

Machine Learning Engineer with 5 years of experience designing and productionizing end-to-end AI solutions, including generative AI (LLMs, RAG) and ML optimization. Proven track record in leading scalable system design, MLOps, and cross-functional collaboration to deliver business impact. Adept at translating complex problems into innovative ML systems, with expertise in AWS/Azure, Kubernetes, and CI/CD pipelines

SKILLS

Machine Learning	Generative AI (LLMs, OpenAI, RAG), NLP, PyTorch, Scikit-Learn, Model Optimization, Model Serving, Online learning, Ranking
MLOps	Weights & Biases, Databricks, MLFlow, Seldon, CI/CD (Jenkins, GitHub Actions, Azure DevOps),
Cloud & Other	AWS (Lambda, S3, DynamoDB, ECS, SageMaker, OpenSearch), Azure (Functions, CosmosDB, App Service, Container Apps), Kafka, MySQL, ElasticSearch, Docker/Kubernetes, Terraform, Bicep, DataDog
Languages	Python, Go, JavaScript
Leadership	System Design, Cross-functional Collaboration, Mentorship

TECHNICAL EXPERIENCE

Machine Learning Engineer 06/2024 — Present
Nationale Nederlanden Rotterdam, Netherlands

- Developed **reusable CI/CD modules** accelerating deployment of AI components, resulting in **65%** reduction in deployment time
- Improved **observability** and API performance by implementing **instrumentation** and monitoring using **Azure monitor**
- Improved governance on Databricks by using **Unity Catalog** and role-based access control (RBAC), ensuring data security and compliance
- Built data pipelines on **Databricks** leveraging **Spark** for model performance monitoring

Software Engineer 12/2023 — 06/2024
See Tickets Benelux Amsterdam, Netherlands

- Reduced CAC** (Customer Acquisition Cost) by **25%** at See Tickets by optimizing CRM-driven marketing strategies
- Leveraged CRM data (DotDigital) to develop customer segmentation models, enhancing targeted marketing campaigns and conversion rates
- Optimised SQL queries, improved existing query performance by **10%**

Senior Software Engineer 06/2023 — 11/2023
Cactus Communications Mumbai, India

- Improved autoscaling of a async **event-driven service** by utilizing caching techniques and predictive scaling policies on AWS
- Developed a modular template to implement **CI/CD pipeline** using **Jenkins** and **github** actions
- Developed and deployed generative AI features** using llama and OpenAI APIs, enhancing user experience and personalization
- Optimized RAG-based recommendation systems using A/B testing, improving user engagement
- Mentored** junior colleagues, fostering a culture of knowledge sharing and **continuous learning**

Software Engineer 08/2020 — 06/2023
Cactus Communications Mumbai, India

- Developed supporting NLP modules to **improve the accuracy of a machine translation model by 15%**
- Moved short lived services to serverless, resulting in **30%** cost savings and better peak load management
- Developed an **lambda pipeline** to scrape data and store in **S3**, and query using **AWS Athena**
- Developed custom module for synthetic data generation and data augmentation using **Apache airflow**
- Automated provisioning of model training infrastructure using **Terraform**, developed reusable modules
- Developed real-time scalable REST API using **FastAPI**, capable of supporting **100TPS**

EDUCATION

Bachelor of Engineering, Electronics and Telecommunications 2016 — 2020
Dwarkadas J Sanghvi College of Engineering, Mumbai University. 7.6/10 GPA

CERTIFICATIONS AND RECOGNITIONS

- iSAQB CPSA-(Certified Professional for Software Architecture)
- Awarded 'Rising star of the year' for excellent performance at Cactus Global.
- 'Machine Learning for business Professionals' by Google cloud - Coursera.

ACTIVITIES

- Head of Technology - D.J Literary committee.
- Engaged in organizing and serving meals to residents, participating in donation drives, and assisting in community outreach programs.