# Ticket Classification and Entity Extraction.README

This project automatically classifies customer support tickets by issue type and urgency level. It also extracts important entities like product names, dates, and complaint keywords from the ticket text. A simple web app lets you input multiple tickets and get predictions and extracted info instantly.

## Features

1. Text cleaning and preprocessing
2. Feature extraction using TF-IDF, ticket length, and sentiment
3. Logistic Regression models to classify issue type and urgency
4. Simple entity extraction (products, dates, keywords)
5. Interactive Gradio web app for batch ticket processing.

## Requirements

Python 3.7 or higher

Required Python packages (install with pip):

* pandas
* numpy
* nltk
* scikit-learn
* textblob
* matplotlib
* seaborn
* gradio
* openpyxl # for reading Excel files

## Setup Instructions

1. **Clone or download this repository** to your local machine.
2. **Install the required packages** by running:

“pip install pandas numpy nltk scikit-learn textblob matplotlib seaborn gradio openpyxl”

1. **Download necessary NLTK data** (this is done automatically in the code, but you can also do it manually):

“import nltk

nltk.download('punkt')

nltk.download('stopwords')

nltk.download('wordnet')”

1. **Place your ticket dataset Excel file** at the specified path or update the file path in the code:

“df = preprocess\_dataframe('/Users/luvgambhir/Documents/ai\_dev\_assignment\_tickets\_complex\_1000.xls')”

“Change this to your file path if different.”

## How to Run

Run the Python script with:

“python your\_script\_name.py”

**The script will:**

* Load and preprocess the data
* Train models for issue type and urgency prediction
* Show distribution plots and evaluation results
* Launch a Gradio web app

## Using the Gradio App

After running the script, a local URL will be shown in the terminal, e.g.:

“Running on local URL: [http://127.0.0.1:7860/”](http://127.0.0.1:7860/\”)

Open this URL in your browser.

In the app:

* Enter multiple ticket texts, one per line.
* Click **Submit**.
* View the predicted issue type, urgency level, and extracted entities for each ticket in a table.