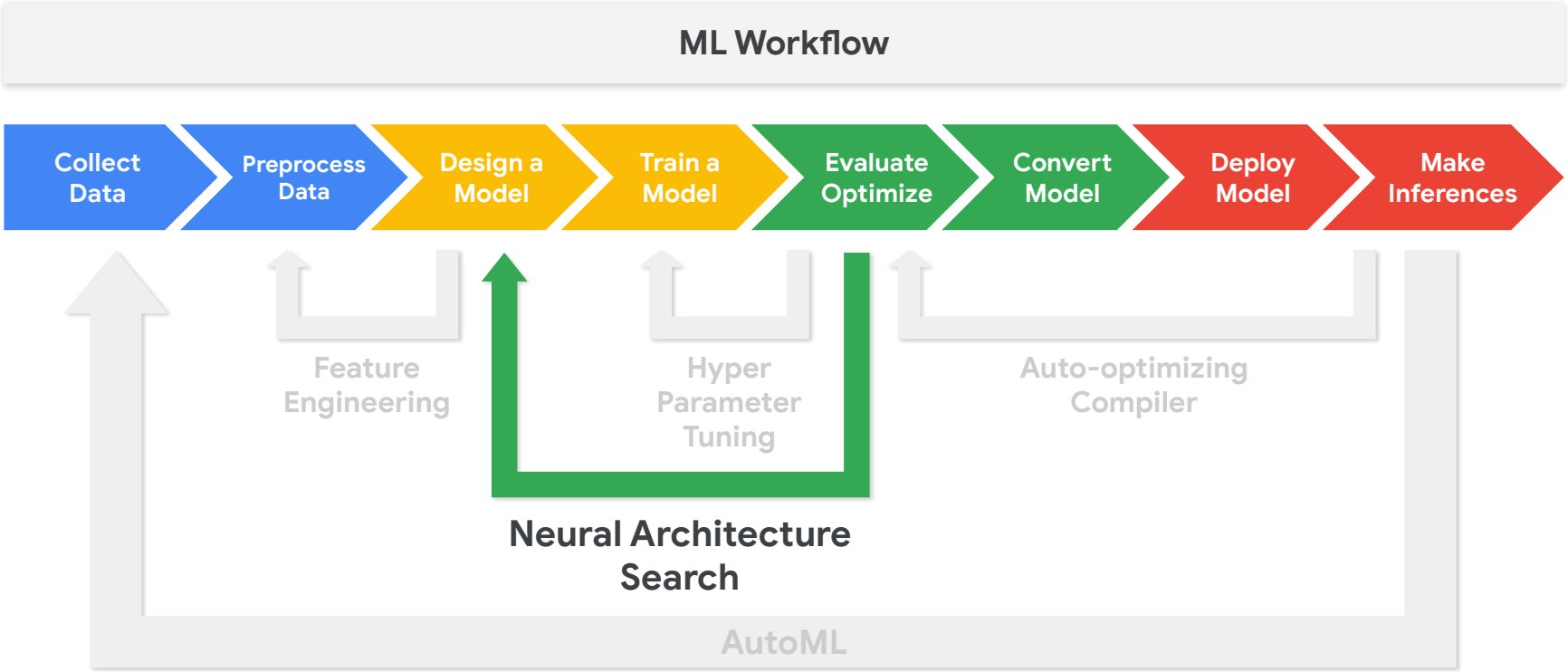
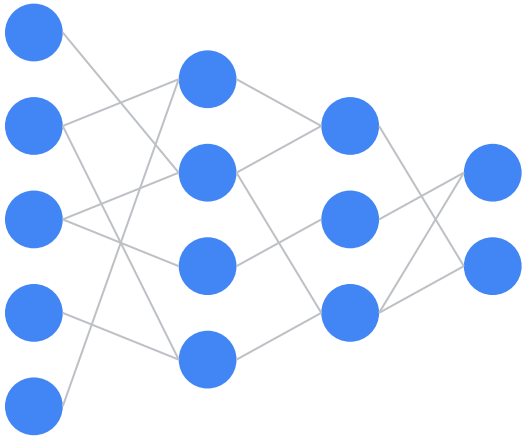
Overview of Neural Architecture Search

What is Neural Architecture Search?

