

SUJAI KARUNAKARAN

Chicago, Illinois 60607 | (669)2-640654 | sujaikarunakaran@gmail.com
[linkedin.com/in/sujai-karunakaran](https://www.linkedin.com/in/sujai-karunakaran) | github.com/SujaiKarunakaran

SUMMARY

Data Science professional with expertise in R, Python, Java, SQL, MongoDB, Hadoop and Tableau coupled with two years of full-time work experience in **Software**, internships in **Analytics and Operations**

EDUCATION

The University of Illinois at Chicago

Master of Science in Business Analytics

Chicago, IL

Expected Dec 2020

Tools Used: Python, R, SQL, Tableau, SPSS, RapidMiner, MongoDB, Hadoop, Hive, Apache Spark, Apache Pig

PES Institute of Technology

Bachelor of Engineering in Biotechnology

Bangalore, India

May 2017

PROJECTS

- **Analyzed online physician reviews to determine the important factors contributing to the star rating.** Using the data scraped from Healthgrades and RateMDs using Selenium Webdrivers, the data was processed using Natural Language Tool Kit (NLTK), sentiment was determined using TextBlob and data was modelled using Anchored latent dirichlet allocation (LDA) topic modelling method. Finally, using star rating as the target variable and machine learning algorithms, important factors contributing to the star rating was identified for different websites.
- **Developed Machine Learning models which identify the factors most influencing a patient to access one's own Online Health record.** Logistic Regression and Random forest learning techniques were incorporated to develop the model. This model was implemented using python like scikit-learn, NumPy, Pandas, Seaborn and matplotlib and tested against the real-world data set from Health Information National Trends Survey
- **Developed a "Propensity to Respond" model to predict the likely behavior of the consumer** using Random Forest, Regularized Logistic Regression, XGBoost (Gradient Boosting Technique) and Stacking. Since the dataset was unbalanced, Synthetic Minority Oversampling Technique (SMOTE) was used to balance the data
- **Visualized Delay Statistics of Envoy Air using Tableau** from a dataset extracted from American Bureau of Transportation Statistics. Designed an interactive dashboard and stories to visualize more profitable routes in terms of money and time Network graphs, LOD Expressions and statistical analysis

SKILLS

- **Coding Languages/Tools:** Python, R, Java, Matlab, SPSS, RapidMiner, Oracle MySQL, Oracle Database, Microsoft SQL Server, Microsoft Access, Tableau, PowerBI, Microsoft Excel, Selenium, HP ALM, MongoDB, Hadoop, Hive, Apache Pig, WordPress, Apache Spark, SDLC (Agile & Waterfall)
- **Analytical Techniques:** Machine Learning based Algorithms, Dynamic Pricing, Network Revenue Management, Network Data Analysis, Data warehousing, Big Data, NLP, Business Intelligence, Twitter Analytics, Time Series Analysis, Unsupervised Learning, Map-reduce, Sentiment Analysis, Topic Modelling and healthcare analytics.

PROFESSIONAL EXPERIENCE

Accenture

Application Development Associate

Chennai, India

July 2017 – April 2019

Tools Used: SQL, Python, HTML, HP ALM, Java, Tableau, Selenium

- Owned and led SQL modules for RTD database, involving multi-million-dollar engagements in telecommunications sector and headed application programming interface testing to attain utmost functional reliability of the product
- Prepared weekly, monthly and sprint reports to evaluate the standard of quality checks using HP ALM tool and built an interactive tableau dashboard
- Assisted an acquisition of RTD database module for Accenture using Oracle MySQL and Oracle Database that helped reduce quality assurance time by 50%
- Developed a webpage using HTML to help fill the knowledge gap pertaining to the functionality of the product and Software development Life Cycle that effectively reduced percentage of non-QOIs from 14% to 9%

Prashaste Education and Management Consultancy Pvt. Ltd.

Management Consultancy Intern

Bangalore, India

Feb 2017 – April 2017

Tools Used: Python, R, SQL, Microsoft Excel

- Improved the performance of sales outlets by identifying the key parameters associated with the profitability of the dealership.
- Perform digital medial analysis aiming at improving digital presence by identifying factors responsible for increasing the digital media presence on the website and other social media platforms

ACHIVEMENTS & OTHER ACTIVITIES

- Published a research paper on "Purification of glycerol produced as by-product of biodiesel production from waste vegetable oil" in "*International Journal of Advanced Research Trends in Engineering and Technology*"
- Received a "**Data Scientist**" Badge from **Tableau**