Rajib Ganguly

FOUNDER: HTTPS://W3EARTH.LINK

Sr. Full Stack | +91 8806688030 | rajibganguly2020@gmail.com |
https://www.linkedin.com/in/rajib-ganguly-0a3031a/ | https://www.kaggle.com/w3rajibganguly

Synopsis: 14+ Years of Experience including Application development, CI/CD pipeline deployment, primarily in the areas of Machine Learning and Predictive Analytics. Experience in executing and providing data-driven solutions to complex business problems. Have an exposure to domains like Financial Banking and Insurance, eLearning, IoT etc

Experience Summary

Have 14+ years of experience in developing various applications across domains.

- Extensive experience in Machine learning algorithms.
- Strong knowledge of Exploratory Data Analysis, Building and Deploying models using different algorithms on large datasets to provide meaningful Insights.
- Experience on model performance measures to analyse and fine tune the model.
- Extensive experience in exploring & implementing the solutions based on research paper.
- Well acquainted with Agile and Waterfall methodology.
- Experience on working closely with customer stakeholders and identifying business opportunities to improve their processes.
- Experience in working with multiple stakeholders from geographically distributed teams for successful development, testing and deployment.
- Extensive experience in providing technical inputs to RFP/RFI and data analytics proposals for business growth.
- Experience on performing multiple roles in the team - Onshore Coordinator, Offshore Manager, Lead, Reviewer, Quality Anchor.
- Experience in managing multiple projects and mentor team members.

Data Science Skills

- Statistical Modelling, Hypothesis Testing, Exploratory Data Analysis, Predictive Modelling, Supervised & Unsupervised Machine Learning, Regression Techniques (Linear, Polynomial, Support Vector Regression, Decision Tree, Random Forest), Dimensionality Reduction Techniques (PCA, FA), Classification Techniques (Decision Tree, Random Forest, Neural Network, Logistic Regression, Linear Discriminant Analysis, K-Nearest Neighbors, Naïve Bayes, Support Vector Machine, Kernal SVM), Data Mining, Model Performance Evaluation, Clustering (Hierarchical, K-Means), Market Basket Analysis/Association Rule, Text Mining (Natural Language Processing), Time Series Forecasting, Deep Learning (Artificial Neural Network, Convolution Neural Network), Edge Detection, Computer Vision, OCR, Mask R-CNN, Yolo Framework, Reinforcement Learning, Cross Validations, Model Tuning (Bagging, Boosting), SMOTE, Operation Research, Optimization, Visualization Dashboard Development.
- Python, NumPy, Pandas, Matplotlib, Seaborn, Plotly, ScikitLearn, Keras, TensorFlow, OpenCV, scikit-image, Tesseract OCR, NLTK, Pillow, Google Colab, R, R Shiny, Google Analytics, SPSS, MatLab, Django, Javascript, SQL & MS Access Databases, Tableau, Power BI, KNIME, Hadoop, Spark, Microsoft Azure, AWS, GCP, .NET, Visual Basic.

Academic and Professional Qualifications

- Bachelor of Technology in **Mechanical Engineering** in **2004** from **Indian Institute of Technology - Roorkee (IIT- Roorkee), India**.
- Post Graduate Program in **Business Analytics and Business Intelligence** in **2018** from **Great Lakes Institute of Management** in academic collaboration with **Stuart School of Business, Illinois Institute of Technology (IIT), Chicago**.

Certifications / Professional Awards

- Microsoft Certified "Azure Data Scientist Associate" in July 2020
- "Tableau Data Scientist" badge issued by Tableau in Apr 2020
- Certified "Computer Vision with OpenCV and Deep Learning" from Udemy in Oct 2020
- Certified "Power BI Masterclass" and "AWS Cloud Basics" from Udemy in Apr 2020
- SAFe certified "DevOps Practitioner" in Dec 2018
- Infosys "Insta Award" in Jan 2018 and Dec 2018
- "Exceptional Performance Award" from a Leading Aircraft Manufacturer in 2012
- Infosys "Best Innovator Award" in 2010
- Infosys "On the Spot Award" in 2008
- Certified "mini MBA program" from SIES College of Management, Navi Mumbai in 2004
- "Vidya Dutta medal" in 2002 and "Dr. K.M. Chandra award" in 2001 from IIT – Roorkee

