

RAJARAM GAUTAM

Data Engineer



832-420-2131



gautam.rajaram15@gmail.com



/in/rajaram-gautam/



/rajaramgautam

SUMMARY:

Proud Army Veteran and Serving Army Reserve as Information Technology Specialist with active secret clearance. Enhancing my skills to be Data Scientist who will add value to your business by helping you make decision based on data. Self-motivated, result oriented, team player, and autodidact data scientist looking to make a difference via data to decision.

TECHNICAL SKILLS:

SQL	Python	Tableau
Matplotlib	Seaborn	Applied Statistics
Machine Learning	Anaconda	Git
Spreadsheet	Storytelling Spark	Deep Learning
Data Warehouse	AWS	Snowflake
Oracle Database	Natural Language Processing	

EDUCATION:

Gathi / Infostretch College Graduate Bootcamp May 2022

Fully immersive, project-based 4-week career accelerator that provides students with 107 hours of expert instruction in applied data science in areas of Data Warehouse Concepts, SQL (PostgreSQL and Oracle), Python, Tableau, AWS, and Snowflake

Codeup - Certificate of Completion March 2022

Fully immersive, project-based 22-week career accelerator that provides students with 670+ hours of expert instruction in applied data science. Students develop expertise across the full data science pipeline (planning, acquisition, preparation, exploration, modeling, delivery), and become comfortable working with data to deliver actionable insights to diverse stakeholders.

American Military University December 2021

Undergraduate Certificate in Oracle Database

SQL, Python, JavaScript, Java

Tribhuvan University December 2011

Master of Physics

(Major in Astrophysics and Atmospheric Physics)

Bachelor of Physics

(Minor Mathematics and Meteorology)

WORK EXPERIENCE:

Information Technology Specialist: December 2020 – Present

Install, operate, and maintain military computer systems and LAN. Assist in planning, configuration, management, and monitoring of WAN and assist in maintaining network equipment.

Business Entrepreneur: October 2018 – August 2021

Operated and supervised convenience store in Belton, Texas. Employee management, daily ordering merchandise based on sales trend. Verified employees pay stuff. Placed work order and verified it is done properly. Worked as Real Estate Agent at free time and managed strip center as owner.

Junior Web Developer March 2018 – September 2018

Built website as customer demands on Joomla Content Management System. Updated various plugins of website for clients. Trained customers on basic operation of website. One of the websites made by me: <https://shewillfoundation.org/>

Health Care Specialist February 2014 – January 2018

Assisted with inpatient and outpatient care. Maintained medical record and ensure medical readiness of the organization. Provided medical support to training events

DATA SCIENCE PROJECTS:

Enron Email Analysis March 2022

Unsupervised Machine Learning Project using NLP and Time Series and determined identifiable trends in email topics over time using topic modeling. Our analysis could help organization to identify key themes of communication between person of interest in early stage of investigating suspicious activity.

Natural Language Processing February 2022

Determined language used in a repository based on the contents of that repository's README file. Web scraped the repo name, programming language, and README contents from 230 repositories on GitHub. Using a Logistic Regression model, we were able to achieve 13.5% accuracy in our predictions, beating the baseline.

Individual Project January 2022

Modeled the New York City Property data using properties features, the Polynomial Model with degree 2 produced the best results with RMSE values of 285585, 283355 and 284915 for train, validate and test dataset. My model performed 9.40% better than the baseline model (when comparing training data set) and 8.00 % better than the baseline model for validate dataset.

Clustering Project December 2021

Exploring Zillow's data, our model predict the drivers of errors in Zestimate with little bit better accuracy than baseline accuracy. Clustering was done based on location features (latitude, longitude, zip code) In most of our models, their performance was better without clustering than with clustering.

Linear Regression December 2021

Created multiple regression models utilizing home features to predict home tax values using Zillow data. Among various models, Polynomial Model with degree 2 produced the best results with RMSE values of 208506, 211676 and 213560 for train, test, and validate dataset. Our best model performed 14.30% and 13.21% better than baseline for training and validate dataset.

Classification November 2021

To Identify driving factors of customer churn from Telco Inc. with help of data provided which is in my GitHub. Our best Random Forest Classifier model performed 7% better than baseline and had an accuracy of 80% on unseen test data.