

Arvind Kumar
Data Analyst
Callisto Infosolutions Pvt. Ltd.
Bangalore
Contact No.- 9538899388
Email-arvinddata29@gmail.com

Professional Summary:

Experienced and driven data scientist with 3.1 years of experience in **advanced predictive analytics** with main focus on consumer analytics in the industry. A candidate with strong background in developing predictive models from scratch using the knowledge of techniques like **Regression, Random Forest, Cluster Analysis, Decision Tree, Correlation, SQL** and programming language- **R and Python**. Commitment to providing support and essential information about trends in marketing domain

Project Profile:

Industry: Insurance services

Role: Data Analyst

Client Description: Auto Insurances client

The project aimed at building a model to predict the losses or the claim the people have made over the years and then finding the factor which impact the company in negative ways so that they can set different premium for different categories of people. The model was build using linear regression algorithm using R and its packages. The **predictive model** was successfully made and delivered to the Client within the allocated time.

Roles and Responsibilities:

1. To understand the Business problem and interact with the client on regular basis to get more and deep insight of the data shared by the client.
2. To do exploratory data analysis on the data which includes to prepare the quality data for model building. This was being done with the help of the **programming language –R**. This includes the calculation of the missing values using **Random Forest Algorithm**.
3. To find out the outliers in the data and use statistical concepts to study the effect of these outliers on the given data.
4. To do the variable selection which are important in developing predictive **model** with good accuracy. This involved usage of dimension reduction algorithms.
5. To draw the insights from the data and develop a predictive equation which will help in finding the factor which are more prone to go for claim and then predicting the claim the people have made over the years.
6. To recommend measurable actions and provide deep understanding of **prededction** using packages of R and Linear Regression to fit the future values on the curve.
7. Rewarded with “**Star of the Project**” within the Delivery Unit.

Industry: Financial services

Role: Data Analyst

Client Description: A leading Bank client

The project aimed at classification (cluster analysis) of certain section of customers which are profitable to the bank and most likely to take the term deposit plan of the bank. Develop a predictive model from scratch using concepts of algorithms (Random Forest, linear Regression) and business problem shared by the Client. The **predictive model** was successfully made and delivered to the Client within the allocated time.

Roles and Responsibilities:

1. To understand the Business problem and interact with the client on regular basis to get more and deep insight of the data shared by the client.
2. To do exploratory data analysis on the data which includes to prepare the quality data for model building. This was being done with the help of the **programming language -R**. This includes the calculation of the missing values and replacement.
3. To find out the outliers in the data and use statistical concepts to study the effect of these outliers on the given data.
4. To do the variable selection which are important in developing predictive **model** with good accuracy. This involved usage of dimension reduction algorithms.
5. To draw the insights from the data and develop a predictive equation which will help in categorizing the customer of the bank and to which cluster it should belong.
6. To recommend measurable actions and provide deep understanding of **clustering** using packages of R and Linear Regression to fit the future values on the curve.

Industry: Technology Company

Role: Data Analyst

Client Description: A leading Satellite TV provider

The project aimed at the **Customer Churn Analysis** where the loyalty of the customer needs to be identified using descriptive analytics and provide the insights how to retain the present existing customers in the company. The project involved a deep understanding of the customer data and analyzing the same with the help of **SQL, Regression, correlation** and programming language – **Python and R**

Roles and Responsibilities:

1. To fetch the data from the database based on the business criterion and load the same in the **excel sheet** for further processing using **python**.
2. To calculate the tenure of the customer and create a new variable which will determine the loyalty of the customer within the company.
3. To select the important variables and deep dive in the patterns followed in the data set to understand the customer sentiments.
4. To find the missing values in the data set using Random Forest Algorithm. This includes testing the training data sets and validating with the test data set.
5. To find the missing values in the data set using different algorithm and check the accuracy with each algorithm.
6. To develop the predictive model (Customer Churn) about the probability that the customer will leave the company and what is the loss made to the company in total of all the **Churns** observed.
7. Received Client appreciation and PRIDE award from the client for valuable contribution to business.

Programming and Development Skills:

- Languages : R,Python
- Machine Learning : Linear Regression, Logistic Regression, Cluster Analysis, Decision Tree, Random Forest, Support Vector Machine, Neural Networks, KNN,Bag of model,Time Series ,text rank algorithm
- Tools : Dplyr, ggplot2, E1071, Caret, Random forest, Jupyter Notebook, Pandas, Microsoft Excel

Education:

B.Tech from Rajasthan Technical University with aggregate of 65 %.

Declaration:

Here by, I do declare that all the above mentioned details are correct.

Arvind Kumar

Bangalore

