

SARA CHOUDHARY

DATA SCIENTIST

INDIA



+91 9311961839



saradune6@gmail.com



Github



LinkedIn

SUMMARY

As a Data Scientist with a background in fintech and hands-on experience in data science and analytics. I specialize in Machine Learning, Statistical Analysis, and Predictive Modeling to generate actionable insights for data-driven decision-making. My expertise lies in optimizing financial processes and enhancing customer experiences through advanced data techniques.



EDUCATION

- Btech in Computer Science Engineering
Raj Kumar Goel Institute of Technology And Management (AKTU)
2021-2024
- Diploma In Mechanical Engineering
Board Of Technical Education, Lucknow
2018-2021



SKILLS

Programming Languages:

- Python
- HTML
- CSS

Databases:

- MySQL
- MongoDB

Frameworks and Python Packages:

- NumPy
- Pandas
- Plotly
- Scikit-Learn
- PyTorch
- Keras
- Web Scraping
- Streamlit
- Flask
- MS Excel

Visualisation:

- Amazon Quicksight
- PowerBi

Machine Learning:

- Supervised and Unsupervised Learning
- Linear and Logistic Regression
- Bagging and Boosting techniques
- NLP

Deep Learning:

- Convolutional Neural Networks (CNN)
- Recurrent Neural Networks (RNN)
- Feature Engineering

Generative AI:

- LangChain
- Hugging Face
- Prompt Engineering
- Open source LLM models
- GPT (Generative Pre-trained Transformer) Series
- Text Generation

Computer Vision:

- OpenCV
- Object Classification



PROFESSIONAL EXPERIENCE

BharatPe

Data Scientist



July 2024 - Present



Gurugram, India

- Developed a model for PAN and Aadhaar name matching using Hugging Face (BERT base/large, Siamese BERT variants) to enhance KYC accuracy.
- Fine-tuned models for name variations, improving performance in handling spelling, abbreviations, and transliteration errors.
- Increased accuracy in identifying mismatches, reducing manual KYC verification.
- Optimized model for real-time processing, balancing performance and inference time.
- Collaborated on production deployment to integrate the model with KYC operations.
- Applied data augmentation to improve robustness for diverse name formats.
- Conducted A/B testing to benchmark and select the best model for deployment.

Rapipay Fintech Pvt Ltd.

Data Analytics



Nov 2023 - July 2024



Noida, India

- Conducted in-depth evaluation of merchants based on Daily Payment Data (DPD) to assess financial health, payment behavior, and risk levels, **providing actionable insights for loan approvals.**
- Developed comprehensive **dashboards using Power BI and Amazon QuickSight** to visualize DPD trends, repayment patterns, and overall merchant performance, aiding NBFCs in data-driven decision-making.
- **Collaborated with non-banking financial companies (NBFCs)** to streamline the loan approval process by using data-backed insights to identify eligible merchants with strong credit histories.
- Implemented scoring models to assess merchants' creditworthiness, incorporating factors such as DPD, transaction volume, and repayment consistency, improving the accuracy and efficiency of loan disbursements.
- Automated the process of **categorizing merchants into high-risk and low-risk** segments based on payment behavior, reducing default risks and enhancing the quality of the loan portfolio.
- Utilized **AWS Athena for querying large datasets** to validate merchant performance metrics, ensuring data accuracy and compliance with regulatory standards.