ARKADEEP SARKAR

47/6 Matangini Hazra Sarani, Aurobindopally, Siliguri, West Bengal, India – 734006
Phone: +91 9749574238/9051443280 | Email: arkadeep1994@gmail.com

EDUCATION

- **Indian Statistical Institute, Kolkata**   *July 2019*
  Master of Technology in Quality, Reliability and Operations Research   Score: **79%**
- **Jadavpur University, Kolkata, India**   *May 2016*
  Bachelor of Engineering, Mechanical Engineering   CGPA: ***9.24/10***
- **Siliguri Baradakanta Vidyapith, Siliguri (**Standard XII**)**   Score: **90.2%**   *May 2012*
- **New St. John's High School, Siliguri (**Standard X**)**   Score: **90%**   *May 2010*

RELEVANT SKILLS

- Business Intelligence: **Spotfire, Tableau, Plotly**
- ETL : **SQL , Spark, Informatica Powercenter Designer, Workflow manager, Monitor**
- ML tools/libraries : **Python, AWS Sagemaker**, **Pandas, PySpark, NumPy, ScikitLearn, PyTorch**
- Machine Learning concepts: **NLP, Supervised learning, Unsupervised learning , Deep Learning**
- Mathematical Concepts: **Multiple Linear Regression, Hypothesis Testing, ANOVA, Probability & Statistics**

PROFESSIONAL EXPERIENCE

Data Scientist at GE Healthcare, Bangalore, India   *August 2019-Present*
Department- ***Data and Analytics***
Major projects and accomplishments:

- ➢ VOC Sentiment Analysis:
  - Was part of VOC (Voice of Customer) *Sentiment Analysis* Project. Got hands on experience in extracting the Sentiment from Customer comments obtained via Surveys using VADER sentiment analysis tool in Python.
  - Tasks included preprocessing of the Comments e.g*: Stopwords removal, lemmatization, valence word detection, calculation of sentiment score* etc.
  - Suggested additions to the lexicon dictionary in order to improve scores for *False Positive and False negative cases*.
  - The outcome of the analysis was presented to the internal stakeholders using *Spotfire* dashboards
- ➢ MDM Spotfire Dashboard:
  - Developed a dashboard for Master Data Management (MDM) Project in *Spotfire*.
  - Was involved in building Information Links to connect to Teradata and Redshift tables
  - Actively collaborated with end users to understand their pain-points and implement the solutions
  - Recipient of '*Impact Award*'
- ➢ Install Base Asset Consolidation ETL :
  - Worked on ETL development tasks on Redshift tables and gained proficiency in *SQL*.
  - Actively collaborated with stakeholders to understand their requirements as per the User stories
  - Involved in developing *Informatica PowerCenter* mappings and running workflows and monitoring them for any anomalies and solving them.

Assistant Manager at **Tata Metaliks Ltd,** Kharagpur   *July 2016-July 2017*
Department- ***Captive Power Plant***
  - ➢ Was responsible for overseeing day-to-day and maintenance activities at the Captive Power plant

OTHER ACHIEVEMENTS AND PROJECTS

- COVID-19 prediction dashboard:
  - Developed a dashboard from scratch for predicting the total number of Deaths, confirmed cases etc for different states of India using the SEIR model of epidemiology using *Scikit-learn* on *Python*
  - Interactive visualizations were developed using *Plotly-Dash.*
- AWS Sagemaker POC:
  - Successfully carried out various test cases such on AWS Sagemaker for the Data Science lab setup.
  - Got exposure to various components of AWS Cloud (AWS Glue, Step Functions, EC2, EMR etc)
- Completed training on '*Spark Basics and Streaming*' and '*Spark: PySpark*' from GreatLearning.

SUMMER INTERNSHIP (2018)

General Electric Aviation, Bengaluru   *May 2018-July 2018*

- *Engine component Damage Prediction Model:*
  The requirement was to model the system, in order to predict the damage in an engine component using sensor data of flight cycles of engines.
    - Initially had to deal with **feature engineering** to choose the most important features out of the 70 explanatory variables after data cleansing by a voting scheme.
    - Two approaches were used: 1) Multiclass classification approach, 2) Regression approach.
    - For Regression approach, **Multiple Linear regression** with **Lasso** and **Ridge** was considered.
      For Classification, **Naïve Bayes, SVM, Logistic regression, Random forest Classifier** etc were considered.

ACADEMIC ACHIEVEMENTS AND PROJECTS

Indian Statistical Institute, Kolkata

- *Analysis of Time series data of Potato prices:*
  Decomposing a time series of weekly average prices of potatoes in different Indian towns/cities using Box-Jenkin's Method (**SARIMA**) to study the nature of variation and hence apply the forecasting techniques using R.
  Guide: Prof. Amitava Banerjee

EXTRA-CURRICULAR ACTIVITIES

- Placement representative for QROR Batch(2017-2019) at ISI Kolkata
- Guitar player (Intermediate)
- Completed Outbound Leadership Course (10 Days) organised by Tata Steel Adventure Foundation at Kuflon, Uttarkashi; Reached Siyari Top (11800 ft)
- Interested in Singing, theatre, travelling

REFERENCES

- ❖ URVASHI DIVYAM, Staff Technical Project Manager, GE Healthcare, Contact no: 9591622266, email: Urvashi.divyam@ge.com
- ❖ AMITAVA BANERJEE, Professor, SQC-OR department, Indian Statistical Institute, Kolkata, email_id:amitava_banerjee2001@yahoo.com