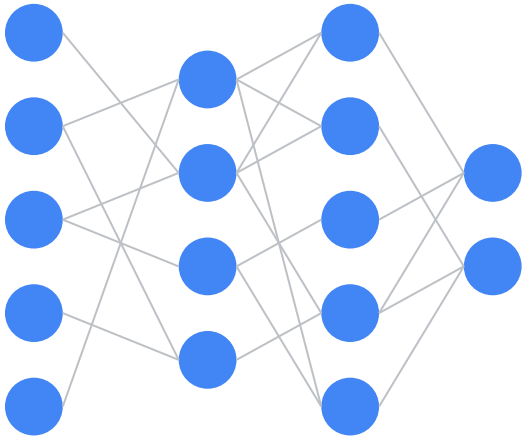


Model **Design** Impacts Accuracy



Accuracy: **84%**

Model **Design** Impacts Accuracy



Accuracy: **90%**

TinyML?

Even less memory

Even less compute power

Also, only focused on *inference*

*model constraints on **memory**, and **latency***

Multi Objective Search

Vanilla NAS

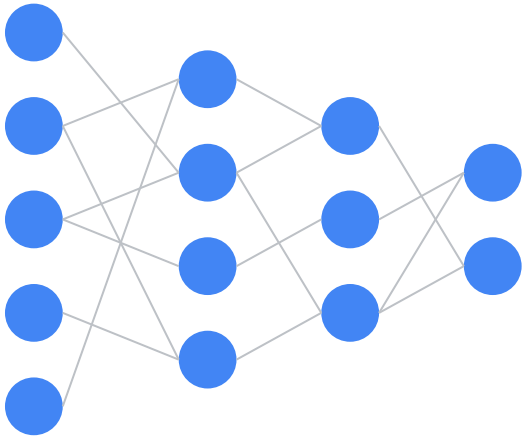
Multi-Objective

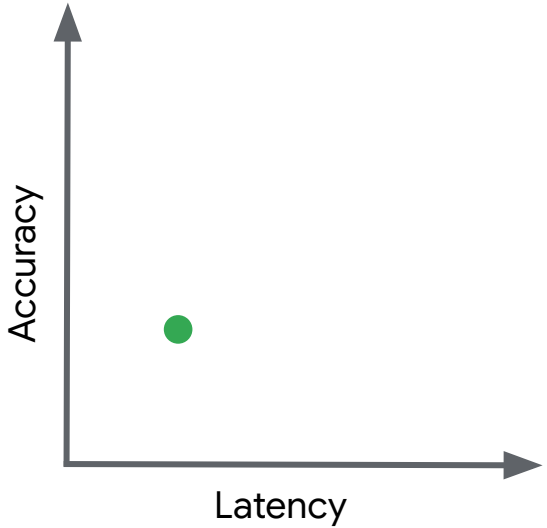
Accuracy

Latency

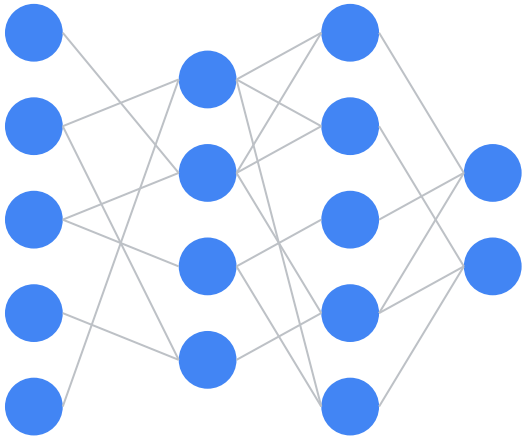
Memory

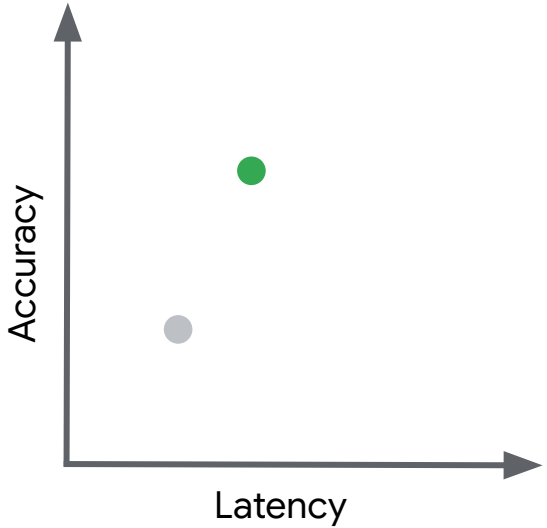
Design impacts **more than** accuracy



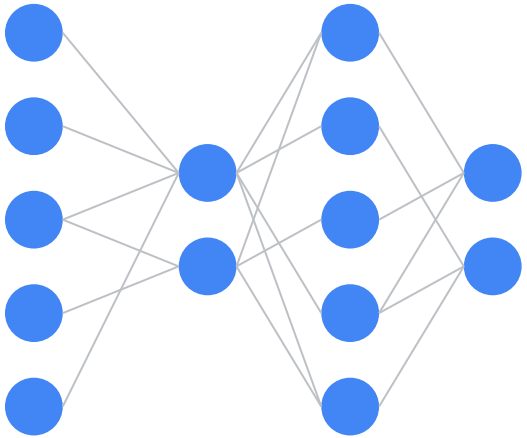


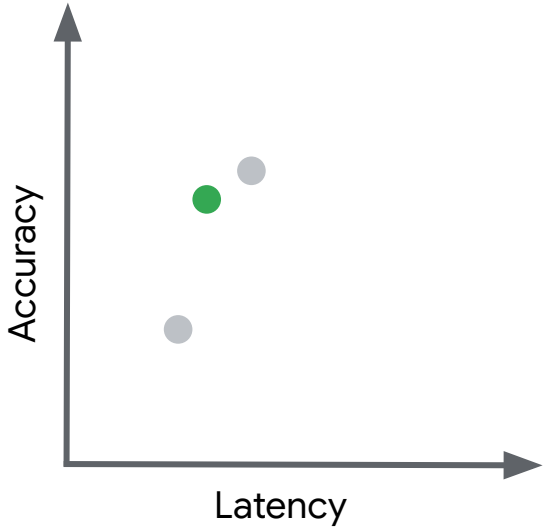
