ASTHA PURI
New York City, NY | [LinkedIn](#)

SUMMARY

Lead Machine Learning Engineer with 9+ years of experience owning and scaling production ML systems in high-traffic consumer and SaaS platforms. Proven expertise in architecting end-to-end ML pipelines (data → training → serving → experimentation), replacing third-party vendors, enforcing latency and reliability SLAs, and mentoring teams. Specialized in personalization, search & recommendation systems, and production-grade Generative AI infrastructure.

CORE MLE EXPERTISE

ML Systems & Architecture
Offline & Real-Time ML Pipelines, Model Serving, Feature Stores, Experimentation Platforms, Latency Optimization

Applied Machine Learning and AI
Recommender Systems, Search & Ranking, Multi-Armed Bandits, NLP, SLMs, LLMs (RAG, Fine-Tuning, Evaluation), Langchain, LangGraph, Google ADK

Engineering & Data
Python, SQL, PyTorch, Tensorflow, PySpark, Kafka-compatible streaming, Elasticsearch, DBT

EXPERIENCE

CVS   New York, NY
Senior Data Scientist, AI   Apr 2023 – Present
- Owned architecture and productionization of personalization and recommendation systems powering the CVS app home screen for millions of users. Led end-to-end development of a **contextual multi-armed bandit** system, integrating feature engineering, offline training, online inference, and A/B experimentation.
- Designed and deployed a BERT-based user search intent classification and typeahead system augmented with LLM-generated (GPT-4o) synthetic data.
- Implemented AI-generated product [summaries](#) with automated evaluation (ROUGE, LLM-based G-Eval), improving SEO by 20% and conversions by 8%. **(patent pending)**
- Architected and deployed a Word2Vec-based [recommendation system](#) with offline batch inference and strict latency SLAs; fully replaced a third-party vendor model, improving CTR by 5%.
- Built and productionized an ML-based product substitution system using semantic similarity and availability constraints, increasing order fulfillment rate by 7% and eliminating vendor dependency. **(patent pending)**
- Led GenAI enablement for 500+ engineers and product leaders, establishing best practices for safe and scalable LLM adoption.
- Supervised multiple data science interns in a project focused on enhancing the website's search experience where they implemented real-time search autocomplete using RNNs and recommendation systems on the company website. Coordinated project scope, milestones, and resources, resulting in the successful implementation of search optimization strategies that improved user engagement and satisfaction.

TWILIO SEGMENT   Los Angeles, CA
Senior Data Scientist   Feb 2022 – Feb 2023
- Designed scalable analytics and traffic forecasting pipelines to support GDPR driven EU data migrations, saving $240K annually through optimized compute and storage allocation.
- Built DBT based data models and pipelines, reducing latency and improving operational efficiency by 30%.
- Partnered with engineering teams to support experimentation and growth analytics across the acquisition funnel.

ORACLE   Broomfield, CO
Lead Data Scientist   Aug 2021 – Feb 2022
- Owned ML pipelines for 9 month churn prediction and customer risk scoring across SaaS products, improving retention by 8%.
- Reduced escalation handling costs by 15% by deploying MIT's DeepMoji for customer sentiment analysis, predicting service request escalation using ensemble models
Data Scientist   July 2018 – July 2021
- Built data pipelines that processed stream events using [Kafka compatible services](#) (with OSaK and Spark applications) to provide real-time analysis of streaming data
- Analyzed time series data from server logs, studied trends across KPIs (restarts, downtimes, SLOs) and built web applications and visualizations for leadership, securing application authentications and data.
- Performed statistical testing (t-test, chi sq test, etc.) to understand product performance and identify areas of improvement.

FORD MOTOR COMPANY   Chennai, India
Business Intelligence Engineer   July 2015 - April 2017
- Built self-service BI reports in Tableau, QlikView and SAP BO for global car parts, maintenance, warranty, and pricing teams. Performed deep dive analysis and presented business insights to executive management.

EDUCATION

UNIVERSITY OF MINNESOTA, Carlson School of Management, MN
Master of Science in Business Analytics (MSBA)   *May 2018*

VIT UNIVERSITY, India
Bachelor of Technology - Electronics and Communication (ECE)   *May 2015*