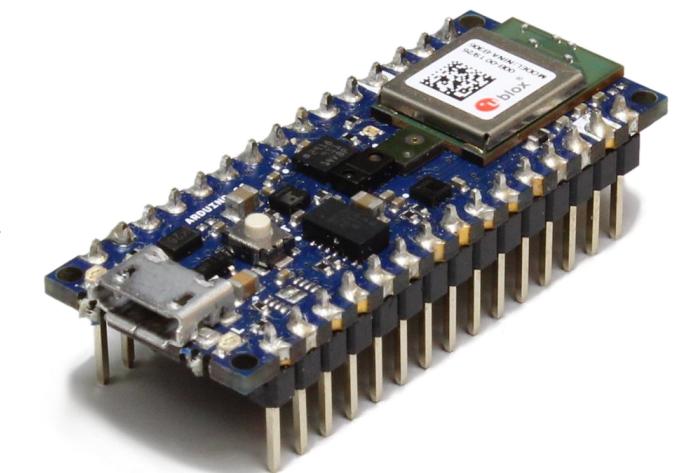
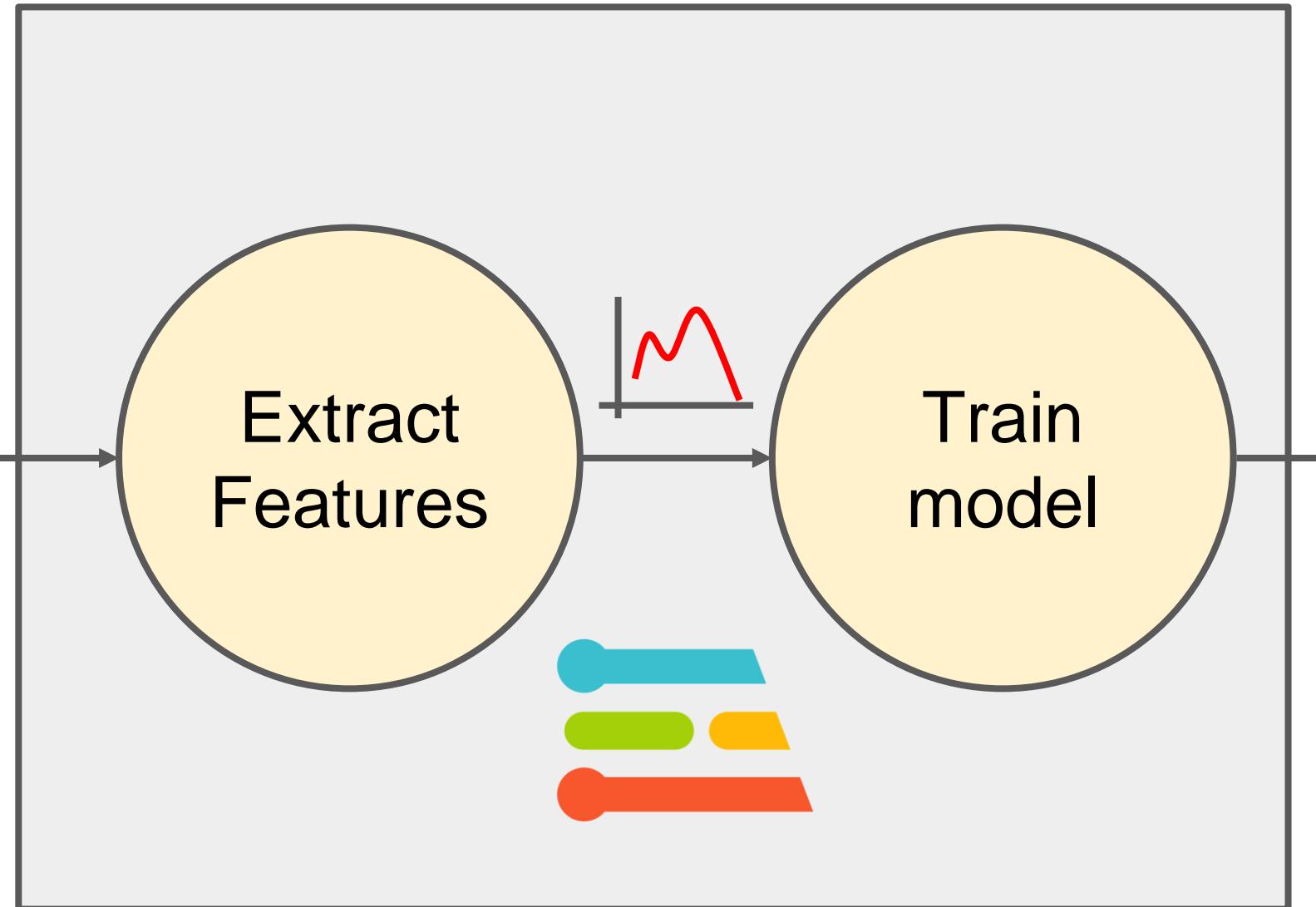
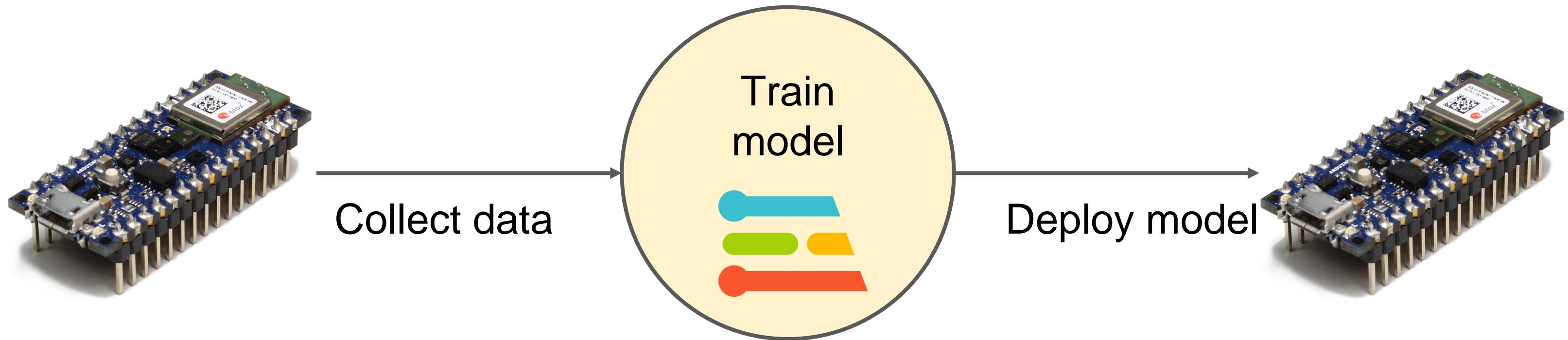
