Computer Science NEA Report

An investigation into how Artificial Neural Networks work, the effects of their hyper-parameters and their applications in Image Recognition.

Max Cotton

Contents

1 Introduction

Artificial Intelligence is a branch of Computer Science that attempts to mimic human cognition in order to perform tasks, understand data and predict outcomes [3]. Machine Learning is a subfield of Artificial Intelligence that uses statistical algorithms, which take as an input training datasets, to produce mathematical models that allow prediction of the outcome of unseen datasets. Deep Learning is a further subfield of Machine Learning that uses Artificial Neural Networks, a process of learning from data inspired by the human brain. Artificial Neural Networks can be trained to operate on, "learn", a vast number of problems, such as Image Recognition, and have uses across multiple fields, such as medical imaging in hospitals.

1.1 Project Aims

This project is an investigation into how Artificial Neural Networks work, the effects of changing the hyper-parameters (such as the shape of the network and the learning rate) used to tune the models, and particularly their applications in Image Recognition. To achieve this, I have derived and researched all the fundamental theory behind the project, using sources such as IBM's online research [4], and developed Neural Networks from first principles without the use of any third-party Machine Learning libraries. I have then implemented the Artificial Neural Networks in the domain of Image Recognition, by creating trained models and have allowed for experimentation in varying the hyper-parameters of each model to provide for comparison and experimentation between different model performances. A Graphical User Interface has been developed to provide a mechanism for a researcher, or interested student, to train and test models and alter key hyper-parameters to explore the effect on performance and results.

1.2 Overview

Developing an Artificial Neural Network has required understanding the fundamental theoretical maths and algorithms underpinning this important technology which has been exciting and challenging. In implementing the network I have focused on an object-orientated approach utilising design approaches such as encapsulation and suitable data structures such as doubly linked lists. I have also researched and set up a development environment utilising tools such as GitHub, Jupyter Notebook, Visual Studio Code, LaTex and automated unit testing etc. At the start of the project, I was lucky enough to interview an expert in the field of Neural Networks and AI and his guidance was valuable in setting the scope of the project and suggesting properties of the network to test. I am particularly pleased with being able to implement the network from the first principle maths. The main learning I have taken from this project is that Artificial Neural Networks require tuning to best address a particular problem - they are not a one-size-fits-all solution - and that this is a combination of research, experimentation and experience.

2 Analysis

2.1 Theory Behind Artificial Neural Networks

From an abstract perspective, Artificial Neural Networks are inspired by the anatomy of the human brain, consisting of layers of 'neurons' all interconnected via different links, 'axons with connecting synapses', each with their own strengths, 'weights'. By adjusting these links and weights, Artificial Neural Networks can be trained to take in an input and give its best prediction as an output [4].

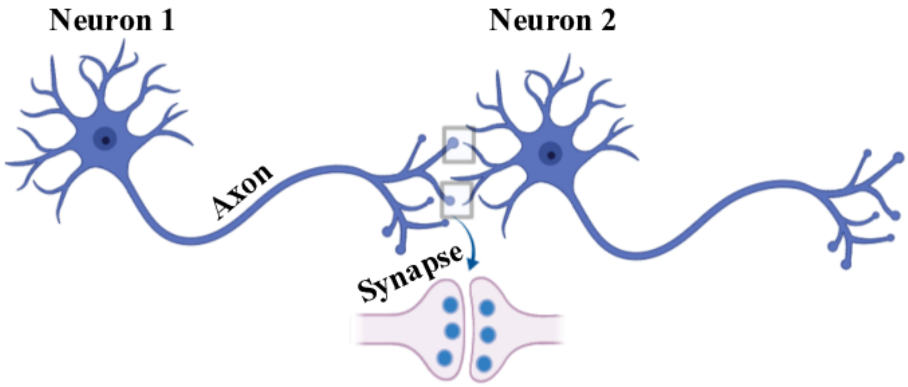


Figure 1: Two connected biological neurons sourced from https://www.researchgate.net [7]

2.1.1 Structure

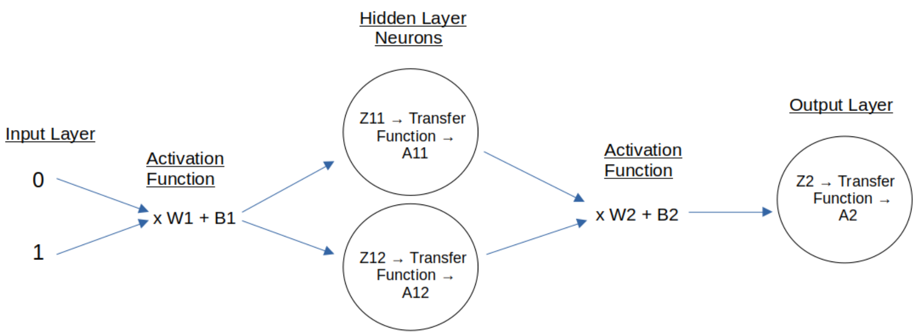


Figure 2: This shows an Artificial Neural Network with one single hidden layer and is known as a Shallow Neural Network.

I have focused on Feed-Forward Artificial Neural Networks, where values are entered to the input layer and passed forwards repetitively to the next layer until reaching the output layer. Within this, I have investigated two types of Feed-Forward Artificial Neural Networks: Perceptron Artificial Neural Networks, that contain no hidden layers and are suitable for addressing simple linear problems, and Multi-Layer Perceptron Artificial Neural Networks, that contain at least one hidden layer. The expanded complexity of the Multi-Layer Perceptron Artificial Neural Network increases the non-linearity in the Artificial Neural Network allowing it to address more complex / non-linear problems.

Multi-Layer Perceptron Artificial Neural Networks consist of:

- An input layer of input neurons, where the input values are entered.
- Hidden layers of hidden neurons.
- An output layer of output neurons, which outputs the final prediction.

To implement an Artificial Neural Network, matrices are typically used to represent the layers, where each layer is a matrix of the layer's neuron's values. In order to use matrices for this, the following basic theory must be understood about them [6]:

- When Adding two matrices, both matrices must have the same number of rows and columns. Alternatively, a single column matrix with the same number of rows can be added by element-wise addition where each element is added to all of the elements of the corresponding rows in the associated matrix.
- In order to multiply matrices, the 'dot product' of the matrices is computed which multiplies the rows of one matrice with the columns of the other, multiplying matching members and then summing up.
- When calculating the dot product of two matrices, the number of columns of the 1st matrix must equal the number of rows of the 2nd matrix. The resulting matrix will have the same number of rows as the 1st matrix, and the same number of columns as the 2nd matrix. This is important in implementing an Artificial Neural Network, as the output of one layer must be formatted correctly to be used with the next layer.
- Alternatively, the Hadamard product of two matrices can be utilised which performs element-wise multiplication of the matrices. For this, both matrices must have the same number of rows and columns.
- Transposing a matrix is also utilised which switches all rows of the matrix into columns and all columns into rows in an output matrix.
- A matrix of values can be classified as a rank of Tensors, depending on the number of dimensions of the matrix. (Eg: A 2-dimensional matrix is a Tensor of rank 2)

I have focused on using Fully-Connected layers, that input values and apply the following functions to produce an output value for the layer:

- **Activation function** which calculates the dot product of an input matrix with a weight matrix, then sums the result with a bias matrix.

