SRIKRISHNA SRIDHAR

srikrishna.sridhar@outlook.com | (812) 955-8935 | Srikrishna-s | srsridh

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, MongoDB, Tableau, MicroStrategy, Hadoop, Git, MapReduce, NLP, A/B Testing
Toolkits:NumPy,pandas,Matplotlib,NLTK,SciPy,scikit-learn,re,OpenCv ,ggplot2 ,caret, dplyr, MLlib

PROFESSIONAL EXPERIENCE

Data Scientist Intern **Jun 2018–Aug 2018**
Domtar Personal Care

- Improved prediction accuracy by 12% by creating new features using holidays, outages and transition between different grades of pulp and paper in production data
- 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
- Identified trends in sales by analyzing the variation of pulp and paper production with holidays
- 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 helping sales team improve sales
- Developed reports to explain trends in production and sales of pulp and paper to a non-technical audience

Software Engineer in Banking and Financial Services **May 2015–July 2017**
TCS Ltd

- Reduced running time of programs by 120 seconds using MIPS reduction (Millions of instructions per second) on critical COBOL programs.
- Implemented FTP-SFTP using JCL to transfer feeds from Mainframe to UNIX environment automatically
- Earned client appreciation by creating an application to uproot manual monitoring by automatically notifying clients via e-mail on the successful transfer of feeds from mainframe to Unix environment

PROJECTS

Movie rating recommendation system in *Python* **Apr 2018**

- Built a recommendation system using collaborative filtering to predict the ratings of 100k and 10million users
- Designed algorithms based on user's gender and movie genre, from the ratings given by top 50 similar users
- Achieved 80% accuracy in predicting the movie ratings of users

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
- Performed dimensionality reduction using Boruta to select the best features to predict restaurant revenue
- Gradient Boosting achieved Root Mean Square Error of 0.3, thus the overall error in prediction was very less

Maps using Artificial Intelligence Search algorithms in *Python* **Sep 2018**

- Designed maps to predict the total distance, time taken and the paths between any two cities in the USA
- Built A*, Uniform, BFS, DFS and IDS search algorithms with distance and time measurements as cost function
- GPS coordinates and length of roadway between cities in the USA were used as metrics for the algorithms
- Uniform search algorithm returned the most optimal path between any two cities, within 4 seconds

Image Classification on Natural Images Data using *HDFS* and *Pyspark* **Dec 2018**

- Implemented Random Forest, Logistic Regression and Gradient Boosting to compare performance with Python
- Classified 6899 images from different categories like Airplane, Car, Cat, Dog, Flower, Fruit, Motorbike, Person
- Run-time was reduced by 2000 seconds and 73% accuracy was achieved using Random Forest

Tweet – Location Predictor in *Python* [Highest accuracy among 200 students] **Dec 2018**

- Predicted the location from which 32000 tweets were posted using a Naïve Bayes classifier
- Achieved 72.5% accuracy by effectively handling stop words, special characters, and missing words

2016 US Presidential election analysis on Election Survey Data in *R* **Mar 2018**

- Fitted Logistic regression models on the post-election survey response of 64000 adults
- Studied the interaction of answers to immigration questions with gender, race , and education
- Analyzed the models to understand the switching of supporters between Barack Obama and Donald Trump