

Partha Pratim Saha

Web, LinkedIn, Duolingo(DET)

E-mail: technical.partha@gmail.com

Interested in pursuing a PhD in Language Models, Natural Language Processing, and Deep Learning

EDUCATION

- **Birla Institute of Technology and Science(BITS) Pilani, Masters, Data Science & Engg.** Pilani, India
GPA: 9.08(First-class Distinction)/10, Courses: Maths, NLP, Data Mining, Machine & Deep Learning Sept'19-Mar'22
[Dissertation Short Paper](#)
- **West Bengal University of Technology, Bachelor Degree, Computer Science & Engg.** Kolkata, India
GPA: 8.49/10, Courses: Probability, Linear Algebra, Data Structures, Algorithms, Artificial Intelligence Aug'06-Jul'10

INDUSTRIAL WORK EXPERIENCE

- **Wipro Technologies, Project : Conversational Dialog Systems Wipro, IBM NLP** Bangalore, India
May'21 - Date
 - **Role:** Developed task-based conversational systems using room reservation Intents, Entities, and booking related Dialog flow and finally deployed the end-to-end project on IBM Cloud
 - **Impact:** Used by 0.3 million people worldwide
 - **Skills/Tools :** Chatbots, Text Generation, NLP, Dialog flow, IBM Watson, NLP, Postman, NodeJS, IBM Cloud
 - **Team handling:** Led 02 deliveries of this project with 05 members each
- **BirlaSoft Limited - Johnson & Johnson R&D, Project: Medical Search Engine** New Delhi, India
Dec'17 - Aug'19
 - **Role:** Developed a search engine using SciSpacy NLP pipeline techniques, e.g. stemming/lemmatization, dependency parse tree generation, parts-of-speech tagging
 - **Impact:** Used by millions of users who use Johnson & Johnson products
 - **Skills/Tools :** Python, Word2Vec, SciSpacy, Fuzzy Search, Flask, Text Mining, Information Retrieval, NLP
- **Infosys Technologies Limited, Project: Cancer Genomics in AI** Chennai, India
Jan'11 - Dec'13
 - **Eole:** Identified top 10 genes that drive [Multiple Myeloma\(MM\)](#) blood cancer using regression and clustering to rank genes responsible for MM using 3 research papers
 - **Impact:** Drug Design to market, Increase in life expectancy
 - **Skills/Tools :** Python, Word2Vec, Machine Learning, Numpy, Pandas, Bioinformatics
- **Infosys Technologies Limited, Project: Genomic Sequence Alignment** Chennai, India
Jan'14 - Jul'15
 - **Role:** Applied Edit Distance and Needleman Wunsch algorithms to determine the minimum no of insertions, and deletions needed for given gene sequences in the FASTA file to identify similarities between those sequences
 - **Impact:** Biological hierarchy determination to identify new species
 - **Skills/Tools :** Python, Pandas, Sequence alignment, Dynamic programming

SKILLS SUMMARY

- **NLP:** Trustworthy and Explainable AI(XAI), Large Language Models(LLMs), Generative AI (GenAI), Transformer, BERT
- **Deep Learning:** Neural Network, CNN, RNN, Attention architecture, Chatbot, Image Captioning
- **Machine Learning:** Regression, Classification, Clustering, Tree-Based Algorithms, Bagging, Boosting
- **Frameworks:** Scikit, Numpy, Pandas, NLTK, SpaCy, TensorFlow, Keras, Flask, NodeJS
- **Tools:** Kubernetes, Docker, GIT, PostgreSQL, MySQL, Matlab, SQLite
- **Domains:** Cancer Genomics, Bioinformatics, Real Estate, Finance, and banking
- **Platforms:** Azure, AWS, IBM Cloud, Watson ML, Amazon EC2
- **Programming Languages:** Python, C, Java, R

AWARDS AND ACHIEVEMENTS

- Reviewer of **NeurIPS 2023** Workshop papers
- Awarded for **ICML 2021** with 100% registration fee scholarship to attend conference and workshop track
- Attended **NAACL 2022** conference as student volunteer
- Awarded for **CLPsych by ACL 2022** with 100% Registration scholarship
- Got selected to attend the **Google Developer's program**, Google India, 2019
- Received **Distinction** in Master Degree program from BITS Pilani, 2022
- **Udacity Bertelsmann Technology scholarship** for AI Track in ML with TensorFlow program, 2021
- Got accepted in **New York University(NYU)** AI Summer School, 2022