- **Transfer function** which takes the result of the activation function and calculates a suitable output value as well as adding non-linearity to the Neural Network. For example, the Sigmoid Transfer function converts the input value to an output number between zero and one - this makes it useful for logistic regression where the output value can be rounded to allow for a binary classification.

2.1.2 How Artificial Neural Networks learn

To train an Artificial Neural Network, the following processes are carried out for each of a number of training iterations called epochs:

- Forward Propagation which is the process of feeding inputs in and getting a prediction (moving forward through the network).
- Back Propagation, the process of calculating the error, known as the loss, in the prediction and then adjusting the weights and biases accordingly.

I have used Supervised Learning to train the Artificial Neural Networks, where the output prediction of the Artificial Neural Network is compared to the theoretical values it should have predicted. With this, I can calculate the loss value of the prediction. I then move back through the network and update the weights and biases via Gradient Descent which aims to reduce the Loss value of the prediction to a minimum, by subtracting the rate of change of Loss with respect to the weights / biases, multiplied with a learning rate, from the weights / biases. To calculate the rate of change of Loss with respect to the weights / biases, I used the following calculus methods:

- Partial Differentiation, which allows differentiation of multi-variable functions, by differentiating with respect to one variable and considering the rest as constants.
- The Chain Rule, where for $y = f(u)$ and $u = g(x)$, $\frac{\partial y}{\partial x}$ can be calculated as: $\frac{\partial y}{\partial x} = \frac{\partial y}{\partial u} * \frac{\partial u}{\partial x}$.

This repetitive process will continue to reduce the Loss to a minimum, if the learning rate is set to an appropriate value

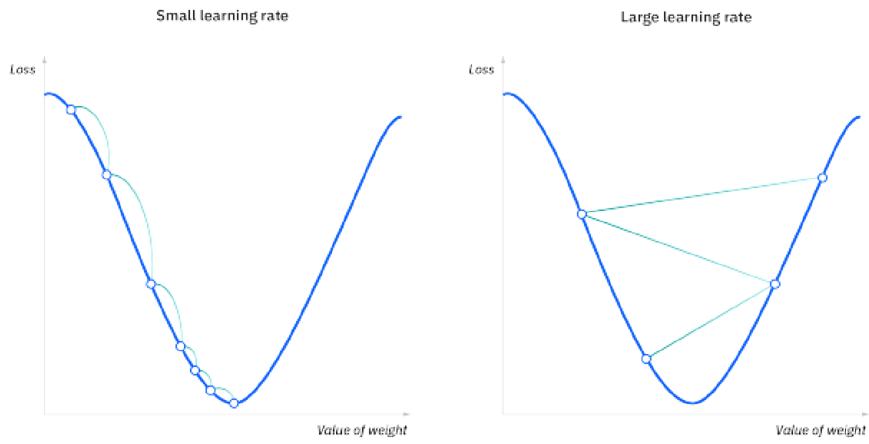


Figure 3: Gradient Descent
sourced from https://www.ibm.com/topics/gradient-descent [4]

During backpropagation, however, some issues can occur, such as the following:

- Finding a false local minimum rather than the global minimum of the function
- Having an 'Exploding Gradient' [3], where the gradient value grows exponentially to the point of overflow errors
- Having a 'Vanishing Gradient' [3], where the gradient value decreases to a very small value or zero, resulting in a lack of updating values during training.

2.2 Theory Behind Deep Artificial Neural Networks

2.2.1 Network Architecture and Training

Figure 4 below shows a simplified view of an Artificial Neural Network with multiple hidden layers, known as a Deep Neural Network, where:

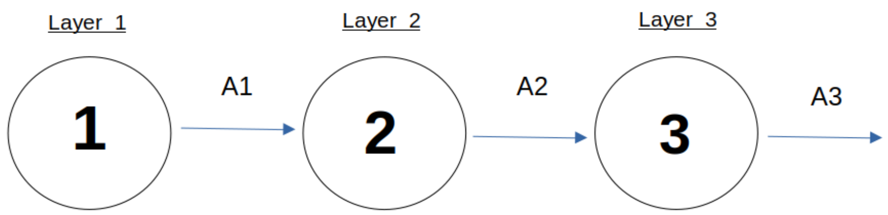


Figure 4: Showing an abstracted view of an Artificial Neural Network with multiple hidden layers.

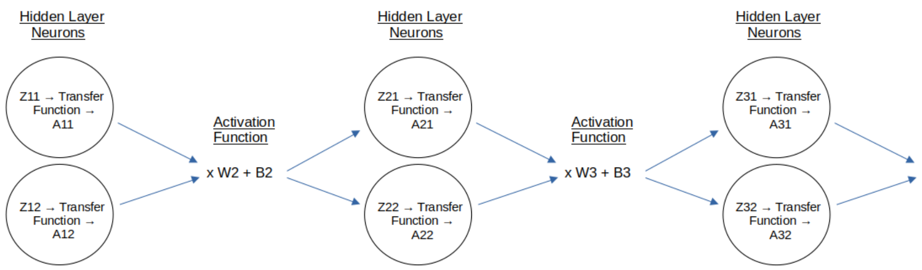


Figure 5: This shows an Artificial Neural Network with multiple hidden layers and is known as a Deep Neural Network.

Figure 5 shows the network in more detail. The layers have inputs x, weight matrix wN and biases BN.

- Each layer takes the previous layer's output as its input X
- An Activation function is applied to X to obtain Z, by taking the dot product of X with a weight matrix W, and sums the result with a bias matrix B. At first when the Neural Network is initialised pre-training the weights are initialised to random values and the biases are set to zeros. $Z = W * X + B$
- A Transfer function is then applied to Z to obtain the layer's output A
  - For the output layer, the sigmoid function, explained in section 2.1, can be used for either for binary classification via logistic regression, or for multi-class classification where the output neuron, and the associated class, that has the highest value between zero and one is selected.
    * Where $sigmoid(Z) = \frac{1}{1+e^{-Z}}$
  - However, for the input layer and the hidden layers, another transfer function known as ReLu (Rectified Linear Unit) can be better suited as it produces larger values of $\frac{\partial L}{\partial W}$ and $\frac{\partial L}{\partial B}$ for Gradient Descent than Sigmoid [6], so updates at a quicker rate.
    * Where $relu(Z) = max(0, Z)$

2.2.2 Training

Training takes place through a series of forward and backward propagations called epochs. The forward propagation generates a potential output and associated error known as the loss. Backward propagation then adjusts the weights and biases in an attempt to reduce this loss.

2.2.3 Forward Propagation:

- For each epoch the input layer is presented with a matrix of input values, which are fed through the network to obtain a final prediction A from the output layer.
- Using the process described in the previous section, backpropagation trains the network by adjusting weights and biases.

