Sharad Raina

Boston, MA | (857) 396-6008 | LinkedIn | raina.sh@northeastern.edu

EDUCATION

Northeastern University **Boston, MA**
Master of Science, Health Informatics Expected Apr 2025
Cumulative GPA: 4 / 4
Relevant coursework: Biostatistics, Database Design, Data Mining and Machine Learning, Collecting Storing Retrieving Data

Guru Nanak Dev Dental College & Research Institute **Punjab, India**
Bachelor of Dental Surgery (B.D.S)

WORK EXPERIENCE

Mass General Brigham **Boston, MA**
Data Analyst Intern Jun 2024 – Dec 2024

- Executed SQL queries to analyze data for 4,000+ patients, generating reports on panel sizes and patient enrollment
- Reduced project timeline by 30%, leading migration of Epic data from SQL Server to Snowflake on Azure Cloud
- Developed Tableau dashboards to track KPIs (enrollment trends, insurance coverage) for 200+ care managers, enabling continuous program monitoring and performance optimization
- Analyzed 100K+ patient records in Excel/SQL to correlate HCC scores with ICMP eligibility, uncovering insurance coverage bias and informing equitable enrollment strategies

Qualtra Pharma Pvt. Ltd **New Delhi, India**
Health Data Analyst Intern Apr 2023 - Aug 2023

- Reduced report generation time by 20%, saving over 5 hours monthly by automating reports with Python and Power BI
- Analyzed 5 years of healthcare data with SQL supporting data-driven decisions for 500+ consumers across multiple regions
- Designed Python ETL pipelines integrating healthcare data from multiple source systems, enabling advanced analytics use cases across several business units

ACADEMIC PROJECTS

Uber Demand-Supply Case Study

- Identified root causes of airport trip cancellations (15% higher than city trips) and driver shortages using time-based request analysis
- Identified peak failure times, with city-airport requests experiencing 40% unavailability during early mornings and late evenings
- Mapped supply-demand gaps (up to 50%) and recommended surge pricing and incentives for critical zones.

An Ensemble Model to Predict the Length of Stay

- Preprocessed 50K+ records (handling missing values, encoding, scaling)
- Trained Linear Regression (RMSE: 2.5 days), Random Forest (RMSE: 1.8 days), and Gradient Boosting (RMSE: 1.6 days)
- Built a stacked ensemble model, achieving RMSE: 1.4 days and R^2: 0.90, outperforming base models.

Income Prediction Using Machine Learning: Naive Bayes, Logistic Regression & Decision Trees

- Trained Logistic Regression (85% accuracy), Decision Tree (83%), and Naive Bayes (78%) on census data.
- Achieved 86% precision with Logistic Regression (top individual model).
- Ensemble model (weighted voting) improved accuracy to 88%, surpassing base models.

ADDITIONALS

- **Programming Language** - Python, R
- **Database Analysis** – Snowflake, Python (NumPy, Pandas), MySQL, Azure Data Studio, SQL Workbench
- **Database Visualization** - Power BI, Tableau, Matplotlib, Seaborn