Nayan Jain

Jaipur, Rajasthan · nayanjain24003@gmail.com · +91 81122 84077 · GitHub · LinkedIn

EDUCATION

VIT BHOPAL University	**Bhopal, MP**
B.Tech in Computer Science · **7.89 CGPA**	October 2022 - Present
Warren Academy School	**Jaipur, Rajasthan**
12th Grade · **85%**	April 2021 - March 2022
Warren Academy School	**Jaipur, Rajasthan**
10th Grade · **90.8%**	April 2019 - March 2020

WORK EXPERIENCE

Netgraph, *Data Science Intern* — Nov 2024 - Jan 2025

- Architected machine learning solutions for customer churn prediction, generating strategic insights that enhanced customer retention protocols and elevated satisfaction metrics.
- Developed and streamlined robust data pipelines (Python, SQL), processing 100K+ records daily while reducing computational overhead by 65%.
- Amplified churn prediction precision from 72% to 89% through innovative feature engineering, hyperparameter optimization, and advanced statistical modeling techniques.
- Delivered actionable intelligence via interactive dashboards and executive reports to key stakeholders; optimized database infrastructure for 40% faster query execution and data retrieval workflows.

TECHNICAL PROFICIENCIES

Programming Languages:	Python, Java, SQL
Frameworks & APIs:	Flask, Django, FastAPI, Apache Spark
Data Science Libraries:	NumPy, Pandas, Scikit-learn, Matplotlib, Plotly, Seaborn
Database Technologies:	MySQL, PostgreSQL, MongoDB
Development & Analytics Tools:	Git, Jupyter, VS Code, Databricks, Tableau, Power BI

PROJECTS

Telecom Customer Churn Prediction System
GitHub Repository ⬈

- **Developed sophisticated ML model** leveraging Decision Trees and SMOTENN algorithms to forecast customer churn with **89% accuracy**, processing and analyzing **50,000+ telecom customer records** utilizing Python and Pandas.
- **Crafted 15+ predictive features** through comprehensive feature engineering from raw customer datasets, identifying **top 10 churn indicators** that boosted model performance by **23%**.
- **Minimized false positive rate by 35%** through systematic hyperparameter tuning, cross-validation techniques, and ensemble learning methodologies.
- **Produced comprehensive technical documentation** including feature importance analysis and performance metrics visualization using Matplotlib and Seaborn; deployed **responsive Flask web application** for real-time churn prediction and business intelligence.

(Technologies: Python, Pandas, Scikit-learn, Flask, Matplotlib, Seaborn)

Walmart Sales Data Analysis
GitHub Repository ⬈

- **Constructed automated ETL pipeline** (Python, SQL) for **45K+ sales transactions** via Kaggle API integration, slashing data ingestion time from 4 hours to 30 minutes through process automation.
- **Cleansed and normalized raw sales datasets** from Kaggle API utilizing Python and SQL, resolving **95% of data inconsistencies and anomalies**, significantly enhancing downstream revenue trend analysis accuracy.
- **Discovered 8 critical revenue patterns** through comprehensive SQL analysis (MySQL/PostgreSQL), uncovering seasonal business trends that informed strategic decision-making and **fine-tuned database queries** achieving **40% faster data retrieval** and analytical workflows.
- **Eliminated 80% of manual data processing overhead** through intelligent automation, scheduling algorithms, and robust error handling mechanisms.

(Technologies: Python, SQL, MySQL, PostgreSQL, Pandas, Plotly)

PROFESSIONAL CERTIFICATIONS

SQL (Intermediate Level)	**HackerRank**
Verification Link ⬈	