**Real-Time Audio Intent Classification**

## **Objective**

The goal of this project is to **capture real-time voice input**, **convert speech to text**, and **classify the intent of the spoken text** using an **open-source NLP model**. This allows us to understand **user intent in real-time** and categorize their requests into predefined categories such as **Shopping, Customer Support, Technical Issues, Billing, and more**.

## **How It Works**

### **1. Capture User's Voice Input**

* The system **records live audio** until the user is **silent for 2 seconds**.
* We use sounddevice to **capture** and NumPy to **process** the audio.
* The recorded audio is **plotted** for verification.

### **2. Convert Speech to Text (STT)**

* We use **Hugging Face Whisper** (openai/whisper-small) to transcribe speech.
* The recorded audio is **directly processed** without saving it as a file.
* The model returns the **spoken text**.

### **3. Classify Intent**

* The **transcribed text** is passed to an **intent classification model** (Serj/intent-classifier).
* We use **Few-Shot Learning** to improve accuracy **without fine-tuning**.
* The model **returns the intent** that best matches the user’s request.

### **4. Display Results**

* The **spoken text** and **detected intent** are printed for the user.
* The system provides a **real-time** classification of user intent.

## **Key Features**

* **Real-time voice input** with automatic silence detection
* **Speech-to-Text conversion** using Hugging Face Whisper
* **Intent classification** with a **pre-trained NLP model**
* **Few-Shot Learning for better accuracy**
* **Works without any fine-tuning or additional training**

## **Example Workflow**

### **User Scenario 1: Online Shopping Query**

**User Speaks:** *"I want to buy a new smartphone. Can you help me choose?"*

**Transcribed Text:** *"I want to buy a new smartphone. Can you help me choose?"*

**Detected Intent:** **Shopping**

### **User Scenario 2: Technical Issue**

**User Speaks:** *"My WiFi is not working properly. The connection drops frequently."*

**Transcribed Text:** *"My WiFi is not working properly. The connection drops frequently."*

**Detected Intent:** **Technical Issue**

### **User Scenario 3: Refund Request**

**User Speaks:** *"I received a wrong item in my order. I need a refund."*

**Transcribed Text:** *"I received a wrong item in my order. I need a refund."*

**Detected Intent:** **Returns & Refunds**

## **How to Run the Project?**

1. **Install Dependencies**

pip install sounddevice numpy librosa matplotlib soundfile transformers torch

2. **Run the script**

python script.py

3. **Speak into the microphone**

* The system will listen until **you stop speaking for 2 seconds**.
* It will transcribe your speech into text.
* It will classify your **intent** and display the results.

## **Final Summary**

✅ **This project enables real-time intent classification from voice input.** ✅ **We use an open-source pipeline (Whisper + Intent Model) to process speech.** ✅ **The system can be expanded for customer service bots, e-commerce platforms, and more.**