OPTIMIZATION OF TARGET VALUES IN AN ARTIFICIAL NEURAL NETWORK

Andrassik, Bhuyian, Yakar, Zauner

STRUCTURE OF A NEURAL NETWORK

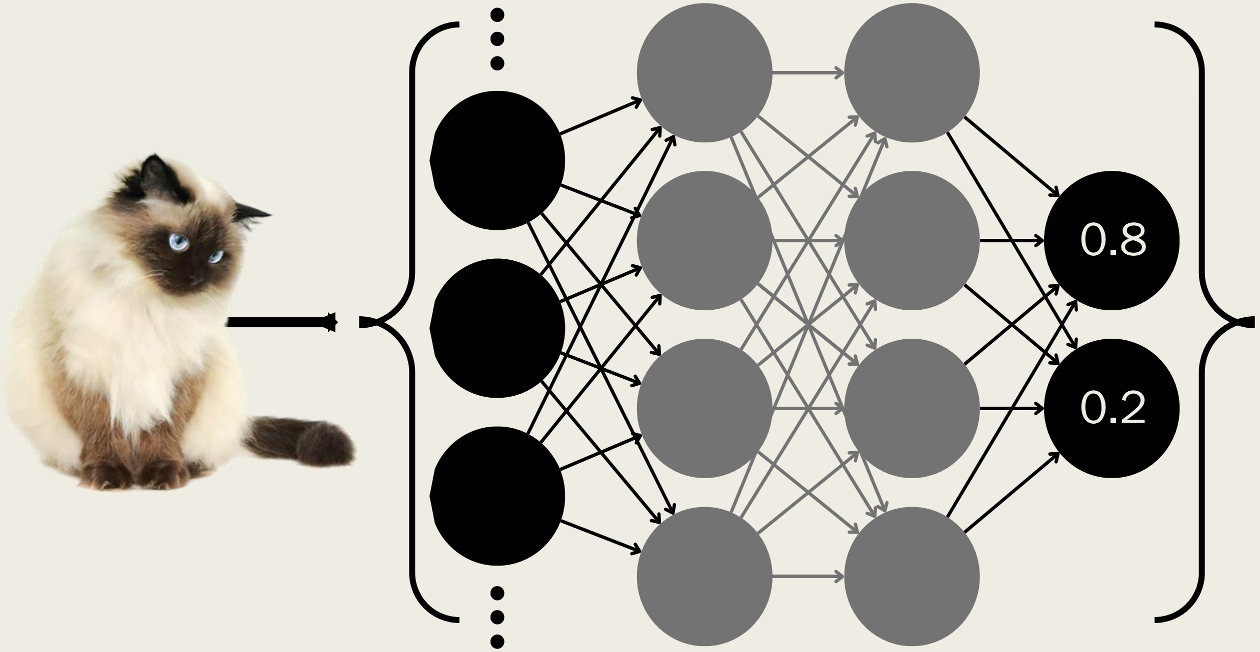
Classification

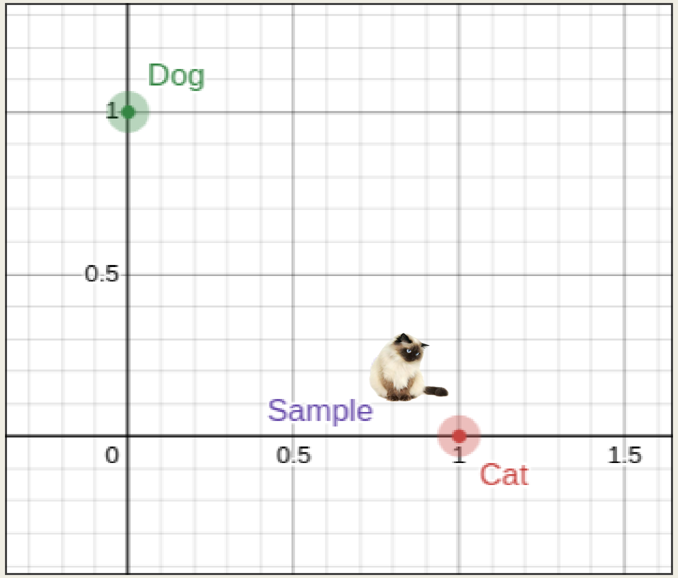
Target values

- *Cat*: (1, 0)
- Dog: (0, 1)

Class/Non-class values

- *Class value: 1*
- *Non-class value: 0*





Cat

What is there to optimize?

Target vectors are made up of class and non-class values.

Focus on finding better class and non-class values.

HOW TO FIND THESE VALUES?

We could…

…simply pick 1 and 0 for class and non-class values.

… manually analyze the activation function and choose values that provide better gradient flow.

… continuously improve values during training.

Adaptive Generation of Class Values

Pre-Training

1. Run all samples through network before training
2. Pick a class-value
3. Average all other values to get non-class value

Training (for every epoch)

1. New class value = Average(old guess, closest in current epoch)
2. New non-class value = Average(entirety of all non-class values)

OVERVIEW OF ADAPTATION METHODS

d_{max} - Adaptation

Definition:

- We introduce a parameter d_{max} to control the separation between class and non-class values.
- The values are set to be as far apart as possible by placing the non-class value at the farthest point in the activation function.

Procedure:

- Calculate the average of the class values.
- If the average is < 0.5, the non-class value is set to 1.
- If the average is $>= 0.5$, the non-class value is set to 0.
- The values are constrained to the range [0,1].

Objective:

- Maximize the distinction between class and non-class values for a clear decision boundary.

ϵ_{min} - Adaptation

Definition:

- We introduce a parameter ϵ_{min} which represents the minimum separation between the average of class values and non-class values to ensure a clear distinction between them.

Procedure:

- Compute the average of class and non-class values.
- If the distance between class and non-class values is smaller than ϵ_{min}, adjust the non-class value:
  - If the non-class average + εmin > 1, subtract ϵ_{min}
  - If the non-class average - εmin < 0, add ϵ_{min}
- Ensure that the adjusted values remain within the range [0,1].

Objective:

- Establish a clear decision boundary by maintaining a minimum distance requirement.

σ - Adaptation

Definition:

- We dynamically adjust the separation between class and non-class values by leveraging the standard deviation of both sets of values.
- The model adapts by using data-driven adjustments.

Procedure:

- Compute the average and standard deviation of class and non-class values, calculate the sum of standard deviations (σ-sum).
- To enforce the minimum required distance, adjust the values:
  - Increase the class value by σ-sum/2.
  - Decrease the non-class value by σ-sum/2.
- If the adjusted values exceed the range [0,1], apply corrective measures:
  - If increasing the class value causes it to exceed 1, subtract σ-sum/2 from the non-class value instead.
  - If decreasing the non-class value causes it to drop below 0, add σ-sum/2 to the class value instead.
- If the computed σ-sum is smaller than the distance between the averages, no further adjustments are needed.

Objective:

- Enable adaptive training by utilizing the output values themselves.


THANK YOU FOR YOUR ATTENTION!