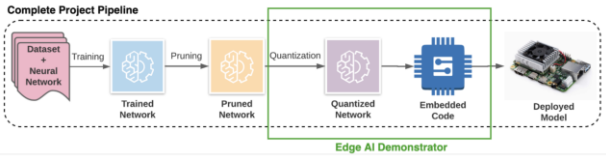
# Edge AI

## MARKET NEED

- Edge AI is predicted to be used in billions of electronic devices to add more advanced AI capabilities to those devices. Certain additional steps are needed to run these memory and computationally intensive algorithms on edge devices.

- This project explores the whole pipeline of an edge AI project from model creation and training, to pruning and quantizing it and finally automatic embedded code generation.

- We explored various existing techniques for each of these steps and also proposed some novel design ideas and pruning techniques.




## TECHNOLOGY SOLUTION

We developed a web-based tool that offers state-of-the-art methods for pruning and quantising a neural network with the following components:

1. *The use of 1D convolution with parallel networks combined in the classification layer to keep the neural network light.*
2. *The use of a magnitude based pruning for each layer.*
3. *The use of a novel scoring method for pruning neurons based on their effectiveness for the task at hand and integer quantization into Q formats.*

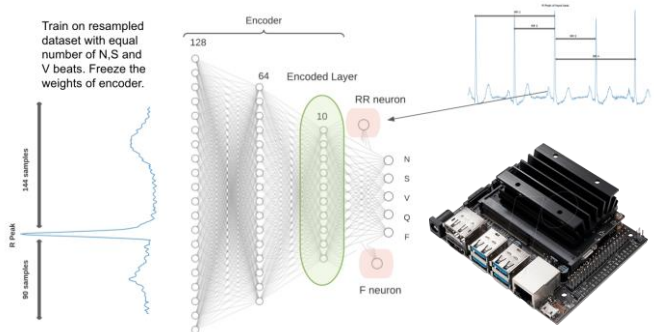
The tool has been extensively tested on a real-world heartbeat signal classification.

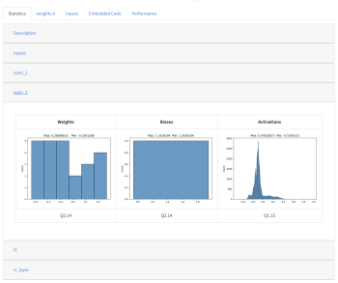
The generated embedded code can be deployed on a variety of embedded systems using ARM processors.

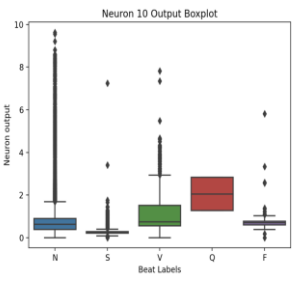
## APPLICABILITY

The tool serves to demonstrate the design of a whole edge AI pipeline.

Those companies working on embedded hardware like microcontrollers, microprocessors, SoC (system on a chip) can use the pipeline to explore how a neural network could be designed and coded into their hardware.






**RESEARCH TEAM**
CeADAR Applied Research Group
UCD

**www.ceadar.ie**

TECHNOLOGY CENTRE
ENTERPRISE IRELAND
IDA IRELAND SUPPORTED

UCD DUBLIN