

## SRIKRISHNA SRIDHAR

 srikrishna.sridhar@outlook.com |  (812) 955-8935 |  Srikrishna-s |  srsridh | <https://srsridh.github.io/> | Austin, Texas

### EDUCATION

Indiana University Bloomington	MS- Data Science	GPA: 3.61/4	Aug 2017- May 2019
Anna University	BE- Electrical & Electronics	GPA: 8.0/10	Aug 2011- May 2015

### TECHNICAL SKILLS

**Languages:** Python, R, SQL

**Others:** Spark, Linux, MySQL, HDFS, MongoDB, Tableau, Hadoop, GitHub, JIRA, A/B testing, Agile, Excel

**Libraries:** numpy, pandas, matplotlib, ggplot2, scipy, sklearn, NLTK, OpenNLP, SparkML, re, spaCy, BI, pytorch, OpenCV

**Machine Learning:** Time series, GMM, Bayes, NLP, Boosting, Random forest, K-means, Bayesian, Confidence Interval

### PROFESSIONAL EXPERIENCE

#### **Data Scientist Intern**

**Jun 2018–Aug 2018**

Domtar Personal Care

- Improved prediction accuracy by 12% using new features created with holidays, outages and transition between different grades of pulp and paper.
- Reduced inventory costs by 10% using Time Series ARIMA and Linear regression to predict the daily, weekly, bi-weekly, monthly and annual production of pulp and paper.
- Reduced production costs of Pulp by 15% by developing a strategy to vary production of Pulp by month.
- Designed business intuitive dashboards for prediction results to help sales team improve sales.

#### **Assistant System Engineer (Banking and Financial services) – TCS Ltd**

**May 2015-Dec 2015**

- Worked on settling of trades for a Fortune 32 company after successful purchase of stocks.

#### **System Engineer/Analyst (Telecommunication) – TCS Ltd**

**Jan 2016-July 2017**

- Used logistic regression to identify possible payment defaulters for a Fortune 130 company.
- Created reports using Tableau to highlight the reasons for default in payments.
- Used KNN to group similar customers and understand their monthly usage statistics.
- Developed reports explaining customer usage statistics and preferences to a non-technical audience.

### PROJECTS

#### **Movie rating recommendation system using collaborative filtering in Python**

**Apr 2018**

- Designed algorithms based on gender and movie genre to predict the movie ratings for 10million users
- Predicted the movie ratings of targets using top 50 similar users and achieved 80% accuracy

#### **Restaurant Annual Revenue Prediction in Python [Kaggle Top 5%](Team of 3 )**

**Mar 2018–Apr 2018**

- Predicted the revenue of 100,000 restaurants in over 50 cities using Gradient boosting, KNN, Linear regression
- Gradient Boosting achieved a Root mean square error of 0.3, indicating very low errors in prediction

#### **Tweet – Location Predictor (32000 tweets) in Python [Highest accuracy among 200 students]**

**Dec 2018**

- Predicted the location from which the tweets were posted using a Naïve Bayes classifier
- Achieved 72.5% accuracy by effectively handling stop words, special characters, and missing words

#### **Stance Detection system on US Airlines Data to analyse the popularity of airlines in Python**

**Mar 2018**

- Extracted tweets from twitter targeting major US airlines and built Random Forest, KNN models
- Classified the sentiment of tweets as positive/negative/neutral and analysed the popularity of airlines

#### **Maps using Artificial Intelligence Search algorithms in Python**

**Sep 2018**

- Designed maps to output the total distance, time and the shortest path between any two cities in the USA
- Built A\*, Uniform, BFS, DFS and IDS search algorithms with distance and time measurements as cost functions
- Uniform search algorithm returned the most optimal path between any two cities within 4 seconds

#### **Image Classification on Natural Images Data(Kaggle) using HDFS and Pyspark**

**Dec 2018**

- Classified 6899 images using Random forest, Logistic regression and Gradient boosting
- Compared to python, run-time was reduced by 33.3 minutes and achieved 73% accuracy using Random forest

#### **2016 US Presidential election analysis on Election Survey Data(64000 adults) in R**

**Mar 2018**

- Fitted Logistic regression models to understand the switching of votes of 2012 Obama supporters to 2016 Trump supporters based on immigration policies, gender, race, and education