Harshit **Singhai**

SOFTWARE ENGINEER

📱 (+91) 9560948115 | ✉ harshitsinghai77@gmail.com | 🖥 fictionally-irrelevant.vercel.app | 🔗 harshitsinghai77 | 💼 harshitsinghai | 🐦 @harshit₇78

Work Experience

Deloitte USI *Gurugram, India*
DATA ENGINEER - STRATEGY AND ANALYTICS - ARTIFICIAL INTELLIGENCE AND DATA ENGINEERING *Feb. 2022 - Ongoing*

- Collaborated with a leading US asset management firm's internal audit team to architect cloud analytics for proactive risk detection
- Developed a RAG app with **AWS Bedrock and LLM**, integrating data ingestion for vector store and similarity search for content retrieval
- Optimized FastAPI and PostgreSQL performance, reducing latency by **40%** and improving response times by **35%**, enhancing user satisfaction
- Eliminated inter-job dependency failures and mismanaged cron jobs by orchestrating cross-account data pipelines with AWS Step Functions
- Architected secure **PII infrastructure** with masking and segregated storage, enhancing compliance, saving **$50k/year in potential fines**
- Saved **52 man-hours** per week by migrating auditor reporting to AWS, automating tasks, and eliminating the need for manual reporting scripts
- Streamlined analytics with **Redshift Spectrum**, enabling direct S3 queries, minimized ETL, and efficient transformations of historical data
- Enhanced data availability through robust pipeline design, achieving **99.8% uptime**, efficient ingestion from **7+** data streams & transactions
- Achieved **40%** S3 cost reduction using S3 Intelligent Tiering and replacing EMR with AWS Glue's autoscaling, reducing AWS cost
- Implemented Glue data quality rules: automate trend checks and anomaly alerts, ensuring consistent and reliable data for governance

Quantive *Indore, India*
SOFTWARE ENGINEER - PLATFORM *Aug. 2021 - January 2022*

- Built microservices pipelines for data transformation and aggregation, enabling real-time anomaly detection and OKR performance analysis
- Optimized microservices leading to **30% reduction** in memory usage & processing time through memory profiling, & enhanced code readability
- Augmented microservice built on Pandas, using efficient vectorized operations to deliver KPI insights **20% faster** than prior batch processes
- Increased data throughput by **40%** using **Celery-based** async processing, RabbitMQ as message broker, and **ClickHouse** for data storage
- Implemented multi-channel anomaly detection alerts via Slack and email notifications, reducing response time to critical issues by **10%**

Deepsource *Bengaluru, India*
SOFTWARE ENGINEER - MACHINE LEARNING *December 2020 - June 2021*

- Built AI recommendation system to guide users towards the most impactful code issues, optimizing their workflow and enhancing code quality
- Created **50k+ personalized ML models**, to predict high-impact code issues based on user's historical data, saving developers' **30 hours/week**
- Leveraged Kubeflow orchestration and Katib for **hyperparameter tuning** on Google Cloud Platform, achieving **20%** faster ML pipeline training.
- Implemented adaptive ML model retraining for continuous code recommendations, adapting to new patterns, and ensuring model relevancy
- Prioritized code quality by deploying only model versions demonstrating superior precision and ROC AUC, driving continuous improvement

Projects

- fictionally-irrelevant.vercel.app/posts/side-projects-2023: Shipped and deployed **20+** user-facing side projects, reaching over **1k users** globally
- https://fictionally-irrelevant.vercel.app : **Authored 72+** in-depth technical blog posts, delving into diverse areas like **AI, Cloud, Data, and web3**
- **MLH Fellowship**: Amongst **144** out of 30,000 applicants; Contributed to **BentoML** Open-Source Python library to build Machine Learning APIs

Accomplishments & Awards

2023	**Cohort Lead,** Deloitte AI Academy Explorer Program: Guided 80+ practitioners on Datacamp ML track	*Deloitte USI*
2022	**Applause Award,** recognized by Senior Leadership for delivering outstanding client outcomes	*Deloitte USI*
2019	**Mentor,** sophomore students for Python: Conducted and managed 30+ labs sessions and tutorials	*Bennett University*
2019	**National Finalist,** Smart India Hackathon: Improvised Kotak Mahindra Bank's virtual assistant Keya	*SIH*

Education

Bennett University *Delhi NCR, India*
BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING (CSE) - 7.98 GPA *Aug. 2016 - Aug. 2020*

- Main coursework: **Data Structures and Algorithm, Database, Artificial intelligence, Machine Learning, Cloud Computing, Big Data**

Programming Skills

Server Side	Python, Go, Node.js, FastAPI, Django, REST API
Data Engineering	Spark, PySpark, Pandas, Airflow, AWS Glue, S3, EMR, Athena, Hive, Presto, SQL, Tableau
Machine Learning	Scikit-Learn, LLM models (OpenAI, Llama 3.1), ChromaDB, Sagemaker, AWS Bedrock, Keras, Kubeflow, MLOps
Database	SQL/NoSQL, PostgreSQL, MongoDB, AWS DynamoDB, Redis
Other	AWS, Docker, Kubernetes, Git, Problem Solving, Data Structures Algorithms