

# SHEKHAR CHORMALE

Phone: +91 9325493919 | Email: chormaleshekhar@gmail.com | [LinkedIn](#) | [GitHub](#) | Pune

## EDUCATION

**Choice College of Pune** **Oct 2021 – Present**  
Computer Application | Bachelor of Computer Application

**Edyoda Digital University, Bangalore** **Oct 2021 – April 2022**  
6-Month Certification | Data Science  
• coursework: Python, Data Analytics, Data Science, Machine learning, NLP, PowerBi, MySQL, Excel

## WORK EXPERIENCE

**Machine Learning Engineer | Ai Adventures Pune** **Jan 2022 – Present**

- Developed and implemented machine learning models using scikit-learn for diverse projects, such as classification and regression.
- Conducted thorough data analysis and preprocessing using scikit-learn's tools to ensure the quality and reliability of input datasets.
- Collaborated with cross-functional teams to integrate scikit-learn pipelines into the software architecture, streamlining the development and deployment process.
- Optimized model performance through hyperparameter tuning, feature engineering, and evaluation metrics.
- Conducted workshops and training sessions to educate team members on best practices for utilizing scikit-learn in machine learning projects.
- Maintained awareness of scikit-learn updates and incorporated new features to enhance the efficiency and effectiveness of machine learning solutions.

## SKILLS & CERTIFICATIONS

- **Programming Languages** - Python, MySQL
- **Tools** - PowerBi, Excel, MySQL Workbench
- **Certifications** – Python, Data Analysis, MySQL, Data Scientist, Machine Learning, NLP
- **Framework** – Django, Streamlit

## PROJECTS

### Fake News Detection

- Project details: With the rise of social media and online news sources, fake news has become a significant issue that can have serious consequences on people's beliefs and actions. This project aims to develop a deep learning model using the Long Short-Term Memory (LSTM) library to detect fake news. **Technology Stack:** Python, LSTM library (Keras, TensorFlow), Streamlit

### Campus Recruitment Prediction

Data Preprocessing:

- Handle missing values, by removing them.
- Encode categorical variables into numerical representations

Exploratory Data Analysis (EDA):

- the dataset to gain insights into the distribution of features.
- Visualize relationships between different variables.
- Identify patterns and correlations that may influence placement outcomes.

Feature Selection:

- Determine which features are most relevant for predicting placement outcomes.
- Select features based on their importance and contribution to the model.

Model Selection:

- Choose a suitable machine learning algorithm for classification, such as Decision Trees, Random Forest
- Split the dataset into training and testing sets for model evaluation

## **Wine QT**

### Tools and Libraries Used:

- Python
- Pandas for data manipulation
- Seaborn and Matplotlib for data visualization
- Scikit-learn for machine learning algorithms
- Grid SearchCV for hyperparameter tuning

### Data Exploration and Visualization:

- Explored the dataset to understand the distribution of features.
- Visualized relationships between quality and individual features using bar plots.

### Data Preprocessing:

- Created a binary classification for the response variable 'quality,' dividing wines into 'good' and 'bad.'
- Applied Label Encoding to convert the categorical labels into numerical format.
- Checked for class distribution in the target variable.

### Model Building:

#### Random Forest Classifier:

- Created a Random Forest Classifier with 200 estimators.
- Achieved an accuracy of 87% on the test set.
- Evaluated model performance using classification report and confusion matrix.

#### Stochastic Gradient Descent Classifier:

- Implemented Stochastic Gradient Descent Classifier.
- Achieved an accuracy of 84% on the test set.
- Assessed model performance using classification report and confusion matrix.

#### Support Vector Classifier (SVC):

- Utilized Support Vector Classifier for classification.
- Obtained an accuracy of 86% on the test set.
- Evaluated model performance using classification report and confusion matrix.

### Model Improvement:

#### Grid Search CV:

- Conducted a grid search to find the best hyperparameters for the Support Vector Classifier.
- Identified optimal parameters:  $C=1.2$ ,  $\gamma=0.8$ ,  $\text{kernel}='rbf'$ .
- Rebuilt the SVC model with the optimized parameters.
- Achieved an improved accuracy of 93% on the test set.

#### Cross-Validation:

- Performed cross-validation on the Random Forest Classifier, resulting in an average accuracy of 89% over 10 folds.

---

## **AWARDS & ACHIEVEMENTS**

### Hacker Rank

- problem solving rank under 20000 in hacker rank

### Achieve ML Engineering Scholarship Program

- Select Amazon machine learning program in this program achieve silver badge