

Rajshekhar Mukherjee (Raj)

Lead Data Scientist (Lead Consultant)
Data Science and AI Group
North American Division
December, 2022



MS in Computer Science and AI
HQP Scholarship Recipient
(Full Scholarship for MS)
University of Ottawa



Data Science SME, Machine Learning Engineer, AI Expert



Rajshekhar Mukherjee (Raj)

Title : Lead Data Scientist
Previous Companies : CGI, Synopsys, Merck, Cigna, SOTI, RBC-Moneris [F100 & F500 Companies]
Education : MS from University of Ottawa (with Full Scholarship)
Email : rajshekhar.mukherjee@gmail.com

Focus Areas

- Data Science, Artificial Intelligence, Machine Learning
- Predictive maintenance, Time series forecasting & Analytics Subject Matter Expert
- Business Intelligence, Finance & Healthcare Data Insights Experience
- Azure, AWS Cloud based ML Solution Delivery Experience
- Computer vision, NLP, Time-series forecast & end-to-end ML Model deployment Solutions
- Dashboard and Analytics platforms Expert
- Natural Language Processing and Text Analytics Experience

Selected Projects

- **Data/AI Transformation:** \$500,000 Merck (US Healthcare) Recommendation System & AI analytics system, \$200,000 Cost saving using Predictive Models for AWS Server Utilization for SOTI
- **Cloud Based ML Prediction: Predictive maintenance** and **ML models** customized to large scale time series data, analyzed insights with closed loop feedback with Supply chain team to model various trends & outliers in past data, improve forecasting results by a margin of ~20% over previous baseline; done on IBM Watson cloud
- **Major AI/ML Capability Addition** : Asset Health – Consuming SME & Data knowledge, building a deep learning based system for an automated Analytics platform to monitor & predict site's asset health using features & incoming data for 3 sites, extended for next phase including 15 sites data using AWS with more than 1TB of data

Professional Experience

A Lead data scientist with 9+ years of work experience in machine learning and data science; Fluent with Python, Linux, SQL, Tensorflow, PowerBI, Cloud; strong Statistics, ML background; strong consultancy background; worked with CGI Inc, Synopsys R&D and SOTI Inc. with experience in leading teams & shipping end-to-end ML/DS projects to production with Cloud, NLP, Predictive forecast, Segmentation, Big data stack; Collaborated with Business Intelligence, Software development, Supply Chain & Management teams. Experience with strategically leading teams and satisfying client's project needs, adapting to new feature(s) and/or architectural/timeline demands, in the Data Science/AI/ML space. Double MS degrees (from Univ. of Ottawa, Canada & Univ of Delhi, India)

Technical Concentration:

Data Science / Machine Learning Stack

- **Software:** Python, PyCharm IDE, Tableau, Power BI, C, SQL, Flask/Django app, Cloud native-Docker/Docker-compose
- **ML/DL skills:** CNN/LSTM, Tensorflow/Keras libraries, Supervised/Unsupervised learning, Time series Forecasting, Statistical modelling, fb Prophet model, Seeq, Databricks
- **ML/DL Cloud:** GCP, Kubernetes, Azure ML, AWS, Databricks, PySpark, Apache Scala, GCP Data Proc

Engagement Delivery Experience:

- International healthcare & eHealth AI/ML projects experience – US and Canada.
- **Planning, leading team and developing** end-to-end CI/CD architecture for various large scale projects in ML/Data Science domain - ranging from **forecasting, demand planning** to **NLP** and Cloud deployment.
- **Extensive consulting experience** in developing **strategic** relationships and long term **vision** with various **clients**, building proposals, **winning** various client projects based on **POC** projects. **Communicating** project **goals** to **Business units**, Higher management and winning **long-term associations**.
- Developing **architecture** and end-to-end projects for **continuously ingesting** & combining **internal** & **external** data sources, training old model regularly and inference on a weekly basis using trained ML based on **Vision, NLP** or **Time series** models.

Project Highlights & Organizations worked for

- Citibank, US :

- End to end ML model using NLP and Time-series forecasting model for predictive analysis ; deployed in IBM Watson, Data ingestion through SQL database, feature selection, python ML forecast model, connected to a Tableau dashboard refreshing daily data

- Merck, US :

- Asset Health – Consuming SME & Data knowledge, building an automated Recommendation system for an Advanced Analytics platform to monitor each site's asset health, feature identification, Advanced dashboards & graphs for 3 sites , extended for next phase to include 15 sites worldwide

- Evicore/Cigna, US :

- ML model for Cancer Regimen Cost prediction ; Python ML classification models, error metrics, class imbalance, multi-collinearity, feature selection, extended to include more projects e.g. NLP analysis of patient's claims data

- Bell Canada :

- Proposal for Predictive analysis to predict phone sales of various segments in stores, and helped identify new store locations given past sales and customer PII data

Projects

- Royal Bank of Canada :
 - Build a forecasting model and a recommendation engine refreshing with daily incoming data, deployed model using Docker and Azure pipeline
- RBC Marketing:
 - End to end ML Forecast and Segmentation model for Loan advancement to both old and new customers- included unsupervised, supervised and time series forecast model building on Big data platform using Apache Spark and Azure Databricks

My Websites :

- <https://rajshekharm.github.io/>
- <https://rajshekharmukherjee.wordpress.com>
- <https://medium.com/@rajshekharmukherjee>

