

Imon Bera

+1 267-466-7408 | ib385@drexel.edu | linkedin.com/in/imonbera13/ | github.com/invcble

Education

Drexel University

Sep. 2023 – Present

Master of Science in Computer Science, GPA: 3.85

Philadelphia, PA

Bengal Institute of Technology

Jul. 2019 – Jun. 2023

Bachelor of Technology in Computer Science and Engineering, CGPA: 9.14

Kolkata, WB

Relevant Coursework: Machine Learning, Deep Learning, Natural Language Processing, Data Structures & Algorithms, Software Design, Object-Oriented Programming, Database Systems

Technical Skills

Languages: Python, SQL, Java, C++

AI/ML Libraries: TensorFlow, PyTorch, Scikit-learn, Hugging Face, OpenAI API

Tools: Git, Visual Studio Code, PyCharm, Jira, Confluence, Swagger/OpenAPI

Cloud Platforms: Google Cloud Platform, AWS, Azure

Databases: MySQL, PostgreSQL, Firebase, MongoDB

Experience

AI Engineer Intern

Tapistro, Inc. | Jun 2024 - Dec 2024 | San Francisco, CA

- Designed and implemented an LLM Agentic RAG framework with periodic web scraping, integrating Function Calling, Code Interpreter, and JSON extractor, optimizing business processes and reducing manual work by 80%.
- Developed a fallback search engine pipeline using Bing Search API v7 and Google's Custom Search API, reducing 'no-answer' error rates by 35%.
- Fine-tuned machine learning models to enhance decision-making and client interactions, leveraging Hugging Face models through vLLM and llama.cpp, achieving 90% cost reduction compared to GPT-4o API pricing.
- Automated testing and deployments via CI/CD pipelines using GCP Cloud Run and Cloud Artifacts, ensuring scalable and reliable AI prompt integration.

Research Assistant

LeBow College of Business | Jan. 2024 – Mar. 2024 | Philadelphia, PA

- Engineered automation solutions using Python libraries such as pyautogui, Tkinter, and PyQt, enhancing accessibility and reliability in healthcare-related research workflows.
- Migrated legacy Perl scripts to Python, achieving a 70% improvement in runtime and automation reliability.

Projects

Voc-Notes: AI-Powered Lecture Transcription Tool

Drexel University | Apr. 2024 – Jun. 2024 | Philadelphia, PA

- Developed a standalone application using Google Cloud Speech-to-Text for real-time transcription and Apache Spark for data preprocessing, ensuring accurate and structured note generation.
- Orchestrated summarization and note generation using Mixtral-8x7b and Llama3-70b LLMs, leveraging PyMongo and MongoDB for efficient data storage and retrieval.

Alzheimer's Classification using OASIS Dataset

Drexel University | Nov. 2023 – Dec. 2023 | Philadelphia, PA

- Leveraged the OASIS Alzheimer's Detection Dataset to classify Alzheimer's progression using Logistic Regression and LDA models, achieving 86% accuracy for multi-class classification.
- Implemented image preprocessing techniques and ensemble modeling to enhance predictive performance.

Sentiment Analysis of Indian Political Tweets

Bengal Institute of Technology | Feb. 2023 – Jun. 2023 | Kolkata, WB

- Created an LSTM model for sentiment analysis, achieving 96% accuracy in categorizing tweets into positive, neutral, or negative sentiments.
- Utilized NLP techniques and Global Vector word embeddings for data preprocessing, ensuring dataset readiness for analysis.

Certifications

- **Reinforcement Learning**, Coursera by Google (March 2023)
- **Google Data Analytics Professional Certification**, Coursera by Google (August 2022)

Activities

- **Course Assistant**, College of Computing and Informatics (Apr. 2024 – Present)
- **CCI Dean's Fellow**, College of Computing and Informatics (Jul. 2023 – Present)
- **Active Member**, Drexel IEEE (Oct. 2023 – Present)