RAVJOT SINGH

+919654446404 ravjot313@gmail.com linkedin.com/in/ravjot03 github.com/Ravjot03 medium.com/@ravjot03

OBJECTIVE

Highly analytical and inquisitive Data Scientist always looking to build data intuition and to get better at using modern analytical tools and cutting-edge statistical methodology. Seeking role with an organization that values quick learning, critical thinking, data driven, and dedication to achieving goals.

SUMMARY

- ✓ Active contributor, and author on Medium and GitHub. Published multiple articles and projects.
- ✓ Experienced at creating data regression models, using predictive data modelling, and analyzing data mining algorithms to deliver insights and implement action-oriented solutions to complex business problems.
- ✓ Identify, analyze, and interpret trends and patterns in complex large data sets.
- ✓ Strong knowledge in statistical analysis and model building using Python and R as well as strong SQL skills.
- ✓ Experience in data visualization and dashboarding using Python (Matplotlib, Seaborn, Bokeh, Plotly); R(ggplot).
- ✓ Excellent in data-driven communication with the ability to wear many hats and work in an end-to-end environment.

TECHNICAL SKILLS

Statistical Programming Languages. Python (NumPy, Pandas, SciPy, Scikit-Learn, TensorFlow, Keras, Scrapy, Requests, OpenCV, PyTorch, Selenium, BeautifulSoup); R (Dplyr, Rvest, Plyr)

Database Management Systems. MS SQL Server; MySQL

Data Discovery/Visualization Tools. Python (Matplotlib, Seaborn, Bokeh, Plotly); R(ggplot); Power BI; Tableau

Statistical Concepts. Regression Modeling, Classification, Clustering, Random Forest, K-Nearest Neighbor, Support Vector Machine (SVM), Time Series Forecasting

WORK EXPERIENCE

VARIDUS – SINGAPORE **May 2020 – August 2020**

Data Science Analyst

Varidus is a Venture Builder firm based in Singapore, helping technology startups scale via introduction to capital, partnerships, and ancillary services.

- ✓ Modelled Merger &Acquisitions data on different startups all across the globe to predict whether the startup would go for a merger or not in the future.
- ✓ Fetched raw M&A data from Kaggle, TrackXN (Andreessen Horowitz Dataset), and CBInsights (AI startups).
- ✓ Identify and interpret trends and patterns in the financial information.
    - Distribution of startups under different categories and their acquisitions category.
    - Funding and investments under different categories across the globe.
    - Amount raised by different startups across the globe.
    - Company stage distribution across the globe.
    - Funding amount vs stage distribution across the globe, etc.
- ✓ Conducted research and analysis to assess the competitive effects, trends, and potential investments in FinTech, Blockchain, AI, Logistics, SaaS, and EdTech startups in the South East Asian countries.
- ✓ Collaborated with the Investment Banking Team at Varidus to generate leads by finding information on Investment Banking Firms based in South-East Asia.
- ✓ Performed web scrapping to capture Potential Family Offices' from all over the world, whose interests lie in Venture Capital Investments.

- ✓ Worked closely with the Venture Capital Team and presented pitch decks of the results and recommendations obtained by performing analysis on data from VentureCapInsights database.
- ✓ Prepared financial and market analysis of specific geographic areas to determine attractiveness of growth opportunities.

SENSEGRASS INC – DELHI, INDIA **April 2020 – June 2020**

Machine Learning Intern

SenseGrass offers 360 farming solutions based on cutting edge technology like Nano-Satellite Mapping, Rover Bots and AI-based mobile & web application that makes farming super-efficient & easy.

- ✓ Researched on remote sensing indices to build systems to source, curate & predict the best sustainable practices to improve crop yield.
- ✓ Retrieved Native Plant Nursery data from Michigan Native Plants database using Web – scraping. Created a robust crawler to scrape all the tabular data systematically.
- ✓ Performed statistical analysis (mean, median, mode, percentiles, variance, etc.) and derived valuable insights from Normalized Difference Vegetation Index (NDVI) Images.
- ✓ Wrote python scripts for data cleansing, eliminating duplicate, and inaccurate data.
- ✓ Worked on different interpolation techniques to evaluate unknown values in Geographic images.
- ✓ Carried out geospatial analysis, including data visualization using Rasterio library.
- ✓ Study and Research work included but not limited to - "Evapotranspiration" utilizing online resources including Food & Agriculture Organization of United Nations (FAO), United States Geological Survey (USGS).

