

INDRAJIT SINGH


DATA/AI/ML EVANGELIST

CONTACT


+91 9591182659

indrajitsingh37@gmail.com

 [NLPBlog](#)

 Pune, MH-411057

 [Linkedin_Profile](#)

 [Github](#)

TOP SKILLS

Compassion

GenerativeAi, Langchain, AWS Bedrock

Data Science (ML/DL/LLM)

Vector Databases

Programming Languages (Python, SQL etc.)

Data Analysis & Visualization

Kafka, Starburst, Distributed Computing

EDUCATION

M.Tech (Data Analytics)

Birla Institute of Technology

2016-2018

Dissertation: Recurrent Neural Tensor
Network for Predictive Escalation

B. Tech (Information Technology)

**West Bengal University of
Technology**

2008-2012

Studied subjects such as Data Structure,
C++, and Automata.

"...you realize that all along there was something tremendous within you, and you did not know it." -
Paramahansa Yogananda

PROFILE

13 Years

Results-driven Senior AI Developer with extensive experience in developing Generative AI solutions using advanced Natural Language Processing (NLP) techniques. Proven track record in leveraging state-of-the-art technologies such as Anthropic Claude, AWS Bedrock, LangChain, LlamaIndex, Weaviate, and Qdrant for complex solution building. Committed to continuous learning and innovation, with a strong background in enhancing model performance and implementing scalable AI services.

WORK EXPERIENCE

GenerativeAi Solutions

Cellstrat, Bangalore, KA

Nov-2023-Present

- Spearheading the development of advanced Generative AI solutions using Large Language Models (LLMs), focusing on both product development and service design.
- Applied advanced Natural Language Processing (NLP) techniques to develop innovative Generative AI (GenAI) models and systems, achieving 30-40 % improvement in content generation efficiency.
- Led the integration of Anthropic Claude and AWS Bedrock to enhance natural language understanding capabilities, streamlining complex AI workflows and reducing development time by 50%.
- Leveraged LlamaIndex to create dynamic retrieval-augmented generation systems, enhancing the accuracy of information retrieval and reducing latency by 10%.
- Implemented Weaviate and Qdrant vector databases to power advanced similarity search and clustering in AI applications, significantly improving search relevance and data accessibility.
- Collaborated with cross-functional teams to analyze requirements and design tailored AI solutions, addressing specific business needs.
- Implemented state-of-the-art LLM-based models to solve complex problems in text generation, dialogue systems, and content creation, contributing to success of some of the big projects at Cellstrat such as MayaAi and FreedomGPT.
- Conducted thorough research and experimentation to enhance model performance and adaptability across various domains, leading to great success.
- Actively participated in code reviews, knowledge sharing sessions, and mentorship programs, fostering a culture of continuous learning and professional development within the team and across different GenAI products at Cellstrat.
- Stayed abreast of the latest advancements in AI by attending conferences, workshops, and seminars, ensuring the application of cutting-edge techniques and best practices.

GENAI PROJECTS

MayaAi- LLM based formula Generation

Developed and launched MAYAI Virtual Lab, a pioneering platform designed to revolutionize formula design ideation and rapid prototyping. The virtual lab offers a sophisticated environment for customized product development, enabling formulators to make data-driven decisions, enhance creativity, and significantly improve the success rate of labor-intensive trials using latest LLMs.

Technologies Used

- Anthropic Claude Sonnet
- AWS CDK & Bedrock
- LangChain, LlamaIndex
- Weaviate Vector Database

Freedom GPT Solution

Led the development of FreedomGPT, a state-of-the-art open-source AI chatbot designed to offer uncensored and private AI interactions. Successfully launched FreedomGPT to the market, achieving 0-300K user adoption in less than 6 months.

CERTIFICATIONS



CellStrat Certified NLP expert

AWARDS

- Super Achiever Award 2023
- The Motivator Award 2022 Q4
- 2022 Bloggathon winner at eClerx
- 3 consecutive appearance in organisation quarterly magazine.

LANGUAGES

English
Hindi
Bengali

WORK EXPERIENCE

Lead Data Scientist

eCelrx, Pune, Maharashtra May-2021-Nov-2023

- Directed the development of VoiceAI and TextAI solutions, using NLP and advanced data engineering techniques.
- Created an automated NLP system, cutting Quality Management processing time by 70%.
- Implemented an NLP pipeline for Speech + LLM-based emotion detection, boosting accuracy by 50% and increasing insurance revenue by 5%.
- Enhanced speech-to-text model accuracy from 67% to 89% through improved data auditing and training.
- Optimized AutoQA code, reducing processing time from 1.5 minutes to 5 seconds.
- Demonstrated Adverse Event Detection POC, securing a \$1M deal in the Pharma domain.
- Developed a Grammar Analyzer, increasing chat audit capacity by 400%.
- Managed a team of 6+ Data Scientists/Analysts for AutoQA, Speech, and VoiceAI.
- Published a whitepaper on "NLP for Business-Driven Applications."
- Provided expert consultation for 30+ projects, addressing diverse business challenges.