Total Work Experience: 16 Years 7 Months

1. ForecastEra Technologies Pvt Ltd   **Duration:** Nov 2020 to Till Date
2. Hexaware Technologies Limited   **Duration:** March 2020 to Nov 2020
3. Infosys Limited   **Duration:** August 2006 to March 2020
4. Ispat Industries Limited (Currently JSW Steel)   **Duration:** July 2004 to August 2006

Project Experience

1. **Opportunity Score and Engagement using ML:** This project is to develop a product which provides opportunity score for open opportunities and sales representative engagement with his/her contacts based on opportunity & engagement related data.

   Domain: S&OP

   - Provided technical inputs and prepared solution approach.
   - Performed exploratory data analysis to get cleaned data.
   - Developed machine learning models and evaluated them using model evaluation metrics.
   - Performed clustering techniques for profiling.
   - Fine tuned the models to improve the result.

   Tools & Techniques: Python, R, Classification and Clustering techniques

2. **Document Classification using NLP:** Client is a leading asset management firm in USA. This project is to create an intelligent document identification and classification framework using NLP and cognitive technologies to automatically classify documents into various categories.

   Domain: Asset Management

   Role & Responsibilities: Data Scientist
   - Provided technical inputs and prepared solution approach.
   - Utilized OCR techniques to extract text from document images.
   - Performed basic text mining and data pre-processing steps to convert unstructured text data into structured data.
   - Developed machine learning classification models and evaluated them using model evaluation metrics.
   - Fine tuned the classification models to improve the result.

   Tools & Techniques: Python, Tesseract OCR, NLTK, Decision Tree, Random Forest, LDA, Azure Blob Storage, Azure ML

3. **Resource Demand Forecasting using Deep Learning:** Client is a Chief Operating Officer of leading IT Service company in India. This project is to forecast resource demand based on business unit, skill, location and grade using deep learning techniques.

Domain: HR Analytics

Role & Responsibilities: Data Scientist

- Involved in workshop with various stakeholders to collect and understand historical data.
- Provided technical inputs and prepared data flow diagram for complete solution.
- Performed data validation, data analysis and prepared data for model building.
- Developed deep learning based LSTM models for multivariate time series data and evaluated them using model evaluation metrics.
- Fine tuned the LSTM models to improve the result.

Tools & Techniques: Python, Exploratory Data Analysis, LSTM, Time Series Forecasting, Optimization

4. **Segmenting the Mining Rocks using Deep Learning:** Client is a leading international mining company in USA. This project is to segment the rocks (which are piled up after the explosion/blast happens in mines) as per specification using Image/Video Analytics & deep learning techniques.

Domain: Mining

Role & Responsibilities: Data Scientist

- Involved in workshop with Client stakeholders and provided technical inputs for complete solution.
- Involved in collecting the real time streaming data and converting into images (frames).
- Performed image preprocessing, image augmentation and image labelling/annotation.
- Developed image segmentation models based on training dataset and evaluated them using model evaluation metrics on train/test datasets.
- Fine tuned the image segmentation models.

Tools&Techniques:.com Python, OpenCV, scikit-image, Pillow, Google Colab, Edge Detection, Clustering, R-CNN, Yolo Framework

5. **Predictive Analytics for Ship Main Engine:** Client is one of the reputed Ship manufacturer in India. The main purpose of this project is to develop Artificial Intelligence (AI)/Machine Learning (ML) application to enable condition monitoring of OPV (Offshore Patrol Vessel) class vessel through anomaly detection, diagnosis, prognosis and predictive maintenance of the Ship main engines.

