SWAPNALI DILIP KADAM
Ph: +91 8329218866 | **E-mail**: swapnalikadam19002@gmail.com
LinkedIn: https://www.linkedin.com/in/swapnali-kadam-38005a286/
Github: https://github.com/swapnali3112 Pune ,Maharashtra

Career Objective: Passionate Data Scientist with proven skills in Python, machine learning, and statistical analysis. Seeking to leverage analytical expertise and hands-on experience to provide actionable insights and drive data-driven decision-making in a dynamic environment.

Professional Experience:

1.Subject Matter Expert – Statistics | Chegg (Freelance)| Feb 2022 – Mar 2024

- Solved 1,000+ statistical problems with a 95% approval rating.
- Expertise in **Hypothesis testing, Probability, Regression analysis**, and **UAT** for educational tools.
- Applied **Python, NumPy, pandas, seaborn** for data analysis and visualization.

2.Data Analyst Trainee | MedTourEasy- Dec 2023

Project: "Analyze Death Age Difference of Right Handers with Left Handers.
Leveraged **Python, R, SQL, Power Bi** for data insights and interactive dashboards.

Key Skills

- **Programming**: Python , SQL
- **Data Tools**:NumPy, Pandas, Power BI, MS Excel
- **Machine Learning**: Scikit-learn, TensorFlow, Keras, XGBoost ,Deep Learning , AutoML
- **Data Visualization**: Matplotlib, Seaborn
- **Cloud**: AWS, AWS SageMaker
- **Big Data Tools**: Apache Spark, Hadoop
- **Databases:** MySQL, MongoDB
- **Frameworks**: Django, Flask
- **Deployment & CI/CD:** Docker, Kubernetes, Jenkins
- **Natural Language Processing (NLP):** spaCy, NLTK, Transformers
- **Soft Skills**: Problem Solving, Critical Thinking, Communication, Adaptability
- **Other**: Agile Methodologies, Git

Projects:

1. Customer Churn Prediction with Machine Learning

Tools/Tech: Python, Pandas, Scikit-learn, XGBoost, AWS SageMaker

Description: In this project, you build a machine learning model to predict customer churn for a subscription-based company. The project includes:

- **Data preprocessing:** Handling missing data, outliers, and feature engineering.
- **Model building:** Using classification algorithms (Random Forest, XGBoost) to predict churn.
- **Evaluation:** Model evaluation using confusion matrix, AUC-ROC, and cross-validation.
- **Deployment:** Deploy the model using AWS SageMaker to enable real-time predictions.

2. NLP Sentiment Analysis for Product Reviews

Tools/Tech: Python, NLTK, spaCy, TensorFlow, Flask

Description: This project involves building a Natural Language Processing (NLP) model to analyze customer reviews and categorize them into positive, neutral, or negative sentiments. Steps include:

- **Text preprocessing:** Tokenization, lemmatization, and vectorization using spaCy.

- **Model building:** Using TensorFlow for building a neural network-based sentiment classifier.
- **Web deployment:** Deploy the model with a Flask app that takes real-time user input for sentiment analysis.

3. Time Series Forecasting for Sales Prediction
Tools/Tech: Python, Pandas, ARIMA, LSTM, Matplotlib
Description: This project involves analyzing historical sales data to forecast future sales. The key aspects include:

- **Data preparation:** Cleaning, differencing, and scaling the time series data.
- **Model building:** Using ARIMA and LSTM models to capture both linear and non-linear trends.
- **Visualization:** Present the results using Matplotlib and Seaborn for trend analysis.
- **Optimization:** Fine-tuning the model to improve forecast accuracy.

4. Real-Time Fraud Detection System
Tools/Tech: Python, Apache Spark, Hadoop, Scikit-learn
Description: In this project, you create a real-time fraud detection system for financial transactions. The project involves:

- **Data ingestion:** Using Apache Spark to process real-time transaction data.
- **Feature engineering:** Creating features like transaction patterns, frequency, and location.
- **Model building:** Training a machine learning model (Random Forest, XGBoost) to detect anomalies.
- **Scalable deployment:** Using Hadoop to handle big data and ensure scalability

---

Education

B.Sc. in Statistics
Shivaji University, Kolhapur | *2015 – 2018*

- **Marks**: 77.44%

Certifications

- **Google Data Analytics Professional Certificate** | Coursera | Jan 2024
- **Data Science with Python** | Simplilearn | Nov 2023
- **Introduction to SQL** | Simplilearn | Dec 2023
- **Python for Beginners** | Simplilearn | Oct 2023

Hobbies
Reading | Drawing | Trekking | Yoga | Music