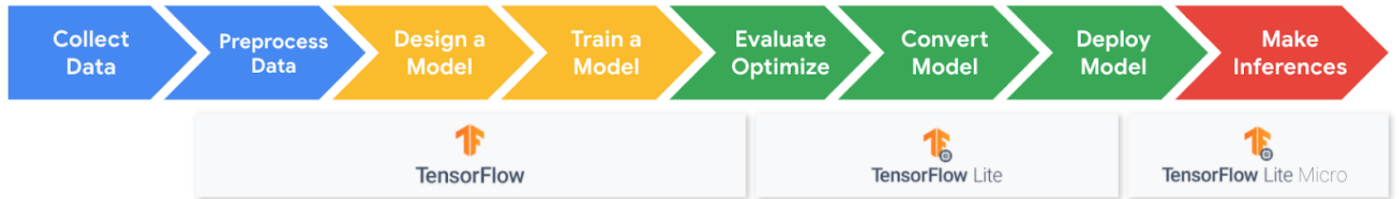


TinyML Application Pipeline

ML Workflow / Pipeline

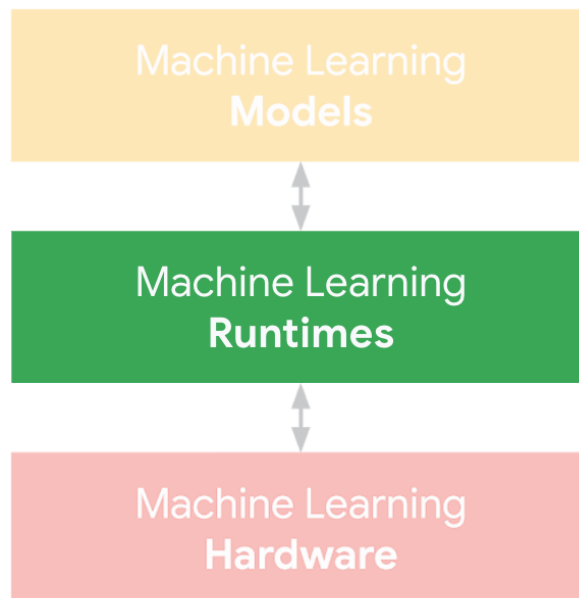


The figure above shows the workflow methodology we will be following for building the tinyML applications in this course. These are broken down into three main stages:

- Step 1: Collect & Preprocess Data
- Step 2: Design and Train a Model
- Step 3: Evaluate, Optimize, Convert and Deploy Model

What's Underneath the Hood?

Often, when learning about machine learning, we forget that models are just a piece of the bigger picture. There is more to machine learning than just the models. These are the underlying ML runtimes and hardware that enable these models to run efficiently. The ML runtimes we will be dealing with are TensorFlow, TensorFlow Lite, and TensorFlow Lite Micro.



One of the goals of our course is to dive into understanding the fundamental differences between machine learning runtimes so that we can deploy the models we prepare efficiently. To this end, it is important to understand the various frameworks we use across all three applications in this course: Keyword Spotting, Visual Wake Works, and Anomaly Detection.

So coming up next is material that highlights the **fundamental differences between TensorFlow and TensorFlow Lite (for microcontrollers)**. You will learn not only the high-order bits but also the mechanics that implement the API calls you learn to make with the code.