Shankhadeep Roy

(+91)9836533804 | shankharay.917@gmail.com | /shankha-deep-roy/ | github.com/shankha06

EDUCATION

Indian Statistical Institute	Kolkata, India
Master's of Technology, Computer Science	*Jun. 2013 – Jul. 2015*
Institute of Engineering and Management	Kolkata, India
Bachelor's of Technology, Computer Science	*Aug. 2009 – May. 2013*

EXPERIENCE

Lead Data Scientist	Bangalore, India
JP Morgan Chase	*Aug 2022 - Present*

- (Present) - Deal Recommendation using multi-modal user-item and next deal prediction
- Enhanced Proposal Answers through tybrid-RAG on local finetuned LLM
- Agentic-RAG for QA Model using mixture of public and private information
- Leveraged Active Learning and Semantic Similarity for Tags to Company
- Large-Scale Company Deduplication through supervised Representation Learning

Senior Data Scientist	Bangalore, India
Airtel X Labs	*Oct 2021 - Aug 2022*

- Loan propensity modeling through representation learning of telecom activity | *pyspark*
- Customer home screen personalization and intent classification | *pyspark, lightfm, fastapi*

Data Scientist	Bangalore, India
SAP Labs Pvt Ltd	*Aug 2015 - Oct 2021*

- Image-Based Clothing Recommendation through multi-task OD model | *docker, torch, fastapi, Kubeflow*
- Real-Time multi-CCTV Facemask detection | *docker, RabbitMQ, fastapi, tensorflow*
- Production Failure Prediction in Engine Manufacturing & Automated RCA | *fastapi, xgboost, lime, shap, Heroku*
- Prediction of Estimated Date of Delivery(EDD) to nearest sorting warehouse | *python, fastapi, keras, tf, xgboost*
- Optimize travel time for entire fleet based on regression models | *fastapi, munkres, GCP*
- Prediction of Award shortlist based on individual contribution
- Prediction of churn for customers and identify incentivize for retention

PROJECTS

Deal Recommendation | *torch, Opensearch, Athena, DynamoDB, Redis*   Mar 2025 – Present

- Developed a hybrid recommendation engine that integrates real-time session prediction model and a multi-modal user preference and item model
- Region-based popular suggestions calculated using ARP and RRP
- Reranking based on brand and category preference
- Currently developing and optimizing the offline model training pipelines and real-time inference APIs

Automated Answering for Proposal | *torch, sbert, Opensearch, textract, EKS, docker*   July 2024 – Feb 2025

- Eeveloped an automated tool that processes user RFPs against a 30K document knowledge base using RAG
- Employed a query-expansion based hybrid-RAG architecture using embeddings from a local fine-tuned LLMs on Opensearch
- Achieved 76% Recall@5 for content retrieval and an 85% ROUGE-L score for the quality of generated responses.

CERTIFICATION

AWS Cloud Practitioner   Jan'23