2.2.4 Back Propagation:

- The Loss value L is first calculated using the following Log-Loss function shown in the equation below, which calculates the average difference between A and the value it should have predicted Y. The average is then found by summing the result of the Loss function [1] for each value in the matrix A, then dividing by the number of predictions m, resulting in a Loss value indicating how well the network is performing. Where
  $L = -(\frac{1}{m}) * \sum(Y * log(A) + (1 - Y) * log(1 - A))$
  and "log()" is the natural logarithm.
- The network is then trained by moving back through the layers, adjusting the weights and biases via Gradient Descent. For each layer, the weights and biases are updated with the following formulae:
  - $W = W - learningRate * \frac{\partial L}{\partial W}$
  - $B = B - learningRate * \frac{\partial L}{\partial B}$
- The derivation for Layer 2's $\frac{\partial L}{\partial W}$ and $\frac{\partial L}{\partial B}$ is shown below:
  - Functions used so far:
    1. $Z = W * X + B$
    2. $A_{relu} = max(0, Z)$
    3. $A_{sigmoid} = \frac{1}{1+e^{-Z}}$
    4. $L = -(\frac{1}{m}) * \sum(Y * log(A) + (1 - Y) * log(1 - A))$
  - $\frac{\partial L}{\partial A2} = \frac{\partial L}{\partial A3} * \frac{\partial A3}{\partial Z3} * \frac{\partial Z3}{\partial A2}$
    By using function 1, where A2 is X for the 3rd layer, $\frac{\partial Z3}{\partial A2} = W3$
    $=> \frac{\partial L}{\partial A2} = \frac{\partial L}{\partial A3} * \frac{\partial A3}{\partial Z3} * W3$
  - $\frac{\partial L}{\partial W2} = \frac{\partial L}{\partial A2} * \frac{\partial A2}{\partial Z2} * \frac{\partial Z2}{\partial W2}$
    By using function 1, where A1 is X for the 2nd layer, $\frac{\partial Z2}{\partial W2} = A1$
    $=> \frac{\partial L}{\partial W2} = \frac{\partial L}{\partial A2} * \frac{\partial A2}{\partial Z2} * A1$

- $\frac{\partial L}{\partial B2} = \frac{\partial L}{\partial A2} * \frac{\partial A2}{\partial Z2} * \frac{\partial Z2}{\partial B2}$
  By using function 1, $\frac{\partial Z2}{\partial B2} = 1$
  $=> \frac{\partial L}{\partial W2} = \frac{\partial L}{\partial A2} * \frac{\partial A2}{\partial Z2} * 1$

- As can be seen, when moving back through the network, $\frac{\partial L}{\partial W}$ and $\frac{\partial L}{\partial B}$ can be calculated for each layer using the rate of change of loss with respect to its output. This is calculated by the previous layer using the above formula; the derivative of the layer's transfer function, and the layers input (which in this case is A1)
  - Where by using function 2, $\frac{\partial A_{relu}}{\partial Z} = 1$ when $Z >= 0$ otherwise $\frac{\partial A_{relu}}{\partial Z} = 0$
  - Where by using function 3, $\frac{\partial A_{sigmoid}}{\partial Z} = A * (1 - A)$
- At the start of backpropagation, the rate of change of loss with respect to the output layer's output has no previous layer's calculations, so instead it can be found with the derivative of the Log-Loss function, as shown in the following:
  - Using function 4, $\frac{\partial L}{\partial A} = (-\frac{1}{m})(\frac{Y-A}{A*(1-A)})$

2.3 Theory behind training the Artificial Neural Networks

Training an Artificial Neural Network's weights and biases to predict on a dataset creates a trained model for that dataset. This model can be used to create predictions based on future data / images inputted. However, training Artificial Neural Networks suffers problems such as Overfitting, where the trained model learns the patterns of the training dataset too well, resulting in poor predictions on new datasets. This can occur when the training dataset does not cover enough situations of inputs and the desired outputs (by being too small for example), if the model is trained for too many epochs on the poor dataset or by having too many layers in the Neural Network.

Another common problem is Underfitting, where the model has not learnt the patterns of the training dataset well enough, often when it has been trained for too few epochs, or when the Neural Network is too linear.

2.4 Datasets

I have utilised a series of open source datasets on which to train and test by Neural Networks.

2.4.1 XOR dataset

As a first step in developing and testing Artificial Neural Networks, I have utilised the XOR gate problem, where the Neural Network is fed input pairs of zeros and ones and learns to predict the output of a XOR gate used in circuits. This takes far less computation time than image datasets and was extremely useful in debugging. It is a good example of a relatively simple problem whilst not being linearly separable.

2.4.2 MNIST dataset

The MNIST dataset [9] is a well known dataset consisting of images of handwritten digits from zero to ten and is commonly used to test the performance of an Artificial Neural Network. The dataset consists of 60,000 input images representing single digits made up from 28x28 pixels with each pixel having a value from 0 to 255. To format the images into a suitable format to be inputted into the Artificial Neural Networks, each image's matrix of pixel values is commonly 'flattened' into a 1 dimensional matrix of values, where each element is divided by 255 (the maximum 8 Bit value) to produce a number between 0 and 1, to standardize the dataset. The output dataset is also loaded which represents the actual value of the number in the image. This is commonly implemented by using an array for each image where the index of the array represesents the number in an image (i.e. a 1 in column 2 could represent a 2, or a 1 if zero indexed).

To create a trained Artificial Neural Network model on this dataset, the model requires 10 output neurons (one for each digit). The Sigmoid Transfer function is then utilized to output a number between one and zero to each neuron - whichever neuron received the highest value is selected as the predicted outcome. This an example of a multi-class classification, where the model must predict one of 10 classes (in this case, each class is one of the digits from zero to ten).

2.4.3 Cat dataset

I have also used a dataset of images of cats sourced from https://github.com/marcopeix [5], where each image is classified as either a cat or not a cat. The dataset consists of 209 input images, made up from 64x64 pixels with each pixel having an value from 0 to 255. To normalise the images into a suitable format to be input into the Artificial Neural Network, each image's matrix of values is 'flattened' into a 1 dimensional matrix of values, where each element is divided by 255 (the maximum 8 bit value) to a number between 0 and 1, to standardize the dataset.

The output dataset represents an array of binary values representing the output of each image (1 for a cat, 0 for not a cat). To create a trained Artificial Neural Network model on this dataset, the model requires only 1 output neuron - representing the chance of being a cat. By using the Sigmoid Transfer function, a number is outputted between one and zero for the neuron, if the neuron's value is closer to 1 it predicts a cat, otherwise it predicts not a cat - this is binary classification.

2.4.4 Theory behind using Graphics Cards to train Artificial Neural Networks

Graphics Cards have been designed essentially to undertake parallel matrix computations utilising many Tensor cores - which are processing units specialised for matrix operations calculating the co-ordinates of 3D graphics, however they can be used for operating on the matrices in the network at a much faster speed compared to CPUs. GPUs also include CUDA cores which act as an API to the GPU's computing to be used for any operations (in this case training the Artificial Neural Networks).

2.5 Interview

In order to gain a better foundation for my investigation, and ensure my project would allow a user interested in Artificial Neural Networks to experiment with the fundamentals of the network and conduct experiments, I presented my prototype code and interviewed the head of Artificial Intelligence at Cambridge Consultants - Will Addison. These were their responses:

- Q:"Are there any good resources you would recommend for learning the theory behind how Artificial Neural Networks work?"

  A:"There are lots of useful free resources on the internet to use. I particularly like the platform 'Medium' which offers many scientific articles as well as more obvious resources such as IBMs'."

- Q:"What do you think would be a good goal for my project?"

  A:"I think it would be great to aim for applying the Neural Networks on Image Recognition for some famous datasets. For you, I would recommend the MNIST dataset as a goal."

- Q:"What features of the Artificial Neural Networks would you like to be able to experiment with?"

  A:"I'd like to be able to experiment with the number of layers and the number of neurons in each layer, and then be able to see how these changes effect the performance of the model. I can see that you've utilised the Sigmoid transfer function and I would recommend having the option to test alternatives such as the ReLu transfer function, which will help stop issues such as a vanishing gradient."

