



EDUCATION

Year	Degree/Exam	Institute	CGPA/Marks
2025	B.TECH	IIT Kharagpur	8.24 / 10
2021	ISC	Julien Day School, Ganganagar	99.5%
2019	ICSE	Julien Day School, Ganganagar	97.6%

INTERNSHIPS

AI Engineering Intern | TATA Consultancy Services Limited, Kolkata [May '24 - Jul '24]
• Extracted **350+** technical documents from PDFs of various Machine Learning textbooks using **PyPDF** and pre-processed them using **spaCy**
• Attained an **F-Score** of **.60** by using various **Zero-Shot** and **Few-Shot** prompts on **Llama-3-70B API** for document-level **relation** extraction
• Improved the F-Score on the validation set by **13.60%** via fine-tuning the **Llama-2-13B** model using **PEFT** on the extracted relation triplets

Data Science Intern | Stanford School of Medicine [May '23 - Aug '23]
• Analyzed the **MCH** dataset with the help of various plots and heatmaps using **Seaborn** and extracted important features from it using **PCA**
• Used a voting ensemble of **XGBoost**, **CATBoost** and **Random Forest Regressor** to predict the values of **6** independent health indicators
• Reduced the **MCRMSE** of the validation set by **14.94%** via tuning the hyperparameters of the **ensemble** models using **RandomSearchCV**

Data Analyst Intern | IIM Udaipur [May '23 - Jul '23]
• Utilized **Fuzzy String Matching Algorithm** to segregate **3** specific domains from over **2,00,000** business reviews and pre-processed them
• Employed **RoBERTa** model and tools like **SentiStrength** to analyze the changing customer sentiments for these businesses over **10** years
• Optimized topic modelling by combining **BERTopic** and **LDA** to yield interpretable **topic vectors** and predict the **Financial Trajectories**

PROJECTS

Multi-Versioned Software Fault Prediction | Bachelor's Thesis Project | IIT Kharagpur [Aug '24 - Present]
• Encoded the features using **Log Transform** to address non-linearity in predicting faults and their **severities** in multi-versioned software
• Achieved the best **R-squared** of **.78** by training multiple **ANN** architectures for the number of detected and corrected faults in test dataset
• Employed various **Bi-LSTM** architectures in forecasting software faults across **Jira** and **Firefox** versions and got a **mean R-squared** of **.81**

E-Commerce Website With Product Recommendations | Walmart Sparkathon [Jul '24 - Aug '24]
• Computed feature vectors of items using **Sentence Transformers** and **ResNet50** and recommended similar items using **Cosine Similarity**
• Integrated **Geolocation** with **Llama-3-70B API** to fetch user **demographics** from product descriptions and recommend customized items
• Developed an interactive web app using **React** and **Django** with user authentication via **django-rest-auth** and specialized product filters

Search Engine For DSA Problems | AlgoZenith Hackathon [May '23 - Jun '23]
• Scraped **10,000+** algorithmic problem statements from **3** dynamic coding sites using **Selenium** and removed the stopwords found in **NLTK**
• Implemented **TF-IDF** algorithm from **scratch** and computed **Inverted Index** of corpus by identifying keywords and counting frequencies
• Built a **Flask** app spawning a Python process and displayed the top **20** questions similar to the given query with decreasing TF-IDF scores

COMPETITION/CONFERENCE

Gold Medal | NSSC Data Analytics 2023 | Exoplanet Habitability Classification [Oct '23 - Nov '23]
• Analyzed **skewness** and **correlations** of multiple features using plots and heatmaps from **Seaborn** and imputed missing values using **KNN**
• Extracted relevant features using **PCA** and addressed the Class Imbalance in the dataset by over-sampling minority classes using **SMOTE**
• Enhanced the **accuracy** of **XGBoost** in predicting Habitability from **99.65%** to **99.80%** by tuning its hyperparameters using **GridSearchCV**

Bronze Medal | Open IIT Data Analytics 2022 | Gender Bias Detection in Text [Sep '22 - Oct '22]
• Classified individual sentences into biased(0) and unbiased(1) labels using **DFS** and pre-processed them while keeping the pronouns intact
• Employed **TF-IDF** and **bag-of-words** for text vectorization and trained an ensemble of **SVM**, **CATBoost** and **Bidirectional LSTM** on them
• Achieved a **Sentence-Pair Accuracy** of **98.98%** using a second **ensemble** layer of the voting model and **BERT** in identifying gender biases

SKILLS AND EXPERTISE

Languages: C | C++ | Java | Python | MATLAB | JavaScript | SQL | LaTeX
Tools: Git | GitHub | GitLab | MS Excel | MYSQL | Google Colaboratory
Libraries: Pandas | Scikit-Learn | OpenCV | spaCy | Selenium
Frameworks: TensorFlow | Keras | PyTorch | Flask | Django

COURSEWORK INFORMATION

Programming and Data Structures | Machine Learning - Foundations and Applications | Linear Algebra, Real and Complex Analysis | Probability and Statistics | Soft Computing | Entrepreneurship Essentials | LangChain for LLM Applications | Knowledge Graphs for RAG

AWARDS AND ACHIEVEMENTS

• Achieved **1st Place** in **CampusPulse 2023** hosted by Kharagpur Data Analytics Group for ideating an AI-driven Library Attendance System
• Secured **3rd Position** in **Innovate4Swadeshi 2023** hosted by **Sane Observer** for building an Innovative LLM-based Content Creation App
• Achieved a General Merit Rank of **652** in West Bengal Joint Entrance Exam 2021 among **1,00,000+** applicants from all across West Bengal
• Attained a Rank among the **top 1.5%** in JEE Advanced 2021 and **top 0.9%** in JEE Mains 2021 among **1 Million** candidates from all over India

EXTRA CURRICULAR ACTIVITIES

• Volunteered as a cadet in the **National Cadet Corps** - 1 Bengal EME COY NCC, IIT Kharagpur and attended its Annual Training Camp 2022
• Participated in various onstage productions and workshops being a part of **Druheen** - Bengali Technology **Dramatic Society**, IIT Kharagpur
• Member of the **Gold** winning Rangoli Team of the Rajendra Prasad Hall of Residence in the Inter-Hall Gymkhana Championship 2023-2024
• Mentored a group of **8** Freshers and Sophomores and helped them navigate through their academic journey and extra curricular activities