Mansi Rathod

mrathod@uw.edu | +1-206-566-8362 | Seattle, WA | Website| linkedin.com/in/mansi-rathod | github.com/rathodmansi

"Diligent & passionate to help organizations through strong analytical thinking & sharp business acumen. With 3+ years strong foundation in statistics & machine learning, proficiently translates business problems into data science solutions"

EDUCATION

University of Washington | Seattle, WA, USA — Sep 2019 - Mar 2021(Exp)

Master of Science in Data Science (MSDS) | GHC Student Scholar for 2020 — **GPA: 3.87/4.0**

Coursework: Machine Learning, Statistic, NLP, Statistical Analysis & Experiments, A/B Testing, Data Visualizations

Teaching Assistant: Statistics & Probability for Data Scientists (MSDS's fundamental subject)

National Institute of Technology | Raipur, India — Jul 2012 - May 2016

Bachelor of Technology in Information Technology | 3rd Rank in class of 90 — **GPA: 8.81/10**

Coursework: Data Structures & Algorithms, Databases Management, Neural Networks, Software Design, Data Mining

SKILLS

- **Programming:** Python, R, SQL (SCOPE, PostgreSQL, MySQL), Linux, C++, Java, HTML, CSS
- **Libraries:** Keras, NLTK, Scikit-Learn, Matplotlib, Seaborn, Plotly, Dash, Numpy, Dplyr, RShiny, ggplot2
- **Tools:** Tableau, GitHub, VS Code, PyCharm, Microsoft Azure SQL, AWS EC2, Hive, Docker, MS Excel

WORK EXPERIENCE

Data Scientist Intern | **Path**| Seattle, WA, USA — Jul 2020 – Sep 2020

- Led ETL process on large immunization data sets (6M+ rows), data imputation & feature-engineering for ML models
- Performed Exploratory Data Analysis (EDA) in R & published Tableau dashboards answering global health question
- Built mixed-effect GLM model to identify impact of different covariates; Designed a Logistic regression model to predict missed opportunities of vaccinations at child and visit level for African countries

Business Analyst | OYO Rooms | Gurugram, India — May 2017 - Dec 2018

- Designed an incentive model in R to align around 2000 sales managers with monthly target achievement of $8 M; Received 'Above & Beyond Call of Duty' award for improving profit ratio by 30% & incentive calculation time by 60%.
- Accomplished demand diversion to low occupancy hotels by updating hotel recommendation algorithm; Created live tracking dashboard using SQL & Tableau which retained 600+ hotel properties from churning.

Data Scientist | MuSigma Inc. | Bengaluru, India — Sep 2016 - Apr 2017

- Built a random forest model (R) for real-time prediction of viral events from Twitter. Leveraged client's Search Engine Marketing (SEM) platform to generate actionable insights such as targeting suggestions for advertisers.
- Developed methodology to classify user search queries into correct business verticals using TF-IDF and LDA Text Analytics Techniques. Automation with R saved 3+ hours of weekly user search queries mappings.

ACADEMIC PROJECTS

Estimating Impact of Asymptomatic Patients in spread of covid-19 | Research Assistant — Jun 2020 – Sep 2020

- Programmed Agent Based Model 'Corvid' on AWS to generate multiple population effects with probability sampling
- Spearheaded calibration values for Basic Reproduction Number (R0) & proportion of asymptomatic patients

Movie Recommendation System (RS) for Netflix| *Kaggle* | Link — Mar 2020 - Apr 2020

- Devised Recommender System (Python) based on 500k Netflix users using collaborative filtering & cosine similarity.
- Produced a user interface to suggest content based on genre & time using Dash (Python) & deployed it on Heroku.

Document Classification using Natural Language Processing & Deep Learning | Link — Jan 2020 - Feb 2020

- Constructed Convolutional neural nettworks model to classify text into 4 categories using TF-IDF & word embedding; Applied randomized grid search & K fold cross validation for hyperparameter tuning & achieved 92% accuracy.

Vehicular Injury Prediction Model for New York City Traffic | Link — Dec 2019 - Jan 2020

- Modeled hypothesis tests, deviance & multi-collinearity tests to estimate association of 19 features in road-collision; GLM Logistic regression model found 70% high risk for pedestrian & 92% risk for motorist among all victim types.