

FARDEEN PATHAN

Pune | fardeenpathan.fp23@gmail.com | +91 9765710945 | linkedin.com/in/fardeen-pathan24

Summary

Data Analyst with strong experience in SQL, Python, and enterprise data workflows, specializing in data quality, reporting automation, and analytical insights to support business, risk, and compliance objectives. Proven ability to build reliable data pipelines, develop performance metrics, and partner with stakeholders to improve data consistency, governance, and decision-making across functions.

Education

DES Pune University, M.Sc. in Data Science

Oct 2023 - Oct 2025

- Coursework: Python, Statistics, Mathematical Foundation, SQL, Machine Learning, Deep Learning, NLP

Poona College, B.Sc. in Mathematics

July 2019 - May 2022

- Coursework: Topology of Metric Spaces, Compact Spaces, Elementary functions, Complex analysis.

Experience

Data Scientist – R&D, Fusion India

Oct 2024 – Nov 2025

- Developed and validated analytical and predictive models to support performance benchmarking and operational risk monitoring, improving system reliability and decision-making accuracy.
- Designed and implemented data quality checks, validation rules, and consistency controls across sensor and operational datasets to ensure reliable downstream reporting.
- Conducted in-depth analysis on high-volume equipment and operational data to identify anomalies, performance risks, and improvement opportunities.
- Automated end-to-end reporting pipelines and dashboards using Power BI, improving transparency, auditability, and reducing manual reporting effort by ~10 hours per week.
- Partnered with engineering and business stakeholders to define reporting standards, KPIs, and data definitions aligned with operational and governance requirements.

Data Analytics Intern, Nigade Software Pvt. Ltd

Jul 2024 - Aug 2024

- Executed ETL workflows on high-volume transactional data to ensure data quality, completeness, and structural consistency for analytics and reporting.
- Performed SQL-based profiling, cleansing, and reconciliation of customer and financial data to support accurate business reporting and operational decisions.
- Developed standardized Power BI reports for business and finance teams, improving reporting reliability and reducing manual effort by 40%.
- Supported documentation of data definitions, metrics, and reporting logic to improve transparency and reuse across teams.

Project

Medical Assistant Bot

- Integrated the Gemini API (gemini-1.5-pro) to provide LLM-powered medical insights, context-aware interactions, and robust fallback responses, significantly enhancing user experience.
- Built a Random Forest Classifier achieving 96% AUC and 92% accuracy for disease prediction, utilizing a large-scale symptom database.
- Utilized Redis for accelerated processing and low-latency caching, optimizing the bot's real-time performance and responsiveness for critical medical queries.
- Fine-tuned proprietary models for precise summarization, delivering accurate and concise medical results.

Publication

- Social Media and Its Impacts on the Human Brain: Integrating Psychology and Digital Behavior: Published in IJIRCCE, Volume 13, Issue 4, April 2025

Technologies

- Programming Languages: Python, SQL
- Libraries: Pandas, NumPy, TensorFlow, Pytorch
- Visualization: Power BI, Looker Studio
- Web Development: Flask, Streamlit, FastAPI
- Databases: MySQL, MongoDB
- Frameworks: PySpark, Hadoop
- Machine Learning & Deep Learning: ML/DL Algorithms, Transformers, NLP
- Computer Vision: LLM (Large Language Models), VLM (Vision Language Models)

Data and Governance Exposure

- Data quality management, validation rules, and reconciliation
- KPI definition, data standards, and reporting consistency
- Operational risk monitoring through analytics
- Audit-friendly reporting and traceability
- Stakeholder collaboration across technical and business teams

Languages

- English
- Marathi
- Hindi