ASHRITH REDDY

MS in Statistics Candidate | Eligible to work in US from May 2022 | Available to relocate nationwide

ashrithssreddy@gmail.com | +1-850-339-5898 | Los Angeles, CA | StackOverflow | GitHub | LinkedIn

Professional Summary

I'm a Data Science professional with 5 years of experience providing data-driven business solutions to Fortune-500 clients across Entertainment Media, Retail and Education industries. I'm skilled in data manipulation & visualization, Big-data technologies, problem-solving and Programming. I'm pursuing MS in Statistics to strengthen the understanding of Statistical Machine Learning techniques and Computational Statistics.

Education

- MS in Statistics, Florida State University — Aug 2020 – May 2022
  - Computational Methods in Statistics I, Statistics in Applications, Time Series & Forecasting Methods, Engineering Data Analysis, Logistic Regression, Linear Models, Applied Machine Learning
- B.Tech. in Electronics and Electrical Engineering, National Institute of Technology Surathkal — Jul 2011 – May 2015

Senior Associate at TheMathCompany — June 2017 - June 2020

- **Price Elasticity and Optimization** for women's value fashion retailer.
  - Price elasticities were calculated to determine the price-demand relationship and drivers of sales.
  - This project involved determining a suitable level of analysis, outlier treatment and other exclusions, product lifecycle identification, selection of model technique and transformation, model building & validation and recommending prices.
  - ElasticNet was employed with log transformations to model for units sold and determine price elasticity. Optimized prices were recommended with various confidence intervals using bootstrapping.
- **Customer Segmentation and Promotion affinity** and for a tween apparel retailer.
  - A major US retailer wanted to understand how their customers perceive different price promotions. This would enable them understand the behavioral traits of customers towards promotions and make informed price promotion decisions based on customer segments.
  - Used K-means clustering technique in Python to segment customers on their product affinity and purchase behavior patterns. Response rate & revenue lift was used to determine segments' affinity for each promotion.
- **Education-Technology projects:** Worked with a leading Edu-Tech company that helps US universities identify and enroll students for graduate medical programs.
  - **Applicant Fit and Yield:** Identified the drivers of an applicant's admission and matriculation. Multiple statistical techniques were tried (like Logistic Regression, Random Forest and SVM) to determine if an applicant will be admitted and will matriculate if admitted.
  - **Applicant Churn:** Real-time identification of applicants who would attrite and nudge them to complete the application. Model steps involved variable importance using Random Forest, recursive feature elimination, Decision tree tuning and grid-search, 5-fold cross-validation, etc.
  - **Applicant Strengths and Weaknesses:** Applicants typically aren't aware of their areas of strengths and improvements when applying to graduate programs. Logistic regression models were built to identify attributes of strong applicants. This enabled the candidate to benchmark their application against successful applications.
- **Customer State Transition** for largest retailer's Customer Analytics team
  - Developed a Markov-chain simulator to assess the impact of promotion (which alters the customers' distribution during acquisition) on revenue and customer count. The Markov chain state here is defined as count of orders made by the customer in a month.
  - This will support the business in decision making on what promotions to offer while acquiring customers and when.
- **Reusable Modules:** Built scalable & automated modules in R and Python for Univariate analysis, Bivariate analysis, conversion of hierarchical data to tabular format, data extraction from leading 8 social media sites/APIs and open sources, optical character recognition, etc.

Trainee Decision Scientist at Mu Sigma — Aug 2015 - Apr 2017

- Worked with US's leading sports broadcaster, to enable serving customized ads based on sports viewers' emotions in addition to achieving better ROI for the advertisers. This analysis helped increase click-through rate by up to 50%.
  - **Prioritization of various events in sports:** Identified major events (e.g. Touchdown, Grand Slam, etc.) occurring in big-6 American sports leagues and ranked them by the impact using frequency and popularity.

- **Behavioral Viewer Segmentation:** Designed an RFM methodology that identifies a viewer's team and league preference, to complement the existing method of geographic location and favorites. It involved analyzing daily customer visit trends and rules generation based on frequency and recentness of visits.
- **Analysis of betting odds:** Extracted odds data for all the games which were analyzed to identify spread or money line thresholds to classify a future game as 'Tight Game', 'Expected Win' and 'Expected Blowout'.

- US's leading retailer: Gains and Losses analysis: Leading US hypermarket chain has seen a decline in cold cereal category sales for the past few years. Transaction data was analyzed to identify the causes of declining sales and the reasons were communicated to the Category Advisory Managers.

Technical Skills

Programming	R, Python, SQL/Hive, Julia
Machine Learning	Regression (Linear, Logistic, Bayesian, Robust, Lasso, Ridge, and ElasticNet), Decision Tree and Random Forest, SVM, KNN, Computer Vision (limited to OCR), Natural Language Processing (limited to text mining), Bagging, Boosting, K-means, Spectral and Hierarchical Clustering.
Big Data	Hadoop, Spark, Google Cloud Platform, EMR, Data Pipeline, CLI, S3
Other Experience	Tableau, R Markdown, tidyverse ecosystem and R-shiny

Leadership and Volunteering

- StackOverflow leaderboard: Contributed to SO and was ranked **top 9% and 7%** for 2020 Quarter 1 and 2 respectively.
- **Open Source Contribution**
  - Developed an R library that enables understanding of dataframes and reduces data manipulation coding time.
  - Developing an R and Python library intended to visualize the network of related Wikipedia pages, using data obtained from MediaWiki API.
  - Developing an R and Python library that *guesses* the most probable date format in an string array storing dates.
- **Mentoring**
  - Designed coursework and taught Advanced R-programming and Python for data analysis to ~250 trainees, which included Analytics professionals of the largest brewer and experienced international audience.
  - Designed the coursework and taught R-programming and Coding best practices (GitHub, modular coding, functional programming etc.) to ~650 inductees.