Shaktiprasad Rudrawar

Data Scientist

My Contact

rudrawarshaktiprasad@gmail.com

+91 86689 47986

Sahakar Nagar , Pune, MH

Hard Skill

- **Language -** Python, Java, MYSQL.
- **Library / Framework -** Scikit Learn , Pandas , Numpy , Flask & Fast API.
- **IDE** - Jupyter Notebook, VS Code/Pycharm, etc.
- **Visualization Tool** -Matplotlib , Seaborn and Power BI.
- **Machine Learning** - Statistics ,EDA, Feature Engineering and Feature Importance , Feature selection, Model building and Evaluation.
- **Deep Learning -** Statistics , Neural Networks, Data Modeling and Evaluation.
- **NLP -** Text Processing and Word Embedding. Lib- NLTK.
- **Manual & Automation Testing** -System Testing, Functional and Regression Testing, SDLC, BDD creation. **Python Automation in Playwright** Etc.
- **API Testing.**
- **Database Testing.**
- **Agile Methodology and Jira Tool use for tracking.**
- **Metrics used -** R square ,Adj R square, MAE, MAPE ,MSE, RMSE, Accuracy, Recall , Precision, F1 Score etc.

About Me

As a detail-oriented Data Scientist with **3 .1 years** of experience. Now looking for an opportunity to grow and find more exposure simultaneously contributing with Python language and Hands-on experience in building and running Machine Learning models.

Professional Experience

2] Machine Learning Project : - Siemens Technology and Services Private Limited.

March 2023 – Present

- **Key Responsibilities and Achievements:**
- **Predictive Modeling:** Built models to predict the net price and discount percentage for low-voltage electrical products in the manufacturing domain.
- **Tools & Libraries:** Leveraged VS Code, Jupyter Notebook, Jira, Confluence, Scikit-Learn, Pandas, NumPy, Matplotlib, Seaborn, and FastAPI for end-to-end development.
- **Data Analysis:** Conducted statistical analysis to extract insights and understand data patterns.
- **Feature Engineering:** Created new features to enhance model accuracy, focusing on feature importance and selection techniques.
- **Machine Learning Model Development:** Implemented a Decision Tree Regressor for prediction and designed K-Fold Cross-Validation to evaluate the model's reliability.
- **Model Optimization:** Architected hyperparameter tuning using Grid Search and Random Search to improve performance.

- **Project:- "Predictive Analytics for Top 10 Product Categories by Revenue and Volume".**
- **Exploratory Data Analysis (EDA):** Performed univariate and multivariate analysis using Pandas, identifying outliers and feature skewness through Seaborn and Matplotlib.
- **Data Preprocessing:** Applied techniques like imputation, transformation, feature selection, and hypothesis testing with statistical tests.
- **Model Building & Tuning:** Built models and optimized them using Grid Search CV and Random Search CV for hyperparameter tuning.

Education Background

- **Python and Data Science - 2021**
  Gamaka AI Institute , Pune.
- **BE - CIVIL- 2011**
  National Education Management and Technology Studies, Pune
- **Diploma - CIVIL- 2008 -**
  SES Polytechnic - Solpaur.
- **Notice period - 15 days.**
- **Language - English ,Hindi, Marathi and Kannada.**

Hyperlink -
LinkedIn profile:
Shaktiprasad Rudrawar - LinkedIn

GitHub profile:
Shaktigit23 - GitHub

1] Machine Learning Project : - Mindchips Consulting.

Jan 2022 - Feb 2023

- **Key Responsibilities and Achievements:**
- **Goal:** Classified patient conditions using drug reviews with natural language processing (NLP).
- **Tools & Libraries:** Worked with Jupyter Notebook, PyCharm, Scikit-Learn, NLTK, Matplotlib, and Regex.
- **Machine Learning Models:** Built classification models using Naive Bayes and TF-IDF Vectorizer, delivering insights into patient drug conditions.

- **Project : - "Disease Condition Prediction Using Machine Learning in the Healthcare Domain"**
- **Objective:** To classify patient conditions based on drug reviews using advanced machine learning and natural language processing (NLP) techniques.
- **Key Contributions:**
- Performed data preprocessing tasks including lemmatization, stemming, and stop-word removal using the **NLTK** library.
- Designed and implemented **Bag of Words (BoW)** and **TF-IDF Vectorization** to analyze drug reviews.
- Built and optimized machine learning models such as **Naive Bayes** for classification.
- Conducted extensive **Exploratory Data Analysis (EDA)** to understand patterns and prepare data pipelines.
- **Impact:** Enhanced healthcare decision-making by enabling accurate classification of patient conditions, contributing to improved treatment recommendations.