

Santanu Banerjee *Senior Software Engineer at Capgemini*

📧 iamsantanubanerjee.github.io 🌐 iamsantanubanerjee in in/iamsantanubanerjee ✉ contactsantanubanerjee@gmail.com

☎ +91 90511 41779 📍 Kolkata, India

SUMMARY

Data scientist with 2+ years of experience in applying advanced analytics, machine learning and deep learning algorithms to extract insights from complex datasets. Possesses a strong foundation in mathematics and a demonstrated aptitude for cutting-edge technologies through internships focused on quantum computing and quantum machine learning.

PROFESSIONAL EXPERIENCE

Capgemini, Senior Software Engineer

Dec 2021 – present | Kolkata, India

- **SecureNet - Risk Monitoring Inspector** (A platform to attack potential privacy concerns)
 - Data Integration and Embedding: Utilized GloVe embedding technique to convert chat and call histories into vectorized representations for efficient storage and retrieval
 - Keyword-Driven Filtering with Cosine Similarity: Implemented a Streamlit UI to map keywords provided by users to corresponding vectors in the database to filter relevant communication records
 - Flagging and Insight Generation: Leveraged the power of GPT to flag the person(s) involved in the potential leakage of confidential information from the filtered documents, providing a summary of the conversation and other relevant insights
- **Data Archival to OpenText InfoArchive**
 - Developed ETL-based data archiving solution using Python scripting to facilitate the implementation of data archival processes, ensuring long-term preservation and access to critical data
 - Collaborated with cross-functional teams to identify and prioritize data to be archived, determine appropriate storage locations and formats, maintained documentation which included policies, procedures and workflows

Aritra Sarkar (Delft University of Technology, QuTech),

Jul 2021 – Aug 2021 | Remote

QIntern (Reinforcement Learning Agent for Quantum Foundations) [🔗](#)

- The project involved using a universal AGI RL framework for simple experiments in QIT. The agents interact with the environment to model the quantum observables
- Received Third Team Award based on best project during the internship

SKILLS

Machine Learning (Linear Regression, Logistic Regression, Decision Trees, Random Forests, Ensemble Techniques - Adaboost, XGBoost, etc., Clustering - Kmeans, DBScan, etc.)

Deep Learning (Activation Functions, Optimizers, Loss Functions, Artificial Neural Network, Convolution Neural Network, Recurrent Neural Network, Transfer Learning)

Large Language Models (LLMs) (LSTM, Bi-Directional LSTM, Encoder and Decoder, Attention Models)

Languages and Frameworks (Python, SQL, PyTorch, Langchain)

Tools and Technologies (Linux, Git, SQL Server, PostgreSQL, Streamlit, Hugging Face Spaces, NumPy and other Machine Learning Libraries, OpenText InfoArchive)

CERTIFICATES & BADGES

Microsoft Certified: Azure Fundamentals [🔗](#)

Issued by Microsoft

Hands On Essentials - Data Warehouse [🔗](#)

Issued by Snowflake

Dataiku Core Designer Certificate [🔗](#)

Issued by Dataiku

Dataiku Advanced Designer Certificate [🔗](#)

Issued by Dataiku

EDUCATION

University of Kalyani, M.Sc. in Data Science

2019 – 2021 | Kalyani

- Elective - Nanoscience and Nanotechnology

Raja Peary Mohan College (University of Calcutta), B.Sc. in Mathematics

2015 – 2019 | Uttarpara

- Pass Subjects - Physics, Computer Science

CERTIFICATE COURSES

Qiskit Global Summer School, IBM

Jul 2021 | Online

- Focus on Quantum Machine Learning

Introduction to Quantum Computing, IBM and The Coding School

Oct 2020 – May 2021 | Online

- Focus on basics of Quantum Computing and Quantum Algorithms

PERSONAL PROJECTS

Quantum Pattern Matching, Paper Implementation [🔗](#)

Apr 2021

- Successfully implemented the paper "Quantum Pattern Matching" by P. Mateus and Y. Omar. Paper Link - <https://arxiv.org/abs/quant-ph/0508237> [🔗](#)