Proteek Kumar Sanyal

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

PROFILE

Holding over 5+ years of experience, I recently graduated from Macquarie University in Master of Data Science. Knowledgeable and well-trained Machine Learning Engineering Intern with foundational industry knowledge and skill in supporting projects, managing documentation, and compiling data. Focused on advancing professional abilities and contributing to project success. Goal-oriented Machine Learning Engineer well-versed in deep learning and LLM theories and accomplished in Healthcare and Deep fake industry application. Versed in design, implementation, and scalability to achieve AI-related objectives. Specializes in Model training, APIs in Python, Cosmos DB collection and creating a small pipelines & frameworks, with deep and conventional learning models.

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

TECHNICAL SKILLS & AREA OF EXPERTISE

Data Handling: Data pre-processing & cleaning, Feature engineering, Data

Deep Learning: Convolutional Neural Networks, (CNNs) Recurrent Neural Networks (RNNs), Transformers

Deployment and Production: Model deployment on cloud platforms, Containerization (Docker), Knowledge of REST APIs for model serving.

Machine Learning Algorithms: Linear and logistic regression, Decision trees, Random forests, SVM, k-NN, Naive Bayes.

Mathematics and Statistics: Statistics, Optimization techniques, Linear algebra.

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 | Truuth   **July 2023 to Nov 2023**

- Targeted development on robusticity, scalability and long-term usefulness of products.
- Completed high-level coding in Python.
- Boosted design and deployment success by incorporating advance machine learning techniques from latest research.
- Identified trends in large datasets with application of analyses to supply chain, scientific research, and other key business areas.
- Informed business decision through successful use of both existing and emerging data science principles.

Project: User Liveness:

- Literature Review on the SOTA of face anti-spoofing.
- Conducted exploratory data analysis to uncover insights and trends in large datasets.
- Fine-tune the solution based on the current MobileNet model, test, and record.
- Integrate the new fine-tuned model with TestApp stream to the Current live app.

Data Scientist | Infinite Computer Solutions (India) Limited   **Aug 2021 to Dec 2021**

- Analysed structured and unstructured data to extract actionable business insights.
- Enhanced data science processes with better storage, integration, and reporting strategies.
- Optimized data collection by designing instruments and applications.
- Manipulated and analysed large and complex datasets using relevant techniques.
- Provided knowledge to interns about the product usage.

Project: Edison Health Link (EHL) platform:
- Developed and implemented deep learning models for medical image analysis tasks, including segmentation, classification, and detection of anatomical structures and abnormalities.
- Utilized convolutional neural networks (CNNs) architectures to extract features and patterns from imaging data.
- Integrated deep learning algorithms with existing imaging systems to provide real-time insights and recommendations during image acquisition and interpretation.
- Implemented rigorous testing procedures and validation protocols to evaluate model performance and generalization across diverse patient populations and imaging modalities.

Data Science Engineer | Tetra Soft India PVT. Ltd.   **Dec 2020 to Jul 2021**

- Developed solutions based on data-driven approaches and good understanding of current and future software needs.
- Created algorithms to handle special user requirements and system-wide needs.
- Built innovative, highly scalable products using latest technologies and approaches.
- Met short- and long-term releases cycles by working effectively within complex development systems.
- Managed and transferred knowledge to the new joiners about the product.

Project: Passport 360:
- Developed predictive models using machine learning techniques such as logistic regression, decision trees, and random forests to assess the likelihood of medical or political evacuation based on a variety of risk factors, including geographic location, geopolitical stability, health indicators, and historical data.
- Conducted feature engineering to identify relevant features and transform raw data into informative variables that capture key aspects of risk factors, such as creating geographic risk indices, encoding geopolitical events, and extracting temporal patterns.
- Interpreted model results and communicated findings to stakeholders, including insurance underwriters, product managers, and executives, by creating clear and insightful visualizations, reports, and presentations to inform strategic decision-making processes.

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.
- Maintained existing software through bug fixes and enhancements.
- Built and tested automation tools for infrastructure provisioning.
- Managed documents and workflow, following established procedures for optimum productivity.
- Explored data to understand and articulate data profiles and limitations.
- Developed and coded software programs, algorithms, and automated processes of cleanse, integrate and evaluate large datasets from multiple disparate sources.
- Identified relationships within data sets using various research techniques.

Projects: Grasp Tool:
- Parsing a document to be in a rule-based format. Used and created deep learning & OCR models to parse data from document if it's an image or text with Python 3.7.
- Created various components and with the help of RabbitMQ to keep a messaging queue structure.
- Creating Web API's in ASP.NET with C# and a keyword query API to find a word in specific or all the document uploaded by the user.
- Creating a Logging framework, to analysis the performance of the GRASP tool. Store, record, and display in case of any component failure, with ELK stack environment.

Voice Platform:
- Developed Micro Services to handle web requests.
- Handle S3 bucket and AWS transcribe services.
- Implement pre-trained BERT model in the current model for better accuracy.
- Highlight the sentiments(mood) of the user, by it statements.