

RATNESH KUMAR KUSHWAHA | BE, ME (Computer Science Engineering)

Email: rat.kush@gmail.com | Phone: +91-7898617389



PROFESSIONAL SUMMARY

Senior Software Engineer with 6+ years of experience in scalable microservices, Python, PySpark, FastAPI/Django, Data Science and Generative AI (MCP, AI Agent). At Pace Wisdom, led LLM-powered platforms using LangChain, Ollama, and vector DBs, and designed FinTech Kafka pipelines reducing latency by 30%. Adept at building robust data and API systems using Spark, AWS, and FastAPI. Delivered high-performance, fault-tolerant microservices and AI applications deployed on cloud (AWS/GCP), and done code reviews and mentored teams for optimal delivery.

SKILLS

Languages/Frameworks: Python, Java, GoLang, Flask, Django, FastAPI

ML & GenAI: TensorFlow, Scikit, CNN, HuggingFace, LangChain, Peft, RAG, Agent, OpenAI, LangGraph

Data Tools: Pandas, NumPy, Postgres, MongoDB, Redis, ChromaDB, FAISS

Cloud & DevOps: AWS (EMR, Lambda, EC2, S3, RDS), GCP, Docker, Git, GitHub Actions, CI/CD

Big Data & Distributed Systems: PySpark, Hadoop, Kafka

ETL & Data Engineering: Scalable ETL pipeline design, data ingestion, transformation, Spark SQL

Others: Agile/Scrum, JIRA, Mentoring, Statistical analysis

EXPERIENCE

Pace Wisdom Pvt. Ltd., Bengaluru | Senior Software Engineer | Apr 2022 – May 2025

Led the development of LLM-powered agent-based chatbots using LangChain and Ollama for intelligent query handling. Designed and reviewed RAG pipelines and NLP workflows integrated with Postgres and FastAPI. Architected FinTech Kafka pipelines, reducing processing latency by 30%. Delivered scalable microservices using Flask/Django, mentored a team of 5, and drove a 55% boost in engineering productivity. Dev and UAT reviews before Prod deployment.

Azilen Technologies Pvt. Ltd. | Contract Application Developer | Jul 2021 – Dec 2021

Migrated healthcare analytics platform from Scala to Python. Developed RESTful APIs and optimized data processing pipelines for extensibility, performance, and TDD-based architecture. Multiple MR code reviews before merging.

Infobeans, Indore | Associate Software Engineer | Feb 2021 – Jul 2021

Built secure NLP-based document classification APIs with email/OTP authentication. Implemented preprocessing pipelines and Hive-backed storage to improve data accuracy.

High IQ, Hyderabad | Solution Architect | Dec 2019 – Feb 2021

Engineered an OCR + CNN-powered PDF classifier on AWS SageMaker. Built scalable batch classification workflows using Parquet/JSON and optimized inference accuracy to 90%.

Emorphis, Indore | Software Engineer | Sep 2018 – Nov 2019

Developed multithreaded Python pipelines for IoT signal analysis. Built TensorFlow-based predictive services with REST APIs, used AutoEncoder

Saras InfoTech (Freelance) | Independent Consultant | Dec 2014 – Aug 2018

Delivered full-stack municipal portals (GRS/UMC) with backend APIs, dashboards, and database solutions. Designed custom ETL pipelines and Java-based reporting interfaces.

KEY PROJECTS

Voice [AI Agent](https://github.com/imratnesh/ask_multimodal) (Whisper, SST, TTS, Steamlit, Audio, Multimodal): This project provides a complete solution for processing audio files through a conversational AI agent. It features a FastAPI backend for processing, a Streamlit frontend for user interaction, and uses a locally-run Ollama model to ensure privacy and avoid paid APIs. https://github.com/imratnesh/ask_multimodal

Social Good Website (LLM, RAG, FastAPI, Postgres, Ollama): Designed and built RAG pipeline using Ollama with 95% accurate regulatory query handling. Integrated finetuning, caching, and fault-tolerant with evaluation (pytest, RagChecker) and monitoring implementations.

Leave Management System (OpenAI, HuggingFace, Django): Rule-based engine + LangChain chatbot for natural language leave queries. Prompt-tuning, REST APIs, and continuous quality tracing and evaluation with different metrics (faithfulness, LLM tests, pytest)

Appreciate App (FinTech, Flask, ML, Kafka, MongoDB): Built ML pipeline for churn/engagement. Scalable Flask APIs, Kafka ETL jobs, LangChain NLP-based insights. Delivered 65% increase in client engagement.

Intelligent Document Classification (SageMaker, CNN, Tesseract): OCR + CNN pipeline for PDF classification (85% accuracy). Used Pandas, NumPy, and multithreading. Deployed using SageMaker with batch monitoring. Evaluated using a confusion matrix.

IoT Anomaly Detection (TensorFlow, Pyspark, Autoencoder): Real-time MQTT/TCP stream ingestion with Autoencoder-PCA models. Used FFT feature extraction, multithreaded ingestion, and applied various statistical techniques. Done anomaly detection on edge device.

EDUCATION

M.E. (Computer Engineering) 2011-13 – SGSITS, Indore (Govt. Autonomous) | CGPA: 7.56

Thesis: Semantic Web Service Composition on Cloud (QoS View)

B.E. (Computer Science) 2006-10 – UEC, Ujjain (Govt. Autonomous) | Percentage: 65.59

Major Project: Smart City

CERTIFICATIONS

- **Big Data Hadoop & Spark (Udemy):** Issued Jun 2025 | PySpark, Hadoop, DB (IISc Bangalore)
 - **5-Day Gen AI Course (Kaggle):** Issued Nov 2024 | LLMs, Vertex AI, Google Gemini
 - **Problem Solving (HackerRank):** Issued Sep 2020 | Python, SQL
 - **Android Workshop (Techfest, IIT Bombay):** Issued Jan 2013 | Android SDK, Java
-