TDT4137 - Cognitive Architectures

Assignment 5

Author:
Martin Skatvedt

18.11.23

Topic 1-A: Connectionism

1.1

One key characteristics of ANNs is planning. ANNs do not plan, as we humans do. However some networks such as RNNs (recurrent neural networks) use LSTM neurons, which can represent sequential decision making as in planning. Secondly ANNs represent information as distributed matrices. They have no explicit symbols, but represent knowledge thought its weights, biases and the patterns which emerge in the system. Lastly they are extremely good at generalizing and approximating complex patterns. Utilizing its non-linear activation functions it can detect ant recognize highly non-linear patterns in data.

1.2

Firstly, connectionism and cognitivist systems represent knowledge and information differently. Cognitivist systems, represent knowledge using symbols, where information is stored as symbols, which can be manipulated. ANNs on the other hand, represent knowledge using distributed weights. The information is stored by the patterns which emerge in the ANN. Secondly they differ in the way they learn. ANNs learn from a iterative process, where each weight is adjusted in each iteration. They can by them self learn and generalize complex patterns. Cognitivist systems often use a rule-based system, where they use pre-defined rules for learning and reasoning.

1.3

Connectionist approaches such as ANN, are heavily inspired by the brains neurons, and try to build artificial neurons. These artificial neurons are connected to other artificial neurons with weights just like biological neurons. They are connected though weights, which can represent the axons, synapses and dendrites. Biological send electo-chemical signals to each other, where good connections result in a high signal. Other features such as CNNs try to model how our brain extracts features from our eyes.

1.4

Some connectionists claim that information is stores non-symbolically in the ANNs weights. They reason that the brain is a neural net, but also a much more abstract symbolic processor.

1.5

In connectionist systems like Artificial Neural Networks (ANNs), intelligence shows up when they learn complex patterns from data, make decisions based on what they have learned and handle new data due to their generalization. They are good at finding patterns and learning from them. Their ability to process lots of information at once, like our brains do, also adds to their intelligence.

Topic 1-B: Artificial neural networks

1.6

It means that an ANN with enough neurons and hidden layers, can approximate any mapping function between inputs and outputs, even though their relationship is highly complex or non-linear.

1.7

The purpose of an activation function is to break the linearity of a neuron. A neurons output is a weighted sum, where a bias is added. Which is a linear function. Activation's functions such as ReLU, will break this linearity. It is this non-linearity which makes ANNs able to deduce complex patterns in the data.

1.8

Firstly the ReLU functions maps all negative numbers into 0, while all positive inputs remain. The sigmoid functions however will map all numbers into the range $[0, 1]$, and looks like an S shaped curve. The output of a sigmoid functions is seen as a probabilistic. Secondly the sigmoid function will map very large or very small inputs close to zero, which can result in vanishing gradients. On the other hand the RElU function can result in dead neurons since all neagtive numbers will outputs as zeroes.

1.9

In supervised learning, supervision is the process of learning from labeled data. Such tasks, can be regression or classification. Each data point is associated with a label or target, such that the model will learn to generalize the data and map inputs to outputs.

1.10

Since an ANN is trained using inputs and labels, each input will have an associated output. The ANN will learn a complex relation between the inputs and outputs. It will learn complex patterns and non-linear relations between its data and labels. It will learn to generalize the data, such that it can classify when given new data.

1.11

In my opinion association making is not the same thing as real intelligence. Association making is the process of relation perception to previous knowledge. Real intelligence however is a complex combination of learning, association making, reasoning and much more. If real intelligence is only defined as association making, how will the agent learn new knowledge, and reason about its knowledge. How will it act if it cannot associate new information with previous information.

1.12

The problem with black boxes, is that us humans can reason about the decision an ANN makes. For example if an autonomous car is controlled by an black box, we can deduce and predict what will happen if it faces a dangerous situation. It may have made its own reasoning which we cant understand because we cant map neurons to decision making.

1.13

Due to the black box nature of ANNs, we cant explain its reasoning. If we look at an example with an AI doctor, which diagnoses patients we face many challenges. Firstly we cannot deduce why it has diagnosed the way it has. Secondly such an ANN would have been trained on historical data, which could be faulty. The historical data could be biased or simply contain errors.

