DNYANESHWAR PATIL
☎ +91 7768881766 📌 Pune, Maharashtra
Email: dnyaneshwarspatil1@gmail.com **GitHub:** https://github.com/dnyaneshwar-ai **Linkdin:** dnyaneshwarpatil09074a249

OBJECTIVE

data scientist with four months of experience as an associate data scientist seeking a full-time role in the data field To leverage my expertise in data science, machine learning, and programming to develop innovative solutions that drive business growth.

EDUCATION AND CERTIFICATION

Bachelor Of Technology from DBATU University Maharashtra with overall 9.01 CGPA	2017 - 2021
Certificate course in full stack data science, Hyderabad	2022
Certificate course in Python, Machine Learning,Excel, Deep Learning, SQL from CloudlyML	2023

SKILLS

- **Languages, libraries, and models:** Python, NumPy, Pandas, Matplotlib, Seaborn, TensorFlow, Keras, SciPy, NLTK, OpenCV, Sklearn, Spacy, XGBoost
- **Data analytics**: MySQL, Kaggle, MS Excel, AWS S3, AWS Athena, AWS Glue, and AWS Lambda.
- **Mathematics and Statistics:** Descriptive and Inferential Statistics, Hypothesis Testing, Linear Algebra, Probability, Conditional Probability, Exploratory Data Analysis (EDA), Data Cleaning, Data Wrangling, PCA, Feature Extraction and Selection, Dimension Reduction, etc.
- **Machine Learning:** Machine Learning**,** Regression Analysis Algorithms, Logistic regression, Decision tree, Random Forest, Ensemble modelling, Principal component analysis (PCA), Cluster Analysis (Clustering), SVM and Kernel SVM, Time series analysis
- **Deep Learning:** Artificial Neural Networks (ANN), Convolution Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM).
- **NLP:** text classification, sentiment analysis, text summarization, name entity recognition, language translation, natural language processing.

EXPERIENCE

CLOUDMANTRA, PUNE, INDIA **JAN 2023 TO JUN 2023**

- Analyse large amounts of data using open-source libraries such as NumPy, Matplotlib, Pandas, and Seaborn.
- Experience of developing Regression and Classification machine learning models using different regression, classification, clustering, deep leaarning algorithems.
- experience using the Data Analytics Platform on AWS using serverless architecture, including S3, Glue, Lambda, and Athena.
- Proficient in using popular object detection libraries like YOLO, OpenCV, and CNN for identifying objects and refining proposals with high accuracy.
- Proficient in various natural language processing tasks like text classification, sentiment analysis, named entity recognition, language translation, text summarization, and text generation.

PROJECTS

1. Employee Attrition Problem:
Problem Statement: To predict the salary of an employee based on the information provided in the dataset. because an employee from a company is leaving due to various reasons. We have to find out the reason and its solution.
Skills and Tools used: Python, Pandas, Matplotlib, NumPy, Seaborn, Sklearn, XGBoost, and statmodels
Responsibilities:

- Used all supervised machine learning algorithms
- Perform EDA and pre-processing to understand the data.
- Used multi-collinearity analysis to understand highly correlated variables and handle them.
- Visualisation is done by Matplotlib and Seaborn by plotting all possible plots.
- Determine the important features using lasso regression, random forest, AdaBoost, and XGBoost, feature important techniques, and apply all Metrics techniques for Regression.

2.Olympics Data (Analysis):
Description: "Perform a comprehensive analysis of Olympic Games data to uncover insights, trends, and patterns, with the objective of understanding the performance and impact of athletes, countries, and sports disciplines over time."
Skills and Tools used: Python, Pandas, Matplotlib, Seaborn, NumPy, Jupyter Notebook, and Tableau

3. Chronic Kidney Disease Classification:

Description: It's a healthcare domain project with real-time applications in healthcare sectors,according to the National Library of Medicine. 9 out of 10 people don't even know that they are suffering from an early chronic disease. In this model, if you give some information about a person, it will predict whether they have chronic Kidney Disease or not with high accuracy.

Skills and Tools used: Python, Numpy, Pandas, Matplotlib, Seaborn, Sklearn, and XGBoost

Responsibilities:

- Clean and transform the data to do the required analysis.
- Perform EDA and pre-processing techniques.
- Plot graphs and charts to understand data using Matplotlib and Seaborn. Convert unbalanced data into balanced data.
- Used supervised machine learning techniques with all metrics parameters
- Bagging and boosting algorithms for best performance, Perform AUC and ROC analysis of the model build.
- Determine the important features using lasso regression, random forest, AdaBoost, and XGBoost.

4. Financial News Headline Sentiment Analysis:

Description: Develop a machine learning model to analyse the sentiment of financial news headlines and classify them as positive, negative, or neutral, aiming to provide investors and traders with real-time insights into market sentiment.

Skills and Tools used: Python, NLP,NLTK, machine learning, text preprocessing, feature engineering, and evaluation metrics

Responsibilities:

- Scraped from CNBC, the Guardian, and Reuter's official websites, the headlines in the dataset reflect an overview of the U.S. economy and Stock market every day for the past year to two years.
- Perform tokanizations, stemming, lemmatizations, stop words, and text preprocessing.
- used the NLTK library to perform NLP operations like feature engineering.
- Perform all vectorization techniques to convert text data into vectors, including sentence editing.
- Used all performance parameters, like the confusion matrix, etc., to check the performance of the model.
- Used machine learning algorithms to classify the result.