Domain: Shipping

Role & Responsibilities: Data Scientist

- Involved in RFP/RFI and provided technical inputs to win the project.
- Involved in technical discussions with Client stakeholders to set up the process and kick start the project.
- Studied the available data of ship main engine system.
- Performed exploratory data analysis before and after maintenance events.
- Performed Diagnostics and developed Prognostic models with Key Performance Indicators.
- Mentored the team members on AI/ML application functionalities.

Tools & Techniques: Python, ScikitLearn, NumPy, Pandas, Matplotlib, Seaborn, Plotly, Django, PostgreSQL, Exploratory Data Analysis, Random Forest, Support Vector Machine, Decision Tree, Linear Regression, Optimization

6. **Warranty Analytics:** Client is a leading Automotive manufacturer in Europe. This project is to develop warranty cost forecast model which can be used by after-sales services team for financial planning.

Domain: Automotive

Role & Responsibilities: Data Scientist

- Involved in workshop with Client stakeholders and set up the process to kick start the project.
- Performed data analysis and visualization.
- Developed warranty cost prediction model to forecast warranty cost for upcoming years.
- Validated the forecasts and fine tuned the model.

Tools & Techniques: Python, AWS, Exploratory Data Analysis, Random Forest, Neural Network, Time Series Forecasting, Exponential Smoothing, Optimization

7. Predictive Analytics of Engine Components: A project to develop Diagnostic/Prognostic

models to determine the condition of tractor engine components and indicate the need for maintenance which helps saving on operational costs, maintaining optimal fuel economy and planning for service/maintenance before failure.

Domain: Automotive

Role & Responsibilities: Data Scientist

- Involved in proposal discussions and provided technical inputs to win the project.
- Performed data analysis and feasibility study of Diagnostic/Prognostic models of engine components.
- Applied clustering on the data for profiling.
- Developed Diagnostic/Prognostic models to predict the health of engine components.
- Developed dashboard to display Key Performance Indicators (KPIs).

Tools & Techniques: Python, AWS, Exploratory Data Analysis, Clustering, Convolution Neural Network, Neural Network, Random Forest, Optimization

8. Supply Chain Management (SCM) Solution: Client is a leading Aero Structure manufacturer in

USA. This project is to develop Supplier Engagement Effectiveness solution to automate manual steps performed by SCM group to realize Planning, Production and Shipping bottlenecks from its vendors at an rajibganguly2020@gmail.com

Domain: Supply Chain, Aerospace

Role & Responsibilities: Data Scientist/ Project Manager

- Developed a centralize Data Management system to manage supplier data.
- Developed a framework to perform following tasks:
  - ✓ Calculate machine/process efficiency and perform capacity assessment.
  - ✓ Generate forward/backward scheduling along with backlog data and capture current status of the part.
  - ✓ Simulate various parameters which reduce backlog and find out recovery strategy to offload the parts.
  - ✓ Information about delay in delivery, stock and daily consumption.
  - ✓ Generate reports and KPIs
- Prepared user manual and documentation.

Tools & Techniques: Python, Tableau, Natural Language Processing, Regression, Exponential Smoothing, .NET, MS Access, Visual Basic, Optimization

9. Persistency Prediction Tool: A project to ascertain profile of persistent & non-persistent customers

through survey of life insurance policy holders across Tier 1, 2 & 3 Cities and compare them with persistent & non-persistent customers to predict factors that can flag potential non-persistent customers at the time of policy log-in.

Domain: Insurance

Role & Responsibilities: Data Scientist

- Collected the survey data through questionnaire for insurance policy holders across Tier 1, 2 & 3 Cities.
- Performed data analysis, data cleansing and data preparation.

- Developed the model to predict the factors influencing the persistent (or non-persistent) at underwriting stage.

Tools & Techniques: R, Tableau, SPSS, Logistic Regression, Linear Discriminant Analysis

10. Telecom Customer Churn Prediction Assessment: Customer Churn is a burning problem for Telecom companies. In this project, data of postpaid customers with a contract is provided which has information about the customer usage behavior, contract details, payment details and customers who cancelled their service. This project it to develop models based on the past data to predict whether a customer will cancel their service in the future or not.

