

YASHANK KOTHARI

kothariyashank@gmail.com | [linkedin.com/in/yashankkothari](https://www.linkedin.com/in/yashankkothari) | github.com/yashankkothari

EDUCATION

K.J. Somaiya College of Engineering

Bachelor of Technology Computer Engineering and Honours in Data Science

Mumbai, India

October 2022 – May 2026

EXPERIENCE

Data Engineer

ICICI Lombard

July 2026 – Present

Mumbai, India

- Engineered scalable ETL pipelines by developing over 100 advanced PL/SQL stored procedures and triggers, successfully elevating data integrity and processing accuracy by 18%
- Overhauled manual data extraction and refresh cycles by implementing automated, schedule-driven Oracle workflows, cutting daily operational data load times by 25%
- Implemented complex Business Intelligence (BI) logic into automated ETL workflows by computing dynamic KPIs, aggregated metrics, and derived columns, leveraging advanced PL/SQL analytical functions to enhance enterprise dashboard reporting
- Resolved critical data synchronization anomalies between operational OLTP systems and the OLAP data warehouse by developing automated data reconciliation scripts, ensuring complete transactional consistency and structural integrity

Data Engineering Intern

Colgate Palmolive

January 2026 – June 2026

Mumbai, India

- Architected a multi-layered Medallion data model (Staging, Curated, Consumption) using dbt and Jinja-templated SQL, optimizing complex multi-granularity joins for enterprise BI reporting in Domo and Sigma Analytics
- Engineered **automated data ingestion pipelines** utilizing Apache Airflow (Python) to transfer high-volume datasets from Google Cloud Storage to Snowflake, ensuring pipeline idempotency and eliminating manual workflows
- Facilitated a major enterprise architecture migration from **SAP ECC to SAP S/4HANA** by remediating legacy tables, updating transformation logic, and conducting rigorous data validation to maintain zero-defect production data
- Developed and trained a natural language query interface using Retool AI, enriching the model's context with domain-specific metadata to empower business teams to query databases independently without writing SQL

AI/ML Intern

Card91

June 2025 – July 2025

Bangalore, India

- Engineered and integrated the Tazama open-source fraud detection engine by converting core rule logic from TypeScript to Python, enhancing Card91's real-time transaction monitoring capabilities with improved processing speed
- Adapted and implemented **30+ transaction rules** across P2P and P2B payment flows, reducing false positive rates and improving fraud detection accuracy to **98.5%**
- Implemented Google Maps API for geolocation-based fraud detection, enabling real-time verification of transaction locations and reducing location-based fraud incidents

PROJECTS

NotesAid | *Next.js, React, TypeScript, Progressive Web Application*

2025

- Launched the beta version on February 24, 2025, achieving **17,000+** views, **1,500+ hours** of active time, and **2,700+** unique users in a single semester
- Made the project open-source, receiving 60+ pull requests and contributions from the community
- Developed a method for efficient content updates and performance optimization, turning it into a responsive Progressive Web App (PWA)

MalNotWare | *Docker, VirusTotal API, Browser Extension, Hybrid Malware Detection*

2025

- Engineered MalNotWare, a hybrid malware detection system that automatically scans downloaded files via a custom browser extension
- Leveraged static analysis by querying the VirusTotal API for known threats and performed dynamic analysis by executing suspicious files in a sandboxed Docker container
- Implemented real-time monitoring for malicious behavior detection with containerized security isolation

OncoLytix | *MobileNet V2, React Native, Expo, Flask, Mobile Application*

2025

- Developed a cross-platform lung cancer detection mobile application using React Native (Expo) and Flask
- Built and trained a MobileNetV2-based image classifier using MediaPipe Model Maker for detecting lung cancer from CT scan images with **87% Accuracy**
- Integrated with a Flask REST API to handle image preprocessing, queueing, and real-time prediction logic
- Designed a user-friendly mobile interface using React Native to upload CT scans and seamlessly display prediction results

CERTIFICATIONS

AI Receptive Software Development – *Clemson University*

2025

Neural Networks and Deep Learning – *DeepLearning.AI*

2024

AWS Cloud Technical Essentials – *Amazon Web Services*

2024

TECHNICAL SKILLS

Languages: Python, SQL, PL/SQL, NoSQL, JavaScript, HTML/CSS, C++
Applications: Apache Airflow, Snowflake, Domo, Sigma Analytics, Grafana
Frameworks: dbt, React, React Native, Node.js, FastAPI, REST API
Developer Tools: AWS (RDS), PostgreSQL, Oracle, Git, Jira, Docker, Linux, VS Code
Libraries: pandas, NumPy, Matplotlib