PetBot

Mahlet Mesfin

Objective

A responsive pet robot that provides light entertainment and interaction without increasing screen time.

Primarily for anyone seeking a fun, interactive desk gadget.

Bandwidth: Reduces network usage because all inference happens locally; only occasional model updates or telemetry need network transfer.

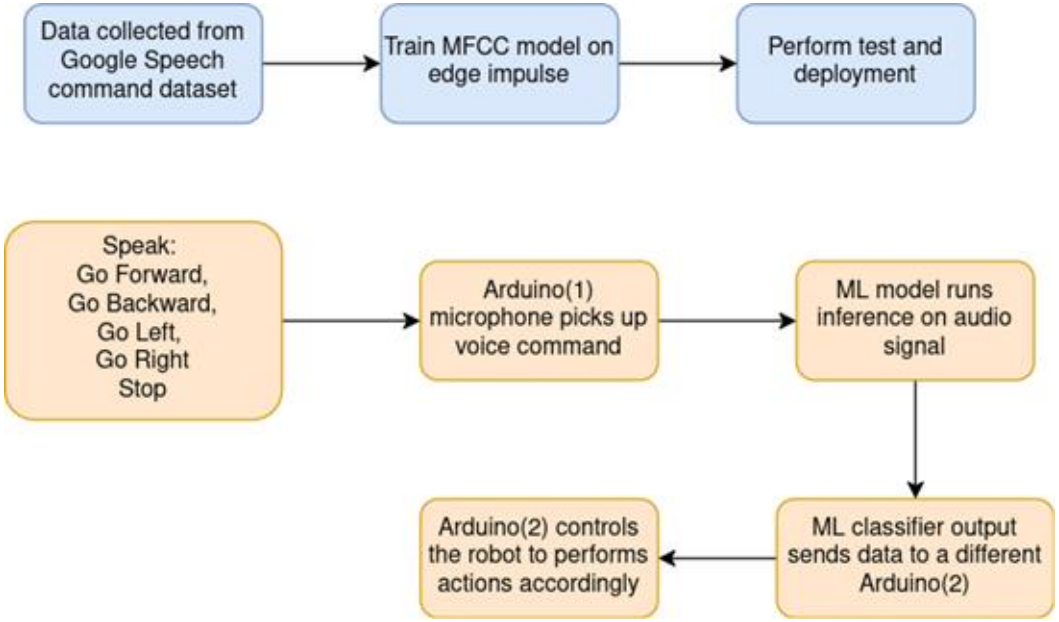
Latency: On-device processing → no cloud latency or dependency.

Economics: cheaper to scale and operate long-term.

Reliability: Fast response time between hearing a command and taking action.

Privacy: All audio processing happens on the device, keeping user data local

Block diagram



Data Collection

- Words: go, stop, left, right, forward, backward, happy
  - Go left
  - Go right
  - Go forward
  - Go backward
  - Stop
  - Happy
- Used google speech command-set
- MFCC: keyword spotting
- Feature extraction and preprocessing: on edge impulse
- Used a Neural network

Embedded systems

Components:

- **Microcontroller:** Arduino Nano BLE 33 Sense (for keyword spotting).
- **Robot Kit:** TokyMaker kit (for movement and actuation).
- **Sensors**: Built-in microphone - Arduino
- **Actuators**: Motors for walking, servo for gestures.

Training Result

Model

Model version: Quantized (int8)

Last training performance (validation set)

