

HARISH RAJ

ASSOCIATE CONSULTANT/ BUSINESS DATA SCIENTIST

Summary

Overall 5 years of experience in various domains with 4 years in Data Science and predictive modeling and 1 year as Amazon Alexa Test Engineer.

- Involved in driving entire project life cycle from gathering business requirement, scoping possible solutions to implementing and presenting actionable insights from analyses to business stakeholders
- Experience in Deep Learning Machine Learning Algorithms(both Supervised and UnSupervised).
- Hands on Experience in using AI platform like **H2O**.
- Experienced in statistical analyses – **Neural Network, XGBoost, Genetic Algorithm, Clustering, Random Forest, Linear Regression, Optimization, Principal Component Analysis, Extrapolation**.
- Experience in Data Manipulation & Converting unstructured data to structured data and using R Libraries **dplyr, MICE, Miss Ranger, Cluster, Randomforest, Parallel** and **user-defined functions**.
- Experience in data extraction from Salesforce database and from various websites using **Web Scraping, Text Mining** and **NLP** (Named Entity Recognition).
- Knowledge in Data Visualization tools like **Tableau**.
- Experienced in Developing and hosting **R-shiny** Application in **Shiny.io** and in **Linux VM**.
- Experienced both as an individual contributor to projects and a team player.

Employment

CenturyLink India Pvt Ltd · **Associate Consultant** · Oct. 2018 to Current

RedHat India Pvt Ltd. · **Business Data Scientist** · July 2017 to Sept. 2018

Amazon Development Center India Pvt Ltd. · **Content Test Associate** · Dec. 2014 to Dec. 2016

Awards

Amazon Development Center India · **Extra Mile Award**
For Outstanding performance in 2016

Apr. 2016

Red Hat India pvt ltd · **Spotlight Award**
Best Performer for April 2017

Contact

✉ Rajlsharish@gmail.com
☎ 9008892008
📍 Bangalore
in [linkedin.com/in/harish-raj-ab1032128](https://www.linkedin.com/in/harish-raj-ab1032128)

Education

Impact College of Engineering & Applied Science 2009 to 2013

Skills

PROGRAMMING

R
Python
SQL

ML ALGORITHMS

Deep Learning
XG Boost
Decision Tree
Clustering
Random Forest
Stacked Ensemble
PCA
Logistic Regression
Linear regression

DATA VISUALIZATION

Tableau

WEB FRAMEWORK

R-Shiny

AI PLATFORM

H2O

Projects

Osmose Steel Towers

Oct. 2018 to Current

At CenturyLink Data Science team, I work for Osmose Utilities Service. It is a US based Utility management firm which provides electric steel towers and wood poles across North America.

As a Consultant ,I'm working as a individual contributor for this project and developed multiple models to predict corrosion state of steel towers.

Roles and Responsibilities:

- Involved in driving entire project life cycle from gathering business requirement, scoping possible solutions to implementing and presenting actionable insights from analyses to business stakeholders
- Architect and implemented a R-Shiny Application that will read in data of Steel towers and apply the data across various models to predict the corrosion state of steel Towers and call out the structures that need maintenance.
- Applied **Miss Ranger/MICE** Statistical technique to calculate missing values from read in data.
- Used probability distribution techniques like **National,Poisson, Bernoulli Distribution** to replace quantitative missing variables.
- Predict future Corrosion state of steel towers for next 50 years.
- Algorithms used: **Neural Network,Random Forest, Genetic Algorithm, XGBoost, MICE, PCA, K-means, Stacked Ensemble.**
- Exploring pain areas in the process, suggesting and implementing solutions for performance improvement and thereby increasing business value.

Real-time Mergers and Acquisition Information Generation

July 2017 to Sept. 2018

As a part of RedhatEDA team, I mainly work on the Sales Data. To expand/retrieve Redhat various services across Customers, Real time information on mergers and acquisition of Redhat customers was required, so that sales team can approach Subsidiary or the parent company to provide services.So a platform was required to continuously monitor and generate information about customers.

Roles and Responsibilities:

- Use Data Mining technique in R to build Web Scraping and generate real time information for sales team using R-Shiny App.
- Responsible for implementing data mining and statistical machine learning solutions to various business problems such as salesforce company name matching, Web scraping real time Information from various sources.
- Architect and implemented various machine learning techniques using R to automate legacy manual processes.
- Conduct data mining, data modeling, statistical analysis, Text Mining & NLP techniques.
- Data cleansing and manipulation for missing information, incorrect details, duplicate data and other related issues before processing further.
- Build various Shiny web apps through R and host them in local server.

Amazon Alexa Skill tester

June 2014 to Dec. 2016

- Extract data of submitted alexa skills from Amazon Redshift.
- Manipulate and clean the data with missing information, incorrect details, duplicate applications and other process related issues before further processing.
- Generate extracted data reports to managers and concerned team members on Submission rates, SLA Rates and any process related information on hourly,daily, weekly and monthly basis.
- Assigning data to team for further processing and making sure team is meeting SLA rate.
- Analyze the status of alexa skills (Echo) success and failure rates.
- Generating daily reports to managers on team productivity, quality and SLA rates.
- Preparing and presenting weekly and monthly management reports on Appstore apps submission rates.
- Analyzed and implemented various enhancement techniques related to process improvement and reducing processing time of apps.

Declaration

Declaration · I hereby declare that all the statements mentioned above are true,complete and correct to the best of my knowledge and belief

