KARTHIK SUNDARAM SARAVANAN

503, Cherry Street, College Station, TX,77840 ▪ +1-(979)-739-7809 ▪ karthik.sun95@tamu.edu
LinkedIn URL: https://www.linkedin.com/in/karthik-sundaram-s
GitHub URL: https://github.com/karthik-sundaram

EDUCATION

Texas A&M University, College Station, TX — May 2019
Master of Science, Industrial Engineering (Data Analysis Concentration) — **GPA:4/4**

National Institute of Technology(NIT), Trichy, India — May 2017
Bachelor of Technology, Production Engineering (Industrial Engineering Concentration) — **GPA: 4/4**

Relevant Coursework: Probability & Statistics, Engineering Data Analysis, Big Data & Databases, System Simulation, Statistical Quality Control, Design of Experiments, Operations Research, Operations/Financial Management

INTERNSHIP EXPERIENCE

Stanley, Black & Decker (a Fortune 500), Houston, TX — May 2018-August 2018 (expected)
(Data Science Analyst Intern)

- *Working on the 'Asset Management' project in the 'Oil & Gas division'*
- *Integrated data from different platforms like SAP using Python & Spark to find the Return on Investment per product.*
- *Wrote Python scripts to compute the Utilization of the products over time using all the customer sales data.*
- *Built a Time Series Forecasting Model using ARIMA concepts and the 'prophet' package in python*
- *Implemented a Market Basket Analysis (a data mining technique) to group the different products & their dependencies.*
- *Mapped the customers & their sales on the world map using an external website's API after generating their coordinates.*
- *Developed a Web Application including various Visualization and Statistics of the Time Series Correlation of sales over time of any 2 products interactively using 'Bokeh' package in python(built over d3.js) from scratch.*
- *Implemented a Machine Learning to reduce the error in the forecast duration of a product rental by 55%.*
- *Processed my models & scripts that needed parallel processing for speed using AWS - Amazon Web Services (S3 & EC2)*

National Institute of Technology, Trichy, India — *May 2014-May 2017*
(Supply Chain Optimization/Analyst - Research Assistant)

- Formulated a Mathematical Model including suppliers, manufacturers, warehouse, retailers, recycling centers and three types of customer demands to aid in supplier selection, reduce cost & pollution, maximize profit and optimize inventory.
- Solved the model formulated using ILOG CPLEX and using meta-Heuristic Algorithms in Matlab.
- Output data analysis & sensitivity analysis was done in Microsoft Excel and Tableau to find the optimum parameters.

Technical University of Munich, Germany **/ Arla Cheese manufacturing plant,** Denmark — *May 2016-August 2016*
(Data Analyst/ Operations Research Intern – DAAD German Scholarship)

- Performed analysis on the water flow data (50+ excel files) of Arla using advanced Excel commands like Vlookup, SQL & Python for analyzing trends & patterns of the water flow schedule. Used Tableau dashboards to represent findings.
- Prepared a scheduling model in CPLEX for maximizing reuse of process water & minimizing intake of portable water.

Centre for Digital Financial Inclusion, India **/Indian Institute of Management**, Bangalore — *November 2015- December 2015*
(Statistical Analyst Intern)

- Worked on questionnaire development and determining the statistical methods that can be used to analyze the data that will be collected to develop a cashless system at PoS in the business transactions of small scale retail shops.

Royal Enfield bike manufacturing plant, Chennai, India**.** — *December 2014-January 2015*
(Industrial Engineering Intern)

- Solved a quality control/supply chain problem in the chassis of a "350cc Thunderbird" model bike by using Microsoft Excel/Tableau & tools like cause & effect diagram, pareto charts/histograms, flow charts to analyze the reason for it.

SOFTWARE SKILL SET

- Languages : SQL ,R ,Python ,C, C++, VBA
- Software : MS Excel, Tableau, Matlab, Spark, CPLEX, Minitab, Simio ,Linux/Bash Scripting (basics)

PROJECT EXPERIENCE (Individual/Self)

Home Depot Product Relevance search- a NLP / text analytics problem in Kaggle

- Aim: "To improve Home Depot's customers' shopping experience by developing a model that can accurately predict the relevance of search results." Solved this problem by data pre-processing (stemming, lemmatizing, spell-checking, stop

word removal) of more than 2,50,000 rows & creating new features for number of words matched & length of query. Then, used Machine Learning algorithms like XgBoost in python & neural networks in RStudio .

Credit Card Fault/Anomaly Detection (Kaggle dataset):

- Used methods like Multi-variate Gaussian, One-class SVM & Logistic Regression with Undersampling on a Kaggle dataset containing transactions made by credit cards in September 2013 by european cardholders to detect anomaly/fraudulent transactions.

Feature/predictor Selection of signals used in a semi-conductor-based manufacturing process

- Performed feature/predictor selection reducing the number of signal features from 591 to less than 15 by data cleaning, imputation and using various machine learning models. With the highest prediction accuracy of 93%, boosting was used to select the 15 important features with the highest sensitivity according to ROC curve and AUC.

Capital One - Data Science Coding Challenge (Data Size: 1.5 million rows, 21 columns)

- Framed predictive models, performed hypothesis testing apart from data pre-processing using Python (scikit learn, numpy, pandas, matplotlib, etc.) as a part of the problem statement which involved more than a million rows of geographical, time stamp & cost related data of taxi rides scraped from the “NYC Taxi & Limousine Commision” website.

BigMart Sales prediction [Analytics Vidhya - Online Project]

- Performed exploratory analysis for around 1500 products across 10 stores in different cities and built a predictive machine learning model in “R” to find out the sales of each product from the test data.