Task 4.1: Name at least 3 different key characteristics of connectionist approaches and explain them. (Explanation is mandatory, only listing terms is not sufficient.) You can focus on Artificial Neural Networks for this task. Some hints: Representation of information. Speed. How to get ’intelligent capabilities into an ANN. Solving logical puzzles. Do ANNs plan?

The connectionist approach when it comes to cognitive modeling make use of large networks of simple computational units specifically nodes in ANN, which communicate by means of simple quantitative signals. This as a way to simulate the way the complex system of the brain reaches the emergent property of general intelligence.

The connectionist approach to learning is the use of higher-level information processing that emerges from the massively-parallel interaction of these nodes by means of their connections, and a network gains the ability to adapt its behavior by means of local changes in the strength of the connections, thus learning through generating these strengthened connections.

The connectionist approach to speed form from the networks ability to run massively-parallel interactions as the connections form in “competition” with each other. This allows for the complete use of the system which propagates up to the higher-level processing, which makes it possible for the network to learn, solve problems, etc. also in parallel.

Task 4.2: Vernon puts connectionism under emergence. Explain why connectionist systems can be subsumed into the family of emergence-based systems.

Vernon puts connectionism under emergence as he classifies connectionist, dynamical, and enactive systems to be grouped together as they argue for the view of cognition as emergent, self-organizing, and dynamical, where all emergence-systems are based to a lesser or greater extent on principles of self-organization**.**

Task 4.3: optional Explain which aspects of intelligence ’emerge’ in connectionist systems.

Task 4.4: Name at least two important differences between connectionism (f. ex. ANNs) and cognitivist systems. Explain each of them.

Cognitivist systems view human cognition to be analogous to symbolic computation in digital computers. Therefore, information storing in cognitivist systems are represented by symbols, just as data is represent in computer memory or on pieces of paper. Connectionist’s claim that information is stored non-symbolically in the weights, or connection strengths, between the units of a network.

Cognitivist systems on adapting focus generally on the structure of explicit symbols and syntactical rules for their internal manipulation, whereas connectionists focus on environmental input and storing this information in a form of connections between neurons as an adaptive strategy.

Task 4.5: optional In which way do ANNs differ from findings about brains and human neural structure? (At least two different aspects are necessary).

The artificial neural network adds layers up, one by one instead of being a part of a network with nodes that don’t add up synchronously, in other words it creates a tree like structure. In the brain, neurons parallelly fire asynchronously with a small portion of highly connected neurons and a large amount of lesser connected ones, with parts having the ability to die off.

In the human brain, there are around 86 billion neurons and 100 trillion synapses that pass on electrical signals. The artificial neural network that are in use today use a much lower amount of neurons usually in the thousands.

The biological neuron fires electrical signals that travel at varying speeds depending on the type of nerve impulse, usually ranging from 0.61 m/s to 119 m/s. In contrast, an artificial neuron emits a signal by the continuous, floating-point number values of synaptic weights, and is almost always faster, by a significant margin.

Task 4.6: What is meant by the statement that artificial neural networks are a universal function approximators?

A universal function approximation means that there is a universal approximation of any function i.e. no matter what f(x) is, there is a network that can approximately approach the result. So the statement that artificial neural networks are a universal function approximators means that ANN has kind of universality where it can in theory solve any mathematical function y= f(x) that can map attributes(x) to output(y) approximately.

Task 4.7: optional Imagine a neural network composed only of neurons where the activation function is simply linear: (output = factor · (weighted sum of input signals)). Can such a network approximate arbitrary functions? (Justify your answer)

Task 4.8: Assume you want to build a classification system with a single neuron. Why must the activation function be nonlinear?

Non-linear means that the output cannot be reproduced from a linear combination of the inputs. The purpose of the activation function is to introduce non-linearity into the network. In turn, this allows you to model a response variable that varies non-linearly with its explanatory variables. The activation function must be non-linearly as a linearly activation function the network will always behave as a single layered perceptron, as the sum would always just give another linearly function.

