Vipul H. Harihar

(917) 214-5091 | vhh2105@columbia.edu | linkedin.com/in/vipulharihar | github.com/virslaan | linktr.ee/vipulharihar

EDUCATION

Columbia University, Columbia Engineering New York, NY
M.S. in Data Science, GPA: 3.6/4.0 Sep 2021 – Dec 2022

- Courses: Algorithms, Applied Machine Learning, Deep Learning, Statistics & Inferential Modeling, EDAV, Computer Systems
- Honors**:** Data Science Department Representative, Professional and Leadership Fellow, Enhanced Teacher Assistant (Algorithms)

DIT University, School of Computing Dehradun, IN
B.Tech. in Computer Science & Engineering, special. in Big Data Analytics, GPA: 9.62/10 (in Top 2 % students) Aug 2016 – May 2020

- Courses: Software Engineering, Artificial Intelligence, Relational Databases, Data Science, OOPs, and Engineering Mathematics
- Honors: Computer Science Program Topper Gold Medal, University Bronze Medal Awards and Scholarship,

SKILLS & CERTIFICATIONS

Languages: Python, Java, R, SQL, HTML, CSS
Python Packages: NumPy, Pandas, SciPy, OpenCV, Scikit-Learn, Keras, TensorFlow, Matplotlib, Seaborn
Databases: MS SQL Server, MongoDB, Google BigQuery, Google Firebase
Data Visualization: Tableau, Power BI, R, Google Data Studio, IBM Cognos Analytics, ggplot2
ML Techniques: Classification, PCA, Sentiment Analysis, Time-Series, Cluster Analysis, Regression
Certifications: IBM Cognos, Data Analysis Using Python (IBM), Applied Data Science with Python-Level 2 (IBM), Microsoft Azure

PROFESSIONAL EXPERIENCE

IBM Bengaluru, IN
Application Developer Feb 2021– Aug 2022

- Managed 2 MySQL databases, performed data modeling & data driven analytics for insights on 5500+ records for client monthly.
- Solved and coded in Java to resolve the data duplication that resulted in saving 182 minutes and improved data flow access time

Associate Systems Engineer Jan 2021– Feb 2021

- Performed about 45 bug fixes on IBM software modules in Python, collected and cleaned data sets.
- Completed more than 23 courses on Data Science IBM training in 29 days and became IBM Golden Employee 2021.

Indian Roads Safety Campaign, Solve Foundation Dehradun, IN
Policy Intern Jul 2020– Dec 2020

- Developed predictive Machine Learning project to classify 75.86 mi^2 dangerous accident-prone area in Dehradun city.
- Supervised 6 team members to collect data and conduct data cleaning, and its predictive analytics for the project.
- Completed the project in 176 days and it got adopted by Non-Government Organization of Uttarakhand - Gram Welfare Society.

IBM Dehradun, IN
Application Developer-Java Full Stack Developer - Intern Feb 2020 – Jul 2020

- Created food delivery app in Spring Boot to solve food crisis in Covid-19 global pandemic outbreak situation for 29 restaurants.
- Led 5 interns to code – Web based restaurants statistics dashboard, performs and compute popularity and impressions in 47 days.
- Trained team of 6 members to perform complex SQL joins on 20 databases that facilitated regression analysis for project.

ONGC, Ministry of Petroleum and Natural Gas, Government of India Dehradun, IN
Summer Scholar - Big Data Engineer Intern May 2019 - Jun 2019

- Scraped open-source data using beautiful soup Python packages on 5 news websites and leveraged filtering on trending keywords.
- Pre-processed big data on Hadoop and used BigQuery (GCD) for 39 days to get insights on oil and its relevant news in media.

DATA SCIENCE PROJECTS

Columbia University New York, NY
Enhanced Credit Card Fraud Detection- Multi Class Classification (Imbalanced Datasets) Feb 2022 – April 2022

- Engineered a 3x faster single-click machine learning solution; automated classification algorithms using Gradient Boosting.
- Programmed data-agnostic codes for Logistic Regression, Naïve-Bayes, Decision Tree, and Random Forest algorithms in Python.
- Structured evaluation metrics: Recall Rate, and ROC-AUC values to detect optimized model for fraud multi-class classification.
- Conceptualized features and their importance in 5 team members to successfully classify imbalanced data of 9062 records.

DIT University Dehradun, IN
Machine Learning based Imaging System for Roads Defect Inspection Aug 2019 - Feb 2020

- Extracted and pre-processed 2000+ images to classify more than 38 potholes using OpenCV library in Python.
- Supervised a 3 members team to complete the project in 7 months and got appreciated by Vice Chancellor on Annual Project Fair.