Domain: Telecom

Role & Responsibilities: Data Scientist

- Performed data analysis, data cleansing and data preparation.
- Developed Classification models and compared those models using Model Performance metrics.
- Interpreted the result from the best model and provided the recommendation.

Tools & Techniques: R, Logistic Regression, KNN (K-Nearest Neighbors), Naïve Bayes

11. Loan Purchase Model Solution: This project involves developing Loan Purchase Model to classify the customers who have a higher probability of purchasing the loan and to devise campaigns with better target marketing to increase the success ratio with a minimal budget.

Domain: Banking

Role & Responsibilities: Data Scientist

- Performed data analysis and applied clustering technique on the data.
- Developed models, interpreted the model outputs and fine tuned the model.
- Performed the model performance measures and provided the Insights.

Tools & Techniques: R, Clustering, Random Forest, Classification and Regression Tree (CART)

12. Product Service Management: This project is to develop regression model for Marketing Research Company to predict Customer Satisfaction based on Purchase, Marketing, Post Purchase and Product Position variables.

Domain: Marketing

Role & Responsibilities: Data Scientist

- Performed exploratory data analysis, cleansing and data preparation.
- Performed dimensionality reduction technique (Principal Component Analysis or Factor Analysis).
- Developed Multiple linear regression model and Validated the model.

Tools & Techniques: R, Linear Regression

13. Hypothesis Testing: A project to perform Hypothesis testing to test the effectiveness of the new incentive payment scheme for sales force.

Domain: Insurance

Role & Responsibilities: Data Scientist

- Formulated and presented the rationale for a hypothesis test.
- Provided descriptive statistical summaries of the data for each model.
- Analyzed the data to provide the hypothesis testing conclusion.

Tools & Techniques: R, Hypothesis Testing

14. Test Correlation Solution: Client is a leading Aircraft manufacturer in USA. This project is to develop Test Correlation solution to process, analyze & correlate the Test & FEM (Finite Element Method) data generated by sensors placed on structural components of Aircraft.

Domain: Aerospace

Role & Responsibilities: Onshore Project Coordinator/ Data Scientist

- Managed Client relationship for **6+ years** during onsite deputation. Direct contribution of **5 MUSD** worth projects.
- Involved in proposal discussions and provided technical inputs to win various projects.
- Involved in technical discussions with Client stakeholders to kick start various projects.
- Analyzed the issues raised by the Client representatives and provided required solutions.
- Coordination and knowledge shared with the Offshore Project team.
- Reviewed offshore deliverables and deployed on Client system.

Tools & Techniques: R, Tableau, Python, Neural Network, .NET, MatLab, CAE tools, Visual Basic, Optimization

15. Structure Analysis and Design: Client is a leading Aero Structure manufacturer in USA. The main purpose of this project is to develop Static and Fatigue Strength analysis of beams, spars and attachments in Aircraft.

Domain: Aerospace

Role & Responsibilities: Data Analyst/ Stress Analyst

- Extracted data from CAE and CAD tools to analyse Aircraft structures using customer processes/methods.
- Validated the analysis using hand calculation and Finite Element Analysis.
- Prepared the analysis report and submitted the status reports to the management on an ongoing

- Mentored the Analysts throughout the project lifecycle.

Tools & Techniques: Excel, Visual Basic, MatLab, CAE and CAD tools, Optimization

16. Predictive Maintenance and Process Optimization: This project is to perform Predictive maintenance, process monitoring & optimization, Six Sigma & Total Productive Maintenance in Sponge Iron Plant.

Domain: Steel, Sponge Iron Plant

Role & Responsibilities: Data Analyst/ Sr. Engineer

- Performed process monitoring, optimization and efficiency calculation for rotary equipment.
- Performed condition monitoring to check the vibration pattern of rotary equipment and finding the anomalies.
- Implemented Six Sigma DMAIC method for "reduction of downtime between bellow failures".
- Prepared reports on different maintenance types e.g. Preventive maintenance, Predictive maintenance, Shutdown maintenance, Breakdown maintenance.

Tools & Techniques: Excel, SCADA system, SAP (PM module), Optimization