

## Academics

2018-2022	B.Tech in Computer Science and Engineering, Presidency University
CGPA	7.7/10.0

## Experience

Present	<b>Software Engineer - AI/ML: Microsoft COSMIC AI</b> <ul style="list-style-type: none"><li>Implemented multilayered remote <b>MCP</b> (Model Context Protocol) servers and client architecture with robust authentication and authorization, integrating <b>Agent2Agent (A2A)</b> protocol to enable interoperable communication between autonomous <b>AI Agents</b> and seamless orchestration of external services.</li><li><b>Finetuned LLMs</b> for specific tasks, optimizing performance and reducing computational overhead through 16-bit and <b>8-bit quantization</b> with fast evaluation.</li><li>Developed and deployed secure <b>RAG</b>-based <b>REST</b> Apps integrated with <b>Azure OpenAI</b> and <b>Azure Cognitive Search</b> for <b>GenAI</b>-powered responses, leveraging Azure AD, JWT, and Managed Identity for authentication, with centralized configuration management and comprehensive monitoring via Application Insights.</li></ul>	Tech Mahindra
Jan '24	<b>Machine Learning Engineer: Evidence Extraction</b> <ul style="list-style-type: none"><li>Developed a <b>Document Question Answering</b> platform using <b>Gen-AI</b> to extract key evidence from audit and compliance reports, reducing manual effort by <b>60%</b> and enhancing the accuracy of generated evidence.</li><li>Built and deployed a Question-Answering pipeline as a REST API endpoint using Django, leveraging Azure OpenAI for vectorization and <b>GPT-4</b> for responses, and using pdfplumber and <b>Azure Document Intelligence</b> for data extraction from complex PDFs.</li></ul>	Freelance
Nov '23	<b>Data Scientist: Customer Experience Intelligence</b> <ul style="list-style-type: none"><li>Developed scalable code to produce precise <b>transcripts</b> from <b>medical audio data</b>, achieving an accuracy exceeding <b>98%</b> through <b>OpenAI Whisper</b>.</li><li>Seamlessly aligned transcriptions with audio files utilizing a large-scale neural network, yielding aligned transcription files suitable for speaker diarization, sentiment analysis, and language detection.</li><li>Implemented <b>Speaker Diarization</b> using <b>NVIDIA NeMo</b>, accurately mapping speakers to corresponding sentences. Additionally, ensured <b>proper punctuation</b> in the transcriptions for clarity and readability.</li></ul> <b>Data Scientist: RT Noise Cancellation</b> <ul style="list-style-type: none"><li>Conducted comprehensive <b>training</b> on diverse datasets to ensure the robustness and efficacy of the model.</li><li>Engineered <b>Django</b> backend infrastructure to streamline audio input processing and machine learning-based noise suppression, achieving a notable <b>60%</b> enhancement in audio quality. Additionally, wrapped up the codebase as a Rest API endpoint for seamless integration with ML processes.</li></ul> <b>Data Scientist: Deloitte Neuroscience</b> <ul style="list-style-type: none"><li>Developed scalable code to analyze animations, fonts, images, word count, bullet points, and content from PowerPoint files, offering tailored improvement suggestions in collaboration with the Neuroscience Research team.</li><li>Utilized BERT for extractive summarization, OpenCV with OCR for text extraction from slide images, and Spacy and NLTK for Named Entity Recognition, improving presentation analysis efficiency by threefold and receiving positive feedback..</li></ul>	Kanini
Aug '22	<b>Data Scientist: Deloitte HOAIML</b> <ul style="list-style-type: none"><li>Orchestrated automation of file uploading processes, facilitating seamless format conversion and restoration post-OCR processing, and elevated table detection accuracy from 60% to over 95% through innovative algorithms and improved documentation clarity.</li></ul>	Kanini
Feb '22	<b>SDE Intern: Conversational Assistant</b> <ul style="list-style-type: none"><li>Led scalable backend development and seamless REST API integration with ReactJS, fine-tuned models for accurate user intent interpretation, enhancing responses.</li></ul>	Kanini

## Technical Skills

---

Languages	Python   C++   C   SQL
Frameworks	Django   Flask   FastAPI
DBMS	MySQL   PostgreSQL   MongoDB
VectorDB	FAISS   Pinecone   ChromaDB
Cloud Services/VCS	AWS   Azure   GCP
Frontend	HTML   CSS   Javascript   Typescript
LLMs	OpenAI   Gemini   Misrtal   Phi   llama   Gemma   Claude
Embedding	text-embedding-ada-003-large
RAG/Agentic	LangChain   LangGraph   Llamaindex   MCP   A2A
Transformers	BERT   T5
Python Packages	Pandas   Numpy   Matplotlib   Scikit-learn   OpenCV   NLTK   Spacy

## Activities & Leadership

---

- **OpenSource Involvement:** Successfully contributed to Khargpur Winter of Code, IIT-KGP for two consecutive years 2018 and 2019. Participated in GirlScript Summer of Code (GSSoC) 2019.