

Prashant Gupta

Portfolio: <https://prashant-gtx.github.io/>

LinkedIn: <https://www.linkedin.com/in/prashant-gtx/>

Email: prashantgtx@gmail.com

Mobile: +91-9029140773

Summary

Versatile data professional with strong foundations in data analysis, data engineering, and applied machine learning. Experienced in building end-to-end data pipelines, performing exploratory data analysis (EDA), developing dashboards, and enabling data-driven decision-making. Practical experience with big data tools, cloud platforms, and AI systems including LLMs and Retrieval-Augmented Generation (RAG). Skilled in transforming raw, large-scale data into reliable analytics solutions and intelligent applications.

Education

- | | |
|---|----------------------------|
| • Centre for Development of Advanced Computing (C-DAC), Mumbai | Kharghar, Navi Mumbai |
| <i>Post Graduate Diploma in Big Data Analytics; Score: 80.63%</i> | <i>Aug 2025 - Feb 2026</i> |
| • SIES Graduate School of Technology | Nerul, Navi Mumbai |
| <i>Bachelor of Engineering - Computer Engineering; CGPA: 7.51</i> | <i>Dec 2021 - May 2025</i> |
| • DAV Public School | Nerul, Navi Mumbai |
| <i>12th Grade - CBSE; Score: 84%</i> | <i>May 2020 - May 2021</i> |
| • DAV Public School | Nerul, Navi Mumbai |
| <i>10th Grade - CBSE; Score: 90.40%</i> | <i>May 2018 - May 2019</i> |

Technical Skills

- **Programming & Databases:** Python, SQL (Joins, CTEs, Window Functions), PostgreSQL
- **Data Analysis & BI:** Power BI, Excel, Tableau, KPI Design, Dashboard Automation, DAX
- **Python Libraries:** NumPy, Pandas, Matplotlib, Seaborn, BeautifulSoup, Regex (re)
- **Big Data & Data Engineering:** PySpark, Spark SQL, Apache Hive, Hadoop, Data Pipelines
- **Machine Learning & AI:** Scikit-learn, CNNs, Transfer Learning, Model Evaluation
- **LLMs & RAG:** Llama 3.2, FinBERT, ChromaDB, Vector Search, Prompt Engineering
- **Other Tools:** FastAPI, Streamlit, Linux, AWS, Git

Projects

- **MarketPulse AI | GitHub:**
 - Built a real-time financial intelligence platform using FastAPI to aggregate financial news, market data, and analytics for retail investors.
 - Designed an Agentic Retrieval-Augmented Generation (RAG) pipeline using ChromaDB and Llama 3.2 with a self-healing mechanism that refreshes stale context using live stock data, reducing AI hallucinations.
 - Implemented automated data ingestion, sentiment analysis (FinBERT), and an alert engine to deliver personalized, sentiment-aware insights and notifications.
- **Detecting Early Stage Knee Osteoarthritis Using Deep Transfer Learning | GitHub:**
 - Developed a deep learning pipeline to classify knee X-ray images into five severity grades (0–4) for early osteoarthritis detection.
 - Applied transfer learning using CNN architectures such as ResNet and VGG to improve feature extraction and model performance.
 - Deployed an interactive Streamlit application enabling image upload, report generation, and hospital recommendations.
- **Consumer Lending Risk Insights Through Data Analytics | GitHub:**
 - Performed end-to-end credit risk analysis on consumer loan data using data cleaning, univariate, bivariate, and multivariate EDA to identify patterns driving loan defaults.
 - Conducted hypothesis testing and correlation analysis to validate key risk factors such as income levels, credit-to-income ratio, employment stability, age group, and external bureau scores.
 - Translated analytical insights into actionable credit policy recommendations including risk-based pricing, stricter approval criteria, credit limit caps, and early-warning monitoring strategies to reduce default risk.

Certifications and Achievements

- Career Essentials in Data Analysis – Microsoft & LinkedIn
- SQL (Basic, Intermediate, Advanced) – HackerRank
- AWS Academy Graduate – Data Engineering
- AWS Academy Graduate – Cloud Foundations
- Data Engineering with Hadoop and Spark – GeeksforGeeks

Positions of Responsibility

- **Administration Head – IEEE SIESGST:**
 - Coordinated 10+ events with cross-functional teams, improving event workflows and participant experience.
 - Streamlined administrative processes, reducing setup time by 30%.
- **NSS Volunteer:**
 - Contributed to multiple community programs related to cleanliness drives, awareness campaigns, and public engagement.

Additional Information

- Strong understanding of data analysis workflows including EDA, KPI development, forecasting, and statistical reasoning.
- Practical experience in building scalable data pipelines and working with big data frameworks for batch and analytical processing.
- Applied experience in machine learning, deep learning, and AI-driven applications using LLMs and Retrieval-Augmented Generation (RAG).
- Skilled at translating business problems into data solutions and communicating insights to both technical and non-technical stakeholders.
- Actively building real-world projects across analytics, data engineering, and AI to strengthen end-to-end problem-solving skills.