|  |
| --- |
| **Profile Summary** |
| A Highly efficient, innovative professional with **2.5+ yrs.** of expr in **Data Science & Analytics**   * Worked with teams, involved in creating full-scale data solutions and industry-specific products * Strong in building cross validated- best **ensemble models** from scratch (Bagging, RF, SVM, glmnet, GBM models) * Expertise in **building predictive models** * Well versed in using appropriate **“Machine Learning”** algorithms * Highly proficient in **Time Series Analysis, Dynamic regression and Forecast** * Build best-fit models based on the data and providing appropriate insights to business problems * Feature selection and engineering * Proficient in understanding and analyzing of **data pertaining to various domains** * Expert in using **all seven layers of “ggplot2”** * Possess excellent project and interpersonal skills * Enjoy being a team player as well as taking leadership * IT faculty at Indian Army training center * DataCamp certified Data scientist |

|  |
| --- |
| **Skills** |
| * **Information Technology**   + Data Science     - **R,** Python**,** SQL, Hive, Pig, Spark & Scala, SPSS, **Time Series Analysis**, Natural Language Processing   + Analyzing and **Visualizing Data with Excel** & Tableau   + Well versed with MS Office Suite( Excel, PowerPoint & Word) * **Business Statistics (Having very strong command)**   + Descriptive Statistics   + Inferential Statistics   + Predictive Analytics |

|  |  |
| --- | --- |
| **Professional Experience** | |
| **First Tech Consulting Data Scientist** | |
| **Project 1** | **Healthcare Analytics** |
| **Description** | This project deals with trauma cases coming to the hospital for treatment. The main objective is to find out the relationship and association between injury types. By using these hidden relations, associations & patterns doctors/physicians recommend different tests and scanning |
| **Responsibilities** | * Cleaned & transformed the data to do required analysis (**By using “dplyr”, “tidyr”** packages) * Performed **chi-square test analysis** to check tests of independence * Used "**apriori**" algorithm **to mine the strong association rules(ARM)** * **Used supervised & unsupervised modelling techniques using R/Python** * Built multiple models to predict injury types & cause of death by using **“caret”** **package** * Created ensemble models with **Decision trees, Random Forest, Bagging** * **Boosting algorithms** to get high model performance**(GBM)** * Performed Classification using **Deep Learning, Artificial Neural Networks** * Performed **AUC & ROC analysis** of models built * Compared classification models with Logistic, Decision trees, Random Forests, & SVM techniques * Handled over plotting by using certain ways with ggplot2 * Created various functions, which will give detailed summary reports * Used **R markdown** to turn analyses into high quality documents, reports, presentations and dashboards |

|  |  |
| --- | --- |
| **Project 2** | **Time Series Analysis & Forecast** |
| **Description** | This project is for visualize time series, trend & seasonality spotting and forecast sales data. In this project, I used/checked variety of algorithms to forecast accurately |
| **Responsibilities** | * Building, visualizing and exploring time series objects * Transformations for variance stabilization * **Box-Cox** transformation * Time series simulation with “arima.sim” * **P/ACF analysis** and deciding with correct method to build the model * Used **naïve, meanf, ses, ets, AR, MR, ARMR, ARIMA, exponential smoothing** to build models * Taken care of proper **residual analysis** * **Increased forecast accuracy** by choosing appropriate ts algorithm * Time series cross validation for best model * **Dynamic regression** to include independent variables in the model * Applied methods to handle more complicated seasonality * Used functions from “xts” and “zoo” to create, import, subsetting, extraction, merging, modifying time series data |

|  |  |
| --- | --- |
| **Project 3** | **Text Analytics** |
| **Description** | This project aimed at distilling actionable insights from text reviews given by employees, such that which company has better work life balance and which has better perceived pay |
| **Responsibilities** | * Analyzed and visualized text data * Created, cleaned and transformed the corpus * Created TDM/DTM matrix * Created different word clouds like **commonality cloud, comparison cloud** * Created pyramid plot, network plot and associations * Performed a hierarchical cluster and created dendrogram for word cluster analysis * **N-gram tokenization** with “weka” for analysis with multiple words |

|  |  |
| --- | --- |
| **Project 4** | **Telecommunication Analytics** |
| **Description** | This project deals with voice connectivity and call drop problems. The main objective is to find out the reasons for call drops and voice connectivity problems. Built a classification regression model to predict a call drop |
| **Responsibilities** | * **Boruta** for feature(s) selection * Used base R graphics for exploratory analysis * Used **ggplot2 for explanatory analysis** * Used data imputation techniques for missing data * Performed **multi-collinearity analysis** to find out and handle highly correlating independent variables * Find out the important variables causing call drops * Performed reduction of variable dimensionality by using **scaling and PCA** * Built a model to predict call drops by using machine learning algorithms * Created ensemble models with boosting algorithms * Used **confusion matrix** to interpret the output * Used **ROC and AUC** for model performance * Handled overfitting in predictive models * Used R markdown to turn analyses into high quality documents, reports, presentations and dashboards |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Education** | | | | |
| **Year** | **Institute** | **Course** | **Key Electives** | **Percentage** |
| **2011-15** |  | PGDM | Marketing and Finance specializations | 73 |
| **2008-13** | ANUCDE, Guntur | B.Com | Commerce, Economics & Computers | 44 |
| **2001-03** | S S & N College, Narasaraopet (AP) | Intermediate | Botany, Zoology, Chemistry & Physics | 61 |
| **2001** | Z.P high school, Sirigiripadu (AP) | SSC | N/A | 79 |

|  |  |
| --- | --- |
| **Personal Details** | |
| **Date Of Birth** | 10th Apr 1986 |
| **Nationality** | Indian |
| **Marital Status** | Married |
| **Languages Known** | **English, Hindi & Telugu** |
| **Permanent Address** | 4-2-206/A, plot-38, Bharathi nagar colony, Bandlaguda jagir, Hyderabad, Telengana-500086 |