

ANKIT PAL

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

SUMMARY

- **Software Developer** with expertise in **Java, Spring Boot, React**, and **cloud technologies (Azure, Databricks, Apache Spark, Hadoop)**, specializing in scalable and secure enterprise solutions.
- **Backend & API Development:** Designed and optimized RESTful APIs, **microservices**, and **multithreaded components**, improving system performance and real-time data processing.
- **Frontend & UI Development:** Built responsive **React** applications with advanced state management (Redux, Context API) for dynamic dashboards and real-time updates.
- **Big Data & Cloud:** Developed data pipelines and analytics solutions using **Apache Spark, Databricks, Kafka**, and **Azure services**, enhancing content personalization and operational efficiency.
- **DevOps & Performance Optimization:** Implemented CI/CD pipelines (Jenkins, Azure DevOps, GitHub Actions) and optimized database performance (SQL tuning, Hibernate, caching) for faster, more reliable applications.

EXPERIENCE

Software Developer (RA), SJSU Research Foundation | California, USA

Oct 2023 - Present

- Led the development of a scalable and secure **Java** and **React**-based Learning Management System (LMS), integrating multiple critical systems and third-party services for seamless content delivery.
- Designed and implemented **RESTful APIs** using **Spring Boot**, enabling smooth integration between the LMS backend and various enterprise systems while ensuring security best practices.
- Developed **multithreaded components** in **Java** for real-time content processing, **improving performance** and responsiveness in course recommendations and personalized learning pathways.
- Implemented a real-time reporting system using **React** and **Spring Boot**, providing instant insights into student engagement, course completion rates, and content effectiveness.
- Built and optimized data pipelines using **Databricks**, **Apache Spark**, and **Hadoop**, enabling dynamic content personalization and improving course completion rates.
- Integrated **Hibernate** for **database interactions**, optimizing storage and retrieval of course materials, student progress, and engagement metrics.
- Implemented event-driven architecture using **Apache Kafka**, enabling real-time processing of student activity logs, engagement tracking, and content updates across distributed systems.
- Deployed and managed cloud infrastructure on **Azure**, leveraging **Azure Functions**, **Virtual Machines**, **Blob Storage**, and **Azure SQL Database** to ensure high availability and scalability.
- Implemented **CI/CD pipelines** using Jenkins and **Azure DevOps**, automating testing, building, and deployment processes for rapid and reliable LMS updates.

Data Analyst - SDE, Optum - UnitedHealth Group | India

Aug 2020 - July 2022

- Developed **scalable software solutions** for insurance claims processing using **Python, Scala**, and **Java**, **improving data processing** efficiency and system reliability.
- Conducted in-depth data analysis using **SQL and Python**, identifying patterns that resulted in operational cost savings of \$120K annually by reducing claim processing inefficiencies.
- Utilized big data technologies, including **Hadoop** and **Spark**, to automate healthcare claims processing and detect anomalies.
- Designed and implemented **RESTful APIs** using **Java** and **Spring Boot**, enabling seamless communication between frontend applications and backend services.
- **Optimized SQL queries** for API data extraction, reducing response times by 40%, improving efficiency in claims processing workflows.
- Implemented **caching** mechanisms (**Redis**, in-memory caching) in backend services, significantly reducing database load and improving API response times.
- Developed and maintained **full-stack applications** with **React**, and **Piral**, on **Azure**, providing real-time analytics and improving decision-making.
- Implemented modular **micro-frontend architecture** with Piral, allowing **seamless integration** of team-specific analytics tools within a unified dashboard.
- Integrated **React state management (Redux, Context API)** for handling real-time updates across dashboards and ensuring data consistency.
- Utilized **multithreading** techniques in Java and **asynchronous** programming in React to improve performance of **data-intensive content** operations and real-time updates.
- Designed and automated **CI/CD pipelines** using **GitHub Actions** and **Docker**, ensuring scalable deployment and continuous integration for React and Java-based applications.
- Actively participated in all **Agile ceremonies**, contributing to sprint planning, daily stand-ups, and retrospectives in a fast-paced enterprise content management environment.

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
Frontend Development	React (Redux, Context API), Angular, Micro-frontend architecture (Piral)
Cloud & DevOps	Microsoft Azure (Azure Functions, VMs, Blob Storage, Azure SQL Database), Docker, Kubernetes, Jenkins, Azure DevOps, GitHub Actions
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

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

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.

Real-Time Ad Bidding Simulation - Ranking, Machine Learning

- Designed an RTB simulation model integrating DSP and auction algorithms, optimizing ad targeting efficiency by 15%.
- Developed ranking and conversion prediction models using XGBoost for improving ad bid strategies, achieving a 12% increase in ROI.

Group Recommendation - Reinforcement Learning – *Recommendation Systems | Python, Tensorflow*

- Developed an advanced **group recommendation** system using **deep reinforcement learning** to optimize decision-making and group satisfaction. The system dynamically improves suggestions by adapting to user interactions and feedback.

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.