Vikash Kumar Prakash |

Pune, India

Data Analytics | Agile-Scrum Story Author | Oracle Core-Tech Technology Specialist

vikashkumarprakash1983@gmail.com +91-9665780336 https://www.linkedin.com/in/vikash-prakash-03850179/

OBJECTIVE

A result oriented Analytics professional looking to implement Analytical tools and techniques to help companies make better decisions, identify key business problems, interpret results, generate insights and recommend strategies on Business Intelligence tools **R**.

Professional Summary

- **8.5** years of **IT experience** in application development and maintenance.
- **2.7 in Machine Learning and Data Analytic using R**.
- Importing data in **R** from **csv, excel, databases** etc.
- **Exploratory data Analysis** (EDA) using **histogram, boxplot,** summary, **table in R**
- **Expertise** in **importing, data cleaning, analyzing, coordinating data, Manipulation of data in R using tidyR, spring, and springr packages.**
- **Outlier identification** using histogram, boxplot, scatterplot, quantile and treatment using imputations by **mean, median**, capping etc.
- Missing value treatment using single imputation and multiple imputation techniques (**MICE** Package).
- Worked on **Advance oracle Pl/Sql** (Collections**, Bulk collect, Analytic Functions in oracle**).
- **Experience** in **Performance tuning / Optimizing SQL Queries.**

TECHNICAL EXPERTIES

- Data Manipulation in **R** using **dplyr, sqldf** and base R.
- **Statistical Analysis: Hypothesis Testing using t-test, ANOVA, Chi-square, Correlation and Regression using R and MS Excel.**
- Regression Techniques such as **Linear Regression**, **Regression Trees**.
- Classification Techniques such as **Logistic Regression**, **Classification Trees**.
- Segmentation techniques: **Cluster Analysis** – **Hierarchical and K-Means**.

- Proficient in ensemble techniques i.e. **Bagging, Boosting and RandomForest**.
- Have a working knowledge of **XGBoost** on **classification** and **regression.**
- Dimensionality reduction using PCA.
- **Proficient** in **Data Visualization** tool **Tableau.**

Experience Detail

iNautix Technology A BNY Mellon Company	Nov-2013 Till Now	Data Analyist
Mphasis	Aug-2011 - Oct 2013	Oracle PL/SC
Zensar Technology	July 2010 – July 2011	Oracle PL/SC

Education Information

B Tech from BPUT University with a CGPA of 6.3

Project Undertaken

Sl No 01	**Period** :Dec 17 Till Now	**Project Name** : **Fraud Detection**
	Technology : R	**Tool** : R Studio

- **Fraud Detection of Payment Transactions using Machine Learning Approach_1.0**

Brief Description:

Anomaly detection approach: Given such a "huge class imbalance" we first considered applying "anomaly detection approach" - unsupervised methods that compute outlier score for each transaction. The idea is outlier transactions generally reflect fraud. Specifically, we applied **Histogram Based Outlier Score (HBOS)** method. More details on method and results can be found here: Histogram Based Outlier Score for fraud detection. However, the HBOS approach produced high false positive rates and one of the feedbacks from business was if we can reduce false positives further.

Sl No 02	**Period** :Dec 15 Till Now	**Project Name** : **Accrual Accounting basis**
	Technology : R	**Tool** : R Studio

❖ **Predicting the Future Operating Cash Flow based on past Accrual Earning.**

The past Accrual based Earning is derived from the income statement prepared on the accrual accounting basis of each client. Concept of matching expenses and revenue in arriving at accrual based earnings reflects the uses of assets in generating revenue.

Roles and responsibility:

- Importing the data from Eagle Accounting system using R by connecting Oracle DB.
- Cleaning the Data from the data set and manipulating the data using the **tidyr, dplyr** and **lubridate** package in R.
- Optimizing the model by checking the Multicollinearity using **VIF** and found correlation between the variable ranges from 0.183 to 0.685.
- Build **Linear Regression Model** on the basis of Income statement earing at all the lags (1 year lag, 2 years lag, and 3 years lag).

Sl No 03	**Period** :Dec 15 Till Now	**Project Name** : **Operating Cash Flow**
	Technology : R	**Tool** : R Studio

❖ **Predicting the Future Operating Cash Flow based on past Cash Flow.**

The cash basis accounting (also known as cash flow accounting) denotes a system of financial reporting which describes the performance of entity in cash terms. It is based on matching periodic cash inflows and cash outflows, free of credit transactions and arbitrary accounting allocations. Cash flows are not infected by measurement problems and thus accelerate the prediction of future dividends and credit and loan payments.

