

DATA SCIENCE PROFESSIONAL

Dynamic and dedicated **Data Science Professional** carrying **7 years** of expertise in **Predictive Modelling, Banking Domain and HR analytics** while generating ideas for new processes and their implementation

A committed professional with 7 years of experience in:

Data Science | Data Analysis | Predictive Modelling | Machine learning | HR Analysis | Statistical Modeling | Hypothesis Testing | Classification/Regression | Clustering | Decision Trees | Natural Language Processing | Random Forest | Navie bayes | Time series | SQL | VBA (Macros) Automation Training & Development

PROFILE SUMMARY

- Deft at undertaking various statistical methods and testing processes such as hypothesis testing, Linear regression, Logistic regression, Clustering, Decision Trees, Random Forest and Natural Language Processing.
- Pragmatic exposure in data ingestion and cleaning using Python Programming. Thorough knowledge of Data warehousing and Data mining concepts.
- Sound knowledge of drawing insights through predictive analysis and data visualization, to drive better decision making through the use of various statistical modelling techniques.
- Experienced in all phases of data analysis spanning definition and analysis of questions with respect to available data & resources, overview of data & assessment of data quality, selection of appropriate models & statistical tests and presentation of results
- Leveraging skill in importing, cleaning, transforming, validating or modelling data with the purpose of understanding or making conclusions from the data for decision making purposes.
- Good at SQL commands (DDL, DML, DCL, Joins, sub queries) and SQL functions.
- Automated the Excel reports using VBA by connecting the SQL server with ODBC connections.
- Enriched experience in automating the process using VBA, SQL and validation techniques for checking model performance and accuracy as well managing financial time tracking compliance and adherence to best practice.

TECHNICAL PURVIEW

Programming Languages	: C, Python, VBA
Databases	: Oracle, SQL Server, MS Access
Analytics Tools	: Python
Analytics Techniques	: Linear regression, Logistic regression, Decision Trees, Random Forest, Cluster Analysis and Natural Language Processing, Hypothesis Testing (t-test, z-test, Chi Square, ANOVA) etc

CAREER TIMELINE

OCWEN financial solutions pvt ltd, Bangalore ▶ Senior Analyst | Jul 18 – Till now

Project: Non Borrower billing (Mortgage) ▶ Client: OCWEN

Business Problem: We need to collect amount from investors not from borrower. So, we need to identify which borrower will pay amount for next month based on last one year data using classification model

- Performed data cleaning, data integration(Extract data from Oracle database) and data preparation to enable further analysis
- Developed the model for **Monthly billing** using **Logistic regression, Decision Tree and Random forest** using python.
- Provide the data to operational team by segregating the data based on probability.
- Operational team will take further approach, how to get amount from investors.

Techniques/Tools: Python (Random Forest), Oracle, VBA and Excel

Growth Path: Data Analyst **Since:** Jan'2015- Jul'2018

Project: SMS Structured Data Development ▶ Client: Nielsen

Business Problem: Data should contain structured format with 15 useful variables.

SMS Content: SMS data received, fed into feature extraction engine
Text Processing: Removing stop words, Stemming, Lemmatization
Sentence Classification: Classifying SMS content into promotion, verification, unsuccessful and actual transaction sentence
Keyword Generation: BOW(Bag of words) to generate most frequent words.
Keyword Classification: For classification, we used business term as well as given/identified patterns
Feature Extraction: Matching similar keywords to extract particular features
Quality Check: Randomly selected 307 unique pattern for all Banks/Wallet and validated

Techniques/Tools: Python (Natural Language Processing)

Project: Loan prediction ▶ Client: HSBC | Tool Used: R | Machine Learning Algorithms: Logistic Regression, Decision Tree and Random Forest

- Working closely with peers and stakeholders to access, pull, clean and analyse data from a variety of data sources. Gathering new data, link disparate forms of data, and/or creatively use existing data to propose metrics and present them in a compelling way
- Interpreting data, analyzing results using statistical techniques and providing on-going reports. Developing and implementing data analyses, data collection systems and other strategies that optimize statistical efficiency and quality
- Working backwards from outputs (integrations, services, reports, etc.) to ensure all required data is captured, controlled and validated within upstream processes.
- Involved in building a repository of data integrity challenges for each data element for action planning and also analysis across all data elements looking for common threads
- Safeguarding or protecting the business and user's needs by verifying functionality, accuracy and completeness of the requirements against the original initiating documents

Objective: When a counter party or group comes to investment bank with a business proposal or some credit need, the project aims at analysing the counter party, business and amount at asserts. Internal and external ratings based on the same will release the amount.

Credit details: Includes Balance of current account, duration in months, payment of previous credits, purpose of credit, Amount of credit, No. of previous credits,

Collateral (Savings): Most Valuable available asserts, Value of savings or stocks, Instalments % of available income (How much he paying amount to EMI already have loan, Further debtors)

Counter Party Details: Age, Occupation, current organization living in current house hold.

Credit Score: Internal and external credit ratings (S&P, Moody's and Bloomberg).

IBM India, Bangalore ▶ Business Analyst | Nov 13 - Nov 14

Project: HR Analytics (Separation Team) ▶ Client: IBM | Tool Used: R VBA, SQL, Excel

Description:

Sending daily reports to global countries like(India, KL, China, Budapest, MEA).How many are completed separation from IBM, How many are not completed, How many are Pending, and How many are crossed date but still pending with full and final settlement

Responsibilities:

- Responsible for all project lifecycle phases, from specifications and coding through deployment, testing, debugging, documentation, and maintenance
- Designed and wrote the automation code for database queries and resolution
- Creating link-based module to enable automatic updates for excel workbooks
- Finished all the projects on or before time and within allocated budget
- Very good knowledge in userforms and active x controls from VBA

SCHOLASTICS PORTFOLIO

B.Tech (Information Technology) from Annamacharya Institute of Technology & Sciences, Tirupati
(Jawaharlal Nehru Technological University, Anantapur)

PERSONAL DETAILS

ssDate of Birth: 26th May 1990 | **Nationality:** Indian | **Present Location:** Bangalore, Karnataka