INDIAN STATISTICAL INSTITUTE – KOLKATA, INDIA **May 2019 – July 2019**

Summer Intern

- ✓ Worked on Image processing algorithms in OpenCV (Python).
  - Learned to apply different Geometric transformations.
  - To convert images to binary images using Thresholding.
  - Smoothing images with custom kernels.
  - Applying Morphological and Hough Line transformation.
  - Learned to find image gradients, edges with Canny Edge Detection.
  - Image Segmentation with Watershed algorithm.
- ✓ Hands on Computer Vision techniques including Modelling and Perspective projection, Coordinate transformations, Camera calibrations, Epipolar geometry, Harris corner formulation, Histogram of Oriented Gradient (HOG) feature etc.
- ✓ Implemented Modelling, Rendering, Animation, Fast algorithms, Boundary filling algorithms, Polygon meshes, Hidden-Surface removal, Phong illumination model, Gouraud and Phong shading algorithms in Computer Graphics.
- ✓ Started Machine Learning by hands-on session in Python packages ranging from Numpy to Pandas and made visualizations using Matplotlib. Implemented k-Nearest Neighbors (k-NN) Classifiers in Supervised Learning and K-Means Clustering in Unsupervised Learning. Applied k-NN Classification Model on Hand-Written Digits Dataset and Image Segmentation using k-Means clustering. Also, implemented Principal Component Analysis (PCA) for dimension reduction in unsupervised learning. Moreover carried-out Linear and Logistic Regression models on different datasets.

PROJECTS (Independent Study)

1. Image Segmentation using K-Means Clustering in Unsupervised Learning

2. Binary Classification of Handwritten 0's and 1's using Logistic Regression
3. Image Segmentation in Satellite Images using Neural Networks
4. Neural Networks on IRIS Dataset
5. 3-D Models using 3DF ZEPHYR Software in Computer Graphics
6. Face Detection Using OpenCV in Python
7. Real-Time Edge Detection Using OpenCV in Python
8. Hand-Written Digits Classification Model using KNN in Python
9. Parkinson's Disease Model in Python
10. Pulsar Star Prediction Model in Python
11. Heart Disease Prediction Model in Python
12. BMI Classification Model using KNN in Python
13. Linear Regression Model on Concrete Dataset in Python
14. Diabetes Prediction Classification Model in Python

CERTIFICATIONS

- ✓ **Introduction to Data Science in Python (University of Michigan)** *(COURSERA):* Course Credentials
- ✓ **Web Scraping in Python with BeautifulSoup & Scrapy Framework** *(UDEMY)*: Course Credentials
- ✓ **Computer Vision with OpenCV library in Python** *(UDEMY)*: Course Credentials
- ✓ Certificate of Participation in **Artificial Intelligence and Machine Learning Workshop** at *IIT-ROORKEE*
- ✓ Certificate of Participation in **Artificial Intelligence and Machine Learning Workshop** at *IIT-DELHI*
- ✓ **Correlation and Regression in R** *(DATACAMP)***:** Course Credentials
- ✓ **Data Visualization with ggplot in R** *(DATACAMP)***:** Course Credentials
- ✓ **Merging DataFrames with pandas** *(DATACAMP)***:** Course Credentials
- ✓ **Working with Dates and Times in Python** *(DATACAMP)***:** Course Credentials
- ✓ **Exploratory Data Analysis in Python** *(DATACAMP)***:** Course Credentials
- ✓ **Data Visualization with Seaborn** *(DATACAMP)***:** Course Credentials
- ✓ **Supervised Learning with scikit-learn** *(DATACAMP)***:** Course Credentials
- ✓ **Manipulating DataFrames with Python** *(DATACAMP)***:** Course Credentials
- ✓ **pandas Foundations** *(DATACAMP)***:** Course Credentials
- ✓ **Cleaning Data in Python** *(DATACAMP)*: Course Credentials
- ✓ **Python Data Science Toolbox (Part 2)** *(DATACAMP)*: Course Credentials
- ✓ **Python Data Science Toolbox (Part 1)** *(DATACAMP)*: Course Credentials
- ✓ **Intermediate Python for Data Science** *(DATACAMP)*: Course Credentials
- ✓ **Importing Data in Python (Part 2)** *(DATACAMP)*: Course Credentials
- ✓ **Importing Data in Python (Part 1)** *(DATACAMP)*: Course Credentials
- ✓ **Intro to Python for Data Science** *(DATACAMP)*: Course Credentials
- ✓ **Intro to SQL for Data Science** *(DATACAMP)*: Course Credentials

EDUCATION

B.Tech in Electronics & Communication Engineering — **2017 – 2021**
GURU TEGH BAHADUR INSTITUTE OF TECHNOLOGY, GGSIPU — CGPA: 8.1

Higher Secondary School Certificate — **2016 – 2017**
S.S. MOTA SINGH SR. SEC. MODEL SCHOOL — Percentage: 79%

Secondary School Certificate — **2014 – 2015**
S.S. MOTA SINGH SR. SEC. MODEL SCHOOL — CGPA: 8.6