- Q:"What are some practical constraints of AI?"

  A:"Training AI models can require a large amount of computing power, also large datasets are needed for training models to a high accuracy which can be hard to obtain."

- Q:"What would you say increases the computing power required the most?"

  A:"The number of layers and neurons in each layer will have the greatest effect on the computing power required. This is another reason why I recommend adding the ReLu transfer function as it updates the values of the weights and biases faster than the Sigmoid transfer function."

- Q:"Do you think I should explore other computer architectures for training the models?"

  A:"Yes, it would be great to add support for using graphics cards for training models, as this would be a vast improvement in training time compared to using just CPU power."

- Q:"I am also creating a user interface for the program, what hyper-parameters would you like to be able to control through this?"

A:"It would be nice to control the transfer functions used, as well as the general hyper-parameters of the model. I also think you could add a progress tracker to be displayed during training for the user."

- Q:"How do you think I should measure the performance of models?"

  A:"You should show the accuracy of the model's predictions, as well as example incorrect and correct prediction results for the trained model. Additionally, you could compare how the size of the training dataset effects the performance of the model after training, to see if a larger dataset would seem beneficial."

- Q:"Are there any other features you would like add?"

  A:"Yes, it would be nice to be able to save a model after training and have the option to load in a trained model for testing."

2.6 Project Objectives and Requirements

Based on the interview above in section 2.5, the following high-level objectives were formulated:

2.6.1 Objectives

Objective ID	**Description**
1	Learn how Artificial Neural Networks work and develop them from first principles
2	Implement the Artificial Neural Networks by creating trained models based on image datasets
2.1	Allow use of Graphics Cards for faster training
2.2	Allow for the saving and loading of trained models
3	Develop a Graphical User Interface
3.1	Provide controls for hyper-parameters of models
3.2	Display and compare the results each model's predictions

2.6.2 Requirements

The following sets out the steps that must be taken to accomplish the above objectives:

ID	Description	Satisfied by	Tested by
1	Learn how Artificial Neural Networks work	Page 4	N/A
2	Develop Artificial Neural Networks from first principles		
2.1	Provide utilities for creating Artificial Neural Networks	Page 37	Page 86
2.2	Allow for the saving and loading of trained models' weights and biases	Page 41	Page 86
2.3	Allow use of Graphics Cards for faster training	Code not included in report	Page 123
3	Implement the Artificial Neural Networks on image datasets		
3.1	Allow input of unique hyper-parameters	Page 48	Page 105
3.2	Allow unique datasets and train dataset size to be loaded	Page 48	Page 112
4	Use a database to store a model's features and the location of its weights and biases	Page 64	Page 81
5	Develop a Graphical User Interface		
5.1	Provide controls for hyper-parameters of models	Page 53	Page 73
5.2	Display details of models' training	Page 53	N/A
5.3	Display the results of each model's predictions	Page 96	User Tested
5.4	Allow for the saving of trained models	Page 96	Page 79
5.5	Allow for the loading of saved trained models	Page 59	Page 78

3 Design

3.1 Introduction

The following design focuses have been made for the project:

- The program will support multiple platforms to run on, including Windows and Linux.
- The program will use python3 as its main programming language.
- An object-orientated approach is used in the design of the project.
- An option will be given to use either a Graphics Card or a CPU to train and test the Artificial Neural Networks.
- SysML will be utilised in the design of the architecture and class diagrams

3.2 System Architecture

The project is architected using object-orientated design principles, it was then implemented in Python. A SysML block diagram [2] showing the key architectural components is shown in the figure below with detailed SysML class diagrams shown in section 3.3.1.

As shown in the Figure 6 the User interacts with the software through a User Interface (UI) which is written in tkinter. The UI classes interface with the Model block which contains all the classes necessary for representing the Artificial Neural Network models. As such functional separation is maintained between the user interface elements and the Neural Network representation.

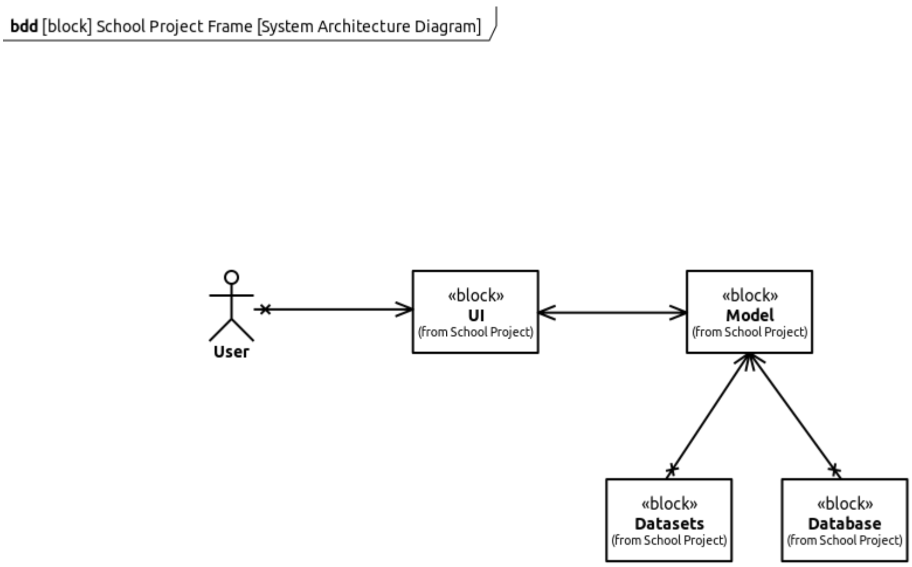


Figure 6: High level system architecture

3.3 Class Diagrams

3.3.1 UI Class Diagram

The classes utilised to implement the UI are shown in Figure 7. As can be seen the School_Project_Frame inherits the tk.Frame class - the tkinter base class which provides the basic frame functionality. The School_Project_Frame then contains a series of UI interaction specific Frames.

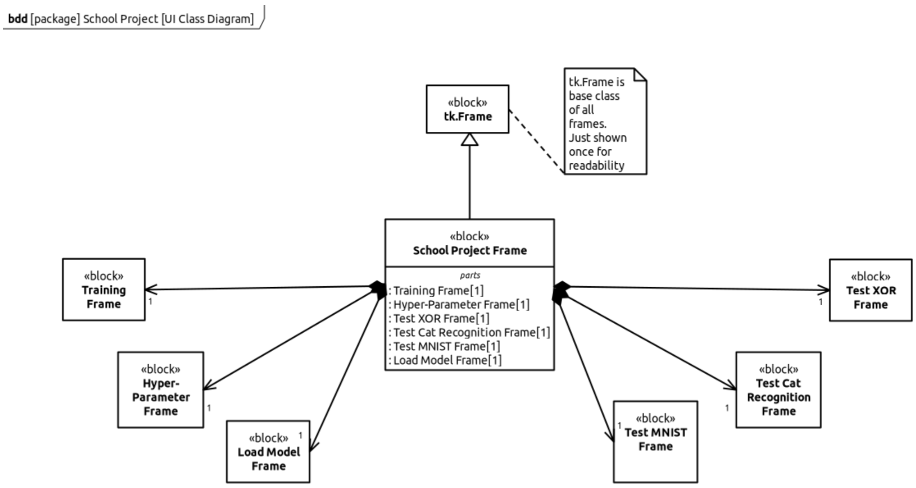


Figure 7: UI SysML Class Diagam

3.3.2 Model Class Diagram

The Model class diagram is shown below in figure 8. An AbstractModel class contains a linked list of FullyConnectedLayer classes which represent the layers within the Artificial Neural Network.

