

TECHNICAL PROFICIENCIES

- ✓ **Programming languages:** Python (Flask, FastAPI), C++, SQL
- ✓ **Machine Learning:** **Supervised Learning** (Linear/Logistic Regression, Decision Tree, Random Forest, SVM, KNN)
Unsupervised Learning (K-Means Clustering, Market Basket Analysis),
Ensemble Models (Random Forest, XGBoost, LightGBM, VotingClassifier, StackingClassifier)
- ✓ **GenAI Technologies:** Prompt Engineering, LLMs, Langchain, Agentic AI (BrowserUse + Playwright)
- ✓ **Cloud Services:** Azure Functions, Azure Cosmos DB, Azure Databricks, Azure Event Grid, Azure Logic Apps
- ✓ **Data & Processing:** PySpark, NumPy, Pandas, Matplotlib, Seaborn, scikit-learn
- ✓ **Model Optimization & Evaluation:** GridSearchCV, RandomizedSearchCV, Cross-Validation, Confusion Matrix, ROC-AUC
- ✓ **Automation & Tools:** Power BI, GitHub, GitHub Actions, Postman
- ✓ **OCR & Web Scraping:** OpenCV, BeautifulSoup

EDUCATION

B.Tech
National Institute of Technology Silchar
07/2019 – 05/2023 GPA: 8.68

12th standard (CBSE Board)
Gurukul Grammar Senior Secondary School
03/2016 – 03/2018

CERTIFICATIONS

- [AI-102](#): Microsoft Azure AI Engineer Associate Certification
- [AI-900](#): Microsoft Azure AI Fundamentals Certification
- [AZ-900](#): Microsoft Azure Fundamentals Certification
- [Google Data Analytics Certificate](#)
- [SQL Certificate - HackerRank](#)
- [Python Certificate - HackerRank](#)

WORK EXPERIENCE

Optum Global Solutions Private Limited | Gurugram SOFTWARE ENGINEER | 09/2023 – Present

PROJECT: Contact Insights | 09/2023 – 10/2024

- **Automated the data flow pipeline** for Optum Bank's Customer Support data. The automated system consisted of loading Customer Support chats from **Azure Databricks**, extraction of relevant data using **Python** and **SQL**, tagging chats into pre-defined categories by building a **XGBoost multi-class classification model** with **GridSearchCV** tuning, and storage to **Cosmos DB**. Deployed into an **Azure Function** pipeline, and scheduled a Databricks job to run at regular intervals, to ensure that Cosmos DB is consistently updated with latest data. The automated system improved tagging accuracy by approx. **12%** over manual efforts.
- **Implemented a real-time automated Evaluation Workflow** by training **Python**-based **XGBoost regression models** to assign quality scores (Accuracy, Relevance, Fluency, Helpfulness) to individual chatbot interactions, as well as predict Coherence and Resolution Effectiveness scores for the entire conversation, triggered via **Azure Functions** and **Azure Event Grid** for event-driven processing, with results stored in Cosmos DB to enhance the quality of our chatbot responses.
- **Developed an automated, on-demand Newsletter Generation system** using **Azure Functions**, **Python APIs** and **Prompt Engineering** to provide summarized insights into Customer Support issues, evaluate resolution effectiveness, propose strategies to address root causes, and incorporated visualizations (e.g., pie charts, bar graphs). Enabled dynamic update of Newsletter with selected chats to provide dynamic insights, thus increasing stakeholder engagement by approx. **80%**.