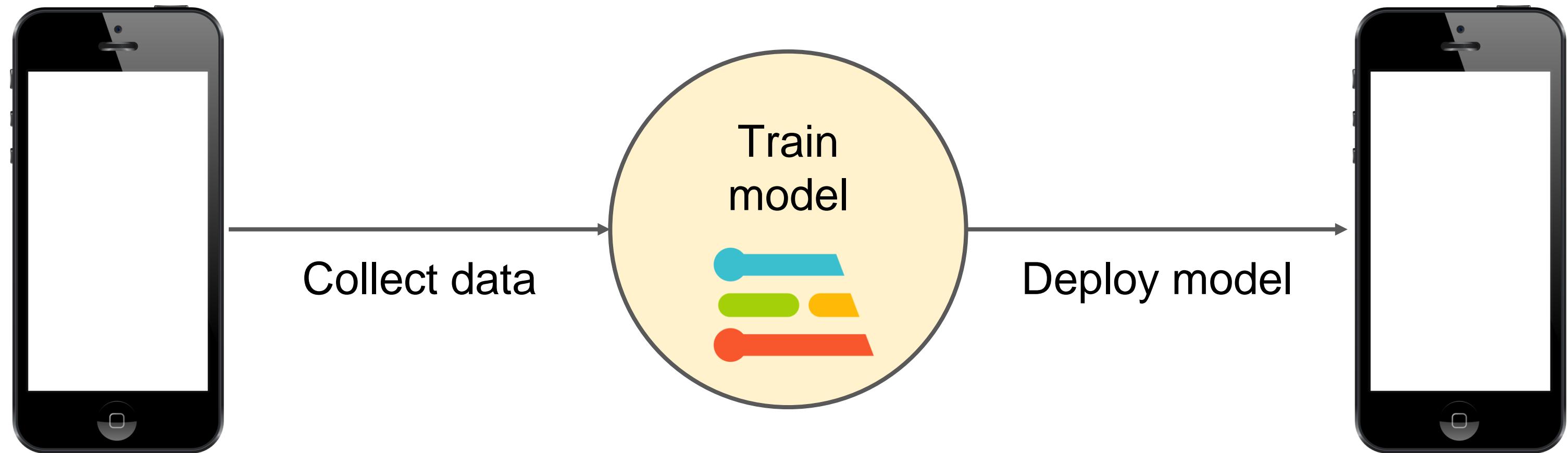
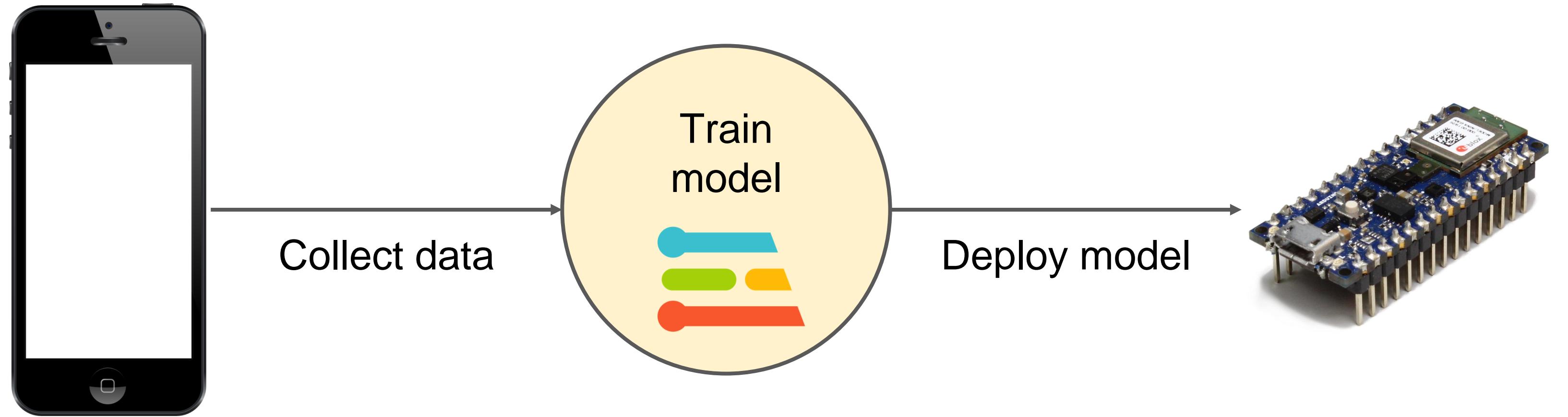