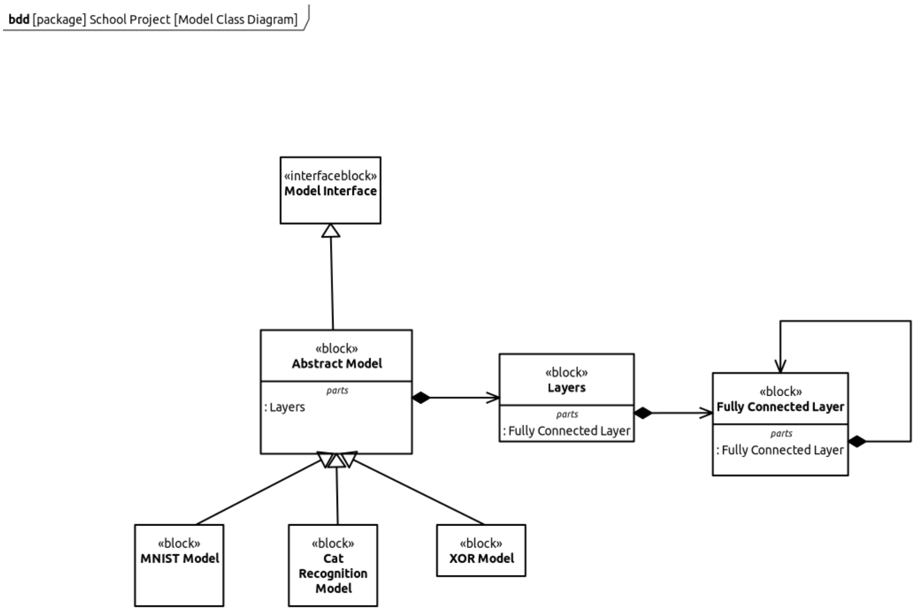


Figure 8: Artificial Neural Network and Layers SysML Class Diagram

3.4 System Flow chart

The system flow is shown in Figure 9 below. When the program is initialised the Home Frame is first displayed. The left-hand route - *Hyper-Parameter Frame* and *Training Frame* - provide the functionality to set hyper-parameters and train a new model. This model can be saved and loaded, following the right-hand route, if desired before entering the *Testing Frame* where the model can be applied to testing data and the model performance assessed. The ability to save models is important as it allows for model training on a computer with a GPU and utilisation on a non-GPU-accelerated computer. This allowed an efficient build test cycle as I could build models on a desk-based, GPU accelerated PC, and undertake testing / demonstration on a laptop.

3.5 Algorithm implementation

The detail of algorithms underpinning the Artifical Neural Network are outlined in section 2.1. To implement the algorithm within the project each layer within the Artificial Neural Network is represented as an instance of a *FullyConnectedLayer* class each of which:

- Contains a weight and bias matrix.
- Has a *transfer_function* which be selected by the user to be either Sigmoid or ReLu but can be easily extended to other functions.
- Has a *forward_propagation* function which takes an input matrix which it then multiplies by the weight matrix and sums the result with the bias matrix, this is then put through the *transfer_function*.

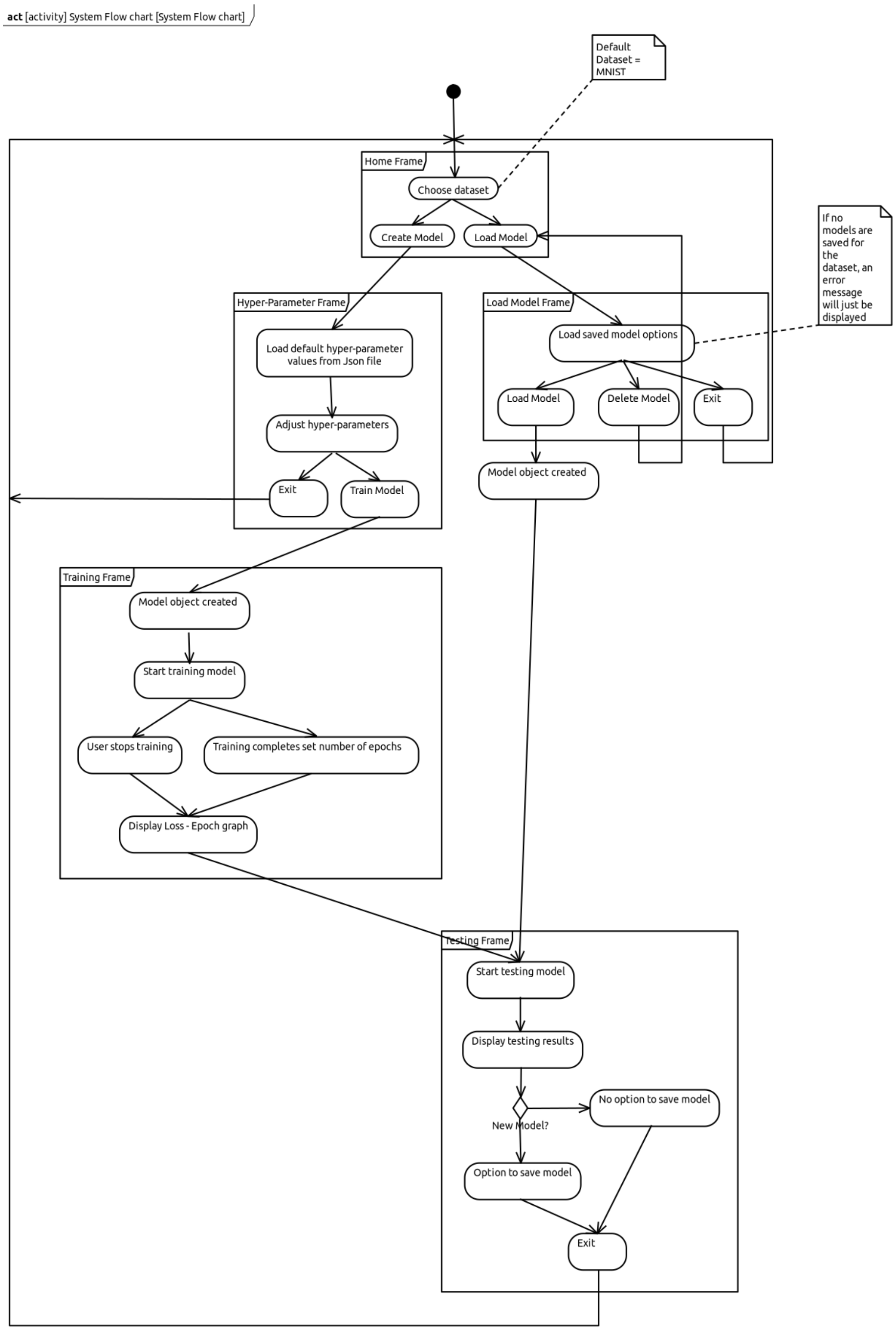


Figure 9: System flow chart

- Has a *backward_propagation* function which passes the loss with respect to each layers output allowing adjustment of the weights and biases.

Each *FullyConnectedLayer* class is held within a doubly linked list contained within the *Abstract Model Class*, which is a parent class of the data-specific classes *MNIST Model*, *Cat Recognition Model* and *XOR Model*. The use of a doubly linked list allows for a user defined number of hidden layers from 0 upwards and allows easy propagation of the list.

