Gaurav Rajesh Parikh

Email: gaurav.rajesh.parikh@duke.edu
Portfolio: gparikh100.github.io
Mobile: +1 (984)377-1861
LinkedIn: linkedin.com/in/gr90/

Education

- **Duke University/ Duke Kunshan University** — Durham, NC/ Kunshan, China
  B.S. Computer Science and Statistical Science — *Aug 2019-Present*
  - **Cumulative GPA**: 3.86/4.00; Dean's List with Distinction, Davis UWC Scholar, Full Merit Scholarship
  - **Coursework**: Data Science Competition, Machine Learning, Algorithms and Data Structures, Asset Pricing and Risk Management, Financial Derivatives, Linear Algebra, Probability and Statistics.
  - **Extracurricular**: Product Management at Duke Applied Machine Learning.

Skills Summary

- **Languages**: Python(proficient), JAVA(proficient), SQL(moderate)
- **Frameworks**: NLTK, SpaCy, numpy, pandas, matplotlib, Beautiful Soup, sci-kit learn
- **Product**: Product Analytics, A/B Testing, Design Thinking, Data Visualization

Research

- **Duke University, Almost Matching Exaclty Lab** — Durham, NC
  Student Researcher — *Jan 2022 - Present*
  - **Collaborative Research**: Co-authored study on "Impact of fare reinstatement and subsidies on ridership", modeling data on transit ridership with over 12.3 million rows and 112 columns describing public transit rider demographics across Seattle for 60-day period. Presenting in JSM 2022 (Statistics Conference) and to Causal Inference for Social Impact Lab at Stanford University.
  - **Causal Inference**: Research under Dr Cynthia Rudin and Dr Sudeepa Roy on Adaptive Hyper-box matching for treatment effect estimation. Ongoing work on estimating causal treatment effects for randomized control trials to isolate treatment effects using explainable adaptive binning framework as opposed to existing non-adaptive fixed-bin methods.

Experience

- **Duke University, Venue and Production Management** — Durham, NC
  Student Worker — *May 2022 - Present*
  - **Leadership**: Managed team of 15 ushers to fulfil front of house responsibilities including ticket sales, artist reception and guest services for performances with up to 1200 patrons.
  - **Communication**: Oversaw pre-event briefings with technical crew, event staff, performers, and security to ensure events go smoothly.
  - **Event Management**: Trained to lead safety protocols and facilitate evacuations across multiple venues in case of emergencies.
- **National Stock Exchange (NSEIT)** — Remote/ Mumbai, India
  Data Science Intern — *May 2021 – Aug 2021*
  - **Research**: Investigated statistical models for anomaly detection in algorithmic trade log data to identify instances of deviation from registered strategy and violation of trade participant norms such as non-algorithmic trading activity from registered algo-trader.
  - **Model Design/Metrics**: Identified and engineered 20 covariates influential in model learning including Order to Trade Ratio, Trade Asymmetry, Activity Density to flag trades resulting in market manipulation.
  - **Impact**: Compared Isolation Forest, DBSCAN and Local Outlier Factor models on sample data using k-Fold cross-validation for model training to achieve test accuracy of 94.23%
- **Stones2Milestones** — Gurgaon, India
  Data Science Intern — *May 2019 – Aug 2019*
  - **Market Research on Pain Points**: Extracted Freadom (EdTech flagship offering) App usage data from 50+ video interviews and 400+ survey responses and carried out data preprocessing for user data (over 16,000 rows) using NumPy and Pandas.
  - **A/B Testing**: Evaluated A/B testing data to create customer segments based on user age group, linguistic proficiency, parental intervention across 7 geographies in India, South Korea and Vietnam.
  - **Synthesis for Stakeholders and OKR setting**: Consolidated findings and created data visualizations using Tableau on app impact to present to investors, parents, and stakeholders. Findings led to a change in app monetization strategy through launch of a 3-tiered subscription model.

Projects

- **Mean Variance Analysis**: Python Project using yfinance, plotly, scikit-learn and Pandas for stock selection through bootstrap sampling on returns and calculation of Sharpe ratio.
- **Mutual Fund Exploratory Data Analysis**: Python Project to compare different mutual funds' net asset value over time and calculate correlation of returns. Used Beautiful Soup to scrape data on Mutual funds in India, and Plotly and matplotlib to create visualizations.
- **Sentiment Classification using Document Embeddings**: NLP project using NLTK, spaCy, pandas and NumPy to identify sentiment of dataset into positive or negative classes using Stemming, Lemmatization and by comparing Cosine distance from train set of word vectors.
- **K Means Clustering**: A Java visualizer for the k-means clustering algorithm that can randomly generate points or can read them in from a csv file and apply different metrics to determine clusters.