Dheeraj R P

Senior Software Engineer, Bengaluru
dheerajrp66@gmail.com
+91-7259931932

Professional Summary:

- *Python*, *PySpark developer*, *Data Science professional with* technical hands-on development expertise of 2 years and 8 months.
- Worked for an open source library in python, which is used by users all over the world on the daily basis.
- Worked on various projects in the field of machine learning, deep learning and natural language processing.
- Closely worked with data scientists to know the process of real time industry project scenarios.
- Experienced in creating blockchain-based applications using slither, a static smart contract analyzing tool.

TECHNICAL SKILLS:

- **Programming Languages:** Python, ReactJs(Beginner)
- **Big Data Technologies:** Hadoop, Apache Spark, Apache Sqoop, Apache Airflow
- **Visualization Tool:** Tableau
- **Database Tools:** MySql, PostgreSQL
- **Artificial Intelligence:** Machine Learning, Deep Learning, Natural Language Processing
- **Operating Systems:** Windows, Linux
- **AWS:** AWS EMR, AWS S3
- **Deployment Tools:** Heroku, Dockers
- **API Frameworks:** Flask, FastAPI

WORK EXPERIENCE:

Delixus Software India pvt Ltd, Bengaluru

February 2020 - PRESENT

❖ **DATA ENGINEERING**

➢ Built the platform on *Spark* to get the data from *AWS S3* from different vendors. Then the data is decrypted, transformed, cleaned and pushed to a dataframe for further analysis purposes.

➢ For the transformations of the data, all the data from the vendors is mapped to the required fields and pushed to the *Spark dataframes*.

- ➢ For the cleaning of the data, different UDFs are used in order to validate the data and put the data in the standard format, hence maintaining the consistency in the data.
- ➢ Performed the aggressive data analysis, identified bad data patterns and produced a unique key for a similar record of data coming from various vendors.
- ➢ All the results are then loaded to *MySQL* tables which were consumed by the elastic search.
- ➢ Utilized *AWS EMR* for deployment and testing for the entire process.

❖ **PYTHON DEVELOPMENT**

- ➢ Contributed to the open source python library - pyTenable which is an interface into Tenable's platform APIs.
- ➢ Developed web application backend component using *python-flask* for analyzing solidity smart contracts using the APIs of static solidity analyzing tool slither, further pushed to Postgres database.
- ➢ Created an application like JARVIS, a virtual assistant that automates some small works that we usually do on the computer.
- ➢ Created an application for generating passwords for use in different services to minimize the time to create a password for the user.
- ➢ Performed analysis on slack message data to provide yearly, monthly, and daily report metrics for the messages and teams by utilizing *PySpark*

❖ **MACHINE LEARNING**

- ➢ **HEART ATTACK RISK PREDICTION:**
  - ■ Created a web tool that analyses the chances of heart diseases based on several parameters.
  - ■ Performed cleaning and exploratory data analysis to get the clear picture of the data using pandas.
  - ■ Trained the data using algorithms like logistic regression, support vector classifiers, random forest classifier, decision trees, XGBoost and gradient boosting algorithms from which random forest model gave the highest accuracy.
  - ■ Created a supervised learning model and used the model with highest accuracy for the predictions.
  - ■ Performed hyperparameter tuning to boost the accuracy of the model.
  - ■ Using flask, exposed the APIs for different functionalities.
  - ■ Deployed to the Heroku cloud where we can actually predict the chances of heart diseases.

➢ **WATER QUALITY PREDICTION:**

- Created web tool that predicts the water quality based on several parameters.
- Performed cleaning and exploratory data analysis to get the clear picture of the data using pandas.
- Trained the data using algorithms like decision trees, logistic regression, K-nearest neighbors, support vector classifiers from which support vector classifiers gave the highest accuracy.
- Created a supervised learning model and used the model with highest accuracy for the predictions.
- Performed hyperparameter tuning to boost the accuracy of the model.
- Used FastAPI, for exposing the APIs for different functionalities.
- Deployed in Heroku cloud where we can actually predict the water quality.

➢ **SENSOR FAULT DETECTION (IN PROGRESS):**

- Tech stacks used: *Python, FastAPI, Machine Learning Algorithms, Dockers, Apache Kafka, MongoDB*
- Infrastructures required: *AWS S3, AWS EC2, AWS ECR, Terraform, GitHub Actions.*

➢ **RESTAURANT RATING PREDICTION (IN PROGRESS)**

- An application that helps restaurants in Bengaluru to know their rating from the customers given in the zomato.

❖ **DEEP LEARNING:**

➢ **SPAM-HAM CLASSIFIER - SENTIMENT ANALYSIS:**

- Created a web tool that predicts whether the given message is spam or not spam(ham).
- Performed cleaning and exploratory data analysis to get the clear picture of the data using NLTK libraries.
- Trained the data using algorithms like logistic regression, naive bayes and long short term memory (LSTM) which is a deep learning approach.
- Created a supervised learning model and used the model with highest accuracy for the predictions.
- Used Flask for exposing the APIs for different functionalities.

EDUCATION:

C-DAC, ADVANCED COMPUTING TRAINING SCHOOL, Bengaluru

Post-Graduation Diploma in Big Data Analytics: Aug 2019 - Jan 2020

- **Coursework:** Machine Learning, Big Data Technologies, Data Visualization - Analysis and Reporting, Data Collection & DBMS, Object Oriented Programming with Java8, Statistical Analysis with R, Python, and Linux Programming
- **Project:** NCERT based Question Answering chatbot.
    - Considered a physics text text book of class 12 prescribed by NCERT for building the chatbot.
    - Further classified every physics terminology with respect to their patterns and the context.
    - Trained the model with an artificial neural network.
    - Used tkinter for creating a frontend for the chatbot.

University of Mysore, Mysuru

Master of Science in Statistics: 2017-2019

- **Project:**
    - A Study on Awareness about Government Schemes among the students of University of Mysore, Mysuru.
    - Modeling and forecasting of Exports of Onion from India.

PROFESSIONAL CERTIFICATION:

- Accredited for the following Tech Neuron courses offered by iNeuron.ai.
    - NLP Foundations
    - FastAPI
- Udemy
    - Python Programming for Data Science
- Coursera
    - Data scientist toolboxelastic search.