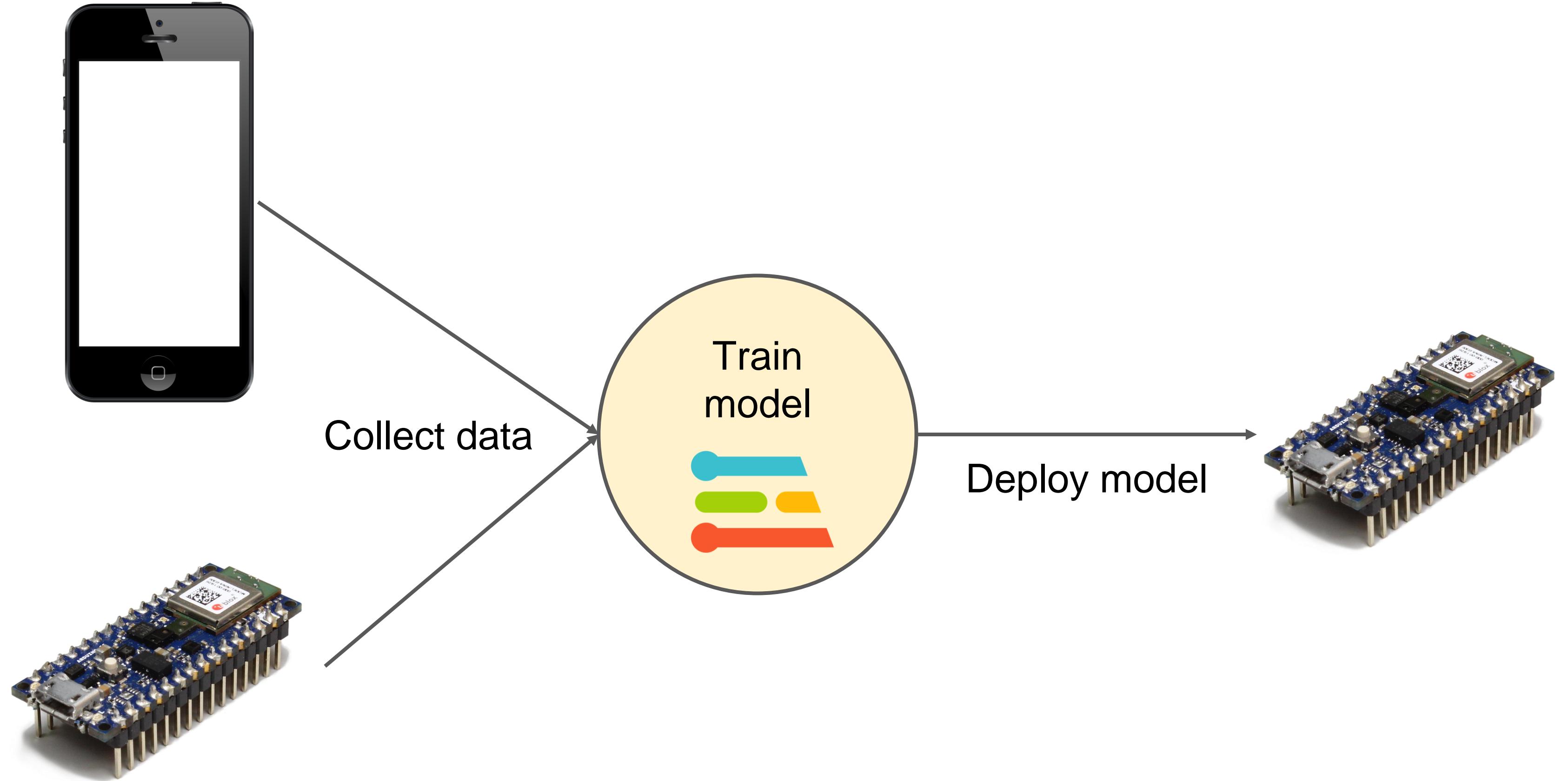


Collect
data

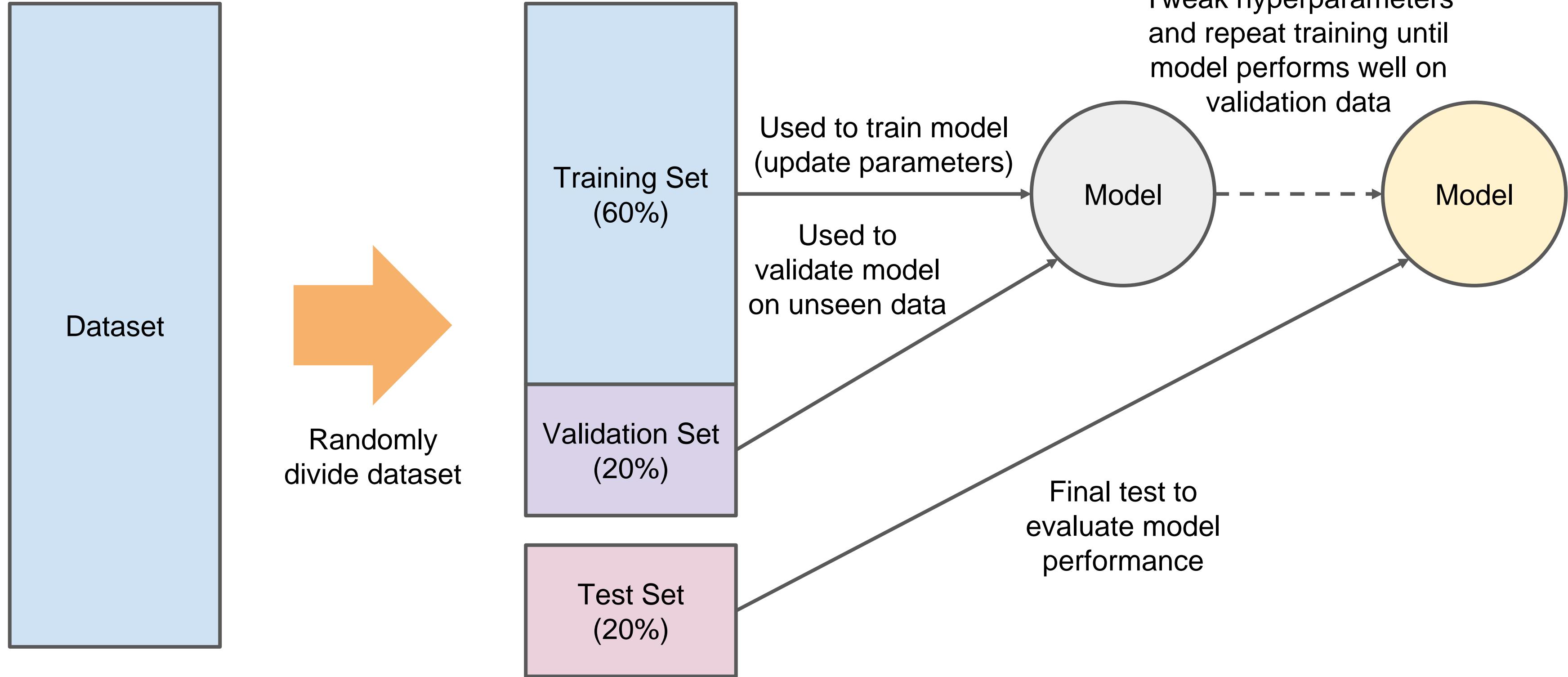




