RAJASHRI DEKA | rajashrideka07@gmail.com | (+91)-6000797881 | rajashrideka | rshdeka | Portfolio

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, RAG, Langchain, Agentic AI (BrowserUse + Playwright)
- **Cloud Services:** Azure Functions, Azure Cosmos DB, Azure Databricks, Azure Cognitive Search, 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 Optum Bank's Customer Support data flow pipeline*:** Automated the extraction of key information from chats loaded in Azure Databricks, and tagging into pre-defined categories utilizing **XGBoost** multi-class classification model with **GridSearchCV** tuning. Deployed into an Azure Function pipeline, and scheduled a Databricks job to run at regular intervals.
  - ***Tech Stack***: **Python, SQL, ML (XGBoost), Azure Cosmos DB, Azure Databricks**.
  - ***Benefit***: Increased operational efficiency by improving tagging accuracy by **12%** and reducing manual analysis workload.
- ***Automated real-time Chatbot Evaluation Workflow*:** Developed a RAG workflow triggered via Azure Event Grid to evaluate user-assistant interactions in real-time, using metrics like similarity, fluency, relevance scores, etc., with results stored in Cosmos DB.
  - ***Tech Stack***: **Python, Gen AI (RAG, LLM, Prompt Engineering), Azure Functions, Azure Cosmos DB, Azure Event Grid**.
  - ***Benefit***: Improved chatbot response quality, enhancing customer satisfaction.
- ***Automated on-demand Newsletter Generation*:** Developed a system using Generative AI to generate newsletter with dynamic insights on Customer Support issues, incorporating visualizations like pie charts and bar graphs. Enabled dynamic updates for selected chats.
  - ***Tech Stack***: **Python, SQL, Gen AI (LLM**: Azure OpenAI, **Prompt Engineering), Azure Functions, Azure Databricks**.
  - ***Benefit***: Boosted stakeholder engagement by **80%**, aiding in timely decision-making.

PROJECT: AI Advancements | 11/2024 – Present

- ***Automated Code Documentation Generation for GitHub*:** Implemented a GitHub Action that triggers on new PRs and leverages Generative AI to auto-generate code documentation for new and existing files in PRs, and commits the documentation directly to the repository.
  - ***Tech Stack*: Python, Gen AI (LLM**: Azure OpenAI, **Prompt Engineering), Azure Functions, GitHub Actions.**
  - ***Benefit***: Maintains 100% up-to-date documentation, and aiding new team members understand the codebase efficiently.
- ***Automated UI Browser Testing by utilizing an Agentic AI framework*:** Utilized Langchain, BrowserUse, Playwright (as fallback) to auto-execute test scenarios on UI. This setup utilizes Generative AI to dynamically generate automation scripts requiring minimal modifications.
  - ***Tech Stack***: **Python, Prompt Engineering, Langchain, BrowserUse, Playwright**.
  - ***Benefit***: Increased automated test coverage by 60%, significantly reducing manual testing efforts.

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.