Bhavish Kumar

315-832-0993 | bhkumar@syr.edu | linkedin.com/in/bhavish-kumar/ | github.com/bhavish2207 | Portfolio

Qualified Data Scientist with a strong passion for building solutions that power data driven decision making. With 3+ years of experience working for a pure play Data Analytics firm focused on using Data Science techniques to aid businesses.

EDUCATION

Master of Science in Applied Data Science, Syracuse University, Syracuse, NY (3.95/4.0) May 2021
Relevant Coursework: Statistics, Data Warehouse, Big Data Analytics, Data Mining, Deep Learning, Cloud Management(AWS, GCP)

Bachelor of Technology in Electronics and Communication Engineering, Cochin University, India April 2016

TECHNICAL SKILLS

Languages: SQL, R (caret, ggplot2, tidyverse), Python (scikit-learn, PyTorch, keras, matplotlib), R Shiny, PySpark, C#, HTML/CSS
Concepts/Techniques: EDA, Machine Learning & Deep Learning, ETL, Data Warehousing, Statistics, Dashboarding
Tools/Softwares: Tableau, Power BI, Alteryx, MS SQL Server, Hadoop/Hive, MS Excel (Vlookup, Pivot Tables, Analysis ToolPak)
Soft Skills: Critical Thinking, Requirement Gathering, Business Communication, Decision Making, Consulting, Problem Solving

EXPERIENCE

Data Scientist at iConsult Collaborative at Syracuse University, Syracuse, New York August 2020 - Present

- Enhancing quality of healthcare by partnering with stakeholders of a major hospital and developing Machine Learning models to make patient level ailment predictions
- Developing an end to end Alteryx Data Pipeline to execute entire CRISP-Data Mining/Machine Learning operation

Mu Sigma Business Solutions Pvt Ltd, Bangalore, India May 2016 - June 2019

Decision Scientist

- Generated an annual 200% increase in overall customer value proposition score of a fortune 10 eCommerce client by formulating a competitor analytics framework to help gauge performance against eCommerce market leader
- Devised PySpark-SQL queries on survey & Hadoop/Hive data sources for evaluating monthly KPI metric scores using formulated analytics framework that quantifies dimensions of customer's eCommerce experience into metrics
- Visualized metric scores on a RShiny dashboard to help Chief Officers and 2500 managers to track monthly progress
- Collaborated with client to perform iterative Big Data exploration, statistical & ad-hoc analysis based on business understanding to improve accuracy of estimated metric scores, thus opening scope for framework automation
- Managed a team of 4 and built Data Marts, Data Pipelines through ETL operations with PySpark & SQL to automate process of assessing metric scores which decreased evaluation & dashboard refresh time by 27-man hours

Trainee Decision Scientist

- Collaborated with Human Resource team and gathered functional requirements to develop business process web applications by leveraging a tool that uses .Net/C# on server side, HTML/CSS for frontend and MS SQL for backend
- Aided HR team to conduct business processes by developing candidate hiring & attendance tracker applications from end to end, starting from data modelling & data warehousing at back-end to UI design at front-end

ACADEMIC PROJECTS

Will Car Be A 'Lemon'? April 2020 - May 2020

- Helped Auto dealers avoid risk of buying faulty cars ('Lemons') at auto auctions, which cannot be resold later
- Led a team of 3 in a strategic direction by implementing CRISP-DM methodology to construct ensemble Machine Learning models (GBM & Random Forest), to help auto dealers accurately predict if a car is going to be a 'lemon'
- Analyzed model performance and fine-tuned Machine Learning models to maximize Accuracy & Recall
- Created a web application (RShiny Dashboard) containing insightful visualizations and a ML prediction module

Hospital Readmission Prediction April 2020 - May 2020

- Aided hospitals to prevent post discharge readmission of diabetes patients caused due to ineffective treatment
- Administered a team of 4 that operated PySpark & devised Machine Learning models to help hospitals predict if a patient is going to get readmitted after discharge, to aid hospitals proactively alter treatment to avoid readmission
- Performed Data Wrangling, Data Preprocessing and identified most important features using PCA technique
- Developed a Machine Learning model using 'Gradient Boosted Trees' algorithm & fine-tuned it to maximize 'AUC'

LEADERSHIP

- Tech Lead, E-Board member 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