** (CS 772 | Prof. Pushpak Bhattacharyya) (Jan'23-April'23)
 - Developed a machine learning model that leverages image inputs to disambiguate ambiguous words effectively.
 - Curated an innovative dataset and manually annotated it for model evaluation, tailored for testing and validation.
- **POS Tagging Using Encoder-Decoder and FFNN** (CS 772 | Prof. Pushpak Bhattacharyya) (Jan'23-April'23)
 - Employed Encoder-Decoder architecture and Feed-forward Neural Network for POS Tagging of a give sentence.
 - Used Universal tagset, compared results of the two approaches and their analysed confusion matrices.
- **Mental Health Prediction Using Machine Learning Models** (CS 725 | Prof. Preethi Jyothi) (July'22-Nov'22)
 - Analyzed tech professionals' personal life dataset for mental health prediction; contributed to innovative research.
 - Leveraged diverse models; highlighted high-accuracy selections, enhancing mental health prediction accuracy.
- **Ghar Dhundo** (CS 699 | Prof. Bhaskaran Raman) (July'22-Nov'22)
 - Developed an application as a comprehensive platform to assist individuals in finding their ideal homes.

- HTML, CSS, JS and Bootstrap have been used in the frontend while Python with Django and PostgreSQL have been used for the robust backend. Implemented features like login-signup, dynamic pricing and wishlist.
- **Development of a Multi-threaded Web Server** (July'22-Nov'22)
(CS 744 | Prof. Mythili Vutukuru)
 - Developed a webserver with multiple threads, capable of handling many clients simultaneously over TCP sockets.
 - Conducted performance optimization to overcome bottlenecks, resulting in a smooth and swift user experience.
- **Linux Shell Functions Implementation** (July'22-Nov'22)
(CS 744 | Prof. Mythili Vutukuru)
 - Implemented functional properties of a Linux shell, including background and foreground process handling.
 - Incorporated forking, reaping, process management, and signal handling along with efficiency testing.
- **Generating Context-based Word Vectors Using CBOW and Skipgram** (Jan'23-May'23)
(CS 772 | Prof. Pushpak Bhattacharyya)
 - Compared CBoW and Skip-gram embeddings in word analogy task to assess their performance.
 - Utilized the Gutenberg corpus to train CboW and Skip-gram models.
- **Dynamics of Challenges Faced in Building a Startup** (Jan'23-April'23)
(CS 752 | Prof. Om Damani)
 - Examined complex challenges in startup ecosystem—social, economic dynamics. Employed accurate models to pinpoint optimal conditions for robust growth using SFDs, CLDs and other parameters.

RESEARCH WORKS

- **Current Research Works**
Currently working in the field of Natural Language Processing and AI at the CFILT Lab, Dept. of CSE, IIT Bombay.
 - Upcoming research paper submission include, **Survey of Triple to Text and Text to Triple Generation**, encompassing problem analysis, dataset overview, and different state-of-the-art methods for the said objective.
- **Previous Research Works**
Previously worked in the field of Computational Biology. During this time period, published 5 Conference Proceedings, 4 Journal Articles and 1 Book Chapter. Some of the significant ones are mentioned below. (July'19-July'22)
 - **Identification and computational analysis of mutations in SARS-CoV-2**, T. Dey, S. Chatterjee et al., Computers in Biology and Medicine, Elsevier, Vol 129, 104166, Feb 2021.
 - **Novel Algorithms for In Silico Peptide Vaccine Design with Reference to Ebola Virus**, S. Biswas, T. Dey et al., in 2020 International Conference on Computer, Electrical & Communication Engineering (ICCECE), IEEE, 2020.

TECHNICAL SKILLS

- **Programming & Scripting Languages:** C++, Python, C, HTML, CSS, JS
- **Tools & Libraries:** MATLAB, \LaTeX , Git, Bash, Tensorflow, Transformers, Scikit-learn, NLTK

POR & EXTRA-CURRICULARS

- **Student Companion, Institute Student Companion Programme (ISCP), IIT Bombay** (July'23-present)
Working in a team of 235+ coordinators, ensuring a smooth transition of incoming first-year PG students. Completed Mentor Training Programme conducted by Student Wellness Center towards better mentoring.
- **Interview Coordinator, Institute Placement Cell, IIT Bombay** (Dec'22)
Part of a team responsible for streamlining the placement process for 1800+ students. Acted as point of contact for students and several firms/universities collaborating the recruitment process.
- **Teaching Assistantship | CS 626 - Speech, NLP and The Web | Prof. Pushpak Bhattacharyya** (Autumn 2023)
Evaluated assignments and course projects throughout the semester. Invigilated exams and conducted seamless availability of course materials to the students.
- **Teaching Assistantship | CS 230 & CS 232 - Digital Logic and Computer Architecture | Prof. Biswa Panda** (Spring 2023)
Supervised labs for the course, evaluated lab assignments and lab quiz papers. Also designed tutorial papers and solutions, evaluated quizzes, midsem and endsem papers and helped in grading.
- **Teaching Assistantship | CS 310 - Automata Theory | Prof. G. Sivakumar** (Autumn, 2022)
Contributed to designing tutorial, invigilated and evaluated the answer papers and helped in grading of the course.

AWARDS & ACHIEVEMENTS

- Secured **All India Rank 21** in GATE Computer Science Exam of 2022 out of 77257 candidates.
- Selected for **Summer Research Fellowship 2020** at **Jawaharlal Neheru Centre for Advanced Scientific Research**.
- Selected as **JBNSTS Senior Scholar Awardee** in the year of 2018.

HOBBIES

Watching and playing sports, Listening to music, Reading books, Swimming, Video-gaming.