

Partha Pratim Saha

Natural Language Processing(NLP) Researcher, Wipro

 Website  LinkedIn  Github  technical.partha@gmail.com

Research interests: Language Models, NLP and multimodality, Conversational Systems, Explainable AI

EDUCATION

- **Birla Institute of Technology and Science(BITS) Pilani, Masters, Data Science & Engg.** Pilani, India
GPA: 9.08/10, Courses: Data Mining, Deep Learning, Data Science, NLP, ML, Statistical Methods Sept'19-Mar'22
Dissertation Research: I proposed a novel **Mixed-model federated learning architecture** using trade-off between faster execution time in centralized learning and high security in de-centralized model under the supervision of **Dr. SP Vimal**.
- **West Bengal University of Technology, Bachelor Degree, Computer Science & Engg.** Kolkata, India
GPA: 8.49/10, Courses: Maths, Algorithms, Artificial Intelligence, Data Structures, Algorithms, OS Aug'06-Jul'10

WORK EXPERIENCE

- **Wipro Technologies, Project : Conversational Dialog Systems Wipro, IBM NLP** Bangalore, India
AI Lead Data Scientist May'21 - Date
 - **Role:**
 - **Developed a conversational system** by leading and managing a team of 5 members
 - **Implemented 50+ custom intents, entities** and dialog flow with IBM Watson
 - **Proposed confidence score** to extract meaning followed by asking clarifying questions
 - **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
- **BirlaSoft Limited - Johnson & Johnson(J&J) R&D, Project: Medical Search Engine** New Delhi, India
Data Scientist Dec'17 - Aug'19
 - **Role:**
 - **Developed a search engine for J&J**
 - **Used SciSpacy NLP pipeline** techniques
 - Implemented stemming/lemmatization, dependency parse tree generation, parts-of-speech tagging, Scibert-Base as the Transformer model
 - **Impact:** Used by **over 1 billion people** who use Johnson & Johnson products
 - **Skills/Tools :** Python, Word2Vec, SciSpacy, Fuzzy Search, Flask, Text Mining, Information Retrieval, NLP
- **Infosys Technologies Limited, Project: Genomic Sequence Alignment** Chennai, India
Senior Systems Engineer Jan'14 - Jul'15
 - **Role:**
 - **Applied Edit Distance and Needleman Wunsch** algorithms
 - Determined the **minimum no of insertions, deletions and substitutions needed** for given gene sequences
 - **Identified similarities** between DNA sequences
 - **Impact:** **Biological hierarchy determination** to identify new species
 - **Skills/Tools :** Python, Pandas, Sequence alignment, Dynamic programming
- **Infosys Technologies Limited, Project: Cancer Genomics in AI** Chennai, India
Senior Systems Engineer Jan'11 - Dec'13
 - **Role:**
 - **Identified top 10 genes that drive Multiple Myeloma(MM)** blood cancer
 - Applied **Regression and Clustering algorithms**
 - **Implemented 3 research papers:** [this](#) [this](#), and [this](#) to rank genes that are responsible for MM
 - **Impact:** **Drug Design** to market, **Increase in life expectancy**
 - **Skills/Tools :** Python, Word2Vec, Machine Learning, Numpy, Pandas, Bioinformatics

TEACHING EXPERIENCE

- Teaching Assistant(TA)
NLP Applications (AIMLCZG519)(remote), AI ML Masters program, BITS Pilani [Winter 2024]
Deep Learning (DSECCZG524), AI ML Masters program, BITS Pilani [Fall 2021]
Deep Reinforcement Learning (AIMLCZG512), AI ML Masters program, BITS Pilani [Spring 2021]

AWARDS AND ACHIEVEMENTS

- Reviewer of **NeurIPS 2023, AAAI 2023**
- Awarded for **ICML 2021** with 100% registration fee scholarship to attend conference and workshop track
- Attended **NAACL 2022** conference as a 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

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