# Manish Gupta

manish.gupta797@gmail.com| +919632557347 | Bangalore


## Software Engineer & Certified Data Science Professional

*Competent software engineer and certified data scientist. Looking for long-term career in data driven technologies. Self-motivated & learnt data science platform with a professional degree from a premier institute. Well versed with statistics concepts, data visualization, and data analysis (in R, python, SQL, Advanced SQL) & modelling with working knowledge on Big Data platform (Hadoop and Spark). Proficient in full stack development of enterprise applications on java platform currently being core member of Oracle Sales cloud CRM SaaS based application. Followed Agile based development throughout career.*

## TECHNICAL SKILLS

| | |
|---|---|
| **Tools & Languages:** | Python, R, C , C++, Java, SQL , Hadoop, Spark, Jupyter IDE, R Studio |
| **Statistics/ML:** | Linear/Logistic, Regression, SVM, Ensemble Trees, Random Forests, Neural Networks ,Clustering -K –means, Hypothesis Testing, Statistics. |
| **Data Visualization:** | Python's scikit-learn matplotLib , R GGplots , Tableau |
| **Enterprise Framework:** | Java Server faces(JSF), Oracle's Application Development Framework(ADF) for Business model to create database structures and UI. Django for Python based web development. |

## KEY SKILLS

| | | |
|---|---|---|
| **Supervised/Unsupervised Learning** | **Business Analysis & Strategy** | **Software Development life cycle** |
| **Data Mining** | **Predictive Analytics & Modeling** | **Data Visualization** |
| **Model Evaluation** | **Full stack Development** | **Application Performance Analysis** |

## EDUCATION

PG Diploma in Data Science | IIIT-Bangalore May-16 to April-17 | 3.73⁄4 CGPA

B.Tech in Information Technology I College of Engineering Roorkee, May -2006 to June-2010 | 71.06%

## PROFESSIONAL EXPERIENCE

**Senior Application Engineer | Oracle | April 2014 – Present**

**Application: Oracle's CRM Suite Development**

- Understand and developed end-to-end full stack product enhancements in Java (UI, SOAP).
- Brought SQL performance improvements of the product by analyzing complex SQLs, Memory Analyzer, Thread Analysis and optimizing the code.
- Work in an agile environment for product development.
- Work and collab with different teams across CRM for development activities and part of few home grown frameworks based on java.
- Trained in writing automated test scripts in Selenium for UI test cases and Junits for model tests.
- Self-coded tool in python (one such example) for easy to write auto upgrade scripts to upgrade complex Meta data files which otherwise would have taken longer time to develop. Multiple teams leveraged the tool across product, which, significantly reduce engineer script development time.

**Senior Systems Engineer | Infosys | March 2011 – March 2014**

Part of core business RBS logic team (banking domain), delivered bank application UI in java (Java front end) for 2 Years then move to Retail domain and worked on developing a social collaboration suite (Java full stack) for Retail giant Darden for 1 year.

**Systems Engineer Trainee | Infosys | Sep 2010 – March 2011**

Trained in Computer fundamentals, algorithm design, data structures with specialization in Java / Java Advanced J2EE with CGPA of 4.81/5

## KEY DATA SCIENCE PROJECTS (Github link provided for projects code and results)

**Objective:** Create a credit risk model to identify the right customers for a finance organization CredX
**Tech Stack**: R
**Solution:** Identified the factors affecting credit risk, create strategies to mitigate the acquisition risk, and came Out with the financial benefit of *489 Million dollars* using the predictive model

**Objective:** Predicted the likelihood of approval of credit card customer applications
**Tech Stack**: **R**
**Solution:** Created the predictive model with 78.29% accuracy by building and evaluating logistic regression model

**Objective**: Predict the particular telecom customer will churn or not.
**Tech Stack**: **R**
**Solution**: Predicted telecom customers likely to churn with *75.8%* accuracy by analyzing 7000+ customers' data Identified best model out of KNN, Naïve Bayes, Logistic, and SVM.

**Objective**: Forecasts the sales and demand of five most profitable market segments of a retail store
**Tech Stack**: **R**
**Solution**: Model prepared to predict the sales and demand

**Objective**: Risk assessment of customer loan.
**Tech Stack**: **R**
**Solution**: Determine the driving factors behind loan default. The company can utilize this Knowledge for its portfolio and risk assessment. Specifically, the company wants to determine which driver variables are having the most influence on the tendency of loan default.

**Objective**: A digital media company (similar to Voot, Hotstar, Netflix, etc.) had launched a show. Initially, the show got a good response, but then witnessed a decline in viewership. The company wants to figure out what went wrong
**Tech Stack**: **Python**
**Solution:** Identified critical parameters for lower viewership and prepared linear model to drive business to focus on parameters, which enable them for good viewership and generate revenue.

**Objective**: Real estate Company that has a dataset containing the prices of properties in the Delhi region. It wishes to use the data to optimize the sale prices of the properties based on important factors such as area, Bedrooms, parking, etc.
**Tech Stack**: **Python**
**Solution:** Prepared linear model to predict the price of the property and identified which factors contributes to the price most.

**Objective**: Analyze the Aadhar data (collected from UIDAI sample data sets) for various demographic Parameters. The analysis at hand involves basic data preparation, processing and understanding. Further, forecast the effects of certain information on the overall Aadhaar number generation

**Tech Stack**: SparkSQL, Hadoop

**Solution**: Analyzed the Aadhar dataset and collect the results using Spark SQL . The results can found Here.