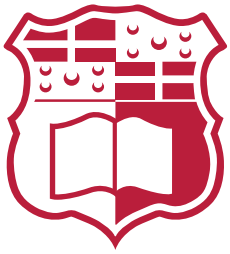

L-Università ta' Malta
Faculty of Information &
Communication Technology

Advanced Computer Vision Individual Assignment

Matthias Bartolo

B.Sc. It (Hons) Artificial Intelligence (Third Year)

Study Unit Code:	**ARI3129**
Study Unit:	**Advanced Computer Vision for Artificial Intelligence**
Lecturer:	**Dr Dylan Seychell**
Date:	**17 October 2023**

Contents

1 Neural Networks & TensorFlow Playground

1.1 Task 1: Understanding Activation Functions

1.1.1 Question

1. Create a neural network with just one hidden layer.
2. Experiment with various activation functions like ReLU, Sigmoid, and Tanh.
3. **Objective:**
    - Note the number of epochs for the model to converge with each activation function.
    - Observe if some activation functions fail to model certain patterns in the data.

1.1.2 Answer

The created neural network with one hidden layer utilised for this experiment and which can be seen in Figure 1.1 has the following properties:

1. Hidden Layer consisting of four Neurons
2. Learning Rate set to 0.03
3. Problem Type set to Classification
4. Regularisation set to None
5. Dataset set to Circle

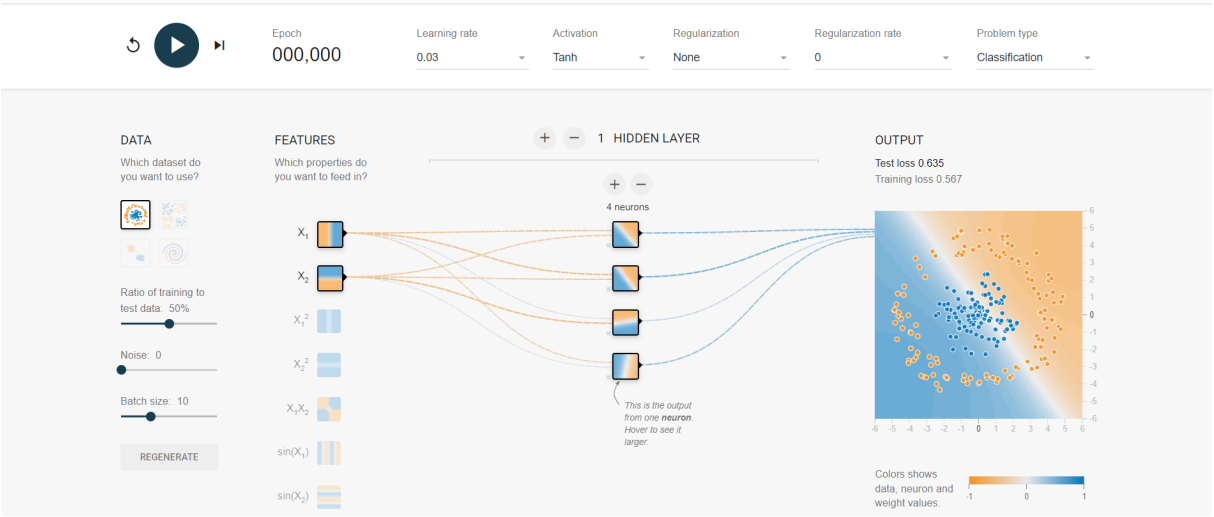

Figure 1.1 Created neural network with one hidden layer.

