

# Harshad Yadav

## Python Developer | GenAI & ML Engineer | Full Stack Developer

Mumbai, India | [harshadyadav20@gmail.com](mailto:harshadyadav20@gmail.com) | +91 8879672921 | **LinkedIn:** <https://in.linkedin.com/in/harshad-yadav-208> | **GitHub:** <https://github.com/harshad208/> | **Portfolio:** <https://harshad-yadav-portfolio.netlify.app>

## Professional Summary

Experienced developer with a strong background in Python, backend systems, machine learning, and GenAI applications. Proven ability to design and deploy end-to-end systems integrating LLMs, RAG pipelines, real-time data processing, and full-stack web applications.

## Professional Experience & Projects

### PYTHON DEVELOPER

Nextgen Clearing Pvt Ltd | Mumbai, India

Nov 2024 - Present

**DCH Computation** : Designed and implemented large-scale data ingestion workflows, creating multiple dataset collections in MongoDB and Apache Pinot. Utilized PySpark for efficient preprocessing and Kafka for real-time data communication, resulting in faster processing and improved pipeline performance.

**SIMPLIFY** : Designed and improved a SIM-to-IMSI assignment system, optimizing large-scale data processing using Apache Pinot and PostgreSQL for OLAP indexing. Built Apache NiFi flows to automate file ingestion and processing from diverse sources, including AWS and OSS, enhancing data pipeline reliability and performance.

### FULL STACK DEVELOPER

Cere Labs Pvt Ltd | Mumbai, India

Jun 2021 - Oct 2024

**Credit Modeling Bot** : Built and deployed an anomaly detection system for financial documents using Python, FastAPI, RabbitMQ, and MySQL. Applied real-time alerting and automated issue handling to improve risk control and operational efficiency. Demonstrated skills in backend architecture, inter-service messaging, and DevOps deployment.

**Salary Slip Extraction Bot** : Developed a bot using Python, pandas, and regex/NLP techniques to extract employee and earnings data from salary slips. This improved payroll analysis and decision-making speed. Showcased expertise in data parsing, automation, and workflow integration.

**Invoice Extraction System** : Trained a BERT model using MLflow to classify document types, enhancing invoice extraction accuracy. Collaborated across teams to deploy the model into a production pipeline using scikit-learn, Transformers, and model management tools. Applied core ML/NLP, MLOps, and collaborative problem-solving.

**KYC Validation** : Automated customer information extraction using advanced ML algorithms and deployed a validation bot. Reduced manual effort and improved compliance using data preprocessing, classification models, and automation scripting in Python. Emphasized AI-driven compliance and workflow optimization.

**Billing Management Module** : Designed a user-centric billing module aligned with organizational financial workflows. Conducted in-depth requirement analysis to architect a solution that streamlined billing processes, reduced manual overhead, and improved financial management efficiency.

## Skills

**Languages & Web:** Python, SQL, Bash, HTML, CSS, JavaScript, TypeScript, Flask, FastAPI, React, Next.js, REST APIs, OAuth, JWT

**ML & GenAI:** Pandas, NumPy, scikit-learn, TensorFlow, PyTorch, Transformers, Hugging Face, LangChain, LlamaIndex, ChromaDB, LoRA, QLoRA

**DevOps & Tools:** Docker, Docker Compose, GitHub Actions, Jenkins, SVN, Postman, Pytest, Airflow, Prometheus, Grafana

**Data & Cloud:** Kafka, RabbitMQ, Apache NiFi, PostgreSQL, MongoDB, MySQL, Redis, PySpark, AWS, Azure, GCP

## Education

### BE in Computer Engineering

Vidyalankar Institute of Technology, Mumbai.

Graduated in May 2021

### Diploma in Computer Engineering

Pravin Patil College of Diploma Engineering & Technology, Thane

Graduated in April 2018