

ANKIT PAL

+1 (408) 759 4021 | ankit.pal.sde@gmail.com | github.com/apal9569 | linkedin.com/in/ankit9569/

EDUCATION

San Jose State University , <i>MS in Artificial Intelligence</i> California, USA	GPA: 3.63 / 4.0	Dec 2024
National Institute of Technology , <i>B.Tech in Electronics and Comm</i> Delhi, India	GPA: 7.49 / 10	Aug 2020

EXPERIENCE

Software Developer (RA) , <i>SJSU Research Foundation</i> California, USA	Oct 2023 - Present
<ul style="list-style-type: none">Led the development of a scalable Spring Boot-based Learning Management System (LMS), integrating multiple critical systems.Designed and implemented RESTful APIs with GraphQL, improving query flexibility and reducing data retrieval latency by 30%.Built event-driven architecture using Apache Kafka, enabling real-time processing of student activity logs and engagement tracking.Utilized MongoDB for scalable storage of course materials, student progress, and engagement metrics.Built and optimized data pipelines using Databricks and Apache Spark, improving dynamic content personalization efficiency by 35%.Deployed and managed cloud infrastructure on Azure, leveraging Azure Cosmos DB, Azure Functions, and Blob Storage for high availability and scalability.Implemented CI/CD pipelines using Jenkins and Azure DevOps, automating testing, building, and deployment, reducing manual effort by 60%.	
Data Analyst - SDE , <i>Optum - UnitedHealth Group</i> India	Aug 2020 - July 2022
<ul style="list-style-type: none">Developed scalable backend services for insurance claims processing using Spring Boot and Kafka, improving data processing efficiency.Implemented event-driven architecture using Kafka, enabling real-time anomaly detection in healthcare claims, reducing fraudulent claims by 20%.Deployed and managed microservices using Docker and Kubernetes, enhancing system reliability and reducing downtime by 35%.Built big data pipelines using Hadoop and Spark, automating claims validation and reducing processing inefficiencies by \$75K annually.Designed and optimized high-performance MySQL databases, reducing query execution time by 50% and cutting storage costs through efficient indexing and partitioning.Built batch processing jobs using Spring Batch, improving the efficiency of large-scale claims data ingestion and processing.Implemented unit and integration tests with JUnit, reducing production defects by 30%.Automated CI/CD pipelines with GitHub Actions and Docker, ensuring smooth deployment of Java-based microservices.	

SKILLS

Programming Languages	Java, Python, Scala, JavaScript, TypeScript, SQL, HTML, CSS
Frameworks	Spring Boot, Hibernate, React, Angular, Piral
Big Data & Data Engineering	Apache Spark, Hadoop, Databricks, SQL optimization, ETL pipelines
Backend Development	Java EE, JPA, FastAPI, Multithreading, RESTful APIs
Cloud & DevOps	Microsoft Azure (Azure Functions, VMs, Blob Storage, Azure SQL Database), Docker, Kubernetes, Jenkins
Databases & Storage	PostgreSQL, MySQL, MongoDB, Oracle, Redis, Azure SQL Database, Hibernate (ORM)
Event-Driven Architecture	Apache Kafka, Multithreading, Asynchronous processing
Testing	Selenium, TDD practices, CI/CD automation with Jenkins, GitHub Actions, Azure DevOps

PROJECTS

AWS ETL Data Pipeline on YouTube Data – AWS Glue, AWS S3, AWS Lambda, AWS Athena

- Developed **ETL significant data pipeline** to manage, streamline, and analyze structured and semi-structured data of more than 1M YouTube videos collected using YouTube API
- Loaded data to AWS S3 using AWS CLI, converted JSON data to Apache Parquet format, automated Lambda by adding an S3 trigger, used AWS Glue as a data processing engine, and created an interactive dashboard using AWS Quick sight.

Generative AI Chatbot - *vLLM, Vector RAG's & Knowledge Graphs*

- Created an ELT pipeline to transform 20+ MOU PDFs into **Knowledge Graphs**, **fine-tuning Llama-2 7b LLM** for domain-specific expertise in labor relations. Led the development of 4 query engine models integrating LLM capabilities, deploying them on **AWS SageMaker** for real-time, reliable query responses.