The algorithm to train the Artificial Neural Network iterates over a user-defined number of epochs. Each epoch starts with the presentation of data to the input layer. The doubly linked list of layers is then traversed calling the *forward_propagation* function on each. When the output layer is reached the derivative of the *Log-Loss* function is calculated alongside the absolute *Loss* for debugging and learning rate plotting.

Following the calculation of the derivative of the *Log-Loss* function, the doubly linked list is then iterated in reverse back through the layers calculating the loss with respect to the weights and biases of each layer. This process is then used to adjust the weights and biases in each layer using the specified learning rate in an attempt to reduce the loss value.

3.6 Data Structures, Techniques and File Structures used

3.6.1 Data structures

The following data structures are utilised in implementing the Artificial Neural Network:

- Standard lists for storing data, for example storing the shape of the Artificial Neural Network's layers.
- Tuples where tuple unpacking is useful, such as returning multiple values from methods.
- Dictionaries for loading the default hyper-parameter values from a JSON file.
- Matrices to represent the layers and allow for a varied number of neurons in each layer. To represent the Matrices I will use both numpy arrays and cupy arrays.
- A Doubly linked list to represent the Artificial Neural Network, where each node is a layer of the network. This allows traversing both forwards and backwards through the network, as well as storing the first and last layer to start forward and backward propagation respectively.

3.6.2 Techniques

Some techniques used include:

- Object-oriented design including inheritance, abstract classes and interfaces
- Encapsulation
- Get methods
- Abstract and Static methods

- Decorators to wrap functions and modify behaviour
- tkinter for user interface design
- SQL for database access
- Type Hinting
- Docstrings
- Generators
- Breaking down the project into subpackages and modules
- Raising and Handling of Exceptions
- Threading

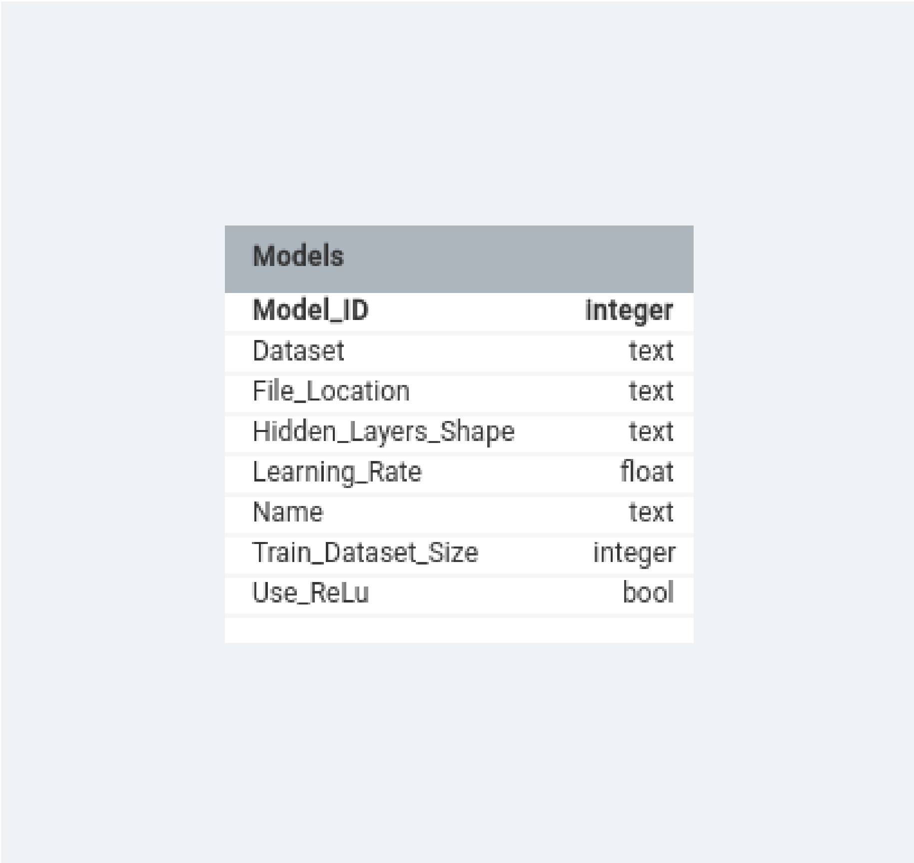
3.6.3 File Structure

I have used the following file structures to store necessary data for the program:

- A JSON file [8] for storing the default hyper-parameters for creating a new model for each dataset.
- Image dataset files are stored in either a compressed archive file (such as .pkl.gz files) or of the Hierarchical Data Format (such as .h5) for storing large datasets with fast retrieval.
- Weights and biases of saved models are stored as numpy arrays in .npz files (a zipped archive file format) in a 'saved-models' directory, which allows compatibility with the standard numpy library.

3.7 Database Design

The following Relational database design is used for saving models, where the dataset, name and features of the saved model (including the location of the saved models' weights and biases and the saved models' hyper-parameters) are saved. The Model_ID field is the primary key.



Models	
Model_ID	**integer**
Dataset	text
File_Location	text
Hidden_Layers_Shape	text
Learning_Rate	float
Name	text
Train_Dataset_Size	integer
Use_ReLu	bool

Figure 10: Database design

The following unique constraint is also used so that each dataset can not have more than one model with the same name:

```
UNIQUE (Dataset, Name)
```

The following constraint is applied to ensure no attribute is left empty:

```
NOT NULL
```

3.8 Queries

Below are some example queries used for interacting with the database:

- Query the names of all saved models for a dataset:

```
SELECT Name FROM Models WHERE Dataset=?;
```

- Query the file location of a saved model:

```
SELECT File_Location FROM Models WHERE Dataset=? AND Name=?;
```

- Query the features of a saved model:

```
SELECT * FROM Models WHERE Dataset=? AND Name=?;
```

3.9 Human-Computer Interaction

Below are the designs of each tkinter frame in the User Interface. The flow of the frames is described in Figure 9. The final implementation of the frames is shown in the technical solution section.

- The home Frame design which acts as the entry point to the program, the implemented version is shown in Figure 22.

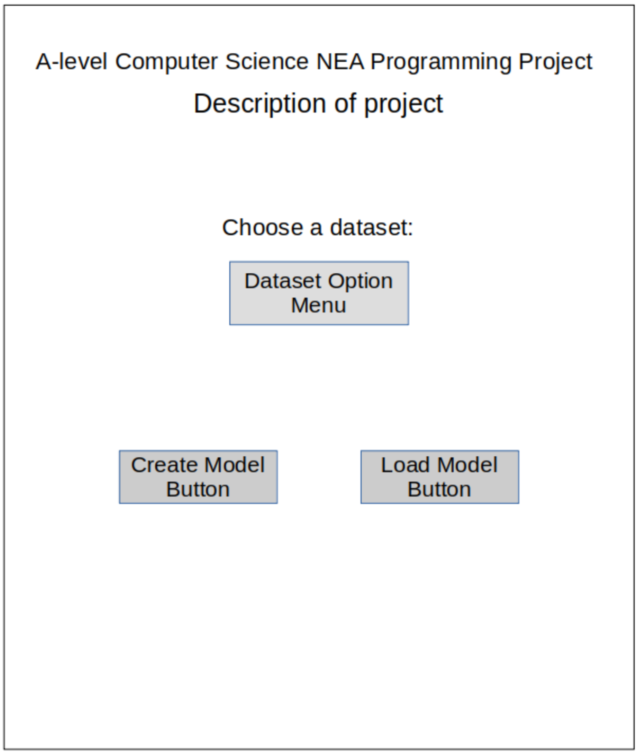


Figure 11: Home frame design - the entry point to the program

- Hyper-Parameter Frame design which allows setting of the Hyper-parameters, the implemented version is shown in Figure 17.