PROJECT: AI Advancements | 11/2024 – Present

- **Automated FHIR Test Data Creation** by developing a solution using **Python**, **Azure Functions**, and **Prompt Engineering** to generate FHIR bundles including key healthcare resource-types (Patient, Condition, Encounter, Appointment, Observation, Service Request, Medication Request, Allergy Intolerance). The API allows test data creation with **dynamic customization** by providing the ability to override any desired parameter with user-provided values. This solution significantly accelerated QE test cycles by reducing manual efforts for test data creation by approx. **95%** (from 45 minutes to 3 minutes per patient).
- **Automated UI Browser Testing** by utilizing an **Agentic AI** framework with **Langchain**, **BrowserUse**, and **Playwright**. Leveraged **Prompt Engineering** to dynamically generate Playwright automation functions (requiring only slight modifications to accommodate specific UI elements) based on test cases, and integrated with the **Langchain model** (AzureChatOpenAI). Utilized BrowserUse to instruct the AI Agent on executing test cases, with Playwright serving as a fallback mechanism to extend the agent and facilitate custom function calls. This solution increased automated test coverage by approx. **60%**, reducing manual efforts by the QE team.

SarvM.AI System Private Limited 🔗

DATA SCIENCE INTERN | 07/2022 – 01/2023

- **Recommendation Algorithms:** Analyzed and optimized recommendation algorithms for food commodities.
- **Delivery Optimization:** Determined the proper order in which packages should be allocated for delivery. Implemented K-Means clustering to group delivery points based on geographical proximity, and utilized the 2-Opt algorithm to iteratively improve delivery routes and find the **most efficient routes for deliveries**, minimizing the total travel distance and re-ordering the delivery data according to the optimized routes.
- **Market Basket Analysis:** Identified relationships between items in transactions and constructed recommendations based on these relationships. Implemented the recommendations, leading to a **30%** increase in sales from related items.
- **Customer Order Data Analysis:** Predicted **user re-orders** by employing machine learning models like: Logistic Regression, Random Forest, XGBoost Classifier, and LightGBM Classifier. Achieved a re-order prediction accuracy of approx. **96%** with **XGBoost Classifier** model. Deployed the model using **Flask**, enabling better inventory management and customer satisfaction.
- **Item Pricing Prediction:** Predicted **prices** of food items in India using historical price data for appropriate item pricing by employing machine learning models like: Linear Regression, Random Forest Regressor, and XGBoost Regressor. Achieved an accuracy of approx. **98%** with **XGBoost Regressor** model.

Motlay Innovation Pvt Ltd 🔗

DATA SCIENCE INTERN | 05/2022 – 07/2022

- **Web Scraping:** Worked on Web Scraping in Python to extract data from various websites utilizing libraries such as BeautifulSoup, ensuring clean and structured datasets for analysis.
- **Automated Data Extraction:** Developed Python scripts to automate data extraction from Government ID cards (like Voter-ID card, Aadhar card, PAN card) using **OpenCV OCR**, and stored the extracted information for further use.
- **Company Data Analysis:** Analyzed data using **SQL** queries to gain insights into company sales statistics. Created an informative dashboard on Power BI for providing quick sales insights to support data-driven decision making.
- **Report Generation:** Worked on generating insightful reports and visualizations using **Power BI**.

PROJECTS

- **Heart Disease Prediction site** ([Website](#), [Source code](#))
 - Applied Machine Learning models: Logistic Regression, SVM, Random Forest Classifier, XGBoost Classifier, LightGBM Classifier, and Ensemble Models to predict if a person is affected by a heart problem or not. Utilized **RandomizedSearchCV** for hyperparameter tuning and achieved an accuracy of approx. **87%** with **XGBoost/LightGBM Classifier** models.
- **Item Pricing Prediction** ([Source code](#))
 - Applied Machine Learning models: Linear Regression, Random Forest Regressor, and XGBoost Regressor to predict prices of food items in India using historical price data for appropriate item pricing. Utilized **RandomizedSearchCV** for hyperparameter tuning and achieved an accuracy of approx. **98%** with **XGBoost Regressor** model.

ACHIEVEMENTS

- **Optum Financial AI Hackathon | 22/07/2024 – 02/08/2024**
 - Worked on a Chat Assistant prototype to assist support representatives in the card re-issuance process for MICROSOFT.
 - Received 3 **Diamond Awards** (highest rated award in Optum) for exceptional performance. 🔗
-