

嘿，YouTube，在此视频中，我将向您展示如何使用Python中的免费开源软件包快速将任何音频转换为文本，称为Whisper。我将显示我安装了它，展示了如何运行它的示例，并将其与现有库进行比较。因此，首先，您可能想去我们在这里查看的Whisper [Get](#) [Get](#) [Get](#) [Got](#) Hub存储库，并提供有关如何安装它的说明。现在，当您只需安装名字窃窃私语时，请记住一件事，它不会安装正确的版本。我们想从此GIT存储库中安装。因此，只需以此PIP安装命令并在您正在运行Python的环境中运行。他们在这里还提到您需要安装FFMPEG。有一些说明可以做，但是我已经在计算机上安装了。现在我已经有了耳语安装，让我们做一些可以对此进行测试的音频。所以我要说一些习语。成语通常很难让模型理解。即使这只是对文本的讲话。这会很有趣。我很想成为云9，这不会伤害苍蝇。我就像一条水从水中出来，在天气下像小提琴一样合身。让我们保存下来。让我们将其保存为波浪。他们确实有说明说明我们如何在命令行安装后直接从命令行运行。我要向您展示如何使用它们在这里显示的Python API。所以这真的很简单。我们只是进口耳语。然后，我们将创建我们的模型，这就是我们要加载。称为基础的模型。然后仅使用此模型对象，我们在音频文件上运行转录。所以我将其命名为习语。让我们使用Wave版本。我们希望这返回结果。现在，我注意到以前运行时，由于Kuda的半张量和浮动张量，我会遇到此错误。我能够解决这个问题。因此，这是要牢记的。如果它对您不起作用，则可能需要将浮点16设置为跌落。而且您可以在这里运行后看到它已经将语言检测到英语，然后这个结果对象中有几种不同的方法，但是我们想进入的内容只是文本，我们可以看到它看起来是外观就像结果是好的一样，我很想在云上九点钟，这不会伤害苍蝇，我会像一条水中的鱼一样，这确实使这条鱼从水中弄得一点点就像小提琴一样适合，也许我没有说出足够清楚的另一件事要知道的是，当您第一次运行时，它将必须下载基本模型。因此，您可能会看到一个进度栏正在贯穿，您必须下载该模型。它说，当您运行此抄录时，实际上它将占用30秒的音频文件并在其上进行预测。现在，您还可以采用另一种方法，这是一种较低级别的方法，实际上创建了模型，然后创建音频对象并进行模式修剪。因此，您只需确保此音频块仅为30秒即可。秒或它将由30秒拍拍它，因为这是该模型期望具有输入的长度。然后，它正在制作日志鼠标频谱图。它正在检测语言，我们可以在此处解码，并提供更多的选择。如果我运行此单元格，请再次获得此错误，现在我可以在解码选项中设置该错误，FP16等于故障。实际上，这一次看起来一切正确。我会像一条水从水中出来。并适合小提琴。就是这样。我只想将其与现有类型的模型进行比较。这样做的流行库是语音识别库。我们运行语音识别库的方式是导入它，然后

创建此识别器对象，然后我们可以将其加载到我们的音频文件。之后，您可以采用识别器对象，并且有几种不同的识别方法。我们将使用Google认可，让我们看看结果是什么。因此，看起来它没有添加任何标点符号，而云九则不同。我很想在云上九，作为一个不会伤害苍蝇的小马。但是要记住的一件事是，这实际上是在使用Google语音识别API。耳语库，您实际上下载了模型，这是您的使用。我还建议您查看使用此代码发布的耳语纸。他们还详细介绍了模型的训练以及使用的架构。耳语确实使用了许多不同语言。他们说的性能根据语言而变化。因此，您可以在GitHub回购库上进行此处，其中有一个图，显示哪种语言实际上对这里的酒吧表现最佳。较小的更好，较大意味着其性能更糟。因此，该模型使用的语言数量仍然令人印象深刻。