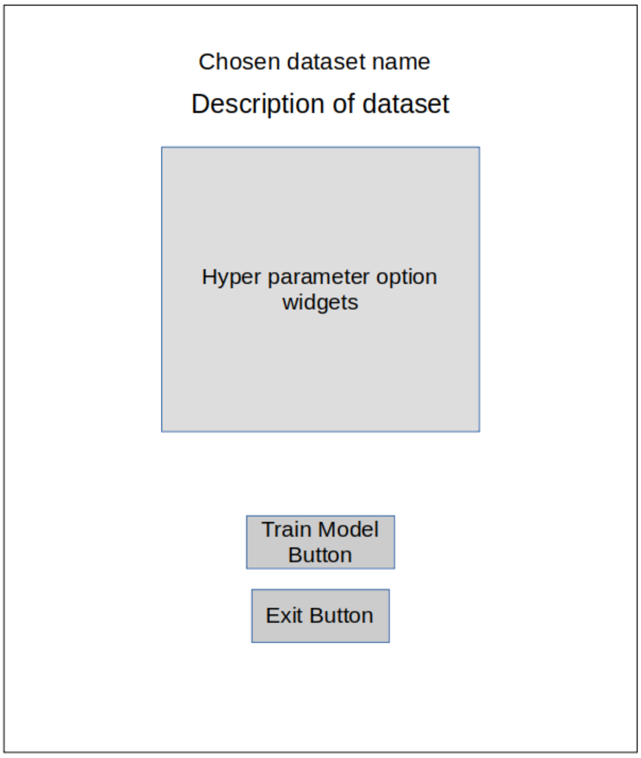


Figure 12: Hyper-Parameter Frame design which allows setting of parameters

- Training Frame design:
  - During training, the following is displayed on the Training Frame, the implemented version is shown in Figure 18

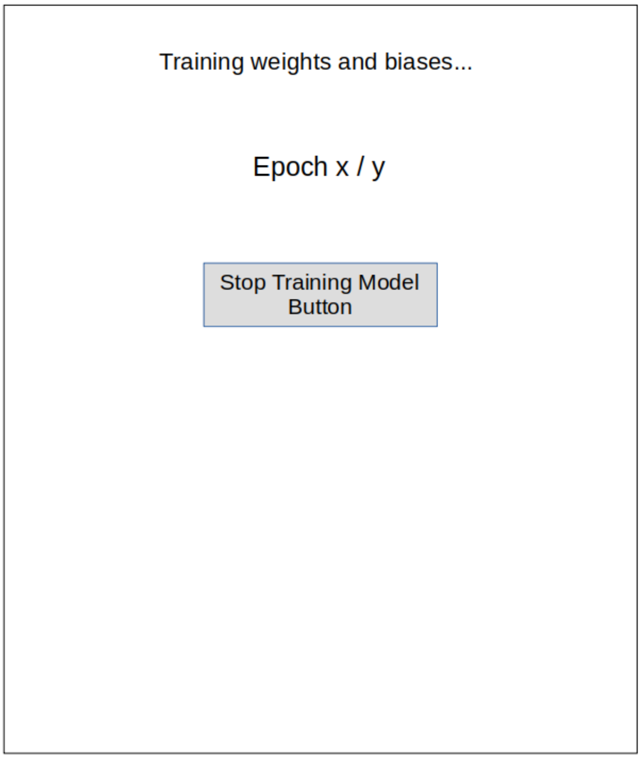


Figure 13: Training frame layout during training

– Once training has finished, the following is displayed on the Training Frame, the implemented version is shown in Figure 19.

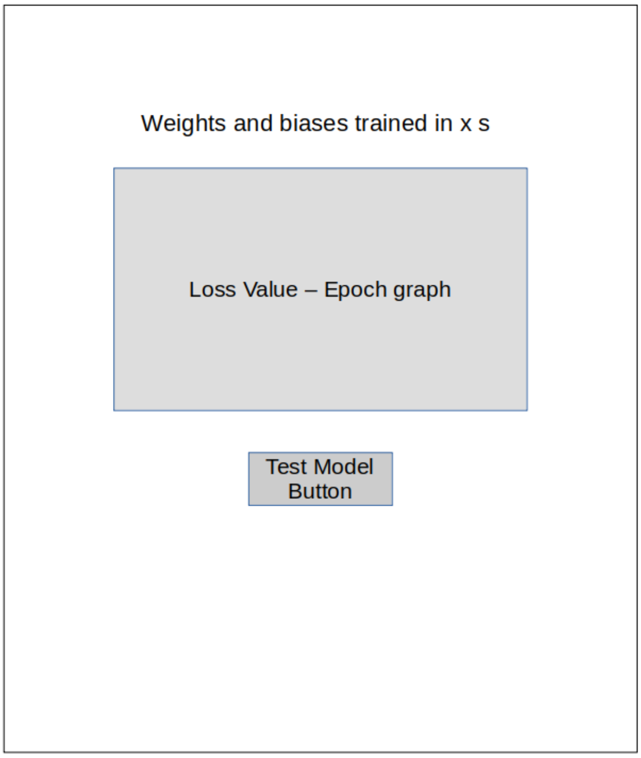


Figure 14: Training frame layout after training

- Load Model Frame design, the implemented version is shown in Figure 20.

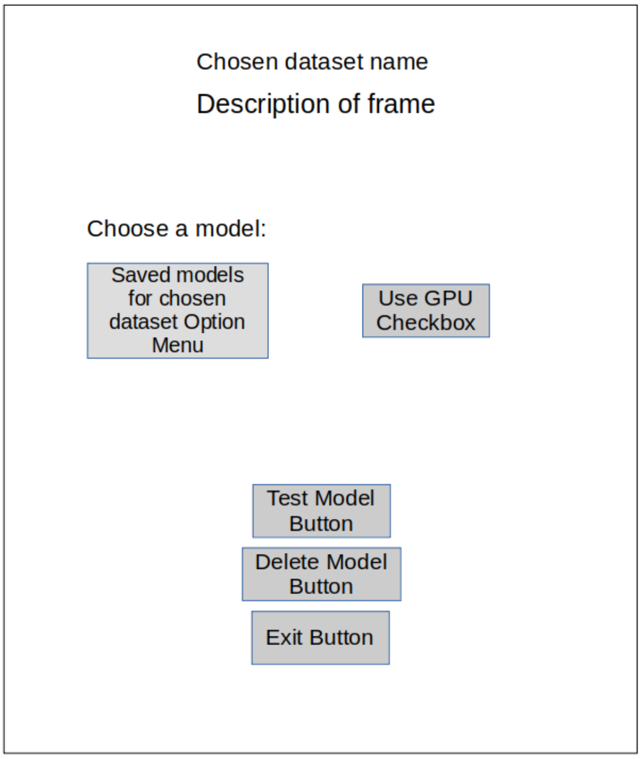


Figure 15: Load model frame design

- Test Frame design, the implemented version is shown in Figure 34.

