

# MAMATA ANIL PARAB

North Chicago, IL | 806-853-1089 |

[parabmamataanil@gmail.com](mailto:parabmamataanil@gmail.com) | [linkedin.com/in/mamata-parab](https://www.linkedin.com/in/mamata-parab) | [github.com/MamataAnilParab](https://github.com/MamataAnilParab)

## TECHNICAL SKILLS

---

**Programming Languages:** Python, SQL, R   **Tools & Technologies:** Pipeline Pilot, SharePoint

**Big Data & Analytics:** Spark SQL, Data Mining, Statistical Analysis, Image Processing

**Data Visualization:** Spotfire, Tableau, Matplotlib   **Cloud Platforms:** Amazon SageMaker

**ML/Algorithms:** CNN, XGBoost, ResNet50, Linear Regression, Decision Trees, Random Forest, SVM, Logistic Regression, Canny Edge Detection

## PROFESSIONAL EXPERIENCE

---

### AbbVie (North Chicago, Illinois) – Data Scientist

*Aug 2022 – Present*

- Spearheaded the redesign and development of ACES Lite, a lightweight, on-demand data visualization platform built in Spotfire, by collaborating with senior leadership to modernize the legacy ACES system; implemented dynamic data loading to enhance performance and ensure data privacy, leading to a 40% increase in user engagement and enabling the detection of 26 previously unrecognized anomalies, while also launching a SharePoint support hub and driving adoption through beta testing and user-driven enhancements.
- Executed the development of an anomaly detection solution during a competitive hackathon by collaborating with a cross-functional team to preprocess 60,000 labeled images and extract features using Canny edge detection and ResNet50; trained 8 binary classification models with CNN and XGBoost on Amazon SageMaker, achieving an average accuracy of 89%, including 92% accuracy for debris detection, delivering a high impact and scalable solution.
- Implemented Python scripts to replace legacy Pipeline Pilot protocols used for database creation and data transformation workflows, addressing performance bottlenecks and frequent failures; enhanced integration with modern technologies resulting in faster execution, simplified maintenance, and improved system reliability.
- Automated a manual monthly reporting workflow by partnering with the Pre-Approval Access (PAA) team to consolidate third-party data from sources like Knowledge Notebook, Central Lab Testing, and Drug Shipment records; performed data harmonization and developed a Spotfire dashboard that processes extracted CSVs from a centralized location, enabling scalable analytics and reducing reporting time from hours to minutes
- Designed and implemented a Python pipeline to extract Pharmacokinetic and Plasma Concentration tables from hundreds of unstructured PDFs of Clinical Study Reports (CSRs), identifying key parameters and storing output in a centralized database, replacing a manual, error-prone process with a reliable solution for downstream analysis.

### Consortium for Research and Analysis (Lubbock, TX) – Graduate Assistant

*Sep 2021 – Jan 2022*

- Leveraged Spark SQL to efficiently extract, aggregate, and process over 16 million data points from distributed CSV files and developed interactive performance and revenue dashboards using Tableau and Python (Matplotlib), reducing data processing and analysis time by 45%.

### Terna Hospital and Research Center (Mumbai, India) – Clinical Data Analyst

*July 2019 – May 2021*

- Translated complex clinical requirements into mathematical models and applied multivariate analysis using supervised machine learning algorithms like Linear Regression, Decision Tree, Random Forest, and XGBoost on a dataset of 4 million patients, enabling physicians to accurately determine optimal anesthesia dosage.
- Conducted model selection and performance evaluation using 5-fold cross-validation, effectively reducing overfitting and improving test accuracy by 3%.
- Integrated the machine learning model into the clinical workflow, optimizing system scalability and performance, which improved anesthesiology decision-making and enhanced patient safety outcomes.

### Four Eyes Research Center (Pune, India) – Biomedical Analyst

*June 2018 – May 2019*

- Implemented Python automation scripts to optimize temperature control and fluid mixing during abdominal cancer surgeries, ensuring precise regulation and preventing burns, thus enhancing patient safety.

- Integrated real-time audio-visual alarm systems into the device to detect abnormal conditions, enabling immediate alerts to medical staff and reducing response time by 70%, thereby minimizing risk and enhancing clinical safety.
- Designed and optimized SQL queries to manage large-scale data, deploying the solution on Google Cloud to ensure high-performance data handling, scalability, and seamless integration with analytics workflows.

#### **Ninad Research Lab (Mumbai, India) – Graduate Research Trainee**

*July 2015 – Jan 2017*

- Acquired, synthesized, and mined relevant datasets for projects including breast cancer detection, RBC classification, and climate trend analysis for weather forecasting, using both supervised and unsupervised machine learning algorithms.
- Evaluated model performance using metrics such as Confusion Matrix, F1 Score, RMSE, and Cross-Validation to ensure accuracy, robustness, and generalizability across datasets.

### **EDUCATION**

---

**Master of Science in Data Science | GPA – 3.89**

*June 2021 – May 2022*

**Master of Technology | GPA – 3.9**

*Aug 2017 – May 2019*

**Bachelor of Engineering | GPA – 3.48**

*Aug 2011 – June 2015*

### **ACADEMIC PROJECTS**

---

**Detection of presence of Escherichia-Coli in water [Python]**

*June 2017 – Oct 2017*

- Devised a system for water quality analysis leveraging supervised machine learning algorithms like logistic regression, SVM, and decision trees to detect the count of Escherichia-Coli in drinking water with an accuracy of 91.42%.

### **PATENTS AND PUBLICATIONS**

---

- Parab, M.A., Mehendale, N.D.: "Red blood cell classification using image processing and CNN." bioRxiv (2020).
- Parab, M.A, Mehendale, N.D.: "Face Recognition-Based Automatic Hospital Admission with SMS Alerts." SN Computer Science (2021).