



Rupesh Suryawanshi

Data Scientist

A punctual and hardworking student, very interested in learning new stuff and research further to unknown information. Looking forward to adapt and learn from training in the new environment, also meeting new friends and to work together with them as a team.

Contact

Phone

9172293641

Email

rsurya12.rs@gmail.com

LinkedIn

<https://linkedin.com/in/rupesh-suryawanshi-a17263260/>

GitHub

<https://github.com/surya1255>

Address

C-802, Sai Paradise,
Punawale, Pune.

Education

2022

Bachelor of Engineering - GPA 9.21
D.Y. Patil Collage of Engineering, Pune

2019

Diploma in Engineering - 76 %
J.T. Mahajan Collage of Engineering,
Jalgaon

Data Visualization

- PowerBi
- Tableau
- Seaborn
- Matplotlib

Experience

Aug 2022 - Feb 2023

Ai Variant I Bengaluru

Data Science Project Intern

- Building predictive and decision models using machine learning algorithms for various business study cases with optimizations and evaluation techniques.
- Analyzed large and imbalanced datasets by performing machine learning techniques to provide data insights, optimistic solutions and strategic directions for making data driven decisions.
- Built Machine Learning Models in python Worked on Various Python Libraries Including NumPy, Matplotlib, Pandas, Seaborn, Scikit-Learn, NLTK and many Other.
- Worked with text data with text mining techniques including text extraction, cleaning, pattern identifications, categorizations, and entity recognitions

Projects

Bankruptcy Prediction

- The objective of the analysis is to predict that the company goes bankruptcy or not considering the different features of the company
- Analyzed and understood each of technical details related to business case, done EDA and model using python. Explained the data insights, outcomes, and experimental findings with teams.
- Used different classification algorithms, finally we have achieved 100% test accuracy and 100% train accuracy in Random Forest classifier after tuning the hyper parameters.
- Presented overall model interpretations, reports to senior lead and deployed the model using Streamlit.

Hotel Review Classification

- The objective of the analysis is to predict that the customer has been satisfied or not and the rating based on review. Model gives Positive if the expected rating is ≥ 3 out of 5 or else Negative.
- The data here is text so preprocessing like removing punctuations, numbers, extra white spaces is done before performing EDA.
- Analyzed the data with several EDA methods and tf-idf vectorization of text is done to perform the model building

ML Algorithms

- Linear Regression
- Logistic Regression
- Decision Tree
- Random Forest
- SVM
- Naïve Bayes
- KNN
- Clustering

Professional Skills

- Teamwork
- Active Listening
- Time Management
- Problem Solving
- Fast learner
- Adaptive

Laanguages

- English
- Hindi
- Marathi

- Several classification techniques are used to classify the ratings based on the review and finalized logistic regression model i.e having train accuracy of 94.7% and test accuracy of 92.6%
- Presented overall model incorporations, reports to senior lead and deployed the model using streamlit.

Air Quality forecasting(CO2 emissions)

- The objective of the analysis is to forecast Co2 levels for an organization so that the organization can follow government norms with respect to Co2 emission levels.
- Analyzed the data with several EDA methods and made data stationary using differencing and moving average methods.
- Used several data driven and model driven algorithms to see the result. Got the best accuracy using Arima model.
- Presented overall model incorporations, reports to senior lead and deployed the model using streamlit.

Certifications

Data Science Professional

ExcelR | Pune

Data Science Internship

Ai Variant | Bengaluru

Programming Languages

Python

- Basic knowledge of Data type and variable, functions, and Looping statements
- ML libraries like Scikit-learn, Scipy, Pandas, Numpy etc.
- NLP libraries like NLTK, Spacy, TextBlob.

MySQL

- Good understanding of concepts like Constraints, Tables.
- Excellent knowledge of writing SQL queries
- Good understanding of SQL concept like Grouping, sub-Queries,
- Good knowledge of joins, statements.