# Shreyas Garga, Data Scientist/Machine Learning and AI Engineer

Phone – (703)-986-6648   Mail – garga.shreyas@gmail.com   LinkedIn - www.linkedin.com/in/shreyas-garga   GitHub - sheygarga

## Summary

Driven data analyst with a Bachelor's in Neuroscience and experience in business and social analytics.  Has created numerous automated and responsive data analytics applications and solutions to identify, analyze, and interpret trends and patterns. Excellent research and analytical skills.

## Skillset

| | |
|---|---|
| Python | Pandas, Numpy, SciPy, MatPlotLib APIs, Tweepy, Vader, Seaborn, PySpark, PyTorch, NLTK, SpaCy, Gensim, Scrapy, |
| Database | MySQL, MS SQL Server, MongoDB, SQLAlchemy, Flask |
| Visualization | Tableau, PowerBI, Excel, Bootstrap, Plotly, Bokeh, D3, ChartJS, Leaflet, Carto |
| Web | Javascript, Angular, JQuery, NodeJs, React, CSS |
| Big Data | R, TensorFlow, Spark, Hadoop, NLP, MRJob, Keras, SKLearn, AWS, GCP, Azure |

## Employment

### Octo/United States Citizenship and Immigration Services                                     Fairfax, VA

Scrum Master, Team Lead ML/AI Development, Data Analyst, Support                     Aug 2019 – Current

**Project:** Automation of application approval process leveraging Artificial Intelligence and Machine Learning.

**Technologies**: Python, R, SQL, Databricks, Jupyter, AWS, S3

**Responsibilities**:

- Prepare data for analysis using data validation, data cleansing and data quality assessment.
- Analyze data using LASSO Regression, correlation matrices, and stepwise subsets for data analysis.
- Identify most important variables and determine weights for assessing application risk level.
- Define optimal algorithm for risk assessment.
- Select model for machine learning using K-means clustering, decision trees and neural net.

**Project:** Data Analytics to Optimize Workflows and Minimize Backlog

**Technologies:** Databricks, Tableau, SAS, SAS Viya, SAS EG, Oracle OBIEE

**Responsibilities:**

- Extract, transform and load structured and unstructured data to prepare it for data analysis.
- Utilize K-means clustering to find correlations between cases, decision trees based on agents' protocols for case evaluation, and neural networks to process the numerous variables and factors.
- Create interactive dashboards for employee evaluation, decision-making, trend analysis, and drill-down reporting.
- Establishing and maintaining best practices for a variety of data analytics tools available to our users

### Administrative Office of the US Courts                                     Washington DC

Lead Data Analyst – IT Security                                     Oct 2018 –Aug 2019

**Technologies**: MySQL, Excel, Tableau, Tableau Prep, d3, PlotLy, and Numpy

**Responsibilities**:

- Design and create dynamic and interactive Excel and Tableau Dashboards for the executives to help analyze threats, identify trends and make strategic decisions.
- Import Security Threats data from the judiciary units across the country into MySQL and write Python scripts to convert data into intelligent information.
- Present findings to the management and prepare reports for the executives to present in Federal and State Level regulatory meetings.
- Responsible for prompt and accurate analysis to help ensure cybersecurity for the judiciary.

### Avyuct                                     Herndon, VA

Developer and Data Analyst                                     May 2017 – Oct 2018

**Project**: Anomaly Detection

**Technologies:** COCO, RCNN, AWS, Sagemaker, TensorFlow, Python

**Responsibilities**:

- Set-up environments on Amazon Web Services  to enable development and testing.
- Develop Python Scripts to convert datasets to TFRecords
- Integrate pre-trained COCO Datasets for Transfer Learning.
- Configured pipeline to tune model and training parameters.

**Project**: Skills Gap Analysis

**Technologies**: Neo4j, Python, NLP

**Responsibilities**:

- Develop Python Libraries for dependency parsing using NLP
- Write scripts to clean and process dependency parsing results and load to Graph database.
- Customize properties of nodes and relations to draw inference and make recommendations.

**Project**: Sentiment Analysis

**Technologies**: Twitter, Python, Tweepy, Pandas, NumPy, TensorFlow, Keras

**Responsibilities**:

- Extracted training data and relevant tweets using Pandas, NumPy, and Tweepy.
- Trained the model using TensorFlow and Keras
- Transformed the script to accept inputs from the company website for trial usage

## Education

**George Washington University**: Data Analyst/Data Visualization Certificate

**George Mason University**: BS in Neuroscience                     Aug 2013- Dec 2017