Sayan Banerjee

+91 9804816854 | sayan.tubai@gmail.com | Portfolio

EDUCATION

University of Leeds	Leeds, UK
Master of Science in Data Science and Analytics	*September 2023 – December 2024*
Institute of Engineering and Management	Kolkata, India
Bachelor of Technology in Computer Science and Engineering	*August 2016 – July 2020*

TECHNICAL SKILLS

Programing Languages: Python, R, C, C++, Java, JavaScript, SQL, NoSQL.
Frameworks and Libraries: Flask, NumPy, Pandas, Scikit-Learn, Matplotlib, Seaborn, Pytorch, BeautifulSoup.
Databases: MySQL, MSSQL, SQLLite, MongoDB, Cassandra.
Developer Tools: Git, DVC, Docker, Kubernetes, Azure Portal, VS Code, PyCharm.
Knowledge Area: Data Analysis, Machine Learning, Deep Learning, Software Development.

EXPERIENCE

Graduate Research Assistant — July 2023 – October 2023
University of Leeds — *Leeds, UK*

- Utilised Generalised Additive Models and Convolution Neural Networks to fast estimate critical stellar parameters like magnetic obliquity angle, inclination angle, and physical compactness using simulated X-ray lightcurves and phase-resolved spectroscopy data generated from magpies package.
- Applied Principal Component Analysis to reduce the high-dimensional astronomical datasets, enabling efficient and accurate modeling of stellar characteristics while preserving essential data features.

System Engineer — October 2020 – August 2023
Tata Consultancy Services — *Kolkata, India*

- Engaged in Azure cloud infrastructure management and environment build activities for a manufacturing execution system web application in healthcare domain for 10+ factories.
- Collaborated with cross-functional teams to design and implement a new containerized version of the existing system using Azure Kubernetes Services thereby increasing the scalability and availability of the product.
- Created automation scripts using Python's Selenium web driver for environment build tasks ranging from installation of custom codes to modeling different manufacturing processes on the web application thereby reducing installation time by 50%.
- Developed Python scripts to connect to MSSQL databases and run scheduled stored procedures, ensuring data integrity and data processing efficiency.
- Optimized existing SSRS reports reduc ing processing time by 30% and improving overall system performance.
- Contributed to Agile Sprints by developing unit and functional tests for over 15 functionalities, accompanied by comprehensive documentation.
- Spearheaded the creation of a new hire training program for the project, leading to a 40% improvement in on-boarding efficiency.

PROJECTS

Image Caption Generator | *Python, Pytorch, NumPy, Flask* | GitHub

- Developed a web-based image caption generator using pre-trained ResNet-152 CNN encoder and an RNN decoder for generating descriptive captions for images trained on a subset of the COCO dataset.
- Achieved an average BLEU score of almost 60%, indicating close alignment to reference captions.

BigMart Store Sales Prediction | *Python, Pandas, Scikit-Learn, Flask* | GitHub

- Developed a sales prediction web application for BigMart stores and deployed it on AWS EC2 instance.
- Utilised KMeans clustering and Random Forrest Regressor achieving the highest cluster R2 score of 78%.

Algerian Forrest Fire Prediction | *Python, Pandas, Scikit-Learn, Flask* | GitHub

- Developed a web application to predict the temperature and occurrence of forest fire in the Algerian forest cover.
- Utilised Support Vector Machine Regressor to predict temperature achieving an R2 score of 65% and Random Forest Classifier to predict forest fire achieving an accuracy score of 96%.