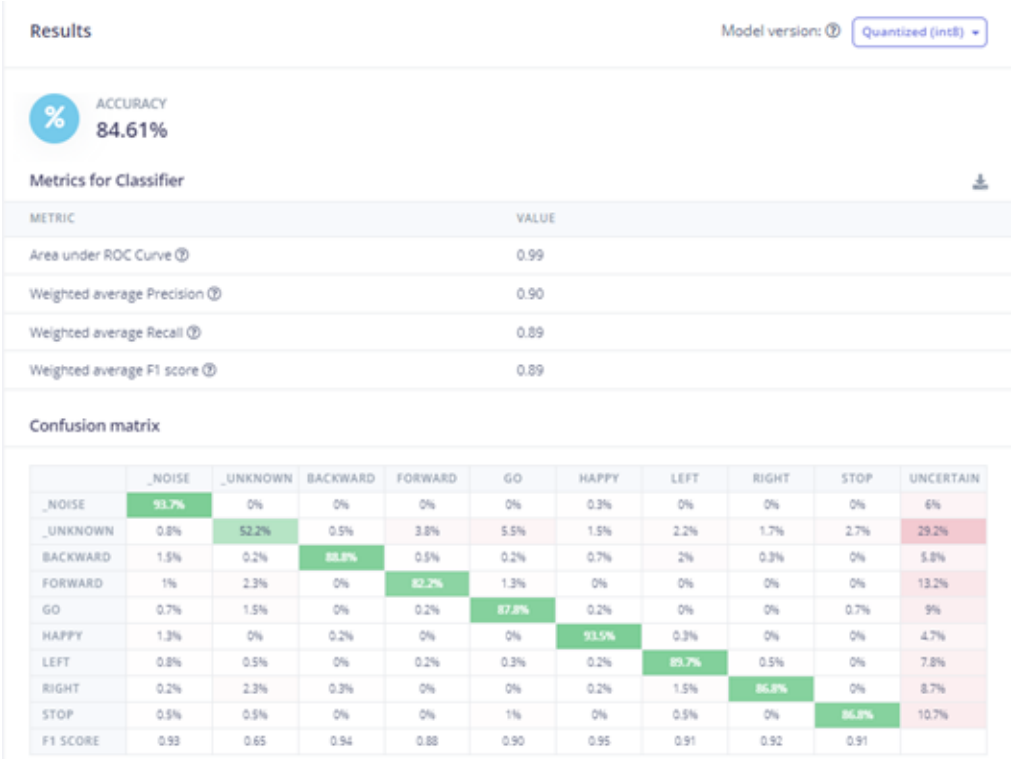
ACCURACY
88.4%

LOSS
0.49

Confusion matrix (validation set)

	_NOISE	_UNKNOWN	BACKWARD	FORWARD	GO	HAPPY	LEFT	RIGHT	STOP
_NOISE	98.4%	0.6%	0%	0%	0.2%	0.4%	0.2%	0%	0.2%
_UNKNOWN	3.9%	71.0%	1.8%	4.3%	7.8%	2.9%	3.9%	2.0%	2.3%
BACKWARD	1.3%	5.1%	87.4%	1.1%	0.2%	1.7%	3.0%	0.2%	0%
FORWARD	1.6%	8.5%	0.2%	87.3%	2.0%	0%	0.4%	0%	0%
GO	1.7%	7.2%	0.7%	0.7%	85.2%	0.2%	0.7%	0.2%	0.4%
HAPPY	2.1%	2.7%	0.9%	0%	0%	93.8%	0.5%	0%	0%
LEFT	1.2%	3.4%	0.6%	0.2%	0.6%	0%	93.4%	0.6%	0%
RIGHT	1.4%	9.5%	0.2%	0%	0.6%	0.4%	2.5%	85.3%	0%
STOP	1.5%	4.4%	0.2%	0%	0.8%	0.6%	0.6%	0%	91.7%
F1 SCORE	0.92	0.68	0.91	0.90	0.87	0.93	0.91	0.91	0.94

Test Result



Results

Model version: Quantized (int8)

