

PURVA JAVALE

United States | +1(502) 650-1425 | purvajavale@gmail.com | LinkedIn | Portfolio | GitHub
Work Authorization: H4 EAD – No sponsorship required

CAREER PROFILE

Software Engineer with 7+ years of experience building Big Data solutions in the FinTech sector. Skilled in designing scalable ETL pipelines, distributed processing with Spark, Hadoop, SQL, and developing solutions on AWS. Led initiatives in mainframe modernization, regulatory reporting automation, and performance tuning. Seeking U.S.-based or remote roles where software engineering intersects with data-intensive systems.

PROFESSIONAL EXPERIENCE

Hexanika **Pune, IN**
Software Engineer - Big Data **February 2015 – January 2022**

Key Projects & Contributions:

Smart Join Platform - Data Processing Framework

- Led design and development of Smart Join, a big data ingestion and transformation platform using Apache Spark, Hadoop, and AWS.
- Architected scalable workflows for data transformation and rule-based automation, enabling 60% reduction in manual processing and full data lineage for compliance.
- Created reusable data cleaning and mapping frameworks, improving regulatory reporting accuracy and efficiency by 40%.

End-to-End Automation

- Built and optimized Spark-Hadoop ETL solution, automating job scheduling and orchestration for high-volume regulatory datasets.
- Achieved 100% process automation and 60% reduction in processing time through performance tuning and workflow refactoring.
- Developed tools and APIs for observability, usability, and auditability across ETL components.

Spark on Mainframe – Proof of Technology (IBM z/OS)

- Executed a Spark Proof of Technology on IBM z/OS, integrating with WebSphere and IBM Studio for legacy-to-modern data flow.
- Enabled hybrid architecture supporting data cleaning and rule-prediction models on mainframe-sourced datasets.

Smart Reg Platform - Regulatory Reporting Framework

- Collaborated with domain experts to develop regulatory rule-based engines for reports like FRY9C, Call Report (FFIEC), HMDA-LAR, and AML-SAR.
- Collaborated with ML team to integrate generative models to enhance rule chaining, prediction, and anomaly detection in reporting pipelines.
- Delivered fully traceable reporting flows with transparent rule application and change lineage.

Infrastructure & DevOps Enablement

- Installed, configured, and managed multi-node Hadoop clusters, and provisioned AWS cloud infrastructure (S3, EC2, EMR).
- Led internal Proof of Technologies to assess performance and scalability of new Spark jobs.
- Built internal Power BI dashboards, Hive views and SOAP APIs to monitor and improve reporting for business and engineering teams.

Agile Delivery & Team Leadership

- Mentored junior data engineers, conducted code reviews, and documented best practices.
- Created architecture diagrams, handoff documents, and hyper-care guides to ensure stability during deployment and transitions.
- Served as Scrum coordinator, managing sprint planning, epics, and boards using JIRA and Microsoft Project Planner.

TECHNICAL SKILLS

Languages: Java, Python, SQL, .NET Core

Big Data: Hadoop, Spark, Hive, HBase, Kafka, Airflow

Formats: TXT, CSV, XML, JSON, AVRO, ORC, PARQUET

Cloud & Infrastructure: AWS (S3, EMR, EC2), Linux

Database: MySQL, SQL Server

Tools: Power BI, Team Foundation Version Control, GitHub, Eclipse, Visual Studio, PyCharm, Microsoft Project, Jira, Microsoft Office

EDUCATION & CERTIFICATIONS

Bachelor of Engineering, Information Technology - Pune University (2014)

Big Data & Cloud: Master's in Big Data (TrendyTech), Databricks Fundamentals, AWS Databricks Platform Architect (Databricks), AWS Certified Cloud Practitioner (AWS)

Data & Analytics: Excel for Data Analytics, Java Performance Tuning

Project Management: Agile Project Management (Google)

AI & Emerging Technology: Generative AI (LinkedIn), IBM Watson (IBM)

Legacy Systems: IBM Mainframe (IBM)

KEY ACHIEVEMENTS

- Awarded as Star Performer of the Year for delivering excellence in projects and a Well-Done Achiever for successfully developing and presenting Proof of Technology for Spark on IBM z/OS at Hexanika.
- Secured second place for presenting innovative and impactful Fintech solutions at Hexanika's Hackathon.
- Delivered a guest lecture on Big Data concepts and real-world architectures to help students align with industry trends at MIT Pune.
- Organized and managed Hexanika's annual offsite, fostering cross-team engagement through technical and team-building events.