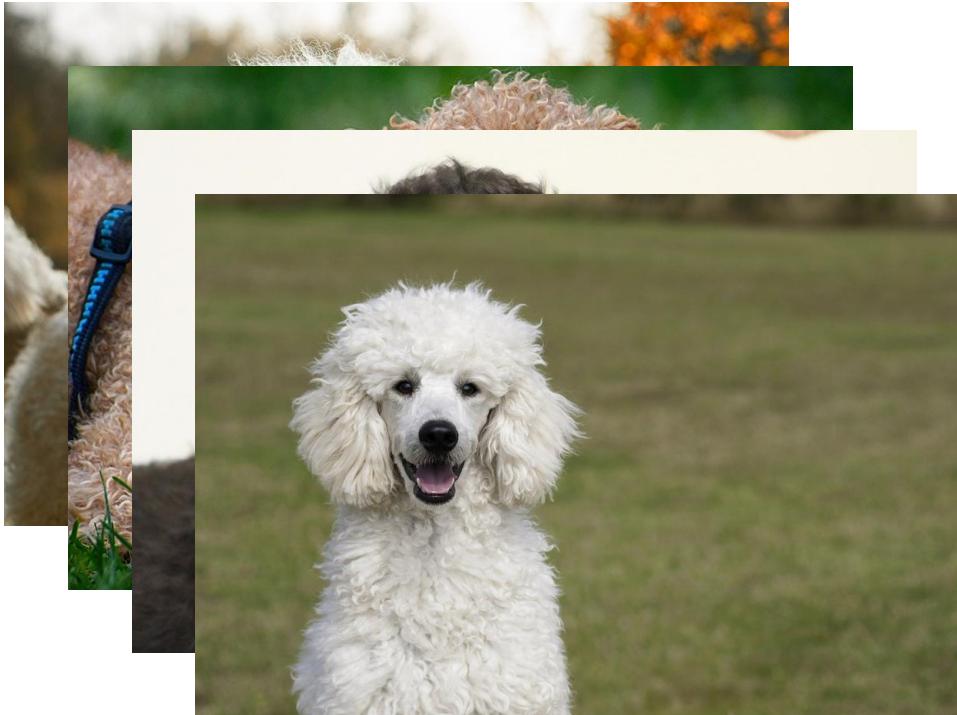




Holdout Method



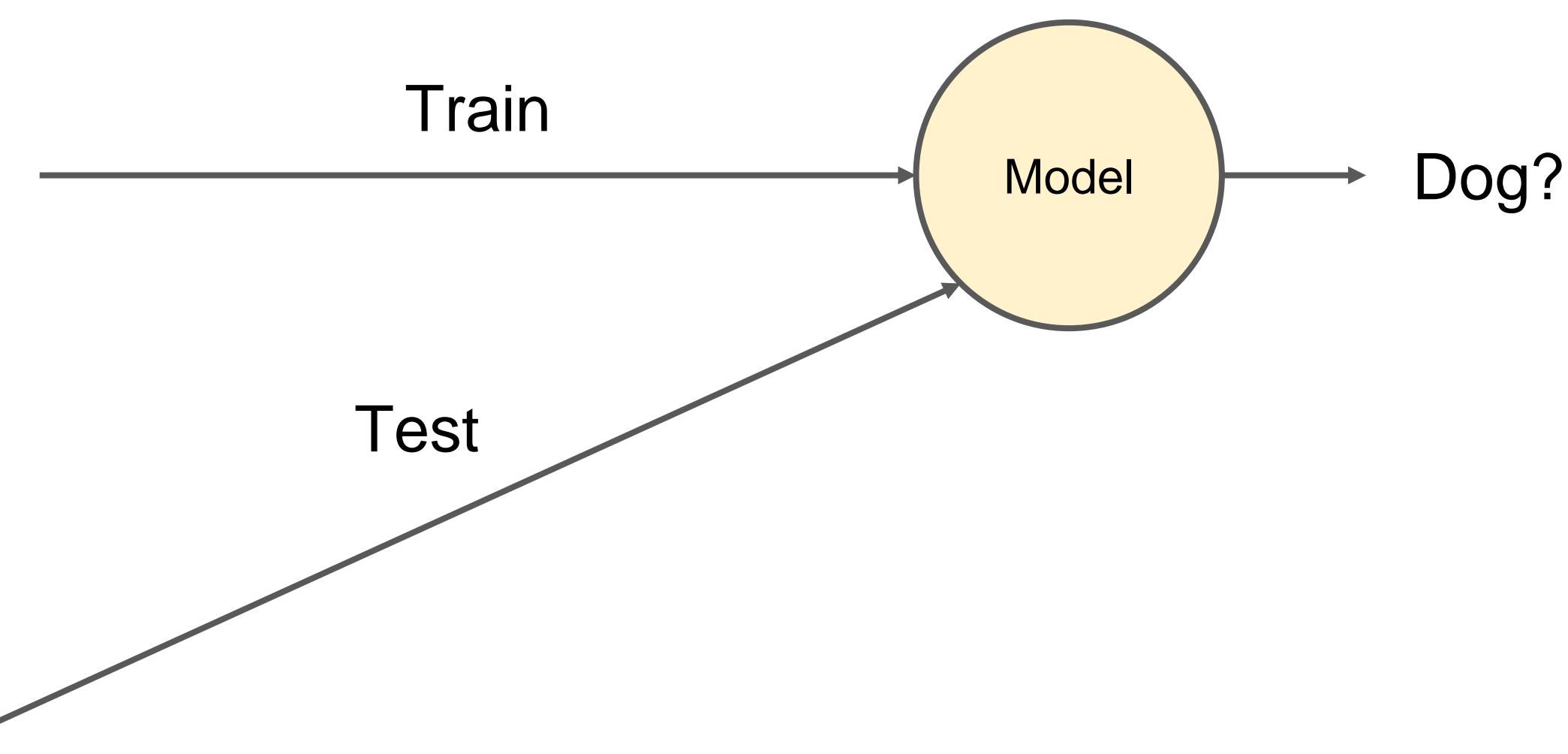
Think about your data!

