SHAUNAK G ROY

Chennai, India · mrpolymathematica@gmail.com · +91 6289135261 ·
https://github.com/shaunthecomputerscientist

EDUCATION

IIT Madras	India, Chennai
BS in Data Science and Applications Data Science *GPA: 8.5*	April 2022 - August 2026

EXPERIENCE

iNeuron	remote
Ai engineering intern	April 2024 - June 2024

- Worked on opensource Data Analysis application leveraging llms. (EDA GPT)

- Created new workflow called LangGroupChain that uses topological sort to break a questions into subquestions and arrange each node (question) in topological order. This graph is passed to llm for answering complex questions in multiple iteration.Implemented advanced rag combination (multiquery, ensemble, similarity with threshold) and agentic tools giving a comprehensive set of abilities to the llm for answering general questions from the internet and pdfs.

- implemented self correcting code agent that gives code response needing data analysts to code less.Used combination of supervised learning (classification) and generative ai to achieve best results.

PROJECTS

Mushify : The Music App   https://github.com/shaunthecomputerscientist/Mushify
End to End flask application with ai features (rag chatbot). Front end is written in vanilla javascript with no frontend framework used. Has proper authentication, authorization and payment system with razorpay.

EDA GPT : Your Personal Data Analysis Assistant
https://github.com/shaunthecomputerscientist/EDA-GPT
An advanced open source data analysis solution leveraging llms.

Project Jarvis : Your Personal Assistant
A multimodal ai agent that serves as personal assistant with voice interaction or interaction through terminal. This ai agent can perform web scraping, video transcription, opening or closing desktop apps and folders. It has multiple agents each specialized with specific tasks. The query is classified into specific task using classification algorithms and then sent to the respective agent.

SKILLS

Programming and Scripting Languages:	Java, Python, Javascript, C++, Bash
Web Frameworks and Databases:	Streamlit,vue,flask,fast-api, Mongo,Postgres,MySql,Sqlite, Faiss,Chroma
Data Manipulation, Analysis, Visualization:	pandas, numpy, sci py, matplotlib, seaborn, plotly
Machine Learning:	Supervised and Unsupervised algorithms, sklearn, tensorflow
Mathematics And Statistics:	Linear Algebra, Advanced Calculus, Optimization, Advanced Statistics
Other Skills:	docker, Azure, gcp, dask, web scraping, oops

COURSES

Data Structure and Algorithms Using Python
learnt to make data structures from scratch and use them in algorithms deriving everything from first principles. Graph algorithms, searching and sorting algorithms, dynamic programming and string matching algorithms are covered in the course.   8/10

DataBase Management Systems
Learnt about mathematical foundations of managing data. Learnt representation of relation data in ER diagrams, normal forms, concurrency control and b trees.   8/10

Linear Algebra, Multivariate Calculus, Statistics
Learnt about mathematical foundations of machine learning purely from mathematical rigor with proofs. 8/10

Machine Laerning Foundations
Learnt about machine learning optimization techniques. From Linear programming to convex optimization techniques like KKT and lagrange multipliers along with proofs, dual and primal and EM algorithm is covered.   8/10

Machine Learning Techniques and Practice
Course dives into both theoritical aspect of classical ml algoritms and practicals using sklearn, pandas, numpy, seaborn.   8/10