**Lakshmi Panguluri**

Mail ID: lakshmipanguluri65[@gmail.com](mailto:amarnadhdatascientist@gmail.com)

Mobile: +91 9502032749

## **SUMMARY**

* Lakshmi is a **Data Scientist** with **4+ years** of experience with primary focus on **data analysis, predictive modeling using Libraries, R Functions and Python libraries, Data Cleansing**.
* Strong analytics knowledge, innovative thinking to extract new meaningful and actionable features from data.
* Proficient in building **Predictive modeling** like, Classification, Clustering/Segmentation for Supervised and Unsupervised models.
* Strong experience in dealing with performance, scalability, distributed caching, API, and open-source technologies.
* First experience with NLP.
* Extremely working experience in data analysis, obtain insights from data then choose appropriate Machine Learning/Data Mining algorithms like Regression, Logit, Probit, Decision Tree, Naïve Bayes, KNN, Random Tree/Forest, Support Vector Machines, Time Series (ARIMA, Holt’s Winter, etc…) and Neural Networks.
* Proficient in uses cases in **insurance industry** such as Workforce management, Fraud detection, Price quotation, Customer segmentation, Customer life cycle prediction, Text mining, Recommendation engine etc.
* Experience in **R libraries** like dplyr, lubricate, ggplot, tidyverse, rJava , forecast, caret, readxl, xts, etc.
* Experience in **R-functions** like sapply, write, read, glm, apply, nnet, lm, forecast and also user defined functions.
* Experience in **python libraries** like Scikit-learn, pandas, matplotlib, numpy, NLTK, Gensim, Whisper etc.
* Experience in public cloud technologies: **Azure Devops**.
* Knowledge on **Core and Advanced Java**.
* Setting up a statistical model and impress upon the clients need and usefulness.
* Have a keen desire to solve business problems and find patterns and insights within structured and unstructured data.
* Passionate about working on **large datasets** and able to communicate high valued solutions to business stakeholders.
* Effectively summarizing and visualizing the data summary results.
* Team player who can successfully lead and motivate cross-functional teams.

**SCHOLASTICS**

|  |  |  |
| --- | --- | --- |
| **Qualification** | **Year** | **Board/University** |
| MCA (Master of Computer Applications) | 2010-2013 | RVR & JC College of Engineering,  Guntur, India |
| BSC (Mathematics, Physics & Computers) | 2007-2010 | JKC College, Guntur, India |

**AREA OF EXPERTISE**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **DOMAIN** | **AREA OF EXPERTISE** | | | |
| **STATISTICAL ANALYSIS** | * Exploratory Analysis * Data Quality * Test of Hypothesis * Design of Experiments | * ANOVA * Regression Models: Functional Models * Classification models * Forecasting models | * Dimension Reduction: PCA * Cluster Analysis * SEMMA   Recommendations   * Associations | * Model Assessing * Model Validation * Model Diagnostics: Validating assumptions and removing violations |
| **MACHINE LEARNING** | * CHAID &CART * Random Forest | * Text Mining * Sentiment Analysis | * Naïve Bayes * KNN | * Neural Networks * Clustering |
| **STATISTICAL TOOLS** | * R-Studio | * Python | * Excel, | * Tableau, Power Point |
| PROGRAMMING | * R | * My SQL | * Python |  |

## **PROFESSIONAL EXPERIENCE**

|  |  |  |
| --- | --- | --- |
| **Organization** | **Year** | **Designation** |
| Infomerica | 2023 June-2023 Sep 18th | Data Analyst |
| Venus Cyber Tech LTD | 2018- 2023 | Data Scientist |

**PROJECT SUMMARY**

**Client**: PEPSICO

**Domain**: Food Services

**Role**: Data Cleansing

**Technologies Used**: **SQL**-Where, group by, joins & **Power BI:** Generating reports & **Python**: **VS Code, Spyder**, Pandas, NumPy & **Azure Devops**: for pushing the code & **Excel**: For Data Mapping.

**Responsibilities**:

1. Extracting the required Data from the source file.
2. Cleaning the extracted data using python: pandas and NumPy
3. Working on Data Mapping
4. Creating Dashboards using Power BI.
5. Getting data from the Database using SQL
6. Push the code using Azure Devops

**Client**: HSBC

**Domain**: Finance

**Role**: Data Scientist

**Technologies Used**: **SQL**-Where, group by, joins & Tableau**:** Generating reports, joins, calculation field, KPI Reports. & **Data Science**: Regression Techniques, Recommendation Engine, NLP, Clustering Techniques. & **Python**: Pandas, NumPy, Matplotlib, Scikit-Learn.

**Responsibilities**:

1. Retention of customers using CLV (Customer Life-time Value) approach
2. Identify the customer types using customer segmentation approach
3. Identify the bad customer using fraud deduction approach
4. Increasing the sales by proposing addition products to customer using recommendation engine
5. Responsible for KPI reports for project.
6. SQL used for filtering, joining the tables and to extract the data
7. Tableau for generating reports, calculation field, joins, KPI reports.
8. Data Science and Python for Statistical modeling and Data Anaysis.

**Client**: Sears

**Domain**: Retail & E-commerce

**Role**: Data Analyst

**Technologies Used**: **SQL**-Where, group by, joins & **Tableau:** Generating reports, joins, calculation field. & **Data Science**: Regression Techniques, Recommendation Engine, NLP,Clustering Techniques, Time series. & **Python**: Pandas,Numpy,MAtplotlib,Scikit-Learn.

**Responsibilities**:

1. Automated the delivery/ shipment charges using predictive modeling
2. Forecasted the volume using time series.
3. Customer identification for new products and recommendation using predictive modeling
4. Reduce the employee attrition using data research
5. Responsible for cleaning the data.
6. SQL used for filtering, joining the tables and to extract the data
7. Tableau for generating reports, calculation field, joins, KPI reports.
8. Data Science and Python for Statistical modeling and Data Anaysis.

**Client**: Philips Electronics

**Domain**: Consumer lifestyle

**Role**: Business analyst

**Technologies Used**: **SQL**-Where, group by, joins & **Tableau:** Generating reports, joins, calculation field. & **Data Science**: NLP & **Python**: Pandas, NumPy, Matplotlib, Scikit-Learn.

**R**: **dplyr**-joins, select, filter & **ggplot, lubridate, sapply,apply,r**

**Responsibilities**:

1. Requirement gathering from stakeholders and prioritization
2. Create user documents and provide user trainings
3. Data management
4. Scheduled and adhoc reports design and development.
5. SQL used for filtering, joining the tables and to extract the data
6. Tableau for ganerating reports,calculation field, joins.
7. R used for Data Extraction and Filtering and generating reports.