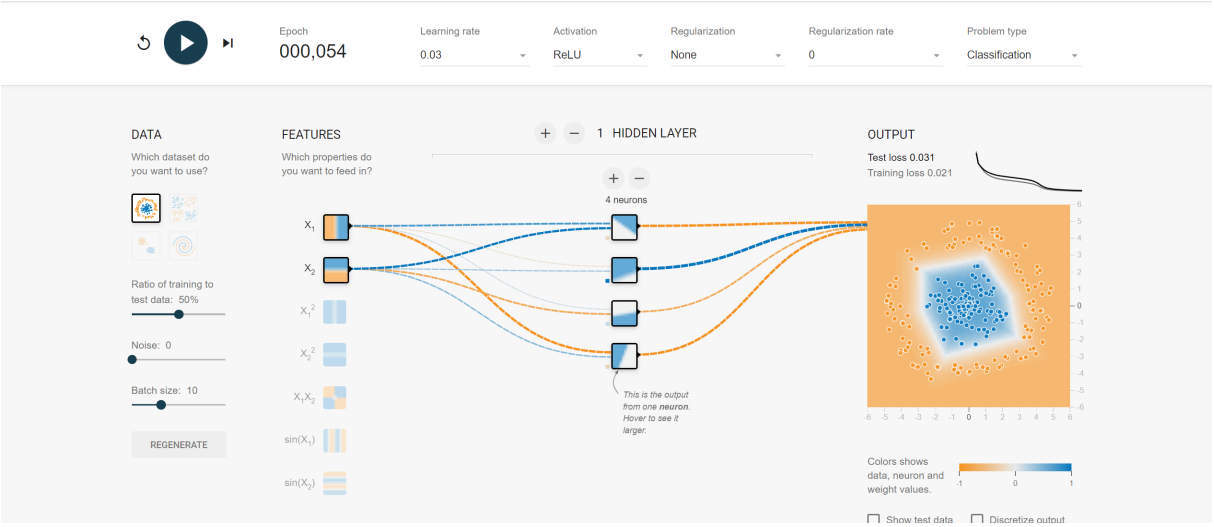

Figure 1.2 Created neural network learned to fit the Circle dataset.

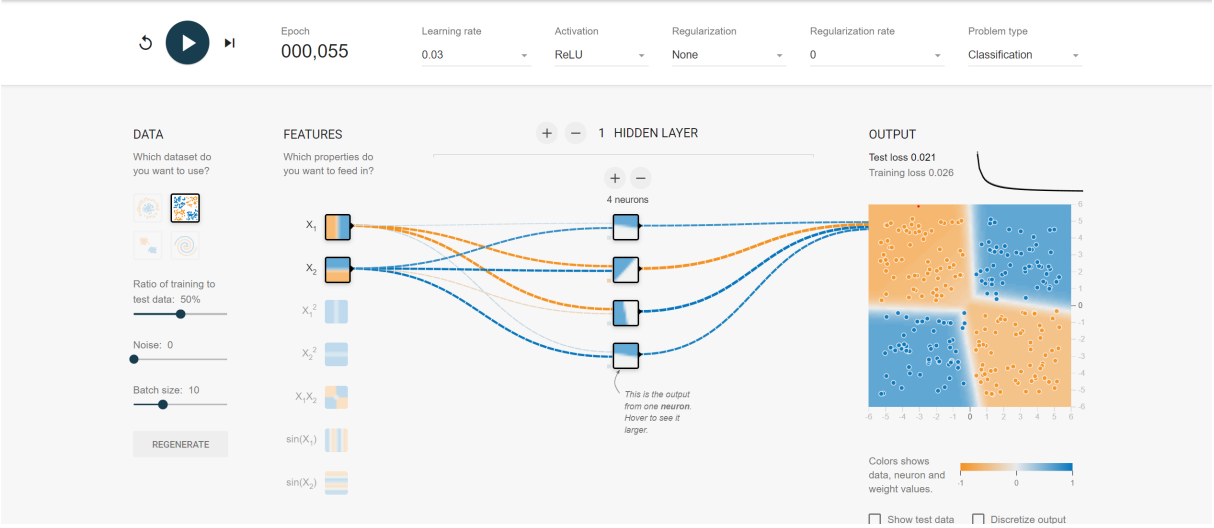

Figure 1.3 Created neural network learned to fit the Exclusive or dataset.

Activation Function	Epochs converged	Converged
ReLU	54	Yes
Tanh	100	Yes
Sigmoid	200	Yes
Linear	∞	No

Table 1.1 Results obtained for changing the Activation Function in the neural network shown in Figure 1.1, using the Circle dataset.

Activation Function	Epochs converged	Converged
ReLU	55	Yes
Tanh	90	Yes
Sigmoid	250	Yes
Linear	∞	No

Table 1.2 Results obtained for changing the Activation Function in the neural network shown in Figure 1.1, using the Exclusive or dataset.

Activation Function	Epochs converged	Converged
ReLU	∞	No
Tanh	∞	No
Sigmoid	∞	No
Linear	∞	No

Table 1.3 Results obtained for changing the Activation Function in the neural network shown in Figure 1.1, using the Spiral dataset.

Activation Function	Epochs converged	Converged
ReLU	10	Yes
Tanh	15	Yes
Sigmoid	20	Yes
Linear	15	Yes

Table 1.4 Results obtained for changing the Activation Function in the neural network shown in Figure 1.1, using the Gaussian dataset.

From the above results in Tables: 1.1, 1.2, 1.4, and 1.3, it was noted that generally ReLU is the best performing function which enables the neural network to converge the fastest on the given datasets. Additionally, it was also noted that Tanh was the second best performing function, followed by Sigmoid, then Linear. It was generally noted, that the best three activation functions always converged, except for the instance of the Spiral dataset, where neither of the models could fit the dataset, due to the fact that only a single hidden layer is being used. Moreover, it was also noted that the Linear activation function did not converge for the Circle and Exclusive or datasets, as the model could not fit the data. However, this was not the case for the Gaussian dataset, as all of the activation functions managed to enable the created neural network to converge.

