

Proteek Kumar Sanyal

0481709247 | proteeksanyal1994@gmail.com | <https://www.linkedin.com/in/proteek-sanyal/> | <https://github.com/proteek-dev>

PROFILE

Data Scientist with over five years of experience in Data Science and AI, specializing in developing and implementing machine learning models to deliver business value. Proven ability to scope, design, implement, and deploy machine-learning and statistical models. Proficient in Databricks, Generative AI, and developing proof-of-concept prototypes. Holds a Master of Data Science from Macquarie University, focused on delivering practical business solutions.

TECHNICAL SKILLS & AREA OF EXPERTISE

Machine Learning Algorithms: Regression, Recommendation systems, Linear & Logistic Regression, Random Forests, SVM, Naive Bayes, k-NN, Clustering (K-means, hierarchical clustering)

Deep Learning: CNNs, RNNs, Transformers, BERT, NLP, CV

Programming Languages: Python, R, PySpark

Python Data Science Libraries: NumPy, Pandas, Matplotlib, Scikit-Learn, Seaborn

Databases & Big Data: SQL, MongoDB, Neo4j, Databricks

Deployment & Production: Model deployment (AWS, Azure, GCP), Docker, REST APIs, Git, CI/CD pipelines

Data Handling: Data preparation, Feature engineering, Data visualization

Mathematics & Statistics: Statistics, Linear algebra, Optimization techniques

Soft Skills: Communication, Stakeholder management, Problem-solving, Team leadership, Mentorship

EDUCATION & QUALIFICATIONS

Macquarie University – Master of Science in Data Science 2022-2023

GLA University – Bachelor of Technology in Computer Science and Engineering 2011-2015

PROFFESIONAL EXPERIENCE

ML Engineer – Intern | Truth

July 2023 to Nov 2023

- Enhanced product robustness, scalability, and long-term usefulness.
- Completed high-level Python coding for Machine Learning models.
- Boosted design and deployment success by integrating advanced ML techniques.
- Extracted trends from large datasets, supporting business decisions.
- Integrated fine-tuned models with TestApp stream for live app deployment.
- Committed to team success by collaborating across departments to deliver practical solutions.

Project: User Liveness

- Conducted SOTA literature review on face anti-spoofing.
- Developed exploratory data analysis insights for large datasets.
- Fine-tuned MobileNet models and integrated them with the current live app.

Data Scientist | Infinite Computer Solutions (India) Limited

Aug 2021 to Dec 2021

- Analyzed structured and unstructured data for actionable business insights.
- Improved data science processes through enhanced storage, integration, and reporting strategies.
- Designed instruments and applications for optimized data collection.
- Utilized Databricks to implement scalable ML models for efficient data processing.

Project: Edison Health Link (EHL) Platform

- Developed deep learning models for medical image analysis tasks, including segmentation and detection.
- Utilized CNN architectures for feature extraction from imaging data.
- Integrated deep learning algorithms with imaging systems for real-time insights.

- Collaborated with cross-functional teams to deliver business-focused solutions.

Data Science Engineer | Tetra Soft India PVT. Ltd.

Dec 2020 to Jul 2021

- Delivered solutions using data-driven approaches and software needs insights.
- Created algorithms to handle special user and system requirements.
- Built scalable products using the latest technologies and approaches.
- Employed Databricks for efficient data processing, enabling accurate predictions.

Project: Passport 360

- Developed predictive models for assessing evacuation risk using machine learning techniques.
- Conducted feature engineering to transform raw data into informative variables.
- Presented model results through insightful visualizations, reports, and presentations.
- Implemented holistic data science solutions using innovative approaches.

Senior Software Engineer - Developer | HCL Technologies

Aug 2016 to Nov 2020

- Drafted software requirements and managed design and testing processes.
- Executed functional and integration testing and software delivery support.
- Built and tested automation tools for infrastructure provisioning.
- Explored data profiles and limitations to identify data relationships.
- Applied analytical skills to solve challenges without prior patterns.

Projects

Grasp Tool:

- Parsed documents using deep learning & OCR models with Python 3.7.
- Created Web APIs with ASP.NET and C#.
- Integrated RabbitMQ for message queuing and ELK stack for logging.
- Developed holistic data science solutions by integrating technical expertise with innovative approaches.

Voice Platform:

- Developed Microservices for web request handling and S3 bucket management.
- Implemented BERT model for sentiment analysis.
- Highlighted sentiments through accurate text-based sentiment analysis.

KEYWORDS: Data Scientist, Machine learning Engineer, AI Engineer, Statistician, Solution Architect, Developer, Innovator

INDUSTRY KNOWLEDGE: Financial Services, Insurance, Healthcare