Roles and responsibility:

- Importing the data from Eagle Accounting system using **R** by connecting Oracle DB.
- Cleaning the Data from the data set and manipulating the data using the **tidyr, dplyr and lubridate** package in **R**.
- Checking the Collinearity and Multicollinearity using **COR** and **VIF** function.
- Build **Linear Regression Model** on the basis of past Cash Flows CFOt-1, CFOt-2, and CFOt-3.

Sl No 04	**Period** :Dec 15 Till Now	**Project Name** : **Segmentation of Trade**
	Technology : R	**Tool** : R Studio

❖ **Segmentation of the Trade information on the basis of the Asset Class across BUY and SELL over the period of time.**

Roles and responsibility:

- Importing the data from Eagle Accounting system using **R** by connecting Oracle DB.
- Cleaning the Data from the data set and manipulating the data using the **tidyr, dplyr** and **lubridate** package in **R**.
- Using **K mean clustering** Segment the Trade BUY and Sell across different Asset class.
- Apply **Scale** function in **R** on the Trade Quantity so that data set has a normal distribution.
- Using **ggplot** function for the graphical representation of the trade in different asset class across the different Sectors

❖ **Analyzing the Traded Cash returns on the basis of Earning over the Time across different Securities and Sectors.**

Roles and responsibility:

- Importing the data from Eagle Accounting system using **R** by connecting Oracle DB.
- Cleaning the Data from the data set and manipulating the data using the **tidyr, dplyr and lubridate** package in **R**.
- Applying the **TS** function in **R** on the Cash returns from each security across the different sectors.
- Using the **seasonplot** function in **R** represents graphically the Cash returns of Trade and the Accrual Earning.
- On the basis of above analysis provide the suggestion to client to invest in which sectors to get better cash returns.

Sl No 05	**Period** :Nov13 till Nov15 Project	**Name** : IMO Dashboard	**DataBase** :Oracle 11g
	Technology : PL/SQl, Pearl	**Tool** : Eagle, Sql Developer	

❖ Provides the inquiry ability to Trade, Reconciliation and Collateral data.

Provides high-level snapshot summaries with the ability to drill down to details of the trade, reconciliation and other details of the trade. Permits clients to customize the display format as either table or chart. Three types of Dash Boards Trade Dashboard, Reconciliation Dashboard and the Collateral Dashboard. Trade Dashboard provides the status of the Trade. The Reconciliation Dashboard provides the status of reconciliation with the custodian.

Responsibilities :

1. Analyzing and Understanding business Requirements for corresponding each of the dashboards.
2. Responsible for changing the existing Events and the process flow of the dashboard if clients request. Adding any new client or services.
3. Developed and Implemented Program Structure using SQL and PL/SQL.
4. Actively involved in tuning of SQL Queries.

5. Creating source, new schemas, feeds, code values, events in the Eagle investment tool.
6. Work involved creation of technical design doc based on FD, coding, unit-testing.
7. Interacting with onsite coordinators as well as the clients frequently.

Machine Learning training Project

Below are the lists of Machine learning Model worked on as part of training /learning.

Model	Description	Git hub Location
Linear Regression	Predicting wine Price	https://github.com/vikashkumarprakash?tab=repositories
Linear Regression	NBA Data Set	https://github.com/vikashkumarprakash?tab=repositories
Logistic Regression	Health Care :Predicting Good or Bad service provided at the Hospital	https://github.com/vikashkumarprakash?tab=repositories
K-Mean	Brain Tumor image	https://github.com/vikashkumarprakash?tab=repositories
Cluster Algorithm (H Clust)	Netflix Data set and doing rating, choice of Movies	https://github.com/vikashkumarprakash?tab=repositories
Sentiment Analysis	Pulled the Tweet of Modi and Kejriwal by calling Twitter API and apply the Sentiment analysis on their Tweets .	https://github.com/vikashkumarprakash?tab=repositories
Logistic Regression	US President election prediction	https://github.com/vikashkumarprakash?tab=repositories
Decision Tree	On the basis of experience and technology creating Decision Tree to predict the salary	https://github.com/vikashkumarprakash?tab=repositories
XGBoost	Apply the XGboost on the Oil Well to get the prediction of oil volume from the wells.	https://github.com/vikashkumarprakash?tab=repositories

Personal Details

Name : Vikash Kumar Prakash

Date of Birth : 24-May-1983

Languages Known : English, Hindi and Oriya

Marital Status : Single