Task 4.9: What are the two most important differences between a sigmoid activation function and the ReLU activation function?

Sigmoid activation function has an (-infinite, infinite) range of acceptable inputs which all return a value between (0, 1) which is particularly useful when working with probabilities. ReLU activation function has an output range of [0, infinite) as it returns zero for all negative values.

Sigmoid activation function is differentiable and monotonic, but the functions derivative is not monotonic. ReLU activation function and its derivative are both monotonic.

Task 4.10: optional Explain the biological inspiration for Convolutional Neural Networks

CNN has one or more layers of convolution units. A convolution unit receives its input from multiple units from the previous layer which together create a proximity. Therefore, the input units form a small neighborhood that share their weights, which resembles the way the brain creates clusters of related neurons to different tasks.

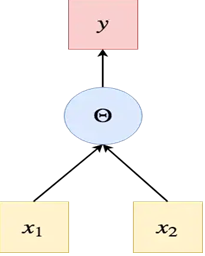
Task 4.11: optional Can, in principle, an ANN be trained to perform the multiplication of two integer numbers of, say, 8 bits length each? Would you recommend to do this ? (justify your answer)

Task 4.12: What is the supervision in supervised learning?

The supervision in supervised learning is the historic- and output-data which the system uses to relate the current input to previously acquired knowledge. One can say that the supervision is the pre-existing knowledge a supervised learning intelligence possesses.

Task 4.13: Give an example of a problem that can be solved by a perceptron. You might sketch this.

A task such as a logical AND-gate can be solved by a perceptron. Where the perceptron’s y will return 1 only for the situation where x1 and x2 both are one. Illustrated below:

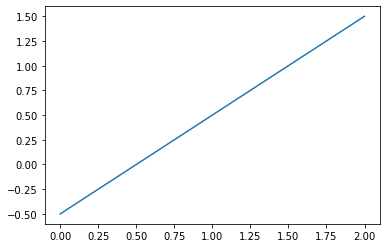


Task 4.14: Give an example of a problem that cannot be solved by a single perceptron.

A logical XOR-gate can not be solved by the use of a single perceptron, because the problem of XOR is not linearly separable, and thus will require more than one perceptron to be solved.

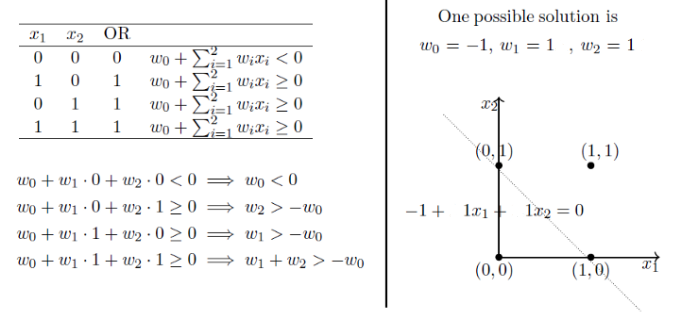
Task 4.15: Implement a simple perceptron. Use the step function for the activation and implement the perceptron learning rule. Test your implementation by learning weights for implementing the OR function. Show your learned perceptron by plotting the decision boundary.

The weights for x1 and x2 have to be the same as they are to be treated equally.



Text

Description automatically generated



Task 4.16: Use your Perceptron model on the data in file assign4data.csv. Show the learned decision boundary. Is this problem linearly separable?

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

As shown by the decision boundary and plots, this problem is obviously not linearly separable as no line drawn can separate the blue from the red, as red encircles blue.

Task 4.17: Can you make the data in assign4data.csv linearly separable? Train the Perceptron again on this the transformed dataset. Show the learned decision boundary.

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

By converting each point to their Polar coordinate value, the points in the inner circle will have smaller radius than the points in the outer circle thus allowing us to transform the points separable by using a linear classifier.