**Data Scientist**

**Name:** Swarna Latha

**Mail id:** swarnalathav.data@gmail.com

**Mobile:** +91 8639042117

**Skills Summary:**

* Around 2 years of experience in software development, deploying and supporting large scale distributed systems.
* Currently working as a **Data Scientist** using **python**.
* Primary technical skills in **Python**, **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) 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 Details:**

|  |  |
| --- | --- |
| **Programming Languages** | Python Programming language |
| **ML libraries** | Pandas, numpy, matplotlib, seaborn, scipy,  scikit-learn and tensorflow |
| **Data Science platform** | Anaconda Distribution |
| **Notebook** | Jupyter |
| **Database** | Oracle, MySql |
| **Defect tracking tool** | JIRA |

**Professional Experience:**

* Aug’19 – till date Working as a Software Engineer at TCS, Banglore

**Qualification:**

* B.Tech from Gudlavalleru Engineering College, Gudlavalleru.

**Projects Details:**

**Project # 1**

**Project Name :** Payment processing

**Role :** Data Scientist

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

scipy, scikit-learn and tensorflow.

Payment solution helps users to analyze and create data stories from large volumes of data that could be available from various sources and channels. This aims to simplify and connect the entire payments world. Creating a dynamic, evolving solution and providing tools for pro-active collaboration amongst all parties. While disrupting and liberating us from today’s static and restrictive payments technology and business models. 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 learnings 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 Modeling** 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

**Name**: Swarna Latha

**Location**: Banglore