

# YASH KAMLESH SHAH

[ykshah1309@gmail.com](mailto:ykshah1309@gmail.com) | +1 (862) 230-8196 | [linkedin.com/in/yash-kamlesh-shah](https://linkedin.com/in/yash-kamlesh-shah) | [github.com/ykshah1309](https://github.com/ykshah1309)

## SKILLS

- **Languages:** PHP (Laravel), JavaScript (Node.js, VueJs), TypeScript, Python (Pandas, NumPy, Scikit-learn), SQL.
- **Data Science & AI:** Machine Learning (XGBoost, TensorFlow), RAG Architecture, LLM Fine-tuning, Prompt Engineering, Statistical Modeling, Time-Series Analysis, Computer Vision.
- **Databases & Cloud:** MySQL (Optimization, Migrations), MongoDB (Optimization), PostgreSQL, DynamoDB, Vector Databases (Pinecone/Milvus), AWS (RDS, S3, Lambda, Amplify), GraphQL.
- **Tools & DevOps:** Docker, Jenkins, GitHub, CI/CD, RESTful APIs, Tableau, PowerBI, Advanced Excel(VLookup/Marcos).
- **Engineering:** Backend Development, Microservices, Serverless Architecture, Data Pipeline Automation, ETL, Agile (Scrum), Cross-functional Collaboration, Security Compliance (GDPR/CCPA).

## PROFESSIONAL EXPERIENCE

**Full-Stack Engineer** | YogoSocial – Agenticue New Jersey, USA | Sep 2025 – Dec 2025

- Engineered a comprehensive analytics pipeline by developing AWS Lambda functions and integrating GraphQL via AWS AppSync, enabling efficient data processing and real-time insights for recruitment workflows.
- Implemented robust CRUD operations and schema changes for the businessMetrics table, ensuring data integrity and facilitating seamless interaction between frontend applications and backend services.
- Developed and integrated frontend pages for analytics visualization and CRUD management, establishing a cohesive full-stack solution that enhanced user experience and operational efficiency for critical business metrics.

**Research Assistant** | MiXR Lab, New Jersey Institute of Technology New Jersey, USA | Sep 2024 – Dec 2024

- Refactored data-driven product workflows to accelerate the analysis cycle by 80%, enabling rapid experimentation and product refinement for complex systems.
- Developed a statistical framework and visualization suite to identify key user engagement levers, impacting product strategy through data-driven prioritization and longitudinal analysis.

**Data Scientist** | Dev Ashish Steels Mumbai, India | May 2023 – May 2024

- Developed and maintained interactive Power BI dashboards to provide comprehensive insights into sales performance, market rate fluctuations, and client order patterns, including volume, steel types, and fulfillment rates, driving monthly, quarterly, and annual revenue analysis.
- Streamlined bill management processes for over [X] clients and [Y] suppliers, enhancing financial transparency and enabling accurate tracking of procurement and sales transactions within a dynamic metal trading environment.
- Implemented robust shipment tracking and cost analysis methodologies, identifying and mitigating additional logistical expenses to optimize supply chain efficiency and improve profitability for diverse metal and steel deliveries.

**AI/ML Intern** | Verzeo India, Remote | Apr 2022 – Sep 2022

- Developed and deployed Linear, Ridge, and Lasso regression pipelines for ROI forecasting on 500K+ financial transactions, achieving 85% prediction accuracy.
- Optimized feature extraction using PCA and TF-IDF, improving sentiment classification F1-score by 10% and reducing pre-processing time by 30% for 100K+ text records.
- Established a robust cross-validation framework (RMSE, F1-score) for unsupervised workflows, improving model reproducibility and accelerating delivery cycles.

## PROJECTS

**Rose: Privacy-First Voice Assistant for Personal Knowledge Management** | GitHub Personal Project

- Built a fully offline Windows service using Vosk ASR and local 7B LLM inference, demonstrating backend development and system integration skills.
- Integrated AI-native features using prompt engineering and RAG-based indexing, achieving sub-2s latency for context-aware recruitment insights.

**Pandora's Box: Conversational Health AI with Empathetic Response Generation** | GitHub Sep 2025

- Developed an AI-driven application with a multi-persona response system, leveraging prompt engineering to transform complex data into personal and actionable insights.
- Pioneered a multi-persona response system with a sophisticated content filtering layer, which successfully blocked 99.4% of non-medical or unsafe queries across a dataset of 5,000+ simulated conversations.

## EDUCATION

**New Jersey Institute of Technology** New Jersey, USA  
Masters of Science in Data Science | CGPA: 3.8/4.0 Graduated: Dec 2025

## CERTIFICATIONS & PUBLICATIONS

[IBM Machine Learning with Python](#) | [HackerRank SQL \(Advanced\)](#) | [Automated Facial Expression Generation \(IEEE 2023\)](#)