CHAPTER 2

Understanding and Building an Application with STT and TTS

This chapter introduces you to the concepts of speech-to-text (STT) and text-to-speech (TTS). It discusses various STT engines, and you build a Python program that records audio. You then graduate to an application that converts whatever you say to text. You also look at the use of various TTS engines and implement them to make a program that repeats whatever you say.

Speech-to-Text Engines

As you saw in Chapter 1, the STT engine is one of the three main components of the virtual assistant, Melissa. This component is the entry point for the software's control flow. Hence, you need to incorporate this piece of code into the `main.py` file. First, you need a sophisticated STT engine to use for Melissa. Let's look at the various STTs available on the Web for free use with your application

Freely Available STTs

Some of the best STTs available on the Internet are as follows:

- *Google STT* is the STT system developed by Google. You may already have used the Google STT if you have an Android smartphone, because it is used in Google Now. It has one of the best recognition rates. But it can only transcribe a limited amount of speech per day (API limitation) and needs an active Internet connection to work.

T. Pant, *Building a Virtual Assistant for Raspberry Pi*, DOI 10.1007/978-1-4842-2167-9_2