SMS Spam Detector - README

# 📱 Project Description

A machine learning project that classifies SMS messages as spam or ham (not spam) using natural language processing techniques and a trained classifier.

# 🚀 Features

* • Text preprocessing (stop word removal, stemming, vectorization)
* • Model training and evaluation
* • Real-time message prediction via command-line or UI (if included)
* • Streamlit web app (if applicable)
* • Easy-to-follow structure

# 📂 Project Structure

sms-spam-detector/  
│  
├── data/  
│ └── spam.csv # Dataset file  
│  
├── models/  
│ └── spam\_classifier.pkl # Trained ML model  
│  
├── notebooks/  
│ └── EDA\_and\_Modeling.ipynb # Jupyter Notebook for training  
│  
├── app.py # Streamlit or CLI app script  
├── preprocessing.py # Text cleaning functions  
├── requirements.txt # Python dependencies  
└── README.md # Project documentation

# ⚙️ Setup Instructions

1. 📥 Clone the repository

git clone https://github.com/your-username/sms-spam-detector.git  
cd sms-spam-detector

2. 🐍 Create a virtual environment (optional but recommended)

python -m venv venv  
source venv/bin/activate # For Linux/Mac  
venv\Scripts\activate # For Windows

3. 📦 Install dependencies

pip install -r requirements.txt

4. ▶️ Run the application

If using a Streamlit app:  
streamlit run app.py  
  
If using command-line:  
python app.py

# 📊 Dataset

The dataset used is the classic SMS Spam Collection Dataset from UCI Machine Learning Repository:  
Link to dataset: https://archive.ics.uci.edu/ml/datasets/sms+spam+collection

# 🧠 Algorithms Used

* • CountVectorizer or TfidfVectorizer
* • Naive Bayes or Logistic Regression (based on your model)
* • Accuracy, Precision, Recall, F1 Score for evaluation

# ✅ Example Usage

You can enter a message like:  
  
"Congratulations! You've won a free ticket to Bahamas!"  
  
And the model will predict:  
  
\*\*Label: Spam\*\*

# 📌 Future Improvements

* • Deploy using Flask/Streamlit on Heroku or Render
* • Use deep learning (e.g., LSTM)
* • Add dataset upload option