Aniket Anvekar

aanvekar25@gmail.com | 201.920.4974 | github.com/aniketAnvekar | linkedin.com/in/aniketanvekar | aniketanvekar.github.io

Summary

Data Scientist and GenAI Engineer with expertise in developing full-stack AI solutions combining deep learning, natural language processing, and cloud-native architectures. Skilled in designing context-aware AI agents, scalable ETL pipelines, and predictive models. Proficient in integrating AI models with production systems via RESTful APIs and orchestration frameworks such as LangChain. Passionate about leveraging AI to automate workflows, deliver measurable business impact, and enable data-driven decision-making in financial and enterprise domains.

Professional Experience

Vanguard – Malvern, PA   Jul 2025 – Present
GenAI Engineer
- Designed RESTful APIs with AWS Lambda, API Gateway, and CloudFront to securely invoke Amazon Bedrock's foundational models.
- Developed a context-aware financial agent using LangChain, FAISS, and Anthropic models that dynamically processed GraphQL queries to generate real-time fund insights.

Kansas Department of Labor – Edison, NJ   Jul 2024 – Jun 2025
AWS Conversational AI Engineer
- Leveraged AWS services like Lambda, Connect, Lex, S3, and Kinesis to develop an IVR system for filing unemployment insurance claims for the state of Kansas.
- Implemented Amazon Bedrock Guardrails to enforce safe, policy-compliant chatbot interactions, ensuring ethical, accurate, and context-aware AI responses.
- Boosted chatbot accuracy by 20% by integrating AWS Bedrock and Llama3 model to validate system and user responses, improving overall user experience.
- Optimized front-end integration with PostgreSQL by developing RESTful Python APIs using Flask.

Eli Lilly & Company – Remote   May 2021 – Aug 2022
Data Scientist
- Conducted a PoC to automate healthcare claims image processing using Tesseract and Azure OCR, reducing manual data entry by 30% and enhancing data validation accuracy.
- Utilized Transformers and NLTK library to extract and summarize data from medical manuals (PDFs), achieving a high ROUGE score of 0.83.
- Improved medical image prediction accuracy with an AUC score of 0.91 using transfer learning with ResNet-50.

Reserve Bank of India – Mumbai, India   Aug 2019 – May 2021
Data Engineer
- Led team of 3 to develop Selenium web crawler on RHEL servers, processing 20+ e-commerce sites to calculate price indices for 8 Indian metropolitan markets using Agile methodology.
- Built social media analytics pipeline using Twitter API and PySpark NLP for GDP sentiment analysis, visualizing insights through Plotly dashboards for executive decision-making.
- Optimized database schema and ETL workflows with business analysts for Apache Hive, reducing weekly data loading time by 6 hours (30% improvement).
- Containerized applications using Docker and Kubernetes, achieving 20% faster processing times and improved scalability in production environment.

Samaritan Scout – Summit, NJ   Jul 2023 – Dec 2023
Backend Lead - Data Scientist
- Developed Python web scraping platform with Beautiful Soup to aggregate 1,000+ NGO volunteering opportunities, reducing manual search time by 80% across tri-state area.
- Improved data extraction efficiency by 20% using GPT-3.5 and Google Gemini prompt engineering, outperforming regex-based keyword identification.
- Architected normalized database schema on Amazon RDS and built automated ETL pipeline using AWS Lambda and S3 triggers, increasing data processing efficiency by 25% and streamlining data workflows.

Skills

Generative AI: LLM, RAG, Prompt Engineering, LangChain, AWS Bedrock, FAISS, GraphQL
Machine Learning/Deep Learning: Regression, Classification, NLP, Time-series, Transfer Learning
Programming Languages: Python, R, SQL, Java, C/C++
Big Data & ETL: PySpark, Hive, Airflow, AWS Glue, Docker, Kubernetes
Visualization: Tableau, QuickSight, Plotly, Power BI
Databases: AWS S3, RDS, DynamoDB, PostgreSQL, MySQL, MongoDB

Certifications

AWS Machine Learning Engineer Associate, AWS Cloud Practitioner, AWS AI Practitioner, Databricks Generative AI

Selected Projects

Legal Document Q&A System
Python, LangChain, RAG, LangGraph

- Built multi-agent RAG system for document analysis, implementing vector search with FAISS and achieving 85% answer accuracy on 5K+ legal queries.

Kidney Disease Classification
Python, TensorFlow, MLFlow

- Developed VGG16 models on 5K+ CT scans achieving 81% accuracy with MLFlow tracking.

Airbnb Data Analytics
Python, Tableau, Machine Learning

- Built Tableau dashboards for rental KPI analysis and developed Random Forest regressor ($R^2 = 0.23$) with LIME explainability, outperforming Lasso and Ridge models.

Education

Rutgers University – New Brunswick   Aug 2022 – May 2024
Master of Science in Computer Science — GPA: 3.65/4.0

University of Mumbai   Aug 2015 – May 2019
Bachelor of Engineering in Computer Science — CGPA: 8.11/10