Deepak Rambarki AI Engineer

rambarkideepak@gmail.com | (940) 843-7994 | USA | LinkedIn

Summary

Data Engineer with 5+ years of experience designing and maintaining scalable batch and real-time data pipelines for high-volume transactional systems. Strong expertise in Python, SQL, Apache Spark, Kafka, Airflow, and dbt, with hands-on experience building cloud-native data platforms on AWS, Azure, and GCP. Proven ability to process 10M+ daily records, model enterprise data warehouses, and deliver reliable, governed datasets supporting fraud analytics, risk reporting, and business intelligence. Adept at collaborating with Risk**,** Compliance**,** and Analytics teams to ensure data quality, security, and performance at scale.

Technical Skills

- **Languages & Programming**: Python, R, Java, C++
- **Machine Learning & Deep Learning**: TensorFlow, PyTorch, Keras, XGBoost, ONNX, BiLSTM, Logistic Regression, CNNs, RNNs, Transformers, Reinforcement Learning, Transfer Learning, Sequence Models
- **AI Models & Techniques**: Fraud Detection Models, Churn Prediction Models, Classification, Regression, Anomaly Detection, Behavioral Modeling, Time Series Analysis
- **Model Evaluation & Explainability**: F1-Score, AUC-ROC, Precision@K, Recall, SHAP, Model Interpretability, Hyperparameter Tuning (Optuna, Grid Search, Bayesian Optimization)
- **Natural Language Processing (NLP)**: BERT, GPT, T5, spaCy, NLTK, Named Entity Recognition (NER), Sentiment Analysis, Text Classification
- **Data Engineering & Processing**: SQL, NoSQL (MongoDB, Cassandra, DynamoDB), Apache Spark, Spark Streaming, Apache Kafka, Hadoop, Data Preprocessing, Feature Engineering, ETL Pipelines, Data Warehousing
- **MLOps & Deployment**: AWS SageMaker, Azure ML, Google AI Platform, SageMaker Pipelines, Lambda, Docker, Kubernetes, MLflow, DVC, CI/CD for ML, Model Versioning, Real-time Inference, Model Monitoring
- **Cloud Platforms & Tools**: AWS, Azure, Google Cloud Platform (GCP), Redshift, Snowflake
- **Version Control & Collaboration**: Git, GitHub, GitLab, Bitbucket, Agile Methodologies, JIRA
- **Ethics & Compliance**: Fairness, Accountability, Transparency (FAT), GDPR, HIPAA, Data Privacy Regulations
- **Visualization & Reporting**: Power BI, Streamlit, dbt

Professional Experience

AI Engineer, *Brex* 12/2024 – Present | Remote, USA

- Designed and deployed a BiLSTM-based fraud risk scoring engine, reducing fraudulent transactions by 19%. Collaborated with Risk and Compliance teams through requirement-gathering sessions to align model goals with business needs.
- Processed and filtered 10M+ transaction records using Python, SQL, and Spark. Extracted features like frequency, merchant category, and location deviation to support feature engineering and prepare data for AI model training.
- Built and trained a BiLSTM sequence model for behavioral fraud detection. Integrated SHAP for model explainability and aligned outputs with Brex's internal audit and model governance standards for compliance transparency.
- Evaluated model performance using F1-score, AUC-ROC, and Precision@K. Achieved 0.87 F1-score and 0.93 AUC through iterative model tuning and validation on real-time transaction data across various customer segments.

- Conducted hyperparameter tuning using Optuna and Bayesian Optimization, refining sequence length, dropout rate, and learning rate. Improved fraud signal detection accuracy while minimizing false positives in high-risk account segments.
- Automated retraining using SageMaker Pipelines and MLflow, triggered bi-weekly based on data drift metrics. Maintained model freshness and compliance by aligning retraining schedule with evolving transaction trends and fraud vectors.
- Deployed models via AWS SageMaker and Lambda for scalable real-time inference on 3M+ daily transactions. Managed 15TB+ of data in Redshift/Snowflake and built monitoring dashboards using dbt for risk and model KPIs.

AI Engineer, Zebronics 01/2020 – 05/2023 | Remote, India

- Designed and deployed a smart recommendation engine for electronic accessories (e.g., headphones, speakers, key boards), leveraging user interaction data to personalize product suggestions. This led to a 28% increase in cross-sell and upsell rates on the e-commerce platform.
- Utilized customer purchase history, browsing patterns, and product metadata to build a hybrid recommendation model combining content-based filtering and collaborative filtering, improving recommendation accuracy by 20% over baseline.
- Developed a demand forecasting model using time series analysis (Prophet, ARIMA) and machine learning (XGBoost, LSTM) to predict weekly product demand across regions, reducing overstock by 18% and minimizing stockouts by 22%.
- Implemented inventory optimization algorithms that dynamically adjusted restocking schedules and distribution priorities based on predictive insights, improving supply chain efficiency and reducing holding costs by 15%.
- Integrated ML-powered personalization features into Zebronics' website and mobile app via REST APIs, enhancing user engagement and resulting in a 12% longer average session duration and a 10% increase in conversion rates.
- Conducted A/B testing on different personalization strategies, optimizing UI/UX and model output, which led to a 6% increase in customer retention and a 4% uplift in overall revenue. Deployed and maintained models using Azure ML, Docker, and Kubernetes, ensuring scalability for over 3M users, with real-time data streaming and model updates via Apache Kafka, Spark, and Azure Data Factory .Deployed real-time scoring pipelines using Azure ML, Docker, Kubernetes, Kafka, and Spark Streaming to integrate churn scores directly into Verizon's CRM systems for proactive customer engagement and campaign personalization. Collaborated with business analysts and data teams, delivering Power BI dashboards that visualized churn trends, model performance, and reduced customer retention costs by 22%.

Education

Master in Advanced Data Analytics
University of North Texas |Denton, TX, USA **08/2023 – 05/2025**

Certificates

- **Python for Data Science and Machine Learning Bootcamp**
- **AWS Fundamentals: Going Cloud-Native**

Projects

MEDCOMPARE – AI Powered Medication Data Benchmarking Tool

- Built a real-time Streamlit tool benchmarking medication data from LLMs using semantic similarity, fuzzy matching, and biostatistics; aligned outputs with FDA and RxNorm standards; supported batch evaluation, scoring, and FHIR exports for clinical review.

Classification with XGBoost + ONNX Deployment

- Developed an XGBoost classifier on 1.2M entries, outperforming baselines; applied scaling, outlier handling, and data prep; converted model to ONNX for real-time edge/cloud deployment, showcasing scalable predictive analytics in IoT and embedded environments.