ACCURACY
84.61%

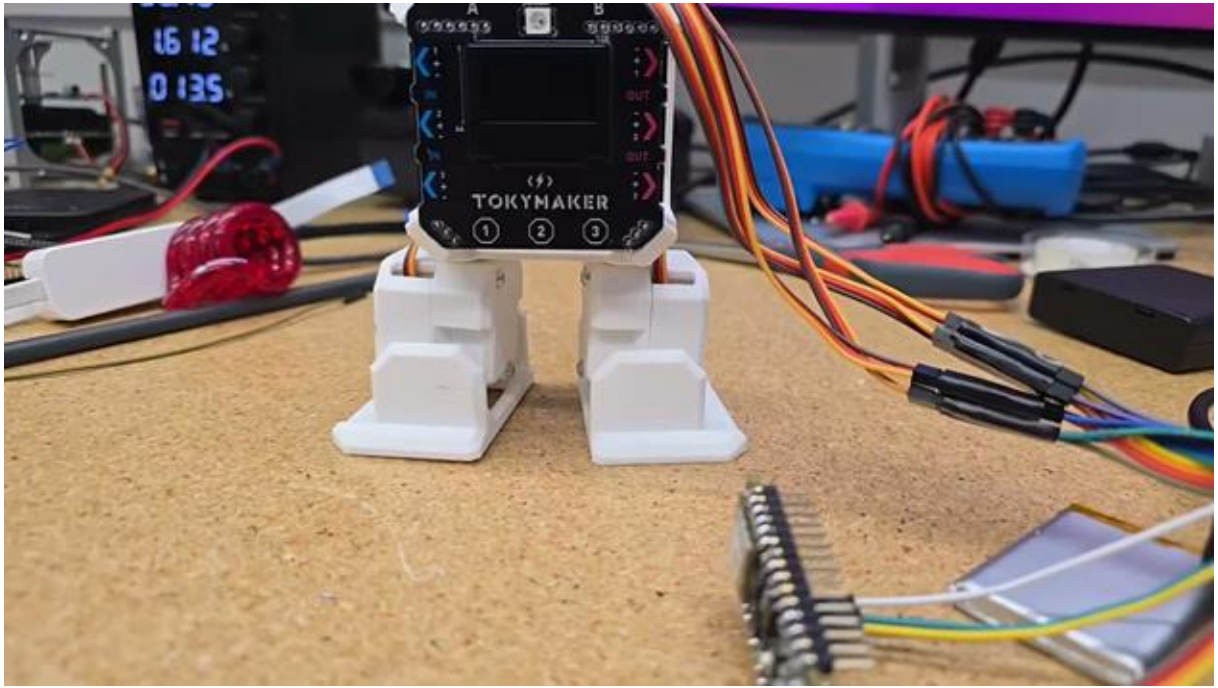
Metrics for Classifier

METRIC	VALUE
Area under ROC Curve	0.99
Weighted average Precision	0.90
Weighted average Recall	0.89
Weighted average F1 score	0.89

Confusion matrix

	_NOISE	_UNKNOWN	BACKWARD	FORWARD	GO	HAPPY	LEFT	RIGHT	STOP	UNCERTAIN
_NOISE	93.7%	0%	0%	0%	0%	0.3%	0%	0%	0%	6%
_UNKNOWN	0.8%	52.2%	0.5%	3.8%	5.5%	1.5%	2.2%	1.7%	2.7%	29.2%
BACKWARD	1.5%	0.2%	88.8%	0.5%	0.2%	0.7%	2%	0.3%	0%	5.8%
FORWARD	1%	2.3%	0%	82.2%	1.3%	0%	0%	0%	0%	13.2%
GO	0.7%	1.5%	0%	0.2%	87.8%	0.2%	0%	0%	0.7%	9%
HAPPY	1.3%	0%	0.2%	0%	0%	93.5%	0.3%	0%	0%	4.7%
LEFT	0.8%	0.5%	0%	0.2%	0.3%	0.2%	89.7%	0.5%	0%	7.8%
RIGHT	0.2%	2.3%	0.3%	0%	0%	0.2%	1.5%	86.8%	0%	8.7%
STOP	0.5%	0.5%	0%	0%	1%	0%	0.5%	0%	86.8%	10.7%
F1 SCORE	0.93	0.65	0.94	0.88	0.90	0.95	0.91	0.92	0.91	

Video



Challenges

- Live testing inference delay
- Doing inference and motion tasks in one core: had to split to two arduinos to smoothly enable the movements
- Testing accuracy skewed by _unknown_ sounds