

Bhavish Kumar

315-832-0993 | bhkumar@syr.edu | [linkedin.com/in/bhavish-kumar/](https://www.linkedin.com/in/bhavish-kumar/) | github.com/bhavish2207 | [Portfolio](#)

EDUCATION

Syracuse University, School of Information Studies, Syracuse, NY
M.S. Applied Data Science

May 2021
GPA: 3.95

Relevant Coursework: Intro to Data Science, Data Analysis and Decision Making, Database Management, Big Data Analytics, Data Analytics, Business Analytics, Text Mining

Cochin University of Science and Technology, Cochin, India
Bachelor of Technology in Electronics and Communication Engineering

April 2016

EXPERIENCE

Mu Sigma Business Solutions Pvt Ltd, Bangalore, India

May 2016 to June 2019

Decision Scientist

- Generated an annual 200% increase in the overall customer value proposition score of a fortune 10 eCommerce client by formulating a competitor analysis framework that helped gauge their performance against the market leader
- Estimated monthly scores using the formulated analytics framework that quantifies dimensions of online customer experience and uses PySpark & SQL on survey and Hadoop data sources for estimating scores of quantified metrics
- Visualized metric scores on a R shiny dashboard to help Chief Officers and 2500 managers to track monthly progress
- Collaborated with client to perform iterative Big Data exploration, regression & clustering analysis based on business understanding to improve accuracy of the estimated metric scores, thus opening the scope for framework automation
- Headed a team of 4 and used Data Marts, Data Pipelines with PySpark & SQL to automate the process of estimating metric scores which decreased the evaluation & dashboard refresh time by 27-man hours

Trainee Decision Scientist

- Collaborated with Human Resources team and gathered requirements to develop business process web applications with the help of a tool that uses .Net/C# on the server side, HTML/CSS for frontend and MS SQL for backend
- Aided HR team to carry out their business processes by developing candidate hiring & attendance tracker applications from end to end, starting from data modelling & database creation at the back end to UI design at the front end

ACADEMIC PROJECTS

Will Car be a 'lemon'

April 2020 to May 2020

- Auto dealers who buy used cars at auto auctions face the risk of buying faulty cars ('lemons') which cannot be resold
- Implemented CRISP-DM methodology to develop ensemble learning classification models (GBM & Random Forest), using the 'caret' package (R), which will help the auto dealers accurately predict if a car is going to be a 'lemon'
- Evaluated the model performance and further fine-tuned the model to yield the best possible Accuracy & Recall
- Developed a web application (R Shiny Dashboard) containing insightful visualizations and a prediction module that uses machine learning models built using 'caret' package to make a prediction about the car based on the user inputs

Hospital Readmission Prediction

April 2020 to May 2020

- Diabetes patients face the risk of getting readmitted to the hospital after discharge, due to ineffective hospital treatment
- Headed a team of 4 that used PySpark & developed Machine Learning models to help the hospital predict if a patient is going to get readmitted after discharge, based on which the hospital can alter their treatment to avoid readmission
- Performed Data Wrangling, Data Preprocessing and identified the most important features using the Principal Component Analysis technique, which was implemented using the 'pyspark.ml' library
- Developed and fine-tuned a classification model using 'Gradient Boosted Trees' algorithm which gave an 'AUC' of 69%

TECHNICAL SKILLS

Languages: SQL, R (caret, ggplot2, keras, yardstick, tidyverse), Python(scikit-learn, pandas, numpy, matplotlib, seaborn, SciPy), R Shiny(shiny, arules, arulesViz, htmltools), PySpark(pyspark.sql, pyspark.ml, pyspark.mllib), C#, HTML/CSS

Techniques: EDA, Regression, Classification, Clustering, Deep Learning, Time Series Forecasting, Association Rules Mining

Tools/Softwares: MS Excel (Vlookup, Pivot Tables, Analysis ToolPak, StatTools, Solver), Tableau, MS SQL Server, Teradata, Hadoop/Hive, Spark, Jupyter Notebooks, Anaconda, GIT, MS Visio, MS Access, Bizapp

Soft Skills: Documentation, Requirement Gathering, Presentation, Business Communication, Decision Making

LEADERSHIP

- E-Board member (Tech Lead) of the Data Science Club at Syracuse University May 2020 to May 2021
- Organized a fund raiser for techno managerial symposium conducted by Model Engineering College March 2015