**DATA SCIENTIST**

**Name T.SUMANTH**

**Phone No** +91 8328040575

**Email-Id** sumanth.datasci@gmail.com

## CAREER OBJECTIVE

To work with maximum potential in a challenging and dynamic environment, with an opportunity of working with diverse group of people and enhancing my professional skills with learning and experience for career growth.

## SKILLS SUMMARY

* Around 3 years of experience in software development, deploying and supporting large scale distributed systems.
* Primary technical skills in **Python** and **Machine Learning**.
* Expertise skill set in **Python** programming especially related to Functional and Object-Oriented Programming.
* Sound knowledge in Python related Data Structures, Functions, OOPS, Exception Handling, modules and Data base connectivity and etc.
* Hands on experience in Python **functional programming** like, pure function, higher order functions, decorators, generators, lambda functions and user defined functions.
* Expertise hands on experience in Python **OOPS principles** like class, objects, methods (instance, static & class), inheritance, polymorphism etc.
* Having experience on essential machine learning libraries **numpy**, **pandas**, **matplotlib**, **seaborn**, **scipy**, **scikit-learn**, **tensorflow** and **keras** etc.
* Understanding business objectives and developing models that help to achieve them, along with metrics to track their progress
* Analyzing the ML algorithms that could be used to solve a given problem and ranking them by their success probability
* Exploring and visualizing data to gain an understanding of it, then identifying differences in data distribution that could affect performance when deploying the model in the real world
* Experience in building and scaling models for data and building predictive maintenance algorithms.

## TECHNICAL SKILLS

|  |  |
| --- | --- |
| **Programming Languages** | Python 3.6 |
| **ML libraries** | Pandas, Numpy, Matplotlib, Seaborn, Scipy,  Scikit-learn and Tensor flow & keras, Keras-OCR, Opencv, ImageAI |
| **Data Science platform** | Anaconda Distribution |
| **Notebook** | Jupyter |
| **Database** | Mysql |

## PROFESSIONAL EXPERIENCE

* Aug ’18 – Till date – Working as a Data Scientist at TCS Banglore

## EDUCATIONAL QUALIFICATIONS

* B.Tech from Jawaharlal Nehru Technological University, Kakinada, in the year of 2018

## PROJECT DETAILS

**Project # 1**

**Project Name :** Drug Recognition

**Role :** Data Scientist

**Technical Environment :** Python, numpy, pandas, matplotlib, Keras-OCR,

ImageAI and Tensorflow.

* Drug Recognition is a Smart Application in order to recognize various Drugs.
* Here we need to recognize various Drugs in the form of Pills, Sheets and in the various forms by using some specific characteristics.
* Here the main aim of this project is, we need to check the Prescription and given medicines are prescribed by the doctor or not, in order to avoid the wrong Medicine Delivery.
* We need to capture the Image of Drugs and we need to Check it out, for this we are using Different Image processing and deep learning techniques
* This aims to give accurate medicine delivery and avoiding wrong medicine Delivery to the patient, it helps and increasing marketing and accuracy.

**Roles & Responsibilities**

* Understanding business objectives and developing models that help to achieve them, along with metrics to track their progress
* Analyzing the **Image processing** techniques and **Text Analysis** that could be used to solve a given problem and ranking them by their success probability
* Using **python libraries(pandas, numpy, matplotlib)** to extract the data into the working environment
* Filtering the image by using various noise detection techniques.
* Various Objects detection Using IamgeAI and Keras-OCR for text processing on the image.
* Creating automated anomaly Drug recognition system and validating with the Prescription.

**Project # 2**

**Project Name :** Alpha Fintech

**Role :** Data Scientist

**Technical Environment :** Python, numpy, pandas, matplotlib, seaborn,

HDFS, scipy, scikit-learn and tensorflow.

* Customer segmentation is a process of where we divide the customers into several groups of individuals that share a similarity in different ways that are relevant to marketing such as gender, age, interests, and miscellaneous spending habits.
* We need to generate the subgroups by using some specific characteristics so that the company sells more products with less marketing expenditure. Generally, we need to categorize our consumer base into subgroups. These subgroups are called segments.
* We need to create the groups in such a way that each subgroup of customers has some shared characteristics. Using clustering techniques, we can identify the several segments of customers allowing them to target the potential user base.
* The technique of customer segmentation is dependent on several key differentiators that divide customers into groups to be targeted. Data related to demographics, geography, economic status as well as behavioral patterns play a crucial role in determining.
* This aims to simplify and connect the entire payments world. It’s an automated machine reasoning platform that translates data from enterprise systems into meaningful insights predictions in the form of narratives without any manual intervention.
* The main features are state of the art Python, Python Libraries, Machine learning algorithms, Predictive modeling, Natural Language Processing Algorithms customized based on analysis types, Narrative Generation Automated narratives for presenting results in the form of a report in readable format, Expert Rule System a complex rule-based system that emulates/mimic the decision-making and reasoning ability of a human being, Deep Learning Using advanced machine learning techniques to enable unsupervised learning’s to improve the accuracy of the analysis.

**Roles & Responsibilities**

* Understanding business objectives and developing models that help to achieve them, along with metrics to track their progress
* Analyzing the **Machine Learning** algorithms and **Predictive Modelling** that could be used to solve a given problem and ranking them by their success probability
* Using **python libraries(pandas, numpy, matplotlib)** to extract the data into the working environment
* Done **Exploratory Data Analysis** to extract insights from the data.
* For Description variable, extracting numerical features used **NLP Techniques(Tokenize, Bag of words, Lemmatization ,TF-IDF)**
* Selecting features, building and optimizing classifiers using **machine learning** techniques.
* Verifying data quality, and/or ensuring it via data cleaning.
* For segmentation the customer used **K-means clustering algorithm**.
* Supervising the data acquisition process if more data is needed
* Finding available datasets online that could be used for training by using **scikit-learn package**.
* For building the model to predict the customers used **Logistic Regression and Decision Tree Algorithms.**
* Defining validation strategies
* Defining the pre-processing or feature engineering to be done on a given dataset
* Training models and tuning their **hyper parameters**
* Analyzing the errors of the model and designing strategies to overcome them
* Better accuracy used **Gradient Descent and Regularization Techniques**(Lasso and Ridge)
* Deploying models to production
* Exploring and visualizing data to gain an understanding of it, then identifying differences in data distribution that could affect performance when deploying the model in the real world
* Enhancing data collection procedures to include information that is relevant for building analytic systems
* Processing, cleansing, and verifying the integrity of data used for analysis
* Doing ad-hoc analysis and presenting results in a clear manner
* Creating automated anomaly detection systems and constant tracking of its performance

## DECLARATION

I am also confident of my ability to work in a team. I solemnly declare that the information furnished above is correct and true to the best of my knowledge.

**Bangalore**

**(T. Sumanth)**