

Kuldeep Sharma

+91-7503767972, kuldeep.sh7972@gmail.com, [GitHub](#), [Leetcode](#)

Summary

A dedicated and innovative Full-Stack Web Developer with hands-on experience in building scalable backend systems, intuitive frontend interfaces, and AI-integrated applications. Proficient in Python, JavaScript, Django, ReactJS, and Node.js. Adept at leading teams, designing RESTful APIs, deploying containerized services, and delivering real-world AI solutions including chatbots, resume screening, and AI interviews.

Work Experience

Python Full-Stack Developer, Select Source International Pvt. Ltd. (Jobma) June 2024 – Present | Gurgaon

- Lead end-to-end development of scalable AI interview platforms with 50–200 concurrent users.
- Designed and maintained APIs using Django and Node.js, and developed responsive front-end interfaces with ReactJS.
- Managed a team of developers, overseeing full project lifecycles, architecture, and deployment.
- Delivered 3 major AI products:
- AskAI – Voice/Text Hybrid Chatbot using RASA, Django, React, and LLMs (e.g., LLaMA, Gamma). Integrated FAQ + Help Center via NLU & Vector DB fallback with real-time streaming answers.
- SkillMatch – AI Resume Screening using LLMs, Apache Kafka, WebSockets, and Django. Handles asynchronous processing and provides live UI updates.
- AI Live Interview – Real-time interviews using Whisper (speech-to-text), TTS models, and LLMs.
- Integrated tools: Docker, Kubernetes, SonarQube, Sentry, Nginx, Redis, JWT Auth, Bcrypt Encryption.
- Deployed services on Google Cloud Platform (GCP) using Kubernetes (K8s) and GitLab CI/CD.
- Contributed to Jobma's tech evolution from Laravel PHP to modern Django + React + Next.js stack.
- Helped clients save 70–80% HR processing time while maintaining 99%+ uptime.

Software Engineer, Plutos One Technology Pvt Ltd. Feb 2024 – June 2024 | Noida

- Built and deployed AI Chatbot solutions with WhatsApp, Google Business Messages and web platforms which have 130K+ active users and available in 17 Indian languages.
- Designed and developed chatbot backends using RASA framework, Django, python and Retrieval Augmented Generation (RAG) system using pre-trained Large Language Models (LLMs).
- Created AI Chatbots for VISA, BOB, IDBI Bank etc.
- Designed and implemented scalable backend solutions with Node.js and Express.js, handling complex business logic and data processing improving response time by 20%.
- Developed frontend applications using React.js, HTML, CSS & JavaScript ensuring high performance and responsiveness.

Software Developer Engineer, i-Neuron Intelligence Pvt. Ltd. Dec 2022 - Jan 2024 | Remote

- Developed Back-end where users can buy products from the platform using Node.js and MongoDB
- Managed the integration and deployment of the application with React frontend.
- Responsible for Schema design and creating an optimized Restful backend application.
- Worked on end-to-end Machine learning projects.
- Performed EDA, data cleaning, feature engineering tasks on the data to make it useful.
- Trained various ML models on the given datasets, done hyperparameter tuning compared and improved their accuracy.

Production Engineer, Laxmi Home Appliances. Jun 2017 - Nov 2022 | Greater Noida

- Managed a team of 10-15 skilled and unskilled workers to ensure on time completion of production targets of home and kitchen appliance products.
- Improved maintenance cycles of machines and equipment to **reduce downtimes by up to 20%**.
- Actively took feedback from the market to improve quality of products and reduced replacement and repair costs.

Skills

- **Languages:** Python, JavaScript, C
- **Frontend:** React.js, HTML, CSS, JavaScript, Next.js
- **Backend:** Django, Node.js, Express.js, REST API, WebSockets
- **AI/ML:** RASA, LlamaIndex, HuggingFace, LLMs, TensorFlow, Scikit-learn, Pandas, Seaborn
- **DevOps & Infra:** Docker, Docker Compose, Kubernetes (K8s), GitLab CI/CD, Nginx, Redis

- **Cloud & Tools:** GCP, Git, Postman, SonarQube, Sentry
- **Databases:** MySQL, MongoDB, SQLite, NoSQL
- **Auth & Security:** JWT, Bcrypt
- **Soft Skills:** Collaborative, Resourceful, Innovative, Team Player, Problem Solver
- **Languages Known:** English, Hindi

Projects

AskAI – Hybrid AI Chatbot for Real-Time Support

- A multi-channel AI assistant that receives commands and responds via voice and text, built to automate customer and candidate support and performs tasks.
- Developed using RASA Framework, Django, React, and integrated with LLMs using RAG.
- Implemented custom fallback mechanism using Vector DB for domain-specific answer retrieval.
- Supports dynamic actions like interview scheduling, interview summaries, and FAQ generation via custom RASA actions.
- Enabled seamless switching between voice and text-based conversation.
- Deployed with Docker Compose, Nginx, SonarQube, Sentry, and hosted on GCP.
- Goal Achieved: Reduced support response time by ~60% and improved user engagement across 3+ platforms.

SkillMatch – AI Resume Screening & Candidate Ranking

- An AI-powered resume screening system that matches bulk resumes to job descriptions and auto-ranks them for recruiters.
- Accepts resume uploads (PDF/DOCX/DOC/ZIP), extracts data using various libraries (pdfPlumber, pypdf, docx, unoconv etc), and matches skills with job role context.
- Used Apache Kafka for background task processing and WebSockets for notifications and live progress updates on the UI.
- Built with Django (API), React (UI), Kafka (async), Redis (caching), and deployed on Kubernetes with CI/CD pipelines (GitLab).
- Offers live result export, shortlisting filters, and detailed AI reasoning per candidate.
- Goal Achieved: Saved 50–70% of manual screening time and enabled instant feedback for 1000+ resumes/day.

AI Live Interview – Real-Time AI Interview Agent

- A fully automated real-time interview system that simulates a human interviewer using speech and language models.
- Conducts interviews using Whisper AI (speech-to-text), TTS models (Coqui TTS), and LLMs for dynamic Q&A.
- Begins with job-specific pre-generated questions, then dynamically generates follow-up questions based on candidate responses.
- Performs real-time evaluation, scoring, and summarization of answers.
- Built with Python (backend), React (frontend), WebSockets (live sync), and Dockerized for deployment on GCP with Nginx, Sentry, SonarQube.
- Goal Achieved: Enabled 24/7 AI-driven interviews, cut down HR load by up to 80%, and maintained >99% uptime for users globally.

Education

- Master of Technology (M.Tech), Greater Noida Institute of Technology, Greater Noida (2022)
- Bachelor of Technology (B.Tech), G.L.Bajaj Institute of Technology and Management, Greater Noida (2017)