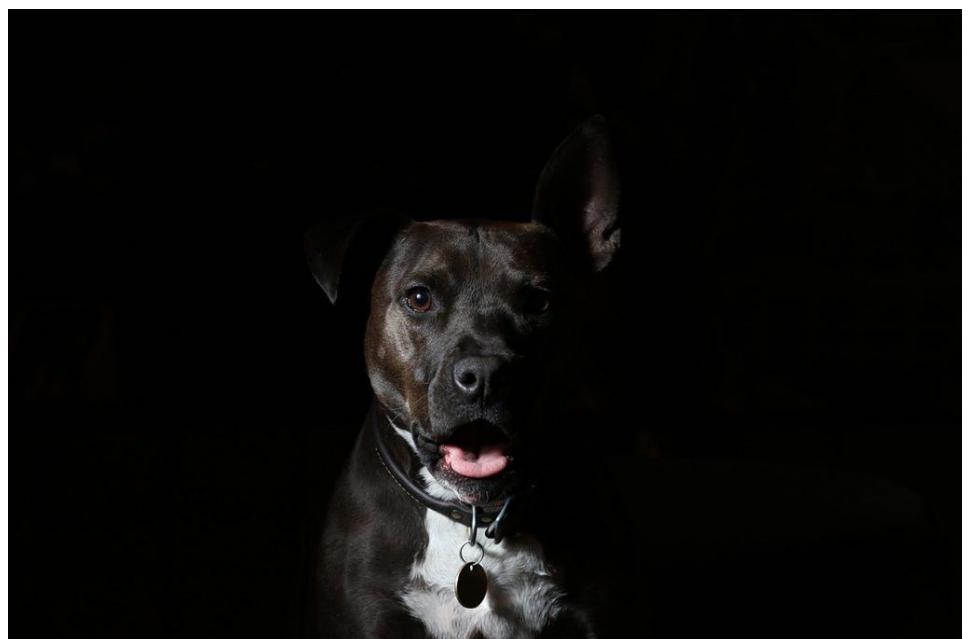
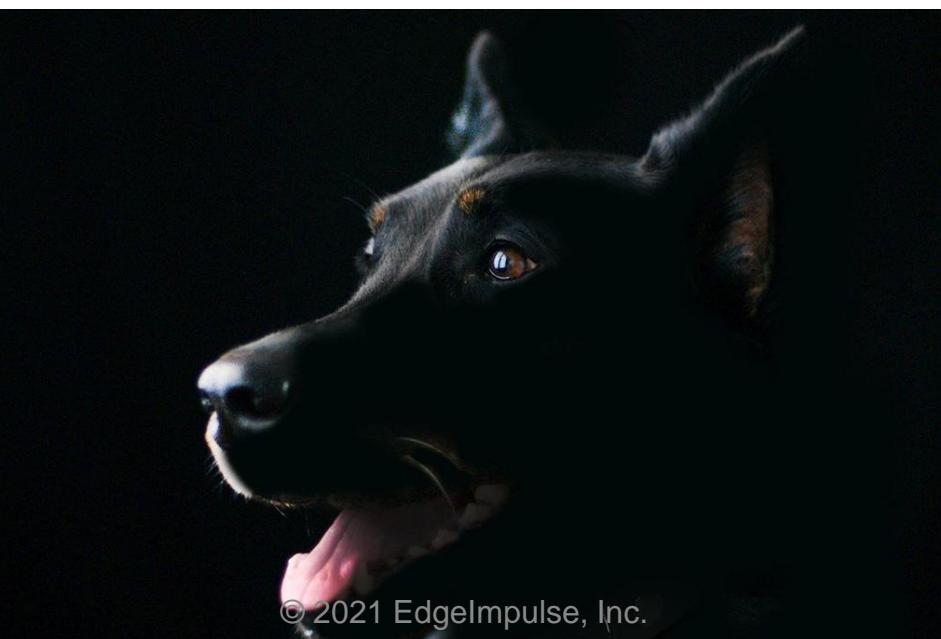
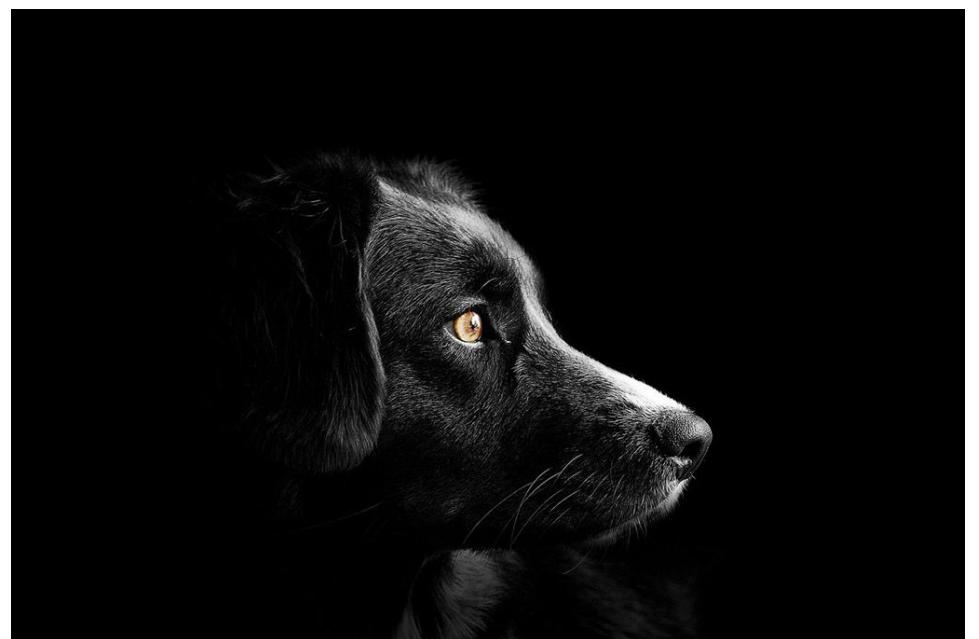