1.2 Task 2: Exploring Learning Rates

1.2.1 Question

1. Use a pre-set neural network architecture for this task.
2. Experiment with learning rates.
3. Objective:
    - Record the number of epochs it takes to converge for each learning rate.
    - Note any oscillations or overshooting in the loss graph for high learning rates.

1.2.2 Answer

The pre-set neural network architecture was set to have the following properties:

1. The default number of hidden layers was set to two
2. Activation Function set to ReLU
3. Problem Type set to Classification
4. Regularisation set to None
5. Dataset set to Circle

Learning Rate	Epochs converged	Converged	Oscillating
10	∞	No	Yes
3	∞	No	Yes
1	26	Yes	No
0.3	19	Yes	No
0.1	20	Yes	No
0.03	50	Yes	No
0.01	100	Yes	No
0.003	300	Yes	No
0.001	700	Yes	No
0.0001	∞	No	No
0.00001	∞	No	No

Table 1.5 Results obtained for changing the Learning Rate in pre-set neural network architecture, using the Circle dataset and ReLU activation function.

From the results obtained in Table 1.5, it was noted that learning rates between 0.1 and 1 present the best learning rates given the specified scenario and Circle dataset. Additionally, it was also noted that high learning rates which were larger than 1 tended to overshoot and displayed oscillations in the output graph. It was also noted that a small learning rate such as 0.0001 and high learning rates did not enable the neural network to converge.

1.3 Task 3: Playing with Layers and Neurons

1.3.1 Question

1. Start with a neural network that has one hidden layer.
2. Incrementally add more layers and neurons to the hidden layers.
3. Objective:
    - Compare the decision boundary's complexity when using few layers and neurons versus
    - more layers and neurons.
    - Note the number of epochs it takes for the model to converge as you increase the complexity

1.3.2 Answer

The pre-set neural network architecture was set to have the following properties:

1. The learning rate was set to 0.01
2. Activation Function set to ReLU
3. Problem Type set to Classification
4. Regularisation set to None
5. Dataset set to Circle

Hidden Layers	Number of Neurons	Epochs converged	Converged
1	1	∞	No
1	3	100	Yes
1	5	58	Yes
2	3	70	Yes
2	5	61	Yes
3	3	100	Yes
3	5	60	Yes

Table 1.6 Results obtained for changing the number of hidden layers in pre-set neural network architecture, using the Circle dataset and ReLU activation function.

From the results obtained in Table 1.6, it was noted that with less hidden layers and neurons, the decision boundary is simple and may struggle to converge, when trying to learn complex data patterns. On the other hand, a large amount of hidden layers and neurons in the neural network, would enable the model capacity to learn complex data patterns resulting in a decision boundary that is more complex and can better suit the training data. Additionally, it was also noted that with more layers and neurons, the number of epochs required to converge is less than when compared to the same neural network with less neurons and layers. Furthermore, when only one neuron was utilised, the decision boundary was only a straight line; therefore, more neurons and hidden layers are needed to make decision borders more complicated.

1.4 Task 4: Optimisation Algorithms

1.4.1 Question

1. Use a consistent neural network architecture for this task.
2. Change the optimisation algorithm from SGD to others like AdaGrad or Momentum.
3. Objective:
    - Record the number of epochs it takes for the model to converge using different
    - optimisation algorithms.
    - Note if specific algorithms are more sensitive to initial conditions or if they plateau more
    - easily.

1.4.2 Answer

Plagiarism Form

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

Declaration

Plagiarism is defined as "the unacknowledged use, as one's own work, of work of another person, whether or not such work has been published" (Regulations Governing Conduct at Examinations, 1997, Regulation 1 (viii), University of Malta).

I /~~We~~* , the undersigned, declare that the [assignment / Assigned Practical Task report / Final Year Project report] submitted is my /~~our~~* work, except where acknowledged and referenced.

I /~~We~~* understand that the penalties for making a false declaration may include, but are not limited to, loss of marks; cancellation of examination results; enforced suspension of studies; or expulsion from the degree programme.

Work submitted without this signed declaration will not be corrected, and will be given zero marks.

* Delete as appropriate.

(N.B. If the assignment is meant to be submitted anonymously, please sign this form and submit it to the Departmental Officer separately from the assignment).

StudentName	Signature
StudentName	Signature
StudentName	Signature
Matthias Bartolo StudentName	 Signature

ARI3129 CourseCode	Advanced Computer Vision Individual Assignment Title of work submitted

17 October 2023
Date