

Garv Sachdeva

garvsachdeva02@gmail.com | +91-8168896007 | linkedin.com/in/garvsachdeva0205 | github.com/garvs0205 |
kaggle.com/garvsachdeva0205 | hackerrank.com/profile/garvsachdeva0205 | leetcode.com/u/garvs0205

EDUCATION

UPES, Dehradun

Expected Graduation: May 2027

B.Tech Computer Science (Specialization in Data Science)

- CGPA: 8.18/10.0

PROFESSIONAL EXPERIENCE

Technoarch Softwares

Noida, Uttar Pradesh

Data Science Intern

June 2025 – July 2025

- Designed and optimized **SQL queries** in MySQL to process and clean **10,000+ business records**, improving data extraction efficiency for analytics workflows.
- Developed **classification and regression models** using **scikit-learn** and **XGBoost**, achieving up to **85% predictive accuracy**.
- Built **3 interactive Tableau dashboards** tracking **15+ KPIs** and automated ETL preprocessing in Python, reducing time and effort by 20%.

PROJECTS

[Amazon Fine Food Reviews](#) – *Sentiment Analytics Dashboard*

June – July 2026

- Developed an end-to-end NLP + BI pipeline: VADER sentiment scoring on 100K reviews (79.4% accuracy vs star ratings), **27 engineered features**, 4-table star schema, **11 DAX measures** including time intelligence and **Pareto 80/20 cumulative analysis**.
- Implemented **14 advanced Power BI features**: AI Decomposition Tree, Key Influencers, Field Parameters, Drill-Through, Row-Level Security, Sync Slicers, Bookmark Navigation with **sub-530ms performance** across 100K rows.
- Uncovered statistically significant sentiment drivers via AI visuals: low helpfulness ratio (0–0.14) increases negative sentiment likelihood **7.39x**; objective reviews averaged **0.395** compound score vs **0.828** for subjective reviews.

[ListingLens](#) – *NYC Airbnb Price Intelligence*

April – May 2026

- Engineered an end-to-end ML regression pipeline on **20,700+ real NYC Airbnb listings**, training and benchmarking 4 algorithms (Ridge, Random Forest, LightGBM, XGBoost); tuned final XGBoost model via RandomizedSearchCV (5-fold CV, 100 fits) achieving **$R^2 = 0.809$ and MAPE = 22.8%**.
- Created a **52 feature engineering system** spanning geospatial signals (Haversine distances to 5 NYC landmarks, K-Means geo-clustering $k=8$), target-encoded **200+ neighbourhoods** with leakage-safe train/test split methodology.
- Built a dual-channel NLP pipeline using **VADER** on **300,000+ guest reviews**, engineering sentiment score, positive %, quality score (0–100), and trend as **XGBoost features** for price prediction.
- Deployed a **full-stack ML web application on Vercel** by implementing a custom **TypeScript XGBoost tree-walker (700 trees, ~50ms inference, zero cold starts)** and an offline NLP preaggregation pipeline storing VADER metrics in SQLite for **sub-10ms review lookups** eliminating all Python backend dependencies at serving time.

TECHNICAL SKILLS

Programming & Databases: Python, SQL, R, C++ | MySQL, PostgreSQL, SQLite, NoSQL

Data Science & ML: scikit-learn, XGBoost, LightGBM, TensorFlow, Keras, Pandas, NumPy, SciPy, Ensemble Methods

Visualization & BI: Tableau, Power BI, Matplotlib, Seaborn, Excel

Statistical Methods: A/B Testing, Hypothesis Testing, Regression Analysis, Hyperparameter Tuning, PCA, Clustering

ML Deployment & Tools: Streamlit, Flask, REST APIs, Docker, MLflow, Git, GitHub, Jupyter Notebook

Cloud Platforms: AWS (EC2, S3, SageMaker), Google Cloud Platform (Vertex AI)

ACHIEVEMENTS

- **Top 10 Finalist** — **Google Cloud Gen AI Exchange Hackathon 2025 (Student Track)**: built a **GenAI-based** youth mental wellness solution, **selected from 275,000+ registrations**.
- **Kaggle Expert** — **Notebooks (5 Medals) & Datasets (3 Silver Medals)**: **Top 3%** in Datasets and **Top 5%** in Notebooks.
- **Gold Badge (SQL)** — HackerRank (2026)
- **SQL 50 Badge** — LeetCode (2026)