Rakshit Sinha

(301)-792-8228 ● rakshit.sinha@marylandsmith.umd.edu ● LinkedIn● Kaggle Profile● GitHub Profile ● Portfolio

EDUCATION

University of Maryland, Robert H. Smith School of Business — **College Park, MD, USA**
Master of Information Systems — December 2022

- Runner-up in "UMD SAC Datathon-2022 organized by Deloitte" at UMD (College Park)
- Runner-up in "Best Cybersecurity AI Project" category in "Info Challenge" at UMD (College Park)
- Best 15 Hacks in Hacklytics-2022 at Georgia Tech
- Terrapin Scholarship, Recipient

R.V. College of Engineering, — **Bengaluru, KA, India**
Bachelor of Engineering, Electrical and Electronics, GPA: 3.8 — May 2018

- Reached Quarter-Finals in All India Datathon organized by IIT-Mumbai's HackTrack-2016
- Won All India National Science Scholarship

TECHNICAL SKILLS

- Tools: MS Power BI, Tableau, Google Analytics, Octoparse, MS Power Apps, KNIME, Git, Jira etc.
- IDE: JupyterLab, VS Code, RStudio, MySQL
- Cloud Platforms: AWS, MS Azure, GCP, Heroku
- Programming languages: Python, R, SQL, PySpark, Bash, HTML, CSS, Javascript
- Quantitative Ability: Statistics, Calculus, Algebra, NLP, Machine Learning, Object detection

WORK EXPERIENCE

Optimal Solutions Group, LLC — **College Park, MD, USA**
Data Science Intern — **May 2022 – Present**

- Developed and Deployed (on AWS) an image classification model for NARA (National Archive and Record Administration) using Tensorflow's NasNet architecture capable of classifying images into 11 different categories with an overall accuracy of 88.56%.
- Developed an application to increase the 'user penetration' of the newsletters published by "Optimal Solutions Group" using advanced NLP techniques (like BERT) to increase the 'click through rate' from 7.4% to 25.8%. The NLP model returned the likelihood of a newsletter being classified as 'spam' and recommended words to be replaced.
- Built a Web crawler REST API capable of web scraping as it crawled from one link to another using libraries like AutoScraper, urllib, fake_useragent, Flask, bs4 etc. Due to the time delay, the API avoided being detected as a bot and efficiently crawling through multiple pages, ultimately giving results in json format.

HCL Technologies Pvt. Ltd. (3+ Years) — **Noida, UP, India**
Analyst — **July 2018 – August 2021**

- Conducted predictive analysis to forecast total memory utilization for servers present in client's environment, leading to reduction in memory bottlenecks and faster processing.
- Developed automated dashboards for coherent display of statistical trends by means of various data visualization tools, leading to better communication of technical knowledge to client.
- Mentored junior analysts for BI report creation, thereby increasing the number of efficient human resources.

PROJECT EXPERIENCE

Detecting Fake Credit Card using scanned images

- Built CNN model to detect whether a credit card is fake original by using scanned images of credit cards.
- Created "filter" using auto-encoder Convolution Neural Network, resulting in recognition of only relevant sections of images in high resolution leading to reduction in image size and memory allocation. Further used NLP techniques to verify whether the information present on the card is authentic or not.
- Decreased process time (for model creation) from three hours to 40 minutes, and increased accuracy from 43.6% to 79.8%.

Predicting potential credit loan defaulters

- Constructed classification model to predict the probability of a person defaulting on a credit loan for a Portuguese Bank. Preprocessed the data using various data cleaning, feature engineering, feature selection and NLP methods.
- Increased balance of dataset by removing high bias towards one target label using "SMOTE method".
- Reduced dataset bias from 96% (towards "default=no" class) to 60%, and increased F-1 score from 19% to 73% by implementing Random Forest algorithm by considering various socio-medical parameters