







Santanu Banerjee

Senior Software Engineer at Capgemini

 github.com/iamsantanubanerjee  [linkedin.com/in/iamsantanubanerjee](https://www.linkedin.com/in/iamsantanubanerjee)
 contactsantanubanerjee@gmail.com  +91 90511 41779  West Bengal, India
 iamsantanubanerjee.github.io

SUMMARY

Senior Software Engineer with over 1 year of experience in data management and archival projects leveraging Python, SQL, OpenText InfoArchive, and ETL solutions. Possesses a strong foundation in data science with an M.Sc. degree and a demonstrated aptitude for innovative technologies through internships focused on quantum computing.





PROFESSIONAL EXPERIENCE

Dec 2021 – present	<div>Senior Software Engineer <i>Capgemini</i></div> <ul style="list-style-type: none">Secure Communication Inspector<ul style="list-style-type: none">Developed a versatile GenAI model with fine-tuned prompts as specialized pipelines for extracting actionable insights from a diverse range of files, including text, PDFs, images, audio and video formatsEngineered an intuitive Streamlit-based user interface, empowering customers to seamlessly upload files or emails with attachments, while the system adeptly detects and categorizes sensitive and non-sensitive data, providing a comprehensive risk assessment for informed business decision-makingImplemented end-to-end deployment of the solution, showcasing expertise in data extraction across multiple file types, ensuring robust customer-centric functionality and enhancing organizational risk management capabilitiesData Archival<ul style="list-style-type: none">Developed ETL-based data archiving solution using Python scriptingResponsible for implementing data archival processes to ensure long-term preservation and access to critical data.Collaborated with cross-functional teams to identify and prioritize data to be archived, and determine appropriate storage locations and formats.Developed and maintained documentation for data archival processes, including policies, procedures, and workflows, and ensure compliance with relevant regulations and standards.
--------------------	--

SKILLS

Core Subjects	Languages	Tools and Technologies	Frameworks
Mathematics, Statistics, Data Science and Quantum Computing	Python, SQL	Linux, Git, Microsoft Office, NumPy, TensorFlow/PyTorch and other Machine Learning Libraries, OpenText InforArchive	Qiskit




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
--	--	---	---

CERTIFICATE COURSES

Jul 2021 Online	Qiskit Global Summer School <i>IBM</i> <ul style="list-style-type: none">• Focus on Quantum Machine Learning
Oct 2020 – May 2021 Online	Introduction to Quantum Computing <i>IBM and The Coding School</i> <ul style="list-style-type: none">• Focus on basics of Quantum Computing and Quantum Algorithms

PERSONAL PROJECTS

Apr 2021	Quantum Pattern Matching  <i>Paper Implementation</i> <ul style="list-style-type: none">• Successfully implemented the paper "Quantum Pattern Matching" by P. Mateus and Y. Omar. Paper Link - https://arxiv.org/abs/quant-ph/0508237 
Nov 2020	AI Learns to play Flappy Bird  <i>Self-Learning Project</i> <ul style="list-style-type: none">• Developed the Flappy Bird game on Python using PyGame that learns to play itself implementing Genetic Algorithm with NEAT-Python

EDUCATION

2019 – 2021 Kalyani	M.Sc. in Data Science <i>University of Kalyani</i> <ul style="list-style-type: none">• Elective - Nanoscience and Nanotechnology
2015 – 2019 Uttarpara	B.Sc. in Mathematics <i>Raja Peary Mohan College (University of Calcutta)</i> <ul style="list-style-type: none">• Pass Subjects - Physics, Computer Science