Naveen Kumar G

**Mob: +91-8182083840**

**E-Mail**:[Naveen.gandhi68@gmail.com](mailto:Naveen.gandhi68@gmail.com)

**Summary**

* Data scientist with 4 years' experience in Machine Learning, SQL, Hadoop and Python. Involved in Python open-source community and passionate about deep Classification learning. a position that allows me to put my analytical, data-oriented skillset to use while also fostering my leadership and creativity.
* Exposure to all stages of CRISP (Cross Industry Standard Process for Data Mining) have depth Knowledge.
* I have grip on Data Extraction using Machine Learning, Deep Learning and I have in depth knowledge of transformation and Loading Data Wrangling and Data Exploration.
* Thoroughly aware about Machine Learning Algorithms like Well-versed in Advanced Machine Learning, Deep Learning, Big Data Processing Frameworks Effective understanding of Dimension Reduction like Principal component analysis.
* Deploy Predicting Models using Shallow models, Black Box technique, Ensemble models and Regression Models.
* Cloud Migration using AWS-Azure-Google Cloud
* Expert in Mathematics and Statistics
* Hands on Experience in Data Visualization tool like Tableau for effective vision for status of features.
* Communicating with Dev, QA, Delivery team related to build and deployment status and concerned issues

**Education:**

# ·B.E, (Visveswaraya Technological University, Belagavi) 2013.

# ·MTech (Visveswaraya Technological University, Belagavi) 2016.

**Certifications**

* Data science using python with R programming from 360DigitTMG, Hyderabad
* Data Visualization using Tableau programming from 360DigitTMG, Hyderabad
* IBM-Biginsights software Training by IBM Bengaluru
* Machine Learning A-Z Hands-On Python and R in Data science course Udemy
* Bigdata and Hadoop for Beginners with Hands-On from Udemy
* Bigdata Hadoop and Spark with Scala from Udemy
* Python A-Z for Data Science with Real from Udemy
* R-Programming A-Z for Data Science with Real Exercises from Udemy
* Data Science A-Z real Life Data Science Exercises Included from Udemy

**Tech Stack**

Pandas, Matplotlib, Tableau, NumPy, RStudio, Scala, Python, Flask, SQL, Spark, PyTorch, TensorFlow, Keras , Bigdata, Hadoop, Mango DB, AWS, Weka, Kaggle, COCO, GitHub.

**Projects**

### **1.Covid-19 Vaccine Sentiment Analysis**

**Description:** Media messages may not always align with science as the misinformation, baseless claims and rumors can spread quickly. This is what we saw with the introduction of the Covid-19 vaccine. In this data science project, we aim to analyze tweets recorded about the Covid-19 vaccine to analyze the sentiments of people for the vaccine.Twitter is a microblogging and social networking platform where users post and interact with messages called “tweets”. With more than 166 million daily users, Twitter is a valuable data source for any social media discussion related to national and global events. So, the dataset for the sentiment analysis task of the Covid-19 vaccine was collected from Twitter.

**Roles and Responsibilities:**

* I have strong knowledge in sentimental analysis algorithms and NLP and capable to solve increasingly complex problems
* Collected 100K tweet's reaction from various vaccinated peoples.
* Extracted valuable information and remove unnecessary data using various preprocessing techniques.
* Determine suitable machine learning algorithm to choose best analysis.
* Designed various time series algorithms to check continuous data like ARIMA.
* Use best visualization tools to represent data tableau is best visualization tool.

**2. Sentiment Analysis and Product Recommendation System in E-Commerce**

**Description:**

The Internet has revolutionized the way we buy products. In the retail e-commerce world of online marketplace, where experiencing products are not possible. Also, in today’s retail marketing world, there are so many new products are appearing every day. During their decision-making process, consumers want to find useful reviews as quickly as possible using rating system. Therefore, models able to predict the user rating from the text review are critically important.

**Roles and Responsibilities:**

* + Coordinated with Project Managers, Development Team and clients for better accuracy in customer review
  + Work with Engineering, Operations, Quality Assurance and Project Management to define and execute projects that implement build model and deployment of model.
  + Collected electronics dataset consists of reviews and product information from e-commerce website were collected. This dataset includes reviews (ratings, text, helpfulness votes) and product metadata (descriptions, category information,price, brand, and image features).
  + Analyse the Complete product review data with team and targeting features.
  + Evaluate the Collected data is in proper format or not if not make it correct format.
  + Extracted Metadata from data like Product description, category information, price,brand.
  + Created data frames same as data merging product reviews and metadata sets using data wrangling technique.
  + Evaluated duplicate and missing values.
  + Investigated descriptive statistics.
  + Extracted valuable data from raw data using preprocessing technique.
  + Removing HTML tags, Removing Accented Characters, expanding contractions, removing stop words, lemmatization, removing special characters.
  + Prepare Exploratory Data Analysis Based on Products.

**Achievements:**

National champion-2015” in "**Big Data and Hadoop**" in IIT Roorkee organized by cognizance.

"Certificate of EXCELLENCE” in big data and Hadoop national championship -2015 context IIT

Roorkee organized by Cognizance.

“Certificate of HONOUR” in Bigdata and Hadoop national championship -2015 context IIT

Roorkee organized by Cognizance.

"Certificate of Merit” in big data and Hadoop national championship -2015 context IIT

Roorkee organized by Cognizance.