DIVYANSHI YADAV

divyanshi4840@gmail.com ◇ Mumbai, India ◇ LinkedIn ◇ GitHub

PROFESSIONAL SUMMARY

Computer Engineering student with backend internship experience in full-stack development and applied NLP. Detail-oriented, writes well-tested code, and picks up new tools quickly.

EDUCATION

Bachelor of Technology, Computer Engineering
K. J. Somaiya School of Engineering — Expected 2028

Relevant Coursework: Data Structures, Analysis of Algorithms, Object-Oriented Programming, Operating Systems

SKILLS

Languages:	C++, Java, Python, HTML/CSS, JavaScript, TypeScript, SQL,
Frameworks & Technologies:	Next.js, React, Spring Boot, FastAPI, PostgreSQL, Redis, Firebase, Playwright
AI/ML:	NLP, spaCy, Snorkel, Named Entity Recognition (NER), GemeniAPI
Tools:	Git, GitHub, Docker, Alembic, pytest, Figma

EXPERIENCE

Summer Intern — June 2026 – July 2026
Hindustan Petroleum Corporation Limited (HPCL)

- Designed a versioned rules engine in Java 21 and Spring Boot on an append-only, event-sourced PostgreSQL (JSONB) data model, enabling instant rollback without mutating history, verified by 66 Testcontainers integration tests.
- Built a full-stack rule management platform with an 11-endpoint REST API, a Next.js and TypeScript UI, and a draft, active and archived state machine for rule changes without code deploys, powered by a deterministic rule evaluator using Java sealed interfaces and pattern matching.

Core Team Member — Aug 2025 – Present
KJSSE CodeCell

- Delivered AI/ML technical sessions to 250+ students while designing branding and marketing collateral for flagship events including HACKX and KJSSE CTF 3.0, mentoring peers on foundational concepts and increasing technical engagement across the community.

PROJECTS

Maintex — Maintenance Work Order Intelligence System

- Built an end-to-end NLP pipeline extracting 7 structured fields from 7,000 industrial maintenance work orders using spaCy named entity recognition (NER), Snorkel weak supervision and a three-tier entity resolver, achieving F1 scores of 0.85 to 0.89 across four evaluated fields on a 250-record held-out gold set.
- Made every pipeline run reproducible with SHA-256 content-addressed IDs and MLflow experiment tracking, served results through FastAPI and a Next.js UI, and backed it with 531 unit, integration and idempotency tests.

Memento — Web Page Version Control & Change Monitoring Platform

- Built a self-hosted platform that tracks web pages over time, storing only changed content through SHA-256 deduplication, with paragraph-level diffs, full-text search and binary search to find when text changed.
- Designed a min-heap crawl scheduler benchmarked at 43x faster per tick than a naive PostgreSQL scan at 10,000 pages, and scaled crawling across Playwright workers with BullMQ job queues and an atomic Redis Lua token-bucket rate limiter.