PARTHVI BHUTANI
pb1967@nyu.edu | 929-326-7945 | parthvibhutani.github.io

EDUCATION

New York University **New York, NY**
*Bachelor of Arts in Mathematics, Minor in Data Science (**GPA: 3.89**)* *May 2021*
• Mathematics Award for Academic Achievement, 2020 (Presented to top five graduating students)
• University Honors Scholar; Dean's List for 2017-18, 2019-20
• Scholarship Recipient - Walter G. Terwedow MD Schlp ($5000 X 4), Maurice G. Kott Schlp ($10000 X 2), Clelia Abele Schlp ($10000)
Relevant Coursework: Probability and Statistics for Data Science (Graduate), Introduction to Data Science (Graduate), Introduction to Machine Learning, Basic Algorithms, Data Structures, Object Oriented Programming, Theory of Probability, Mathematical Statistics, Database Design and Implementation, Honors Linear Algebra, Partial Differential Equations, Combinatorics, Numerical Analysis, Math Modeling, Discrete Mathematics, Intro to Web Design

University of Delhi (*CGPA: 9.045/10*) **New Delhi, India**
Bachelor of Science (Honors) in Mathematics, Minor in Economics *May 2017*

SKILLS

Programming & Software: Python, Java, MATLAB, R Studio, Jira **Web Development:** HTML, CSS, jQuery, Bootstrap, React
Database Management: SQLite, MySQL, MariaDB, MongoDB, HiveQL, Azure, SAS, SPSS, MySQL Workbench, AWS Console
Machine Learning Stack: Scikit-learn, Keras, NumPy, Pandas, SciPy, Theano, TensorFlow, PyTorch, Regex, Matplotlib, Seaborn
Data Analysis Skills: Power BI **,** Pandas, Tableau, Adv. MS Excel, Jupyter Notebook, IBM Cognos, Qlik, Google Analytics, Google Data Studio, KNIME, Looker, Analytics Edge(edx)

PROFESSIONAL EXPERIENCE

HDFC ERGO General Insurance Company **Mumbai, India**
Data Analyst Intern *May 2020 - Aug 2020*
- Designed an unsupervised machine learning model to segment 1.1 million health insurance customers who are yet to claim insurance
- Analyzed various retail factors like net-premium, sum insured, loss-ratio, up-sell, ticket size, location-tier, portability, income, and age
- Explored various underwriting factors such as loading amount, add on cover premium, BMI, marital status, and payment mode
- Executed Hierarchal-DBSCAN clustering algorithm to profile customer segments as profitable, potential-profitable and not-profitable
- Demonstrated the most profitable 200000 customers to target and spur shift to exclusive or premium policies and encourage add-ons

NYU Courant Summer Undergraduate Research Experience Program **New York, NY**
Research Intern advised by Dr. S.R. Srinivasa Varadhan (Abel Prize) *Mar 2018 - Sep 2018*
- Designed Markov Chain Monte Carlo Samplers for Random Sampling of high dimensional probability distributions
- Simulated the design on various probability distributions to check computational efficiency of the algorithm used
- Compared the algorithm to Gibbs Sampling Algorithm and Metropolis-Hastings Algorithm to help fine tune the parameters
- Proved efficiency of Markov process through limiting distribution, convergence and mixing properties using mathematical analysis

NYU Courant Institute **New York, NY**
Grader & Tutor for Calculus 3 (MATH-UA.0123) and Numerical Computing (CSCI-UA.0421) *Jan 2021 - Present*
- Assist instructor with study material, grading, uploading marks, record keeping, and other tasks
- Evaluate assignments, homework and exams; Holding weekly office hours to solve doubts and communicate with 120+ students

ACADEMIC PROJECTS

NYU LibCal Datathon **New York, NY**
Datathon Competitor (First Place) *Feb 2020*
- Analyzed data from L2 Political Academic Voter File - private updated database of every registered voter in the United States
- Developed reproducible pipeline to get samples of data with 600 behavioral, 400 demographic and 91 predictive fields
- Generated national, multi-state and state samples from the pipeline with voter data from 2008 to 2019
- Anonymized the data using elliptic curve encryption and chaos algorithm
- Employed a VoterMapping Map Query Interface to map and plot the transformed and anonymized data using geolocation libraries

Twitter Sentiment Analysis in lieu of COVID-19 **New York, NY**
Academic Project advised by Prof. Brian Dalessandro *Aug 2020 - Present*
- Collected tweets from March 1, 2020, processed and analyzed over 2 million tweets and tracked user geolocation and IDs
- Created functions to assess the subjectivity and polarity of tweets to classify tweets as fact/opinion and positive/negative
- Achieved precision accuracy of 97% and recall accuracy of 98% using a Naive Bayes Classifier
- Analyzed sentiment polarity between verified and non-verified profiles and polarity over time for different states in the US
- Observed strong correlation of negative tweets with spikes of new cases in different states especially in dense regions like NY and NJ
- Investigated user mobility from state to state based on user geo-location and it's correlation and impact on COVID-19 propagation

NYC Department of Health and Mental Hygiene Restaurant Ratings **New York, NY**
Academic Project *May 2019 - Jul 2019*
- Analyzed over 190000 restaurants across NYC and how their ratings changed over time for individual restaurants as well as big chains
- Evaluated roll rate as the probability of transitioning from one grade to another and evaluated roll rate categories over A rated places
- Reaffirmed previous analysis that an A rating remained unchanged 90% of the time especially in well-known or high-rated restaurants
- Examined individual factors-Borough, Food Type, Name, Street and zip-code to determine probability of restaurant receiving grade A