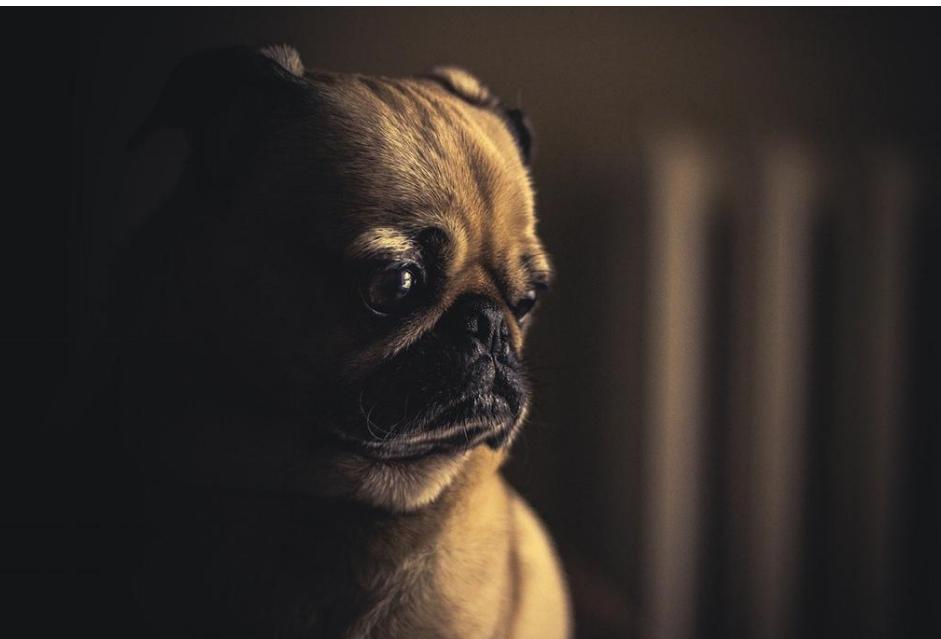
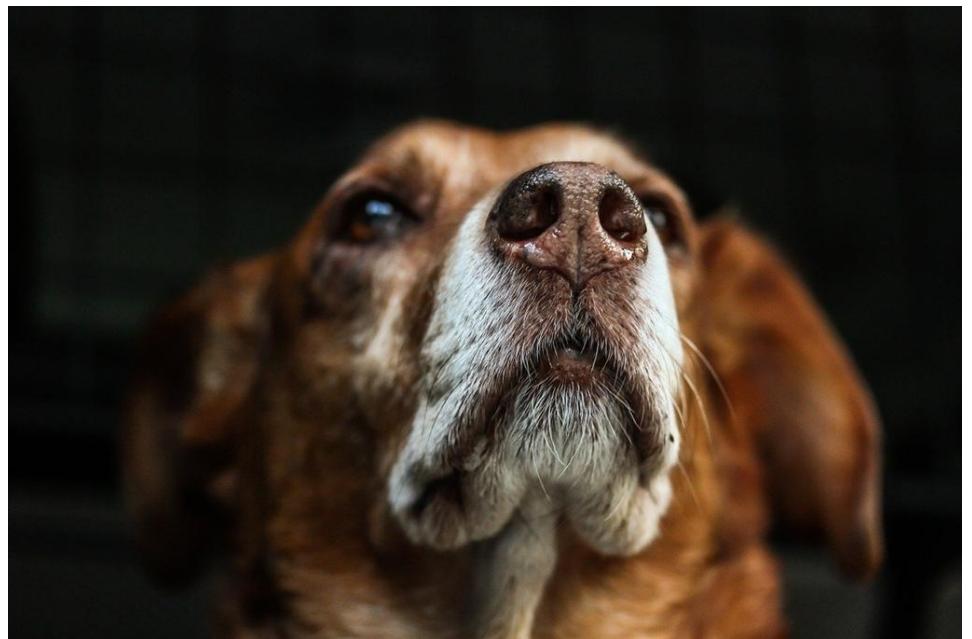
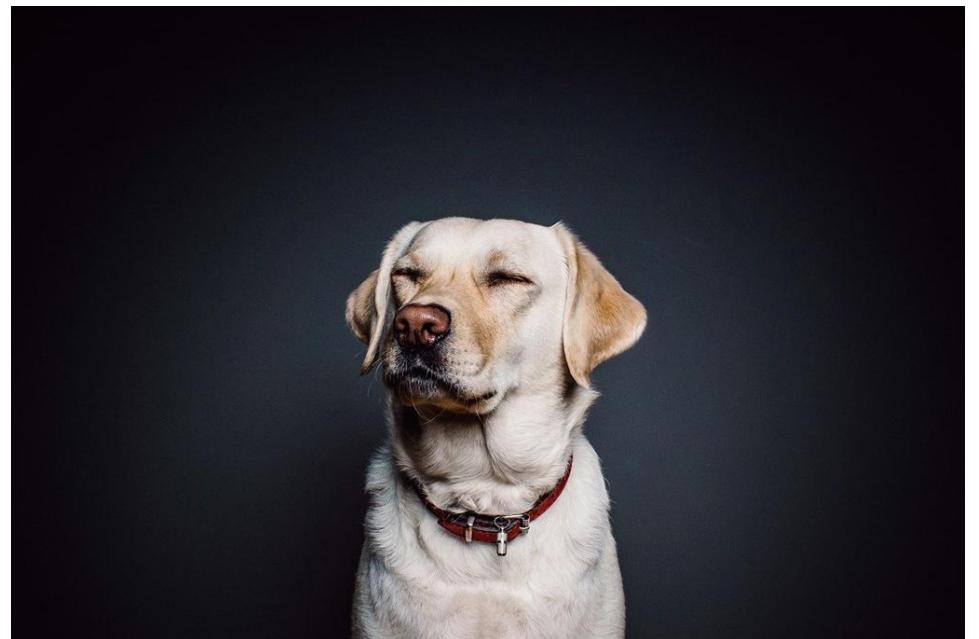


