Sahasra Ranjan

Homepage: www.sahasrarjn.com
Github: github.com/sahasrarjn/

Email: sahasra.ranjan@gmail.com
Mobile: +91-916-2095-738

EDUCATION AND SCHOLASTIC ACHIEVEMENTS

- **Indian Institute of Technology, Bombay** — Mumbai, India
  B.Tech - Computer Science and Engineering; GPA: 9.05/10; Honors GPA: 9.75/10 — Aug 2019 - Present
- Recipient of the 'Thomas Dooie, Class of 1974 **Research Excellence Award**' for the best Bachelor's Thesis 'Novel Gradient-based Optimization Algorithm in Deep Learning' under the guidance of Prof. Preethi Jyothi *(2023)*
- Received the '**Outstanding Paper Award**' as first author at the 61st Annual Meeting of the Association for Computational Linguistics (ACL'23), an honour granted to **less than 1%** of the accepted papers. *(2023)*
- Achieved All India Rank **127** in Joint Entrance Examination, Advanced out of **2,40,000** candidates *(2019)*
- Secured All India Rank **281** in Joint Entrance Examination, Mains out of **1.2 million** candidates *(2019)*
- Recipient of the prestigious **KVPY** Fellowship and **NTSE** Scholarship from the Govt. of India *(2018)*

PUBLICATIONS

1. **Sahasra Ranjan**, Richeek Das, Shreya Pathak, Preethi Jyothi: **Improving Pretraining Techniques for Code-Switched NLP**, accepted and awarded the **Outstanding Paper Award** at *the 61st Annual Meeting of the Association for Computational Linguistics (ACL), 2023*.

2. Alexander Kroll, **Sahasra Ranjan**, Martin K. M. Engqvist, Martin J. Lercher: **A general model to predict small molecule substrates of enzymes based on machine and deep learning**, published in the *Nature Communications journal, 2023*

3. Alexander Kroll, **Sahasra Ranjan**, Martin J. Lercher: **Drug-target interaction prediction using a multi-modal transformer network demonstrates high generalizability to unseen proteins**, submitted to the *Science/AAAS journal, 2023*

RESEARCH AND WORK EXPERIENCE

Rubrik: Cloud Data Software Engineer — June 2023 - Present
Full-Time Software Engineering — Rubrik, Bangalore, India
- Contributed significantly to enhancing Rubrik's core **data management system** for businesses.
- Conducted comprehensive testing and development of Rubrik's advanced software and hardware services.
- Innovatively implemented features to support static versioning data, thereby enhancing testing capabilities.

Improving Pretraining Techniques for Code-Switched NLP — July - Dec 2022
Guide: Prof. Preethi Jyothi, CSALT Lab | B.Tech. Project — IIT Bombay
- Awarded the **ACL'23 Outstanding Paper Award**, bestowed to less than 1% of accepted papers, as the first author of 'Improving Pretraining Techniques for Code-Switched NLP', introducing innovative MLM methods.
- Pioneered the 'Switch-MLM' approach, optimizing MLM by leveraging switch-points within code-switched text.
- Introduced 'Freq-MLM', addressing the scarcity of language identification tagged datasets in code-switched content.
- Enhanced multilingual models with residuals and auxiliary loss, boosting F1 scores by 3-4% in diverse language pairs.

Protein Language Modeling with ESM1b-e2e — Summer 2021, 2023
Prof. Martin J. Lercher, Alexander Kroll | Research Internship — HHU, Dusseldorf, Germany
- Worked on a based Multi-modal **BERT** to create high-dimensional representations for enzymes and chemicals.
- Adapted **ESM-1b** model to build an E2E model which **outperformed** the existing SOTA models for the task.
- Improved accuracy for the task was achieved from **79.8%** to **87.5%** with the new richer representations of enzymes

Snapshot Metadata Garbage Collection — May - July 2022
Software Engineering Internship — Rubrik, Bangalore, India
- Worked on a GC algorithm for a database in a cloud data backup and recovery system to address slow db queries.
- Ideated & implemented a framework to find unnecessary data and remove them from the database periodically.
- Developed a mechanism to track expired **critical data** which were blocking more than **80%** of the DB table.

SKILLS SUMMARY

- **Programming Languages**   C++, Python, Go, Java, Bash, sed, AWK, TypeScript, SQL, VHDL
- **Data Science**   PyTorch, TensorFlow, Keras, NumPy, MATLAB, Octave, Pandas, Matplotlib, OpenCV
- **Softwares and Frameworks**   Git, LaTeX, Docker, Quartus, Django, Angular, React, NodeJS, HTML5, CSS

POSITIONS OF RESPONSIBILITY

Teaching Assistant, *CS725: Foundations of Machine Learning* — Aug 2022 - Nov 2022
Seasons of Code 2021, Mentor, *Web and Coding Club, IIT Bombay* — May 2021-July 2021
Core Member, *Electronics and Robotics Club, IIT Bombay* — May 2020 - April 2021
Team Member, *DevCom - Development Community, IIT Bombay* — May 2020 - April 2021

EXTRACURRICULARS

- Performed Bollywood themed group dance during the Annual Insync Dance Show event *(2022)*
- Completed **Guitar** learning course under Summer School of Cult conducted by Symphony, music club of IITB *(2019)*