

# Sandesh Shankrod

Data Scientist

[\\*\\*\\*\\*\\*24@gmail.com](mailto:*****24@gmail.com) | +91 77\*\*\*\*\*

[GitHub](#) | [LinkedIn](#)

## About Me

---

- Passionate about data science with a strong foundation in leveraging Python and SQL to analyze and derive meaningful insights from Data.
- Committed to learning and applying data science techniques to solve real-world problems.
- Skilled in working with various data tools and platforms to manage and interpret data effectively.

## Technical Skills

---

- **Artificial Intelligence (AI):** Data Analytics, Data management, Database, Predictive modelling, Machine Learning, Natural language processing (NLP).
- **Algorithms :** Linear Regression, Logistic Regression, Naive Bayes, KNN, SVM, Decision Tree, Random forest, Ensemble models, Time Series Analysis, Classification, Clustering, GPT.
- **Generative AI (Gen AI):** Lang Chain, Gemini, LLM.
- **API Keys:** Google Gemini, Meta3, Ollama.
- **Libraries:** Pandas, NumPy, NLTK, Scikit-learn.
- **Data Visualization:** Seaborn and Matplotlib, PowerBI.
- **Data Collection and Storage:** SQL, MySQL.
- **Mathematical Skills:** Probability and Statistics.
- **Development:** Git
- **Deployment:** Streamlit.
- **Platforms:** Anaconda, Jupyter Notebook, Spyder IDE, Visual Studio, Python, Windows.

## Experience

---

### 1. The Baap Company | Intern Data Analyst

#### Responsibilities and contribution:

- Performed Exploratory Data Analysis (EDA) to identify patterns and insights in data, aiding in better decision-making.
- Extracting and transforming data from the ERP system for analysis.
- Performing data analysis to identify trends, patterns, and insights.
- Creating reports and dashboards to visualize key metrics and performance indicators.
- Assisting in improving the data quality and integrity within the ERP system.
- Providing data-driven recommendations to enhance business processes and strategies.
- Creating customized reports based on specific business requirements and user needs.
- Supported new team members in understanding project fundamentals and workflows.

**Key Achievement:** Teamwork with quality & timely project completion.

## 2. Naresh IT | Intern Data Scientist

### Responsibilities and contribution:

- Performed Exploratory Data Analysis (EDA) to identify patterns and insights in data, aiding in better decision-making.
- Executed Variance Inflation Factor (VIF) analysis to detect multi-collinearity among features, leading to improved feature selection and enhanced model performance.
- Updated model features to reduce overfitting, enhancing the model's generalization capabilities.
- Transformed data splitting methods to minimize data leakage and improve model accuracy.
- Contributed to model retraining efforts, resulting in improved overall market growth predictions.
- Assisted in identifying and implementing filters and hyper-parameter tuning, achieving better model performance.
- Assisted in identifying and implementing filters and hyper-parameter tuning, achieving better model performance.
- Delivered comprehensive documentation of the entire pipeline workflow.
- Supported new team members in understanding project fundamentals and workflows.

**Key Achievement:** Teamwork with quality & timely project completion.

## Projects / Open-Source

---

- **Gen Ai Projects | [Link](#)**
  - Engineered an AI-powered application specializing in food photo analysis for estimating calorie content, promoting healthier dietary habits.
  - Designed and implemented an AI bot capable of interpreting invoices across various languages, extracting detailed invoice information seamlessly.
  - Developed a sophisticated generative AI tool focused on PDF enhancement, incorporating analysis, text and image embedding, and automated annotations for comprehensive document processing.
- **House Prediction with 13 Algorithms | Machine Learning | [Link](#)**

AIM: Predict house prices using 13 different machine learning algorithms to identify the most accurate model for real estate valuation.
- **Customer Retention | [Link](#)**

AIM: Develop strategies to enhance customer loyalty and reduce churn through data-driven insights and predictive modelling.
- **Diabetes Prediction Detector | [Link](#)**

AIM: Develop a predictive model using machine learning to assess individual diabetes risk, empowering proactive healthcare decisions and personalized patient management.

## Certification

---

- Full Stack Data Scientist | Naresh IT
- Data Analyst | The Baap Company

## Education

---

**University:** Savitrabai Phule Pune University (SPPU)

July 2018- July 2023

**College:** Smt. Kashibai Navle College of Engineering, Pune (SKNCOE)

**Branch:** Computer Science (Bachelor of Engineering)