

# SUPRIYA SHELKE

## Senior Big Data Engineer

📞 8208547331    @ supriyashelke@gmail.com    🔗 LinkedIn    📍 Mumbai, India

### SUMMARY

I am a Senior Big Data Engineer with over 6+ years of experience in technology. My expertise lies in application development and the implementation of innovative microservices and big data solutions. With a solid background in the banking domain, I excel at managing resources and optimizing project workflows. I leverage advanced programming technologies and demonstrate a visionary approach to drive forward-thinking innovation

### EXPERIENCE

#### Senior Big Data Engineer

##### IDFC FIRST BANK

📅 01/2023 - Present    📍 Mumbai, India

- Designed and scaled robust microservices using Java Spring Boot with Java 17+
- Deployed applications utilizing AWS CI/CD pipelines for efficient delivery
- Implemented Java multithreading to boost application performance and responsiveness
- Optimized Java memory management to improve application efficiency
- Developed and documented REST APIs, integrating Swagger with MongoDB
- Processed and stored data using Apache Kafka with topic chaining
- Utilized AWS services including S3, EC2, IAM, Lambda, and EKS for scalable solutions
- Leveraged OpenSearch, Kibana, Jaeger, and Grafana for log analysis and observability dashboards
- Build Gen AI system using RAG (LLM + vector DB) for contextual response generation for domain specific queries
- Implemented RAG pipeline for getting HR domain policy details
- Implemented document ingestion, preprocessing chunking, splitting
- Implemented NLP techniques for entity recognition, sentiment analysis, and query understanding
- Preprocess user queries using tokenization and high-dimensional embeddings generation using BERT, sentence-transformers. similarity search done based on distance metrics like cosine similarity, Euclidean distance Used ANN vector search algorithms and optimizing done using indexing strategies such as HNSW, IVF, PQ
- Hands-on experience in RAG systems using LLMs like Open AI, GPT-4, and Mistral. Embedding generation using Open AI Embeddings and Hugging Face Transformers
- Worked on python, NLP, Tensorflow, Pandas, numpy libraries Worked on MLFLOW model deployment Built and tested pipeline using LangChain, Fast API
- Familiarity of Streamlit and Gradio UI Worked on scalable RAG pipelines using vector databases Pinecone, FAISS and frameworks like LangChain and Llamaindex

### SKILLS

Python	Java	Microservices	
Spring Boot		Apache Spark	
Grafana	Apache Kafka	Hadoop	
PySpark	HiveQL	MongoDB	
Oracle DB	Snowflake	ML Ops	
RAG	LLM	Vector DB	AWS
Kubernetes	ML	MLFLOW	NLP
PyTorch	REST API	TensorFlow	
Airflow	BITBUCKET	CI CD	

### STRENGTHS

- ✓

**Trusted Team Contributor**  
Trusted team contributor known for a collaborative approach and a continuous improvement mindset
- 🚩

**Inspirational and Innovative**  
Inspirational and innovative thinker committed to addressing challenges positively

### EDUCATION

#### Master of Engineering

##### Savitribai Phule Pune University

📅 06/2013 - 06/2015

#### Bachelor of Engineering

##### University Of Mumbai

📅 06/2008 - 06/2012

### TRAINING / COURSES

#### Post Graduate Diploma Certification

CDAC in Big Data Analytics

## EXPERIENCE

### Big Data Engineer

Kiya.ai

05/2021 - 11/2022    Mumbai, India

- Scheduled ETL jobs with Sqoop, achieving efficient data imports from Oracle DB for UPI, MBL, IMPS, and ATM products
- Configured and scheduled workflows using Apache Airflow, ensuring timely and effective data processing
- Created performance-enhancing scripts in Spark-Scala and PySpark, boosting data processing efficiency
- Integrated Hive table data with Tableau, enhancing reporting and data visualization capabilities
- Executed comprehensive data analysis through aggregation and transformation
- Implemented real-time data processing solutions with Apache Kafka and Spark Streaming, enabling dynamic data management
- Deployed advanced credit risk models utilizing ML Flow, improving predictive analytics capabilities

### Python Developer

Bristlecone India Pvt Ltd

03/2019 - 05/2021    Mumbai, India

- Leveraged advanced Python programming skills to design efficient, scalable solutions in fast-paced environments
- Deployed Flask applications via leading web hosting services, enhancing accessibility and user experience
- Implemented machine learning models for effective mask detection, contributing to safety initiatives
- Automated role assignments using Robotic Process Automation (RPA), achieving significant reductions in processing time
- Conducted comprehensive testing of web applications to boost performance metrics and elevate user satisfaction
- Automated 60+ COSO SAP audit processes utilizing RPA and Python, streamlining compliance workflows

## MY LIFE PHILOSOPHY

I am committed to continuous growth and innovation, approaching challenges with a positive mindset and consistently delivering lasting value in every endeavor

Supriya Shelke

## LANGUAGES

English	
Native	<div><div></div><div></div><div></div><div></div><div></div></div>
Hindi	
Native	<div><div></div><div></div><div></div><div></div><div></div></div>
Dutch	
Intermediate	<div><div></div><div></div><div></div><div></div><div></div></div>