Raghavendran Shankar

906-275-9935 | ragashankar@gmail.com   Houghton, Michigan

EDUCATION

Master of Science in Data Science - Graduate Fellowship Recipient

Michigan Technological University - Houghton, MI   **Aug 2016 – Dec 2017**

GPA: 3.8

Bachelor in Computer & Electrical Engineering

Anna University, Chennai, India   **June 2009 – Apr 2013**

GPA: 3.5

TECHNICAL SKILLS

- **Programming**– R, Python (Numpy, Pandas, Matplotlib), Java
- **BI and Big Data Tools** – Hadoop and Map Reduce, Sqoop, Spark, Informatica Power Center 9.5.1
- **Data Science** – Exploratory Data Analysis, Machine Learning, Tableau visualization, PCA Dimensionality Reduction
- **Database** – MySQL, Hive

PROFESSIONAL EXPERIENCES

Miracle Software Systems, Michigan, USA   **Oct 2017- Present**

- Fetched data through API programming in Python from Social Network data sources such as Bing, Choozle and Facebook to analyze the clickthrough rate of different Ads.
- Worked on AWS to store the data in S3 instance and created tables in Amazon RedShift to store the data.

Data Analyst Intern – Hone Zone Technology Academy, Chicago, USA   **May 2017 – Sept 2017**

- Data Preprocessing and Data Mining techniques to analyze the enrollment and participation of students.
- Implemented Hive Queries on Avro, Parquet, ORC formats to store and partition student information in the form of tables on Hadoop Files System.
- Predicted the average graduation rate across each year using regression algorithms and asses coaching for students.

Larsen and Toubro Infotech, Houston, USA   **Oct 2013- May 2016**

Clients:

Data Scientist - NXP Semiconductors   **Jan 2015 – May 2016**

- Data transfer from Transactional DB to a Data Warehouse system with ETL tools in Informatica Center 9.5.1.
- Analyzed market growth and implemented efficient strategies using exploratory data analysis.
- Identified customer patterns and implemented regression techniques to predict the increase in quarter sales by 6% during next fiscal year.
- Visualized the analysis through plots obtained from the analysis and drawn conclusions.

Data Scientist – Sony Music Entertainment   **Oct 2013 – Dec 2014**

- Analyzed the song features for a specific range of years using R Studio and Python to find out the song successfulness
- Implemented Data Preprocessing techniques like Imputation, Label Encoding and Feature Scaling on the song features
- Performed 10-fold cross-validation to train the model and predict the successfulness based on test data.
- Applied Machine learning models such as Decision trees and Support Vector Machines, Gradient Boosting and Bayesian learning on Supervised dataset and achieved accuracy of more than 80% for each model.
- Evaluated the models based on Information Gain, confusion matrix and Region Operating characteristics curve.

DATA SCIENCE RESEARCH PROJECTS

- Analyzed the Housing prices at Michigan, plotted correlation graphs across the predictors and applied Gradient regressor model with Grid Search and predicted an r2 score to measure high Goodness of Fit as 0.8632.
- Analyzed Twitter pages of pharmaceutical companies and created term-document matrix to visualize most frequent words using Word Cloud and histogram; performed a sentimental analysis of positive and negative tweets.
- Exploration of survival data in automobile accidents and predicted the survival rate based on gender, age and place of boarding.
- Processed big data using Hadoop Ecosystem using Spark, worked on Sqoop to distribute files across in a Multi-node HDFS, performed Hive queries to partition tables for speedy retrieval of data.

linkedin.com/in/raghavendran-shankar   github.com/ragashankar/Projects