SOWJANYA BOJJA

Data Scientist

Email:simplysowj@gmail.com Tel: +91 8073226086 **Location**: Hyderabad, India

Linkedin:https://www.linkedin.com/in/sowjanya-bojja/   Github:https://github.com/simplysowj

Summary:

Passionate and detail-oriented data science graduate seeking an entry- level position to apply analytical skills and contribute to data-driven decision-making. Eager to leverage academic knowledge and practical experience in a professional environment.

Skill Set:

Languages: Python, SQL,JAVA,C,HTML,CSS,JavaScript,Pl Sql database programming language
Statistical Analysis:Matlab,Excel

Libraries/Framework: NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, NLTK,ReactJS,Bootstrap,Streamlit Python library,FLASK.

Tools: Anaconda, Jupyter notebook, VS code, Colab,SQL Server,Toad,GIT,IntelliJ IDE,POSTGRESQL,JAVA Swing .

ML: Data Exploration & Analysis, Data Modelling, Statistics and Probability, Linear Regression, Gradient Descent, Logistic Regression, Regularization, SVM, KNN, Decision Trees, Random Forest, Ensemble Techniques, Bagging & Boosting, Cross-Validation, Cluster Analysis,Tensorflow,SKlearn,Hyper parameter tuning,Experiment tracking and model management using MLflow.

MLOPS,Workflow orchestration with Prefect 2.4

Cloud Deployment Heroku GitHub

NLP: Tokenization, Bag of words, Stemming, Lemmatization, POS Tagging, TFIDF,BERT,word2vec,glove.
Deep learning:ANN,CNN,GAN,RNN

Visualization: Excel,Advanced Excel,Tableau,Matplotlib,Seaborn.

CERTIFICATIONS:

Certified IBM full stack data scientist from Innomatics Research labs(one of the top growing Edu-Tech Company in Hyderabad)

Certificate of training in SQL,Tableau,Advanced Excel in Internshala Training.

Work experience:

July 2022 - November 2022
Data Science Intern in Innomatics Research Labs | Hyderabad, India
Projects:
Quora Duplicate question detector (capstone project)
Tasks
Developed application using Streamlit
Applied machine learning algorithms
Experiment Tracking and Model Management using MLFLOW
Different vectorization techniques including bert
Workflow orchestration with Prefect

Natural Language Processing
Cloud Deployment Heroku GitHub

Projects:

1)NLP project using Streamlit for sentimental analysis on the IMDB Movie reviews tweets dataset, here in this project I used vectorization technique like #bert to convert text data into numerical vectors 2) covid - 19 twitter sentiment analysis with streamlit application using python.

3) MLOps task-Diamond Dataset ML pipeline orchestration and experiment tracking

4) Experiment Tracking and Model Management for Diamond price prediction using MLFlow

5) streamlit app for Diamonds dataset EDA and ML Model

6) Machine learning project - Regression Problem

   Title: Concrete compressive strength

7) Python Flask CRUD Web Application project

8) Portfolio Application Project

9) Python Web Scraping Project with EDA

 10) Open Pub Application using Streamlit python

Education Project:

IPL Data Analysis using SQL(Postgre sql) and Data visualization using Excel.
Analysis of Fitbit fitness tracker data to deliver marketing and business solutions to WeFit and its subsidiaries

Minor project:

Exam seating arrangement system using JAVA with GUI Swing.

Industry Experience:Plsql Developer

       Tools: TOAD 7.6.0.11, PL/SQL Developer, Forms 61, Reports 6i Company TCS

      Client Electronic Arts-India Location Bangalore, India (2010 to 2014) (Nov 2010
           May2014)

      Client Tata TeleServices Limited - India (2008 to 2010) Mar 2008-Oct 2010

EDUCATION

Bachelor of Technology (B.Tech.) | Electronics and communications Engineering

                      JNTU, Hyderabad, TG

Masters in Data Science in Chandigarh University

Have done certification in Innomatics Research institute,Hyderabad

ACCOMPLISHMENTS

I have completed Certification and Internship in Data science in one of the top growing Edu-Tech Company and received a Letter of Recommendation for my best performance