NLP Research @Cellstrat

Cellstrat 2018-2020

- Developed advanced models with PyTorch, Huggingface, Keras, and TensorFlow, enhancing data analysis and prediction.
- Executed multiple projects with large language models (LLMs), demonstrating expertise in advanced natural language processing (NLP).
- Excelled in developing and deploying question answering systems, text summarizers, and TTS solutions on AWS and GCP.
- Attained over 93% accuracy in QA models using ALBERT and RoBERTa-based approaches.
- Published "BERT - Algorithm Behind Better Search Engines" with CellStrat, advancing industry knowledge in search algorithms.
- Created abstract text summarization with LSTM, facilitating efficient extraction of key information from extensive text.
- Implemented Named Entity Recognition (NER) on CRM log data with CRFsuite and LSTM, enhancing data categorization and analysis.
- Mentored and guided Data Scientists at CellStrat, fostering a collaborative learning environment.

Senior CSE

Fivetran, Bangalore, KA 2020-2021

- Successfully implemented on-premises ETL knowledge into cloud-based ETL/ELT tools, such as Fivetran, facilitating seamless data integration and transformation.
- Managed and optimized very large data warehouses and data lakes deployed on cloud platforms, including BigQuery, AWS Redshift, and PostgreSQL, ensuring efficient data storage and retrieval.
- Implemented a Quality-driven approach, focusing on three key principles: Understand, Empathize, and Solve, resulting in enhanced customer satisfaction and improved service delivery.
- Attained a perfect 100% customer satisfaction score for consecutive quarters, reflecting exceptional performance and dedication to meeting customer needs.
- Consistently achieved a 99 % Service Level Agreement (SLA) for resolving bugs.
- Built robust data pipelines on cloud platforms, leveraging tools such as dbt, data flow, and model data warehousing, facilitating successful migration from on-premises to cloud environments for enterprise customers.

NLP PROJECTS

VoiceAi- Speech Analytics Solution

Achieved notable success in VoiceAi project, utilizing speech analytics to enhance customer interactions. Proficient in gathering, analyzing, and interpreting large speech data volumes, and designing deep learning models for valuable insights.

AutoQa- Text Analytics Solution

AutoQA enhances customer service quality through automated assurance. It analyzes agent behavior, identifies improvements, and provides actionable insights for better customer experience. This reduces manual costs while delivering accurate and consistent results, saving time and money.

Project Self-Realization

My personal project : An innovative AI solution guiding individuals in discovering life's purpose. Leveraging advanced AI techniques, it provides tools to explore inner selves and answer fundamental existential questions.

CERTIFICATIONS



CellStrat Certified NLP expert

AWARDS

- Super Achiever Award 2023
- The Motivator Award 2022 Q4
- 2022 Bloggathon winner at eClerx
- 3 consecutive appearance in organisation quarterly magazine.

LANGUAGES

English

Hindi

Bengali

WORK EXPERIENCE

Senior Software Engineer

Informatica, Bangalore, KA 2018-2020

- Collaborated with a top-performing team of ETL Engineers and product specialists, significantly enhancing organisational success through streamlined data processes and advanced data pipelines available at that time.
- Transformed business propositions into quantitative queries, effectively collecting and cleaning data for accurate analysis.
- Architected and developed scalable databases utilising SQL and Spark, enabling seamless ETL processes and efficient data management.
- Evaluated and optimised data pipelines, achieving enhanced workflow and increased efficiency for processing over 50 TB of data daily.

Software Engineer

Informatica, Bangalore, KA 2012-2018

- Demonstrated expertise in enterprise data management, building end-to-end pipelines, APIs, and data warehouses for large banks and DW systems.
- Successfully designed and implemented scalable data architectures, ensuring efficient data processing and storage for high-volume data sets.
- Extensive experience with Spark and other big data technologies, proficient in both streaming and batch processing methodologies.
- Developed ETL tools as a software engineer at Informatica gaining valuable insights into data processing and transformation workflows.
- Acquired comprehensive knowledge of Linux/Unix and Windows operating systems, effectively deploying and maintaining enterprise applications across diverse environments.
- Proficient in working with various databases, including MySQL, MS-SQL, Oracle, MongoDB, and AWS Redshift, enabling seamless integration and efficient data retrieval.
- Implemented multi-stage data pipelines utilizing both NoSQL and SQL large data stores, effectively resolving complex customer issues.