ROHAN MAHAJAN

rmahaj01@syr.edu | www.linkedin.com/in/rmahaj01 | (331)-472-9062

EDUCATION

Syracuse University, School of Information Studies — **Syracuse, New York**
M.S. in Applied Data Science — **May 2020**
Relevant Courses - Data Analytics, Business Analytics, Big Data Analytics, Applied Data Science, Decision Making Analytics.

Savitribai Phule Pune University — **Pune, Maharashtra**
B.E in Computer Engineering — **June 2018**
Relevant Courses – Data Mining Techniques, Database Management, Data Structures, Design and Analysis of Algorithms.

PROFESSIONAL EXPERIENCE

Graduate Research Assistant | School of Information Studies - Syracuse University — **Jan 2020 - Present**
- Provide health resource recommendations to users via web scrapping reddit posts in Python.
- Mapping recommendations based on severity of subreddits using a sentiment analysis classifier and NLP packages.
- Executing topic modelling to generate topics and subjects of interest amongst subreddit content.

Machine Learning Intern | BMC Software — **Dec 2019 – Jan 2020**
- Revamped normalized datasets based on category recommendations using Python scripts.
- Calculated and measured performance metrics on IBM Watson based on IT ticket classification of original and modified datasets.
- Increased classification training data accuracy by 5-10% on revamped datasets.

Machine Learning Intern | BMC Software — **June 2019 – Aug 2019**
- Provided category recommendations to normalize datasets by designing and implementing models for text summarization tasks.
- Used bar chart and pie chart visualizations using matplotlib and seaborn packages to interpret labeled category distribution.
- Performed text conversion through sentence embeddings using Google's Universal Sentence Encoder followed by performing ticket theme identification using K- means clustering with natural language processing and machine learning packages.
- Visualized final dataset distribution based on category recommendations to display comparisons of initial and final distributions.

ACADEMIC PROJECTS

Loan Default Analysis - Big Data Analytics — **Jan 2019 – Apr 2019**
- Predicted a user's loan repayment abilities using variety of alternative data from a Home Default Credit Risk dataset in PySpark.
- Featured engineered new variables for classification followed by data cleaning of invalid values and balancing target variable.
- Used Logistic Regression and Random Forest models with principal component analysis followed by model cross-validation.
- Concluded that educated, employed, asset owning, well qualified borrowers are more likely to default on their loans.

Wine Review Analysis – Data Analytics — **Jan 2019 – Apr 2019**
- Analyzed Wine Reviews provided by customers to determine the attributes that affect a rating of a wine in Python.
- Performed sentence conversion along with visualizing wine reviews through word cloud visualizations followed by using Logistic Regression, Naïve Bayes and Random Forest classification models obtaining an accuracy of 87.99% on the best model.

Restaurant Review Analysis – Decision Making Analytics — **Aug 2018 – Dec 2018**
- Conducted data analysis and visualization on a Zomato Restaurant Review dataset to determine factors and variables affecting restaurant rating using simple, multiple and stepwise linear regression analysis using Microsoft Excel.
- Described summary statistics of variables along with visualizing dataset distribution per variable using bar chart and pie charts.
- Provided business insights where having options of online bookings and deliveries could help restaurants increase ratings.

TECHNICAL SKILLS / CERTIFICATIONS

- **Programming Languages** – Python, Apache Spark, R, SQL.
- **Software** – Microsoft Excel, Power BI, Tableau, Jupyter Notebook, Microsoft Access, Microsoft Visio, Microsoft Visual Studio.
- **Certifications –** Google Analytics. — **July 2019**

LEADERSHIP SKILLS

- **President** | iSchool Graduate Organization(iSGO) E-board - Syracuse University — **May 2019 - Present**
  - Lead a team of 13 students by conducting weekly meetings and collaborating with other student organizations.
  - Budgeted $6000 funds for organizing events and funding other student organizations throughout the academic year.
  - Communicated regularly with other student organizations, administrative offices and staff for collaborating events.
  - Initiated a Peer-to-Peer student mentorship program to provide guidance to incoming first year graduate students.
  - Hosted a Graduate Internship Panel event to prepare first year students for upcoming internship application process.
  - Revitalized the organization's social media presence followed by a new marketing strategy leading to increased involvement.

- **Academic Program Senator** | Graduate Student Organization Senate - Syracuse University — **Oct 2018 – May 2019**
  - Represented the Applied Data Science program in the Graduate Student Senate to discuss issues of graduate students.