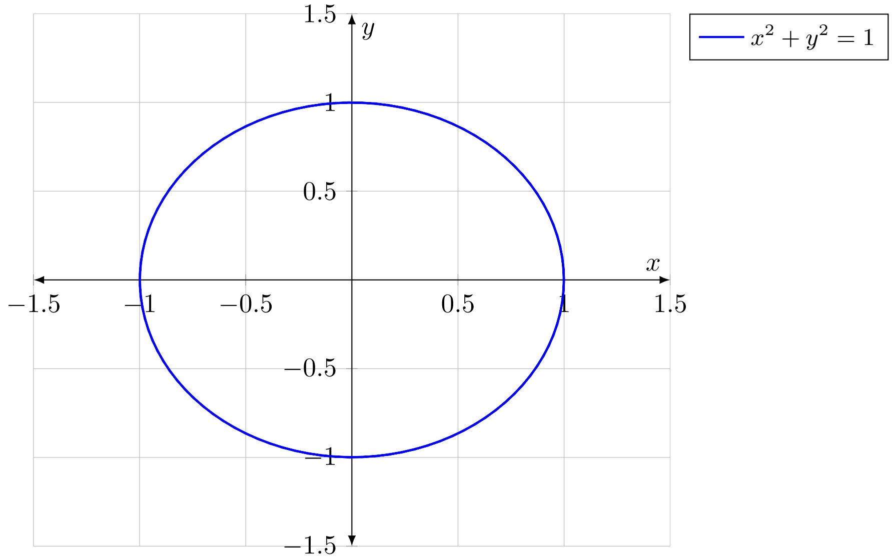
Topic 1-C: Perceptron

1.14

An example of a problem that can be solved by a perceptron is the binary classification of fruits, for example apples and oranges. We could say that the perceptron gets two inputs, the fruits color and weight. The perceptron would then classify if the inputted data was an apple or orange. The classes could be encoded as $apple = 0$ and $orange = 1$.

1.15

If a problem cant be solved by a perceptron, means that its data points are not linearly separable. An example of such a problem would be a dataset where each class is ordered in a circle such as the XOR problem. Another example of such a problem would be to classify if a points satisfies the equation $x^2 + y^2 < 1$. Since the equation represents a circle, the two classes are not linearly separable. We can se the circle plotted below.



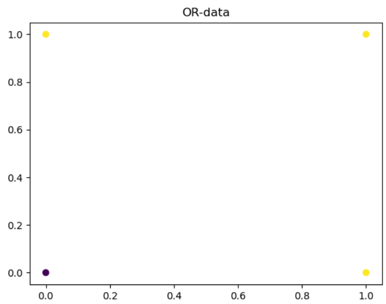
1.16

The perceptron convergence theorem states that if there is a solution the perception will find that solution in a finite number of steps. A solution is meant by that the data points are linearly separable. Linearly separable means that each class in the data can be separated by the other classes using a straight line.

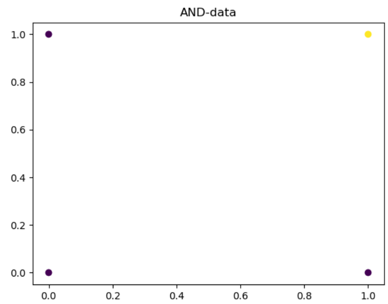
1.17

Using the perceptron model I trained my perceptron on the **AND** and **OR** functions.

Both perceptrons had two weights each and a bias. They were trained on 10 epochs, to ensure convergence and managed to classify the data successfully as we can see in figure 1. The yellow points represent a logic true, and the purple points represent a logic false.



(a) OR model

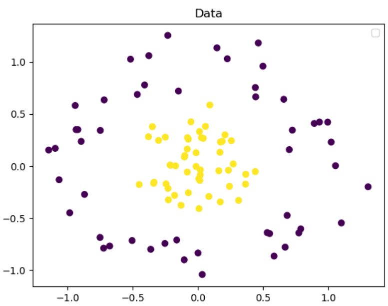


(b) AND model

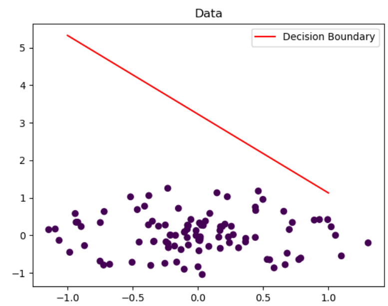
Figure 1: AND/OR Classification

1.18

Using the given I trained my perceptron with two weights and a bias. As we can see in figure 2a, there exists no solution as the classes are not linearly separable. There exists no straight line which can separate the classes. The resulting predictions are plotted in figure 2b, where the red line represents the decision boundary. Since I initialized my weights using a random function the decision boundary would not always converge into the same line.



(a) Original data

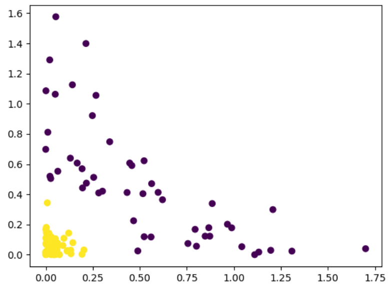


(b) Perceptron classification

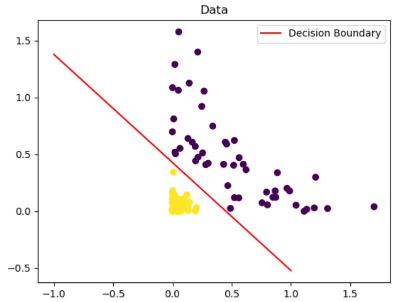
Figure 2: Perceptron trained on given data

1.19

To solve the problem I used a square transformation, where each point in the data would be squared, the result of this can be seen in figure 3a. Here we can see that the data has become linearly separable, and the perceptron managed to find a correct decision boundary as we can see in figure 3b.



(a) Square transformed data



(b) Perceptron classification

Figure 3: Perceptron classification on transformed data