

HARSHIT GAUR

Boston, MA | gaurharshit04@gmail.com | +16173960129 | **Portfolio:** <https://harshitgaur6155.github.io/>
www.linkedin.com/in/harshitgaur-/ | public.tableau.com/app/profile/harshitgaur | www.github.com/harshitgaur6155

TECHNICAL SKILLS

Programming Languages, Databases:	Python, R, SQL, MySQL, MSSQL, PostgreSQL, Databricks, PHP, JavaScript, Laravel
Machine Learning Libraries:	Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Fastai, SciPy
Data Visualization Tools:	Tableau, Microsoft Power BI, Microsoft Excel, R Shiny, Google Data Studio
Cloud Tools and Platform:	AWS, Microsoft Azure, GCP, Amazon RDS, Amazon S3, GIT, JIRA, Jupyter, Shopify

EDUCATION

Northeastern University <i>Master of Professional Studies in Analytics</i> Concentration: Machine Learning and Artificial Intelligence	Boston, MA Sept 2021 - Jun 2023 GPA: 4.0
Guru Gobind Singh Indraprastha University <i>Bachelor of Technology</i> Concentration: Information Technology Awards: <i>magna cum laude</i> ; Silver Medalist in Single's & Double's Badminton Cup	Delhi, INDIA Aug 2013 - Jul 2017 GPA: 3.71

WORK EXPERIENCE

Software Engineer, Mobikasa Private Limited <ul style="list-style-type: none">Achieved translation of audio to lingual using Microsoft Azure Cognitive Services with 67% accuracy within the project of an established US-based hospital chain (BSSNY) to record & track debrief-timeout audio sessions of surgeries and operations.Developed e-commerce marketplace website 'AnytownUSA', deployed on AWS EC2 instance, which served 100,000 US citizens.Accomplished improved customer-oriented sections on the website with 30% increased sales by integrating Google Analytics to monitor traction and spearheading data mining techniques to investigate customers' shopping behaviors and patterns.<ul style="list-style-type: none"><i>Achievements: Got recognized by CEO for technical prowess and adaptability to cater to customer requirements.</i>	Dec 2018 - Aug 2021, Delhi, India
Associate Software Engineer, Appster LLP <ul style="list-style-type: none">Created RESTful APIs with MVC architecture implementation using scrum methodology. Assisted in User Story Analysis.Designed database architectures. Incorporated cryptography for 2-way security in applications, and payment gateways.<ul style="list-style-type: none"><i>Achievements: Considered for the role of Business Analyst. Commended with Best Performer of the Month twice.</i>	Jul 2017 - Dec 2018, Gurugram, India

ACADEMIC PROJECTS

Airbnb Price Estimator (Python, SQL, Streamlit, HTML) (link) <ul style="list-style-type: none">Created dynamic Airbnb price estimator, deployed on HerokuApp, using predictive analytics taking in user input to calculate price.Predicted price using Decision Tree, Random Forest, XGBoost algorithms and LIME, SHAP for interpretability.
Prediction of the US Presidential winning political party (Python, R, SQL) (link) <ul style="list-style-type: none">Built classification machine learning model using Logistic Regression to analyze 14 years data to predict winning party in counties.Found significant factors of swing battleground states at 49% to 51% votes to target election campaigns effectively.Executed data collection and feature engineering to merge 8 datasets. Transformed 5 important variables to augment dimension.Performed exploratory data analysis for statistics and feature selection by Lasso Regularization, Stepwise Selection methods.
New York Tree Census 2015 Analysis using Tableau Dashboard (Tableau, SQL) (link) <ul style="list-style-type: none">Reported major causal factors of tree problems in NYC, examined by tree survival rates and health attributes of dominant 5 species.Detected root and trunk problems being major in trees, measured by analyzing 3 species covering 34% plantation across the city.
Prediction of readmittance of diabetic patients using analysis of HbA1c value (Python, SQL) (link) <ul style="list-style-type: none">Uncovered that readmission of diabetic patients depends on HbA1c value above 7 and primary diagnosis.Achieved readmission predictions of patients using k-NN, Random Forest Classification, Decision Tree, and SVM algorithms.
Deterministic Topic Modeling & Clustering (Weka, SQL, MALLET) <ul style="list-style-type: none">Evaluated improvement in finding set of topics over collections of textual data in document using text mining and text analysis based on topic modeling technique of Latent Dirichlet Allocation (LDA) and K-means clustering technique.
Business Solution to identify inaccuracy in ADCS prediction for company - DocDigitizer (Python, SQL, Microsoft Azure) (link) <ul style="list-style-type: none">Examined that 21% times Automated Data Capture System (ADCS) of DocDigitizer extracts incorrect information.Built predictive model (Decision Tree) identifying incorrect extracted information & factors responsible for inaccuracy of ADCS.