


# SUSHIL CHAVAN

 9309158031

 [sushilchavan2468@gmail.com](mailto:sushilchavan2468@gmail.com)

 [LinkedIn](#)

 [GitHub](#)

 [Portfolio](#)

---

**AI/ML Engineer** with experience building **predictive** and **machine learning models** across NLP, tabular, and **deep learning** domains to solve real-world problems. Proficient in **Python, Scikit-learn, TensorFlow, and SQL** with strong foundation in statistics, **forecasting**, and **model evaluation on datasets up to 40K+ records**. Experienced in developing **end-to-end ML pipelines** and translating data into actionable insights.

---

## EXPERIENCE

### Data Analyst Intern

**Gamaka Intellectual Machines Pune (July 2025 – Dec 2025)**

- Developed **interactive Power BI dashboards** and optimized data models for **KPI tracking** and business reporting.
  - Automated **data preprocessing** and feature preparation workflows using **Python** and **SQL** on structured datasets.
  - Performed **exploratory data analysis** and **validation** to support reliable analytics and decision-making.
- 

## SKILLS

- **Programming & Data:** Python, SQL, Pandas, NumPy, Statistics, EDA, Data Cleaning, Feature Engineering
  - **Machine Learning & AI:** Scikit-learn, TensorFlow, NLP, Time Series Forecasting
  - **Visualization & BI:** Power BI, Tableau, Matplotlib, Seaborn, Excel
- 

## PROJECTS

### 1. LexiMind — Legal Document Intelligence System

 [GitHub](#)

- Built **NLP-based** classification model achieving **~92% accuracy** for legal judgment categorization on **17K+ documents**.
- Generated contextual **embeddings using Sentence Transformers** and implemented **cosine similarity**-based semantic search.
- Designed vector retrieval pipeline to **return Top-5 relevant legal cases** for efficient reference discovery.

### 2. Bank Term Deposit Subscription Prediction

 [GitHub](#)

- Developed **ML models** on **41K+** banking records to **predict customer subscription** (11% minority class).
- Addressed **class imbalance using weighted models** and evaluated performance using **Precision, Recall, and F1-score**.
- Compared **Logistic Regression, KNN, Naive Bayes, and Random Forest** to identify key influencing features.

### 3. Face Mask Detection System (CNN)

 [GitHub](#)

- Developed a **CNN-based computer vision model using TensorFlow** to classify images into three categories: mask worn correctly, mask worn incorrectly, and without mask.
- Implemented **image preprocessing and resizing (128×128)** using Python and OpenCV.
- Evaluated model performance using **validation accuracy and confusion matrix** and built a prediction pipeline for testing on new images.

### 4. Data Analyst Agent (LLM + Automated EDA)

 [GitHub](#)

- Developed an **LLM-powered assistant** for automated EDA and natural language data queries..
  - Supported **CSV, Excel, PDF, and Word** inputs through integrated processing pipelines.
  - Enabled faster insights using **conversational AI with statistical analysis workflows**.
- 

## EDUCATION

**G H Raison College of Engineering and Management, Jalgaon**

2022 - 2026

*B-Tech in Artificial Intelligence & Machine Learning*