D-61 Nigahi Colony
Singrauli, M.P. - 486884
himanshu6513@gmail.com

Himanshu Kumar Singh

https://www.linkedin.com/in/himanshu-s-b108959a/

github.com/Himanshu6513

+91 83196 37167

Education

SRM University, Chennai **May 2014 - May 2018**

Bachelor of Technology in Computer Science and Engineering GPA: 7.97/10

Technical Skills

- **Programming & Scripting Languages:** ***Python***, familiar with C & Java, SQL, Bash, Matlab
- **Platforms & Tools:** Cloud computing (primarily AWS), Linux, Excel automation, Web Automation
- **Data analysis, modelling, visualization:** Pandas, NumPy, Seaborn, Matplotlib, Plotly, Beautiful Soup, Selenium, PySpark
- **Other skills:** Algo trading, Data cleaning, Web scraping

Certifications

Chartered Financial Analyst (CFA) Level 1
Finance for Non-Finance Professionals (Coursera)
Complete Python Bootcamp (Udemy)
Python for Data Science and Machine Learning Bootcamp (Udemy)
Statistics for Data Science and Business Analysis (Udemy)

Professional Experience

Freelancer **September 2020 – Present**

Automating various process involving data mining and task automation as per client requirement on Ministry of corporate affairs website using selenium and various python libraries. And preparing master data of all companies registered in India using data mining libraries.

Orian Research, Pune — Market Research Intern **January 2020 – February 2020**

Preparation of collaterals, reports, samples and company profile writing. Gained experience in market segmentation, company analysis and customization of sample reports as per client's requirements.

L&T Infotech Limited, Pune — Graduate Engineer Trainee **August 2018 – June 2019**

Worked with the Application Engineering Group (AEG). Responsibilities included preparation of insurance forms for Accord Forms for multiple states in the US; testing enterprise resource planning (ERP) features for multiple countries and internal divisions of the firm Marsh & McLennan Companies (MMC).

Academic and Personal Projects

Melbourne housing market analysis and price prediction

Performed data cleaning and feature extraction from a Melbourne housing market dataset consisting of ~35k entries using Pandas and NumPy. Visualized the trends in the housing prices and preferences of residents around the Central Business District using Seaborn and Matplotlib. Implemented a price prediction tool, using the linear regression model, with inputs like year of purchase, size of the house, amenities, etc.

Tracking system for live market environment

Designed and implemented a tracking system to generate feature plots in a live market environment. Used Beautiful Soup web scraping library to gather data such as option writer data, delivery percentage, value-at-risk(VaR) and FII, DII, etc. on a periodic basis. Tracked daily buy/sell data from the NSE website using Apscheduler and Selenium, to generate feature plots using Matplotlib Animation.

Analysis of COVID-19's impact on various stock indices of NSE

Exploratory data analysis of COVID-19's impact on various NSE stock indices (such as NSE50, NIFTY IT, NIFTY FMCG, etc.) with special focus on their price, volume, P/E, P/B attributes. Used Python libraries to provide visual insights of the impact by juxtaposing the indices alongside public data such as open, high, low and close prices, trade volume of the indices, etc.

Credit card approval prediction

Used a dataset consisting of ~27k credit card users' information to perform data analysis and attribute visualization using Python libraries, and generated credit scores for each user based on their attributes. Partitioned the dataset into training and test data to build a ML model based on the random forest algorithm, to predict a user's eligibility for a new credit card with 72% accuracy.

Movie recommendation system

Implemented a movie recommendation system using a correlation factor between movies viewed by the user and all movies available in the library. Performed data cleaning, user and movie data merging and data analysis using NumPy and Pandas, feature visualization using Seaborn and Matplotlib, and finally, created a pivot table with corrwith function to provide movie recommendations.

Category prediction system for Udemy course catalog

Performed data cleaning and analysis using NumPy and Pandas for a dataset consisting of more than 8k Udemy courses with attributes such as price, subscriber count, category, etc. Split the dataset into training and test data to train a ML model based on KNN algorithm to predict a course's category.