**Name:** Manish Kumar Singh

**Contact Details:** 9930609020/8689815281

**Email ID –** [manishsingh7815@gmail.com](mailto:manishsingh7815@gmail.com)

**CAREER OBJECTIVE**

To establish a successful career in Data science. Well experienced in the extraction, analysis and presentation of data for credit and risk using **SQL, R**, **Python and Tableau** technologies.

**SUMMARY**

* Having good Analytical, Mathematical and Statistical skills.
* Proficient in R, Python, MySQL.
* Hands on experience on SQL, Python and libraries (NumPy, Pandas, SciKit learn, Matplotlib, SeaBorn, SciPy), R and packages, Tableau, TensorFlow and Keras.
* Created two Dashborad Data visualization using Tableau
* Actively participating in data science projects on Kaggle and Analytics Vidhya.

**STATISTICAL AND MACHINE LEARNING SKILLS**

* Hands on experience on Linear Regression, Logistic Regression, Decision Tree, Random forest, XGBoost and SVM. (Hands on experience is basically on projects).
* Hands on experience in unsupervised learning using K-means clustering and hierarchical clustering.
* Basic understanding of NLP and Neural Network.
* Knowledge and hands on experience on Deep learning NN, ANN, CNN, RNN, LSTM, Bidirectional Lstm.
* Data visualization in R using ggplot2 and plotly for interactive graphs and in Python using Matplotlib.
* Knowledge of Data visualization tool Tableau.
* Knowledge of Model Deployment on Heroku, Docker, AWS S3 using Sagemaker.
* Basic understanding of creating API through Flask.
* Feature engineering in R and Python, Missing value and outlier handling, transforming variable, creating new variables, Reshaping data using packages like dplyr and tidyr in R.
* Good understanding of ROC, AUC, KS, Accuracy and lift performance matrix.

**TECHNICAL SKILLS**

* Language : R, Python, SQL, Visual Basic
* Platforms : Windows, Linux (Ubuntu)
* RDBMS Tools : Oracle SQL Developer, MySQL
* Tools : R Studio, Anaconda (Jupyter), PyCharm, Visual Studio, Spyder
* Data Visualization Tools : Tableau
* Framework/ API : TensorFlow, Keras, Google Colab
* Quality management tools : HP Quality Centre
* Other Applications : Microsoft Outlook, Office, Word, MS Excel
* Cloud Platform : Heroku, Git, AWS, Docker

**DATA SCIENCE PROJECTS**

1. **Real Estate:** A real estate agency wanted to reduce the negotiation time and improve closure for buyers and sellers of homes by ensuring that both sides were advised well on the potential sale/purchase price of the home. To that end the agency wanted to predict a transaction price for all the houses in its market which would be as close as possible to a price where the transaction would take place. Data regarding all possible variables and qualities of over 10,000 houses sold in the past was collected and analyzed.

**Solution *-*** Created a predictive model using linear regression to arrive at a potential transaction price for all future transactions. Over the next 6 months, negotiation time was brought down from 3 weeks on average to 8 days. Deal closure went up by 28%.

1. **Banking:** A bank in Portugal was rolling out a new term deposits product for its existing customers. They wanted to understand which of their existing customers to call and target for this new product so that ROI would be high. In the past they had connected to their customer base through phone calls for various other products. Results for these previous campaigns had been made available for over 50,000 customers.

**Solution *-*** Used Logistic regression to create a propensity model to predict those customers most likely to respond positively to the new product and the campaign. Model had an accuracy of 80% and increased campaign ROI by 68%.

1. **HR:** A mid-sized IT company with over 10,000 employees needed to plan its hiring and reduce risk of projects getting delayed due to employees leaving. It also wanted to understand why its attrition rate was high and how it could be reduced.

**Solution *-*** Created a decision tree model along with a Random Forest model to predict attrition and also to better understand the factors related to attrition taking place. Random Forest provided greater accuracy. Communicated the findings to the client.

1. **Retail:** A retail giant needs to plan its new store openings over the next 12-18 months and has multiple locations where the stores can be opened. It wants to understand which locations would be the best to open new stores in terms of market size and potential revenues. Relevant data regarding locations, population, sales and revenues of key products for stores opened in the past was provided to be analyzed and for model building.

**Solution: *-*** Created a Random Forest and Logistic regression models to predict if a store should be opened or not in that particular location. Random Forest worked out to be a better model and 23 locations were selected to open new stores based on this analysis.

1. **Tableau project:** Created Dashboard for Sales analysis that which shows category and year from highest to lowest order in terms of vendor item price

Created Dashboard for customer analysis that shows Top N customers for the selected year, number of customers in each region, and customer order frequency

1. **NLP PROJECT:** Created NLP Project using NLTK tool and sentiment analysis on Fake news detection and spam-ham message.
2. Face detection and Object detection using transfer learning and pretrained model like MTCNN for object detection and Face detection VGG
3. **Titanic**: -Kaggle: Prediction on who will survive
4. **Loan Prediction: *-*Analytics Vidhya**: Prediction on if loan will be provided or not.
5. Prediction for a Portuguese bank’s customer’s eligibility for discounts using Random Forest and XGBoost

**Experience:**

NCR Corporation (On 3rd Party payroll)

Database Administrator

June 2018 - April 2020 (1 year 11 months)

Mumbai, Maharashtra, India

**Description** - Worked as Database Administrator at Client location for a Banking Client to provide support and resolve the queries raised and to smoothen the process flow. Data base design, finding data from the data base and making back up of cheque and data recovery of cheque, troubleshooting, software installation, Capacity planning, performance monitoring and Making account of officer to check the cheque and Update the cheque and check the cheque via SQL query and retrieve the data for account holder and many another bank.

Technologies/Tools used - SQL, MS Excel, Outlook, MS Word

**ACADEMIC QUALIFICATIONS**

|  |  |  |
| --- | --- | --- |
| **COURSE** | **BOARD / UNIVERSITY** | **PERCENTAGE** |
| Degree | Mumbai University | 6.48 |
| HSC | MAHARASHTRA | 74.5 |
| SSC | MAHARASHTRA | 80.9 |

**CERTIFICATIONS:**

* **Edvancers:**
  + ‘Certified Business Analytics Professional in R’
  + ‘Certified Data Analytics Professional in MySQL’
  + ’Certified Machine Learning with Python’
  + ‘Certified Tableau Professional’
  + ‘Certified Data Science Expert’

ACHEIVEMENTS:

* Won the Intra College Cricket Tournament 2015.
* Runner up for Intra College Volley Ball Tournament 2017

**EXTRA CURRICULAR ACTIVITES:**

* Sports Head of the College for 2015.
* Organized the Sports Fest, SCORE 2015&2016.

**INTERNSHIP:**

Organization: Century Rayon, B.K. Birla Group of Companies

Duration:13th June 2016 – 26th June 2016

Title: Power House and Instrumentation.

**PERSONAL PROFILE**

* **Gender :** Male
* **Date of Birth :** 7th August 1994
* **Permanente Add :** Plot No - 10, Line - D, Room No - 4, Shivaji Nagar, Govandi, Mumbai – 043
* **Languages :** English, Hindi, Marathi
* **Nationality :** Indian
* **Pan :** FDWPS5823P
* **LinkedIn URL :** <https://www.linkedin.com/in/manish-kumar-singh-2a931b15a/>
* **GitHub Repo URL :** <https://github.com/MANISH7815>

I hereby declare that the information given is best to my knowledge and belief.

Place: