



Jay Singh

Indian Institute of Technology Kharagpur

Phone: +91 8651274328

Email: Jays.iitkgp@gmail.com

Github Profile : [JaySingh79](#)

Linkedin Profile: [Jay_Linkedin](#)

EDUCATION

Year	Degree/Exam	Institute	CGPA/Percentage
2021	B.Tech, Metallurgical and Materials Engineering	Indian Institute of Technology Kharagpur	7.91
2020	Class XII, CBSE	Gyan Bharti Global School, Gaya, Bihar	91.5%
2018	Class X, CBSE	Nazareth Academy, Gaya, Bihar	94%

INTERNSHIPS

INNOVACCR | DATA SCIENCE INTERNSHIP

June'24- Aug'24

- As part of **NLP R&D team**, engineered "Sara" progressing from **GTR-T5 base**, **Flan-T5 XL**, and other models, fine-tuned over custom dataset.
- Leveraged **Langchain**, for **Chain of Thought**, **Few-shot** prompting, with a **RAG** pipeline with **Milvus DB**, multi vector embedding (**COLBERT**).
- Developed a multi-chain model for diverse use case segmentation using **asynchronous** programming with **GPT-4o**, SLMs, and other models.

AFFINE | MACHINE LEARNING INTERNSHIP

April'24- June'24

- Worked with Data Science teams to make custom use case **LLMs**, **fine tuned** the latest **Meta-Llama-3-8B** model as an instruction QnA LLM.
- Engineered a **16** attention head, **500M** parameters model, using mixed-precision for a **35%** reduction in memory and **20%** faster convergence.
- Created a data pipeline with **Apache Spark** for **10TB** of text data, using Byte-Pair Encoding to cut token size by **40%** preserving performance.

CAMBRIDGE JUDGE BUSINESS SCHOOL | RESEARCH APPRENTICESHIP

May'23 - Sep'23

- Building a building a **startup accelerator portal** designing an interactive portal that generates tailored startup ideas for startup enthusiasts.
- Working on integration of AI models within a multidisciplinary team, delivering end-to-end solution from concept, funding, to market launch.
- Spearheaded the development of cutting-edge AI models using OpenAI, Llama3, etc models providing the ideation for aspiring entrepreneurs.

PROJECTS

MEDIUM ARTICLE SEARCH AND SUMMARIZATION AGENT

Jan'24-Feb'24

- Developed a Medium search and summarization tool, enabling real-time retrieval and generation of articles from internet with URL/summary.
- Used langchain, integrated **Anthropic** and other Hugging Face models like **Flan T5 Large**, other models for filtering and generating responses.
- Built an interactive UI with **Streamlit** and deployed over **Render** with minimal latency, for a seamless user experience, with efficient caching.

MUSIC POPULARITY ESTIMATION

Jan'24-Feb'24

- Analyzed a dataset of **227**, reducing dimensionality by **35%** retaining over **90%** of predictive accuracy with regression techniques **0.85** score.
- Trained **Random Forest** and **Gradient Boosting** regression models, with **hyperparameter tuning** for boosting accuracy by **15%** over others.
- Playlist optimization through feature analysis, revealing that higher energy and danceability increase popularity likelihood by more than 25%.

HYBRID NATURAL LANGUAGE EVALUATION

Mar'24-Apr'24

- Developed a model over streamlit to evaluate raw text generation, against language models deployed over **Amazon Elastic Compute Cloud**.
- Utilized **LSTM** architecture in **NLTK**, **Pytorch** and **multi-layer** hidden states, to create a text generation model, training **10,000 large** dataset.
- Engineered a pipeline to reflect **cosine similarity**/other similarity metrics against other models, reflecting errors, suggestions improvements.

COMPETITIONS

OPEN IIT DATA ANALYTICS | IIT KHARAGPUR 2023 ([Link to Certificate](#))

1st Position

- Developed a Next Word Prediction model using Python, **TensorFlow**, and **Keras**, capable of generating word predictions with 86.52% accuracy.
- Engineered textual data having tokenization, padding sequences, and one-hot encoding, reflecting expertise in data preparation for NLP tasks.
- Implemented deep learning architecture with **Embedding**, **LSTM**, Dense layers for RNN-based sequence modeling and language generation.

IMPETUS 5.0 CASE STUDY | UNSTOP ([Link of project](#))

National Semi-Final Round

- Formulated a compelling business model along with targeted marketing plan for **FinBlend**'s new integrated finance platform, FinInvest Pro.
- Strategized personalized offerings, data-driven acquisition, and partnerships to unlock FinInvest Pro's potential across consumer segments.
- Developed a go-to-market and sustainability strategy to enable FinBlend's expansion into lending, investments, and other insurance verticals.

SKILLS AND EXPERTISE

- Courses:** Programming & Data Structures, Linear Algebra, Probability & Statistics, Statistical Learning & Applications, Operational Research
- Languages/softwares:** Python, C++, SQL, Excel, Amazon Sagemaker, Gitlab, TensorFlow, Pytorch, LLM, Pandas, Streamlit, Langchain, Pyterrier
- Certifications :** Data Analytics Foundation by **NASBA** | **DBMS** with MySQL | Career Essentials on **Generative AI** | Advanced Python Udemy | **PyTorch** for Deep Learning | **Langchain:** Developing LLM Powered Applications | **Statistics** for Data Science | Advanced Learning Algorithms

POSITIONS OF RESPONSIBILITY

PUBLICITY AND MEDIA OUTREACH HEAD | SPRING FEST | IIT KHARAGPUR

June'23-Jan'24

- Led a team of 31+55 to organize nationwide events in 16 locations, engaging 850+ colleges, leading to 10,000 participants from across India.
- Enhanced Spring Fest's online presence and engagement by 60%, through big content creators and delivering high-quality content strategies.
- Led 'Prabal: Fighting Stronger on cell at a time' cancer awareness initiative, with Yuva India and DKMS-BMST Foundation for impactful results.

CORE ORGANIZING TEAM MEMBER | SPRING FEST | IIT KHARAGPUR

Aug'22 - Jan'23

- Triumphant sponsorship drive in the Uttar Pradesh, collecting over **INR 1 Lakh** through alumni contributions with a remarkable 200% YOY.
- Led a team of 4 to secure 26 media partnerships with major outlets like **Punjab Kesari**, **Amar Ujala**, and **Lutopia Magazine** for Spring Fest.
- Executed 4-day star shows with co-members, featuring artists like **King**, **Nucleya**, and **Sunidhi Chauhan**, attracting a footfall of over 50,000.

EXTRA CURRICULAR ACTIVITIES

- District-level Badminton Champion during high school from Nazareth Academy, amongst more than **400** different contestants of **38** districts.
- Part of the **Gold winning team** in Intra Hall Tug of War competition at **Pandit Madan Mohan Malaviya** Hall of Residence, IIT Kharagpur.
- 1st runner-up at '**ILLUMINATION**,' a major inter-hall competition/event at IIT Kharagpur, Pandit Madan Mohan Malaviya Hall of Residence.