SIDDHARTH SURESH

ss3zg@virginia.edu | +1 202 378 3623 | Charlottesville, Virginia
LinkedIn: siddharthsuresh92 | Github: Siddharthsuresh92 | Website: siddharthsuresh92.github.io

Education

M.S in Data Science, School of Data Science, University of Virginia (Current GPA: 3.97/4) *July '19 – May '20*

B. Technology in Mechanical Engineering, Shiv Nadar University (GPA: 8.54/10) *August '11 – May '15*

Relevant Work Experience

Graduate Research Assistant, University of Virginia *November '19 – May '20*

- Conducted research & analysis of brand marketing campaigns on consumption behavior using location data
- Identified patterns in food consumption behavior during COVID-19 using location data with ~500 million rows to understand how socio-economic status impacts such behaviors
- *Skills*: *Python, AWS, GeoPandas, Unsupervised Learning, EDA, Data Mining, Hypothesis Testing*

Data Science Intern, Capital One (Capstone Project) *September '19 – May '20*

- Developed a soft attrition metric to improve existing attrition models by identifying churn behavior early
- Engineered new features to better identify customer churn behavior, improving F1 measure by 16% to 0.9
- Created data pipelines and functions for big data to assist training and deployment of prediction models
- *Skills*: *Python, Scikit-learn, AWS, Git, Data Wrangling, Data Mining, Random Forest, Boosted Trees, SVM*

Senior Assistant – Deal Advisory Services, BDO India LLP *April '17 – February '19*

- Evaluated financial statements based on business models of target companies to suggest adjustments in earnings, working capital, net debt and cashflow as part of financial due diligence for clients
- Identified key patterns in revenue/expenses and its driving factors over voluminous data spanning at least 3 financial years
- Developed analysis frameworks to automate financial analysis processes, increasing efficiency by 90%
- *Skills*: *Excel VBA, Pivot tables, Financial Analysis, Reports & Presentations, Client engagement*

Key Projects

Bayesian modeling to evaluate **tactical styles** in **soccer**

- Defined standing classes within the English Premier League table and determined tactical styles for each soccer statistic
- Developed a hierarchical ordered logistic regression model to capture the influence of different tactical styles on season to season shifts in the standing class, using data from 2006-07 to 2017-18 seasons
- *Skills: Python, PyMC3, GraphViz, Matplotlib, Bayesian Hierarchical Models, WAIC (model selection), Data Wrangling*

Movie recommendation system using **Topic modeling** (Latent Dirichlet Allocation)

- Preprocessed a corpus of 35k documents containing movie plots using n-grams, lemmatization techniques and optimized tuning parameters of the LDA topic model to achieve a coherence measure of 0.447 for the final model
- Developed a recommendation system using Jensen-Shannon divergence supporting user interfacing to suggest movies
- *Skills: Python, Gensim, NLTK, spacy, pyLDAvis, Topic Modeling, Natural Language Processing, Text Similarity*

Tableau Dashboard - viral trends on **YouTube** during a **pandemic** (Click here to view on Tableau Public)

- Web scrapped popular videos using YouTube API and python scripts to create a dataset (~50k observations) containing video stats from 18 different countries across a week in April 2020 during the COVID-19 pandemic
- Created an interactive dashboard visualization providing stats on video category, top channels, views, tags to help understand viewing patterns across countries during a pandemic and cultural tastes in different parts of the world
- *Skills: Python, Web API, Tableau, Data Visualization, Dashboarding*

Other skills: R, SQL, Scala, Deep Learning, Time Series, Predictive Modeling, Apache Spark (for more projects visit Github/Website)

Additional Information

- Passed Level 1 of the **CFA® Program**, CFA Institute in August 2018
- **Volunteered** for "Child Rights & You" (CRY) and organized a fund raiser worth $3k
- **Research paper** titled "*Evaluating and Improving Attrition Models in the Retail Banking Industry*" accepted for Systems and Information Engineering Design Symposium 2020 (https://edas.info/p27197)
- **Tutored** 10 students from McIntire School of Commerce, UVA on Topic Modeling using LDA (https://bit.ly/2ytmUuC)
- Selected for a government sponsored **student exchange program** to Netherlands. Selection process included academic & extracurricular merit-based screening of 120 students → interview of 10 shortlisted students → final 4 students
- Other interests: Music (Guitar & Piano), Sports (Basketball & MMA), Chess and Travel