

# SAYESHA ARAVAPALLI

[sayesha.aravapalli@utexas.edu](mailto:sayesha.aravapalli@utexas.edu) | (512) 905-1548 | Austin, TX-78751

[linkedin.com/in/sayesha-aravapalli](https://www.linkedin.com/in/sayesha-aravapalli) | [asayesha.github.io](https://github.com/asayesha)

## EDUCATION

<b>The University Texas at Austin</b>	Master of Science in Business Analytics	GPA : 3.84/4	May 2020
<b>Indian Institute of Technology, Bombay</b>	Bachelor and Master of Metallurgical Engineering and Materials Science	GPA : 8.41/10	June 2017

## TECHNICAL SKILLS

<b>Technologies</b>	Python, SQL, R, ggplot2, Pandas, NumPy, Matplotlib, Seaborn, plotly, Scikit-Learn, NLTK, Vader, Gensim, Matlab, VBA, MapReduce
<b>Tools/Platforms</b>	Jupyter Notebook, RStudio, Tableau, Spotfire, Google Cloud Platform, Databricks, MS-Excel, MS-Powerpoint
<b>Modeling Skills</b>	Linear Regression, Logistic Regression, Lasso, Ridge, Naive Bayes, Random Forest, KNN, XGBoost, ARIMA, PCA, k-means clustering, A/B testing

## EXPERIENCE

<b>Visa Capstone Project Consultant, Austin, TX</b>	January 2020-present
<ul style="list-style-type: none"><li>Building unsupervised models using Graph analytics and DBSCAN to identify fraud patterns in real time to decrease losses in revenue and improve customer relations; Visualized transactions data using Tableau to study customer behavior</li></ul>	
<b>Energy Sector Start-up Student Consultant, Austin, TX</b>	August 2019-November 2019
<ul style="list-style-type: none"><li>Consulted a solar-energy based start-up to identify key drivers influencing customer sign-ups using Google Analytics data and designed Tableau dashboard to stakeholders to help them improve their business model</li><li>Deployed XGboost, resulting in a summary-level data recall of 0.77 and individual-level data recall of 0.86</li></ul>	
<b>Merck &amp; Co. Analyst, Pune, India</b>	June 2017 - October 2018
<ul style="list-style-type: none"><li>Assessed and forecasted in R the impact of market events, such as competitor entry and loss of exclusivity of Merck's drug, on sales using therapy area knowledge and treatment protocol</li><li>Implemented k-means clustering method in python to group patient data and identify market opportunity based on cluster</li><li>Developed market opportunity analyzer and trained 50+ people in marketing teams across LATAM, EMEA and APAC countries to use the tool for analyzing market before entry of generics</li><li>Designed and created global reporting TIBCO Spotfire-based dashboards to track pre-launch and post launch KPIs of patented drug to present actionable insights for global marketing strategy</li></ul>	
<b>DBS Bank Summer Analyst, Mumbai, India</b>	May 2016-June 2016
<ul style="list-style-type: none"><li>Boosted efficiency and reduced stress testing time by automating operational risk estimation module using SQL to deliver an integrated stress testing framework</li><li>Identified opportunity areas to cross-sell products by evaluating cross-sell index of product exposure to clients leading to increase in business; Integrated data from various sources to publish the daily balance sheet of the bank</li></ul>	

## DATA ANALYTICS PROJECTS

<b>Instagram user engagement</b>	Fall 2019
<ul style="list-style-type: none"><li>Investigated ways to boost user engagement of NatGeo Instagram handle using LDA topic modeling on images and TF-IDF on post with engagement prediction accuracy 71.9%; Over-saturated topics have low engagement scores</li></ul>	
<b>GoodReads Recommendation system</b>	Fall 2019
<ul style="list-style-type: none"><li>K- means clustering and user-based collaborative filtering using correlation similarity on Goodread books data improved RMSE of book recommendation; Long tail problem can be solved by recommending mixture of books from the clusters</li></ul>	
<b>Energy consumption Prediction</b>	Fall 2019
<ul style="list-style-type: none"><li>Employed lightGBM on data to predict energy consumption to understand worth of improving building efficiencies to reduce costs and emissions; Feature engineering and outlier removal was performed on the time series data</li></ul>	
<b>US traffic Fatalities</b>	Fall 2019
<ul style="list-style-type: none"><li>Illustrated fatal traffic car crashes data on Databricks to understand causes of motor vehicle crashes and deaths on national highways; Only 27.5% accidents were due to drunk driving; 72% of accidents occurred in clear conditions</li></ul>	

## LEADERSHIP AND ACTIVITIES

<ul style="list-style-type: none"><li>Student editor of Insight, IIT Bombay's monthly newsletter, which publishes articles on campus issues</li><li>Part of 22-member core management team to ideate and execute Asia's largest student-organized science and technology festival</li><li>Mentored 10 students as Department Academic Mentor at IIT Bombay</li><li>Classical 'Bharatanatyam' dancer; Traveler; completed 10-day Vipassana (no talk for 10 days) Meditation course</li></ul>	
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--