Training Data

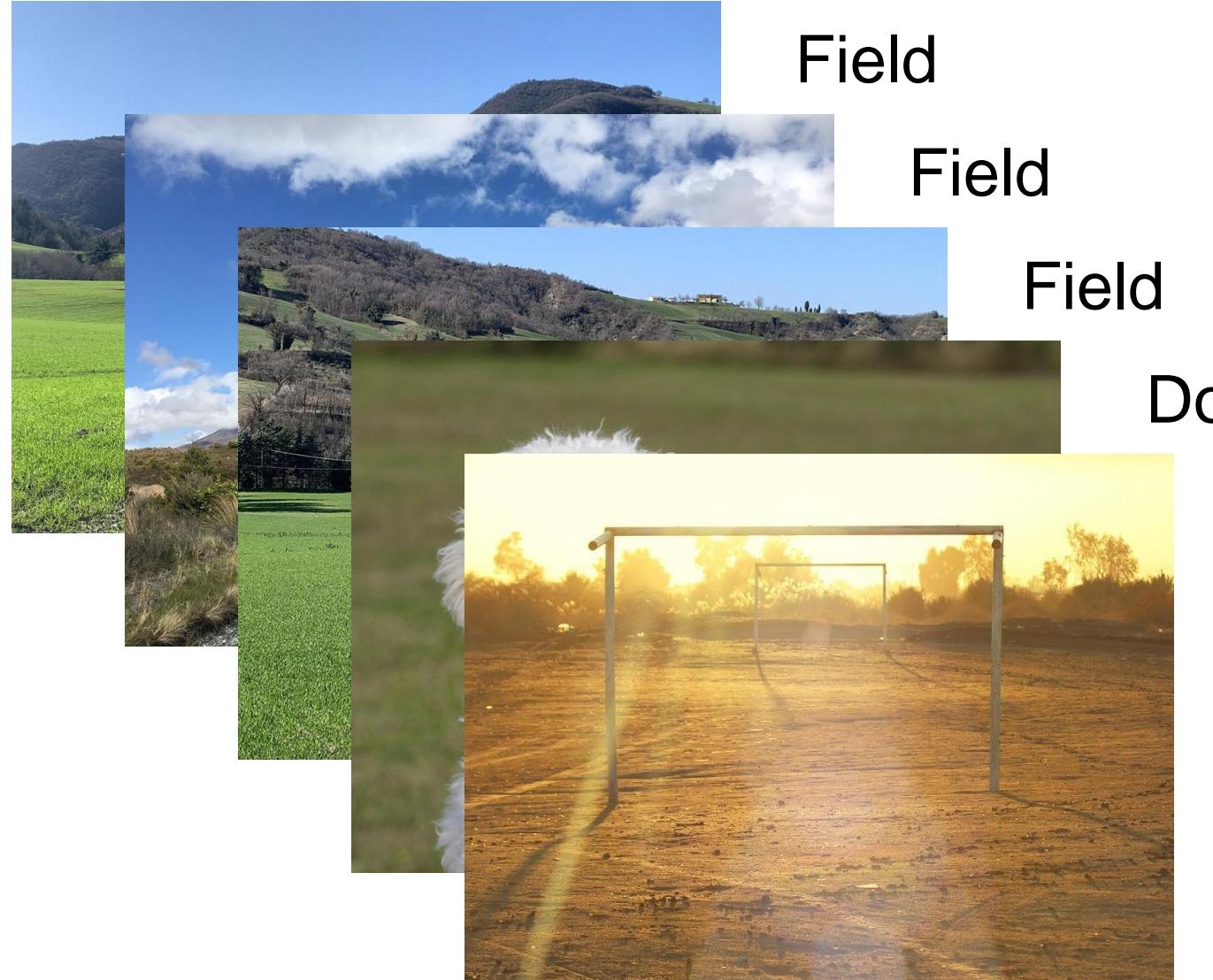


Test Data

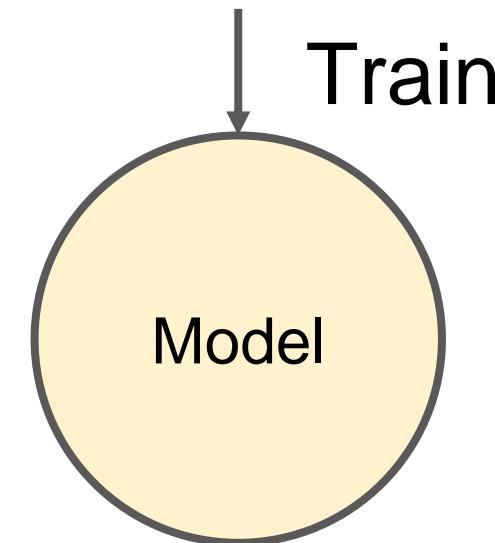
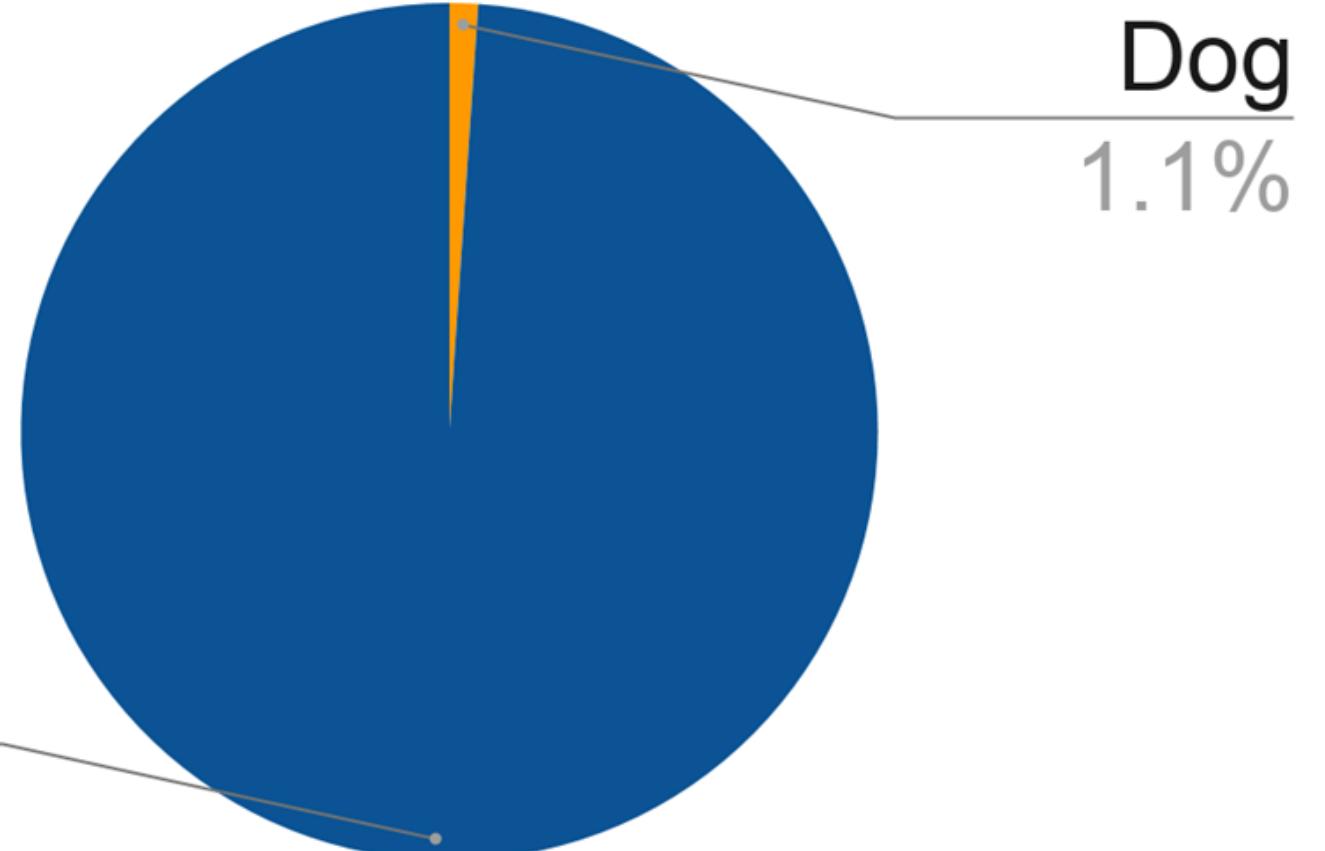




