Harshit Singhai

🛿 (+91) 9560948115 | 🖾 harshitsinghai77@gmail.com | 🏶 fictionally-irrelevant.vercel.app | 🖸 harshitsinghai77 | 🛅 harshitsinghai | У @harshit.78

Work Experience_____

Deloitte USI

DATA ENGINEER - STRATEGY AND ANALYTICS - ARTIFICIAL INTELLIGENCE AND DATA ENGINEERING

- Collaborated with global asset leader's internal audit team, architecting cloud analytics for proactive risk detection and improved compliance
- 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 •
- Built ML-powered fraud detection using SageMaker model, predicting anomalies with 90% accuracy, saving \$2M from averted fraud
- Drove 40% S3 cost reduction and enhanced scalability by shifting data pipelines to serverless AWS Glue with autoscaling, replacing EMR •
- Maximised analytics efficiency by replacing manual data transfers with AWS DMS synchronization, reducing overhead and ensuring consistency
- Built trust with Glue data quality: automated checks on trends, alerts for anomalies, ensuring reliable data for data governance and analytics

Quantive

Indore, India Aug. 2021 - January 2022

Bengaluru, India

December 2020 - June 2021

Gurugram, India

Feb. 2022 - Ongoing

- · 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

SOFTWARE ENGINEER - MACHINE LEARNING

SOFTWARE ENGINEER - PLATFORM

- 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

- 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
- fictionally-irrelevant.vercel.app/posts/side-projects-2023: Shipped and deployed 20+ user-facing side projects, reaching over 1k users globally

Accomplishments & Awards __

- Cohort Lead, Deloitte AI Academy Explorer Program: Guided 80+ practitioners on Datacamp ML track 2023 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

BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING (CSE) - 7.98 GPA

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

Programming Skills_____

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

Delhi NCR, India





Aug. 2016 - Aug. 2020