SHUB JAURA

Email | LinkedIn | GitHub | Website

Summary: A data-driven software engineer with 3+ years of experience in software engineering, data engineering, and machine learning, specializing in developing scalable solutions in Python & Java using object-oriented designs. Experienced in distributed, multi-tiered systems, algorithms, relational databases and optimization mathematics.

FULL-TIME EXPERIENCE

Software Engineer - SDE I

***Amazon**, Vancouver*   June '25 - Present

- Improved system observability by implementing requestId-based advanced logging and fixing missing logs, reducing investigation and root-cause analysis time (Java, AWS, CloudWatch).
- Supported multi-region expansion for 50+ payment plugins across ZAZ and FRA to support Amazon's internal DARU (Diversifying AWS Region Utilization) initiative to improve latency and optimizes costs and capacity.
- Revamped the payments plugins onboarding experience reducing onboarding time.

INTERNSHIP EXPERIENCE

Data Scientist

***Scotiabank**, Toronto*   Sep '24 – Dec '24

- Designed and implemented a scalable RAG chatbot leveraging *LLM* Gemini Pro 1.5 and Vertex AI, integrating distributed processing with PySpark to deliver high-performance NLP solutions that offer low latency.
- Leveraged Google Cloud BigQuery, SQL, and Python to extract financial data and applied prompt engineering (*few-shot, contextual*) with Gemini Pro fine-tuning to generate high-quality financial summaries.
- Leveraged *multi-threading & cloud resources* to parallelize summary generation, reducing time by 40%

Machine Learning Research Assistant

***University of Alberta**, Edmonton*   Sep. '23 – Apr. '24

- Engineered a Temporal Convolutional Neural Networks (TCNNs) model in Python for analyzing time series data from *NASA* turbofan engine sensors, achieving a mean absolute percentage error of 37.87%.
- Optimized hyperparameters using TensorFlow, Scikit-Learn, and Optuna, enhancing model accuracy.

Software Engineer

***ATCO Ltd**, Edmonton*   Jan. '23 – Aug. '23

- Created a full-stack web application using Flask, React, HTML, CSS, and RESTful APIs facilitating file uploads and data input to automate cost savings calculations. Worked in a collaborative, agile environment.
- Validated and help integrate 40+ features and 8 key metrics into Salesforce & performed sandbox testing.
- Designed SQL queries in Oracle Data Warehouse, and automated data integration into PowerBI

Data Engineer

***ONE Insurance Group**, Winnipeg*   May '22 – Aug. '22

- Designed and deployed end-to-end ETL pipeline using *PyAutoGUI, Python and* Azure SDK to extract legacy data, perform transformations with *Pandas*, and load it into *Azure Blob Storage*, reducing manual effort by 90%.
- Integrated Azure Cloud Storage and PowerBI, leading to real time dashboarding and $40K in annual cost savings.

EDUCATION

BSc. Computer Engineering   **Sep. 2020 – April 2025**

University of Alberta   *GPA: 3.82/4.0*

SKILLS

Programming: Python, Java, C++, SQL, MATLAB, R, JavaScript, VHDL, HTML, CSS
Data Science: Neural Networks, TensorFlow, Keras, Scikit-Learn, Pandas, NumPy, PyTorch, OpenCV, A/B Testing
Software Engineering: Flask, Django, React, Android Studio, Unix, Git, Docker, Agile, Scrum, Kanban, Jira
Data Engineering: ETL pipelines, Azure Blob, MongoDB, MySQL, Oracle, SQL Server, Hadoop, Spark, AWS S3

SOFTWARE PROJECTS

LLM ChatBot | *Python, Jupyter Notebook, RAG, LLM, Gradio, PyTorch, GPT*   Mar. '24 – Apr. '24

- Developed a Retrieval-Augmented Generation (RAG) based chatbot leveraging LlamaIndex for dynamic indexing and retrieval from diverse data sources, enabling real-time, context-aware query responses.
- Integrated Microsoft Phi-2 LLM of Hugging Face LLM to enhance scalability and optimize performance within a 4096-token context window for complex NLP applications.

EventLinkQR App | *Android Studio, Java, Firebase, API, Figma, Scrum, Git, Kanban*   Jan. '24 – May. '24

- Built a mobile app in Android Studio using Java, Google Cloud Firebase, and UUID-based sign-in, incorporating QR code scanning for event check-ins, real-time messaging and geolocation verification via Google Maps API.
- Enhanced UX with Figma, following Scrum methodology and utilizing Git for version control.