RAHUL RANJAN PANDEY

Bengaluru, India | rahul.ranjan.pandey.4370@gmail.com | +91-7631921588 | linkedin.com/in/rahul-ranjan-pandey-0a13b0116/

Profile

Driven AI/ML & Data Engineer with over 2 years of experience, primarily focused on developing and deploying advanced machine learning models, AI solutions, and generative AI applications on cloud platforms such as AWS and GCP. Skilled in designing scalable data pipelines and robust ETL processes using Hadoop, Hive, Databricks and Spark to support high-impact AI initiatives, ensuring data security, compliance, and tailored, client-specific innovations.

Skills

Technical Skills: Python | SQL | ETL Pipeline Development | Data Warehousing | Hadoop | Spark | Hive

Machine Learning & AI: Machine Learning | Deep Learning | GenAI | LLMs | AWS | Databricks | GCP | PyTorch | LlamaIndex | Langchain | Data Analytics (EDA using Pandas, Numpy, Matplotlib) | AI Tools | Git | Agile

Soft Skills: Leadership | Communication | Teamwork | Critical Thinking | Creativity | Adaptability

Work Experience

Y MEDIA LABS PVT. LTD. (Now Code and Theory) | Software Engineer [AI/ML]   **Aug '22 - Current**.
- Developed tailored AI/ML models to address unique client requirements, delivering data-driven solutions that boosted operational efficiency by 35%.
- Integrated targeted data analytics and streamlined ETL processes using Hadoop, Hive, and Spark, while leveraging AWS and GCP cloud services to optimize data ingestion and enhance model performance, resulting in a 25% reduction in processing time.
- Collaborated with cross-functional teams to integrate advanced analytics and machine learning insights into robust production systems.

Automation Anywhere | Software Engineer Intern   **Mar '22 - Aug '22**
- Assisted in cloud migration and application testing to support efficient backend processes.

Projects

Project 1: Agentic GenAI-Powered Knowledge Base & Evaluation Framework   **Jan '25 – Present**
- Built a smart RAG-based knowledge assistant that ingests data from Workfront, HubSpot, APIs, and web scraping, followed by ETL and ingestion into a vector DB with TF-IDF boosted retrieval.
- Developed specialized agentic modules (search agent, relevance agent, processing agent) to improve context retrieval, response accuracy, and overall efficiency of the RAG system.
- Designed and implemented a GenAI evaluation framework, combining BLEU, ROUGE, and LLM-as-a-judge methods to assess model outputs on correctness, relevance, instruction-following, and QA performance.

Project 2: AI-Driven Universal Web Scraper for User Insights   **Sep '24**
- Developed an AI-powered web scraper to extract structured data.
- Built an ETL pipeline using Databricks (Spark), Hadoop, and Hive for scalable data warehousing and analysis, reducing processing time by 25%.
- Leveraged PySpark notebooks in Databricks for seamless data transformation and monitoring.
- Delivered user segmentation and sentiment analysis for actionable insights.

Project 3: Event Recommendation System with High-Accuracy Personalization   **Apr '24**
- Built a recommendation engine for event suggestions based on user interests, achieving an accuracy rate of 95%.
- Experimented with similarity metrics, clustering, and collaborative filtering techniques, achieving a 15% improvement in personalization accuracy.

Project 4: Customized LLM for Client-Specific Summarization and Recommendations   **Nov '23**
- Developed a custom LLM solution to provide client-specific summaries, email generation, and resource recommendations based on conversation data.
- Implemented data cleaning, feature engineering, and context-limiting using LlamaIndex.

Education

SJB Institute of Technology, Bengaluru, India   **2018 - 2022**
- Bachelor of Engineering in Information Science and Engineering | CGPA: **9.1**

Mount Assisi School, India
- *ISE (Class XII), Percentage: 74 [2017 - 2018], ICSE (Class X), Percentage: 87.1 [2015 - 2016]*

Languages

English | Hindi

Academic and Extracurricular Achievements

- SOP Award recipient at Y MEDIA LABS PVT. LTD. (Now Code & Theory) for exceptional contributions to projects.
- Earned multiple KUDOS for outstanding teamwork and proactive problem-solving.

Courses and Certifications

- Machine Learning Specialization - Coursera, 2023
- Deep Learning Specialization - Coursera, 2023
- Python for Data Science - Udemy, 2022

Hobbies

Reading Books | Hiking | Playing Chess | Watching Sports (football, cricket, etc.)