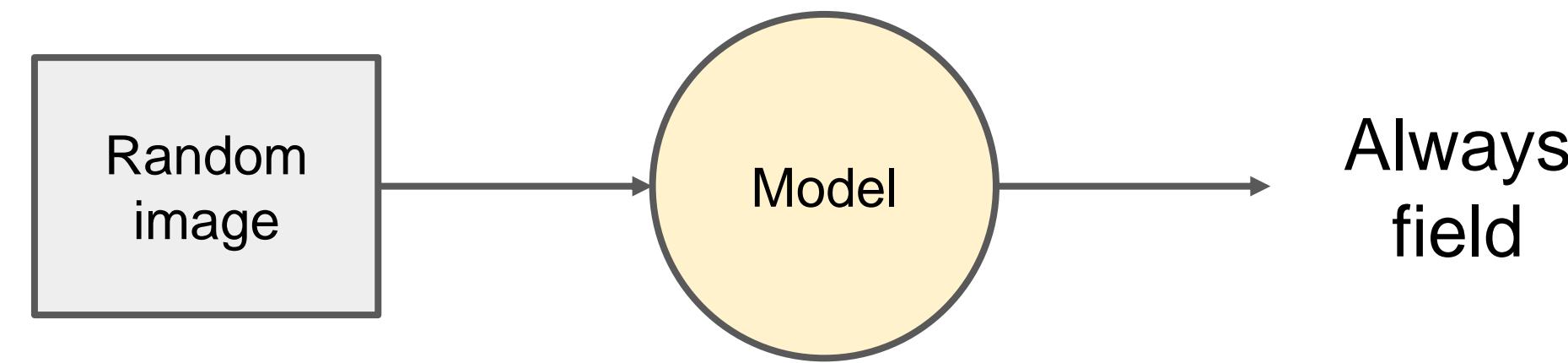
Balanced Dataset



Dataset Classes



Balanced Dataset

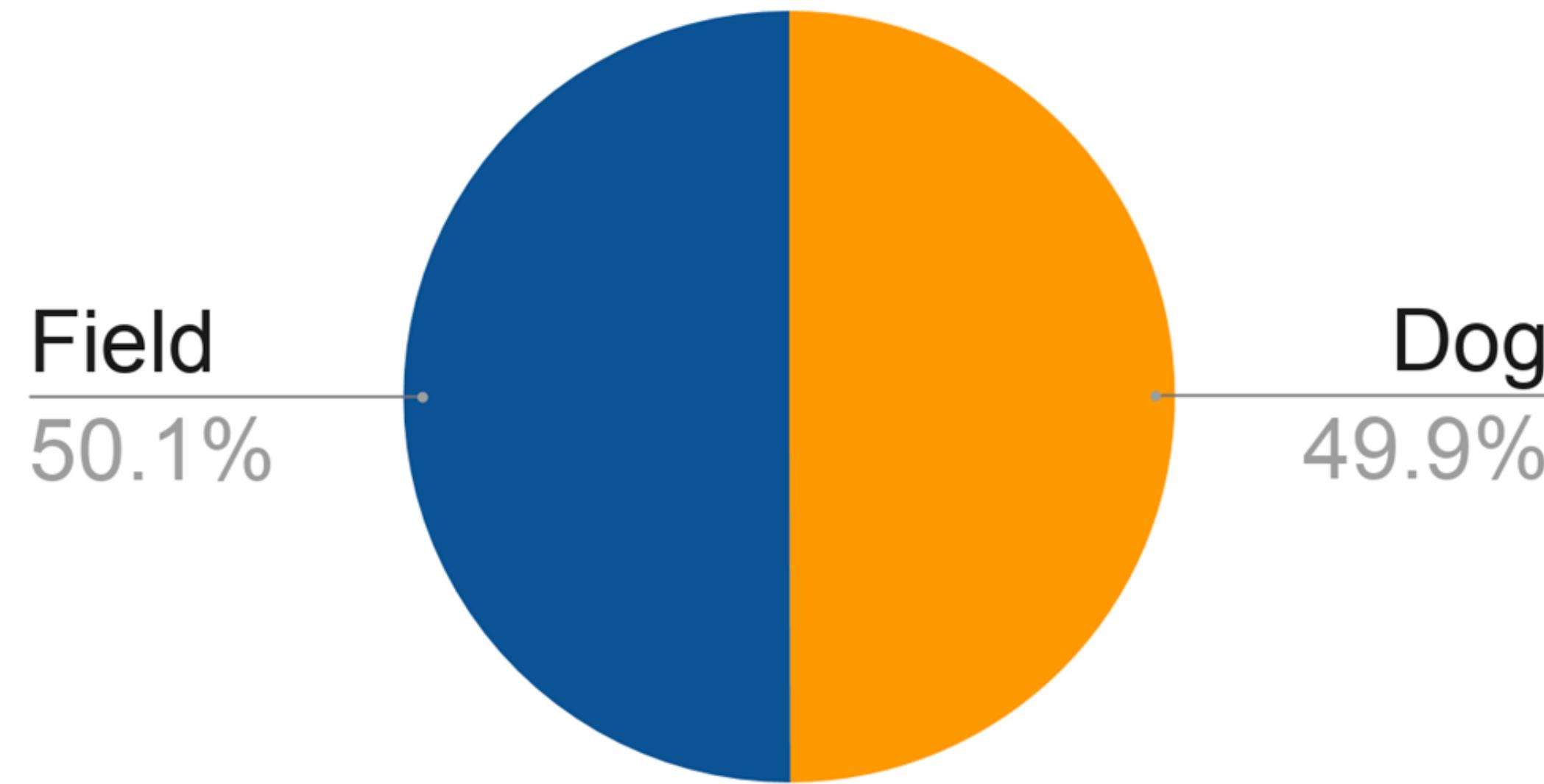


99% accuracy!

“Naive classifier”

Balanced Dataset

Dataset Classes



Anomaly Detection

Email Types

