Harshit Gaur

Data Science and Analytics Professional

✉ harshit.gaur2720@gmail.com | ☐ +91-9582705290
○ harshitgaur2 | in harshitgaur2 New Delhi, India

Skills Summary

Languages:
Python, T-SQL, C++, Javascript

Artificial Intelligence:
Machine Learning, Deep Learning, Fine Tuning LLM, Generative AI

Technologies & Tools:
Azure Databricks, Pyspark, Azure Data Factory, Apache Spark, Azure Data Lake, Azure Synapse, Logic Apps, AWS S3, Redshift, Glue, BigQuery, Hive, Hadoop, AWS Quicksight, IAM, Azure Automation, REST APIs, Confluent Kafka, Machine Learning, Redis, Docker, Flask, Django, Git, Linux, Docker, CI/CD

Data Analytics:
Adobe Analytics, Adobe Target, Adobe Customer Journey Analytics, PowerBI, Tableau, QlikSense, Big Data Analytics, Google Analytics, Looker Studio, Adobe Real-Time Customer Data Platform

Education

Guru Gobind Singh Indraprastha University   August 2018 - July 2022
Bachelors of Technology - Computer Science and Engineering | *CGPA: 8.4/10*
Relevant Courses: Data Structures and Algorithms (C++ and Python), Data Mining and Intelligence, Machine Learning, DBMS, Algorithms and Data Structures I and II (C++), Computer Organisation and Architecture, Compiler Design

Experience

Publicis Sapient   July 2022 - Present
Associate Data Science and Analytics L1

- **Current Role: Big Data Engineer and BI Developer | Nissan NMGB, London, United Kingdom** Collaborated with their analytics team on Nissan UK website. Provided advanced analytics and performance insights for their different hubs like Aftersales, EV Hub, etc.
- **Custom Reporting and Dashboards:** Created custom Adobe Analytics reports/dashboards using Calculated Metrics, Segments, and Virtual Report Suites, delivering insights to optimize campaigns, improve user experience, and enhance conversions. Increased marketing efficiency by 45%.
- **Advanced Data Analysis and Insights:** Conducted user behavior analysis, segmentation, and cohort analysis, improving marketing ROI by 26%, boosting conversions by 60%, and optimizing user journeys.
- **A/B Testing and Optimization:** Partnered with the A/B testing team to set up Adobe Target experiments for optimizing content, layout, and offers. Utilized analytics insights to inform testing hypotheses and re-designed the user interfaces, driving a success rate of 25%
- **Campaign Tracking and Attribution:** Implemented tracking for marketing campaigns, enabling comprehensive attribution and improving channel performance insights.
- **Collaboration with Data Engineering Teams:** Worked with data engineering teams to integrate Adobe Analytics with CRM, CDP, and BI Tools for enriched customer insights and reporting.

- **Unilever, United Kingdom:** Led data instrumentation for the AI Beauty Assessment tool across 13 markets (Lakme, AHC, Dove).
- **Adobe Launch Instrumentation:** Utilized Adobe Launch for custom tracking rules and validated hypotheses through A/B and multivariate testing.
- **Tool Testing and Insights** Increased tool usage by 52% and market penetration, developed dashboards for conversions, engagement, and retention analysis.
- **Recommendations and Optimization:** Contributed to a 32% product upsell with analytics-driven recommendations.

- **Civitas Resources, Colorado:** Led migration of on-premise data to Azure Blob Storage and ADLS Gen2 for Colorado's leading oil and gas producer.
- **Data Streamlining:** Automated data transformation and storage using Microsoft Power Automate, improving operational efficiency.
- **PowerBI Reports:** Developed dynamic Power BI dashboards and reports, providing stakeholders with actionable insights. Reduced emission waste by 40% and increased infrastructure management productivity by 27%.

- **Data Science POV:** Facilitated well-received technical data team workshops to develop real-time Data applications by harnessing Langchain and Streamlit and collaborated in fine-tuning LLMs to develop in-house GenAI use cases for the Sapient Data team, benefiting 100+ members.

Projects

Retrieval Augmented Generation application using Haystack, Weviate and Chainlit　　　　Jul. 2024

Personalized Spotify tops and insights web application　　　　*Python, Git*

- Showcased an application which can store custom documents into vector embeddings and can answer the user-based questions using that knowledge vectors. This was a part of Publicis Sapient Friday Tech Yard.
- Implemented using Haystack and Weviate for vector embeddings. Created the python application using Chainlit.

FUZZY RULE BASED SYSTEM WITH DECISION TREE FOR BREAST CANCER DETECTION　　　　Jul. 2024

Machine learning application to detect early-stage breast cancer　　　　*Python, Decision Trees, MATLAB, Fuzzy Logics*

- Implemented an Expert System for detection of Breast Cancer based on a knowledge based system
- Used WBCD dataset and implemented the rule base of the system using decision tree and used MATLAB, Tkinter(GUI) and decision trees

Select Research Publications

- **Gupta, V., Gaur, H., Vashishtha, S., Das, U., Singh, V.K., Hemanth, D.J.:** A fuzzy rule-based system with decision tree for breast cancer detection. *IET Image Process.* 17, 2083–2096 (2023). Click Here
- **Elsayed, S., Gaur, H., Schmidhalter, M.:** Interpretation of the Influence of Hydrometeorological Variables on Soil Temperature Prediction Using the Potential of Deep Learning Model. *Knowledge-Based Engineering and Sciences*, 4(1), 55–77 (2023). Click Here
- **Vashishtha, S., Gaur, H., Das, U. et al.:** Optifake: Optical flow extraction for deepfake detection using ensemble learning technique. *Multimed Tools Appl* (2024). Click Here

Professional Certifications

- **Adobe Certified Professional - Adobe Analytics Developer**　　　　March 2024 - March 2026
  Certification Number: RVY2PMRCCBR1QWKQ
- **Adobe Certified Professional - Adobe Target Business Practitioner**　　　　February 2024 - February 2026
  Certification Number: 33X40DTKC1FE1ECN
- **Adobe Certified Professional - Real-Time Customer Data Platform**　　　　April 2024 - April 2025
  Certification Number: MJR80MD22MQ4QG3C
- **Microsoft Certified: Power BI Data Analyst Associate**　　　　June 2023 - June 2025
  Credential ID: 7E4BAFFACDE685D
- **Microsoft Certified: Azure Solutions Architect Expert**　　　　April 2024 - April 2025
  Credential ID: 1CEA9AC4E1D0E9E7
- **Microsoft Certified: Azure Administrator Associate**　　　　April 2024 - April 2025
  Credential ID: 590965ADF61B94EA
- **Google Analytics 4 Certified**　　　　April 2024 - April 2025
  Credential ID: 278070515

Leadership and Volunteering

Google Cloud　　　　Dec. 2023 - Ongoing

Facilitator

- Guiding a community of more than 300+ students in their professional development, technical skills and interest in tech.
- Orchestrating workshops, coding sessions, and events to enhance skills and encourage continuous learning.

Awards and Merits

Best Citation Winner - AI&Data COE　　　　Aug. 2023

Publicis Sapient

- Won the award as the Best Citation Winner in the Rewards and Recognition 2023 Cycle Q2.