Design impacts **more than** accuracy



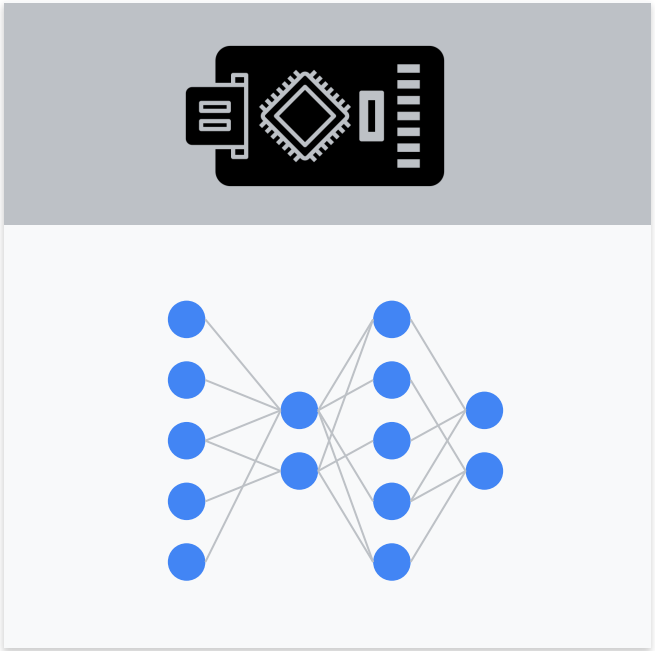


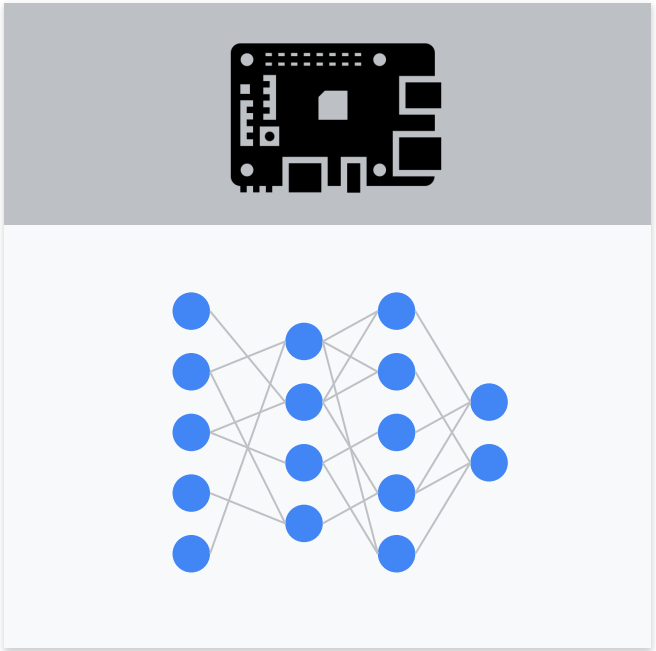
Design impacts **more than** accuracy



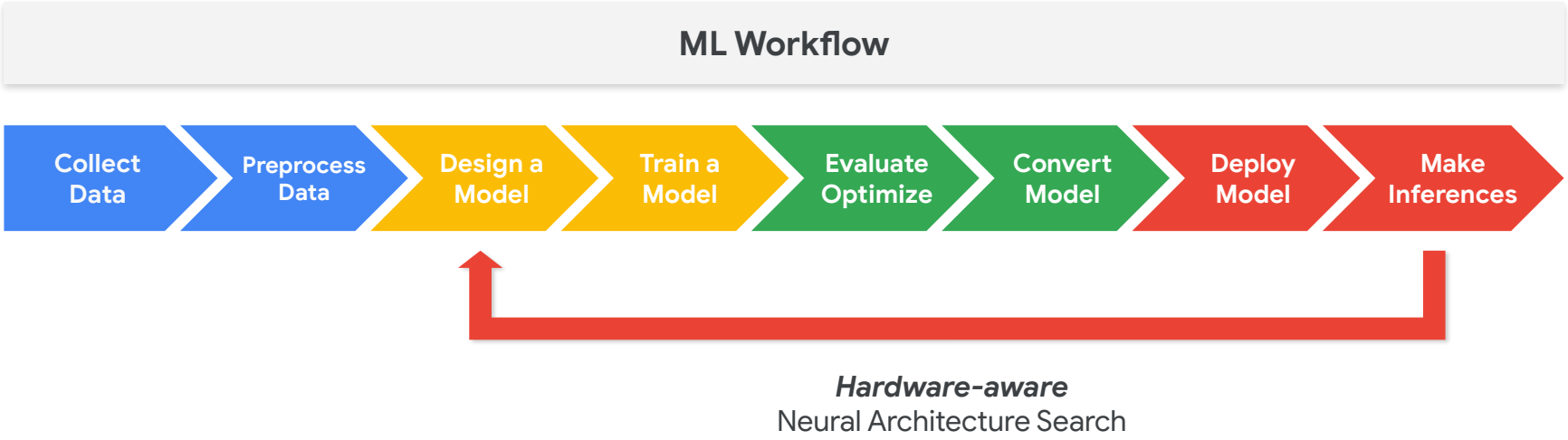


Tailored for **Hardware**





Tailored for **Hardware**



An **Applied** Perspective

Neural Architecture Search: **Stages**