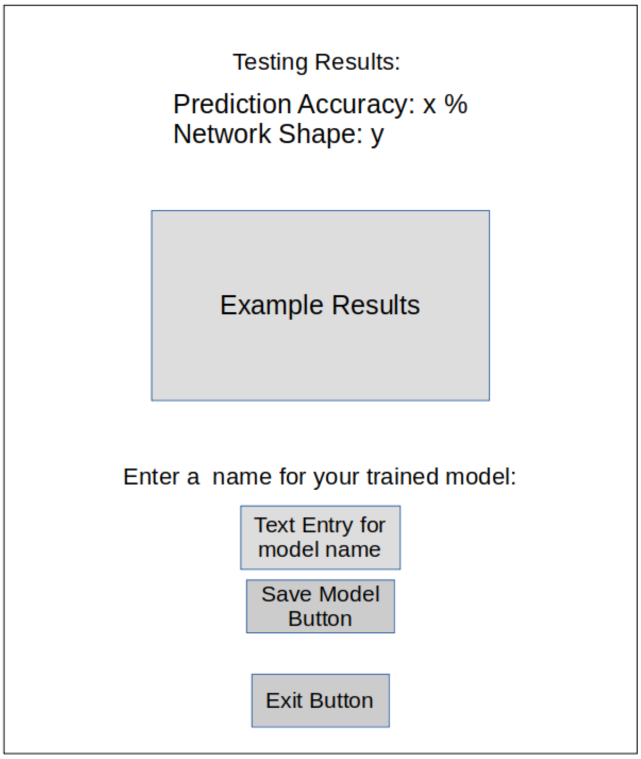


Figure 16: Test frame design

3.10 Hardware Design

To allow for faster training of an Artificial Neural Network, I gave the option to use a Graphics Card to train the Artificial Neural Network where available. I also gave the option to load pretrained weights to run on less computationally powerful hardware using just the CPU as standard. This was implemented by using the CuPy library.

3.11 Workflow and source control

Git and GitHub were used to manage the workflow and provide source control as the project was developed. In particular the following features were utilised:

- Commits and branches for adding features and fixing bugs separately.
- Using GitHub to back up the project as a repository.
- Automated testing on GitHub after each pushed commit.
- Providing the necessary instructions and information for the installation and usage of this project, as well as creating releases of the project with new patches.

4 Technical Solution

4.1 Source file organisation and management

4.1.1 File Structure

The following file structure is used to organise the code for the project, where school_project is the main package and is constructed of two main subpackages:

- The models package, which is a self-contained package for creating trained Artificial Neural Network models.
- The frames package, which consists of tkinter frames for the User Interface.

```
.
|-- Dockerfile
|-- .github
|   `-- workflows
|       `-- tests.yml
|-- .gitignore
|-- LICENSE
|-- notebooks
|   |-- cpu-vs-gpu-analysis.ipynb
|   |-- epoch-count-analysis.ipynb
|   |-- layer-count-analysis.ipynb
|   |-- learning-rate-analysis.ipynb
|   |-- neuron-count-analysis.ipynb
|   |-- relu-analysis.ipynb
|   `-- train-dataset-size-analysis.ipynb
|-- README.md
|-- school_project
|   |-- frames
|   |   |-- create_model.py
|   |   |-- hyper-parameter-defaults.json
|   |   |-- __init__.py
|   |   |-- load_model.py
|   |   `-- test_model.py
|   |-- __init__.py
|   |-- __main__.py
|   |-- models
|   |   |-- cpu
|   |   |   |-- cat_recognition.py
|   |   |   |-- __init__.py
|   |   |   |-- mnist.py
|   |   |   |-- utils
|   |   |   |   |-- __init__.py
|   |   |   |   |-- model.py
|   |   |   |   `-- tools.py
|   |   |   `-- xor.py
|   |   |-- datasets
|   |   |   |-- mnist.pkl.gz
|   |   |   |-- test-cat.h5
|   |   |   `-- train-cat.h5
|   |   |-- gpu
|   |   |   |-- cat_recognition.py
|   |   |   |-- __init__.py
|   |   |   |-- mnist.py
|   |   |   |-- utils
|   |   |   |   |-- __init__.py
|   |   |   |   |-- model.py
|   |   |   |   `-- tools.py
|   |   |   `-- xor.py
|   |   `-- __init__.py
|   |-- saved-models
|   `-- test
|       |-- __init__.py
|       |-- models
|       |   |-- cpu
|       |   |   |-- __init__.py
|       |   |   `-- utils
|       |   |       |-- __init__.py
|       |   |       |-- test_model.py
|       |   |       `-- test_tools.py
|       |   |-- gpu
|       |   |   |-- __init__.py
|       |   |   `-- utils
```

```
|       |   |       |-- __init__.py
|       |   |       |-- test_model.py
|       |   |       `-- test_tools.py
|       |   `-- __init__.py
|       `-- test_database.py
|-- setup.py
`-- TODO.md

19 directories, 50 files
```

Each package within the school_project package contains a __init__.py file, allowing the school_project package to be installed to a virtual environment so that the modules of the package can be imported from the installed package.

Show below are the contents of the frames package's __init__.py for example, which allows the classes of all modules in the package to be imported simultaneously:

```
"""Package of tkinter frames for the main window."""

from .create_model import HyperParameterFrame, TrainingFrame
from .load_model import LoadModelFrame
from .test_model import TestMNISTFrame, TestCatRecognitionFrame, TestXORFrame

__all__ = ['create_model', 'load_model', 'test_model']
```

4.1.2 Dependencies

The python dependencies for the project can be installed by running the following setup.py file (as described in the README.md in the next section). Instructions on installing external dependencies, such as the CUDA Toolkit for using a GPU, are explained in the README.md in the following section.

- setup.py code:

```
from setuptools import setup, find_packages

setup(
    name='school-project',
    version='2.0.0',
    packages=find_packages(),
    url='https://github.com/mcttn22/school-project.git',
    author='Max Cotton',
    author_email='maxcotton22@gmail.com',
    description='Year 13 Computer Science Programming Project',
    install_requires=[
                        'cupy-cuda12x',
                        'h5py',
                        'matplotlib',
                        'numpy'
    ],
)
```

4.1.3 Git and Github files

Git and Github were used extensively to manage the codebase and utilised the following files:

- A .gitignore file for specifying which files and directories should be ignored by Git:

```
# Byte compiled files
__pycache__/

# Packaging
*.egg-info
```

```

# Database file
school_project/saved_models.db
```

- A README.md markdown file to give installation and usage instructions for the repository on GitHub:
  - Markdown code:

```
<!-- The following lines generate badges showing the current status of the automated testing (Passing or Failing) and a Python3 badge correspondingly.) -->
[![tests](https://github.com/mcttn22/school-project/actions/workflows/tests.yml/badge.svg)](https://
[![python](https://img.shields.io/badge/Python-3-3776AB.svg?style=flat&logo=python&logoColor=white)]

# A-level Computer Science NEA Programming Project

This project is an investigation into how Artificial Neural Networks (ANNs) work and their applications in Image Recognition, by documenting all theory behind the project and developing applications of the theory, that allow for experimentation via a GUI. The ANNs are created without the use of any 3rd party Machine Learning Libraries and I currently have been able to achieve a prediction accuracy of 99.6% on the MNIST dataset. The report for this project is also included in this repository.

## Installation

1. Download the Repository with:

   - ```
     git clone https://github.com/mcttn22/school-project.git
     ```
   - Or by downloading as a ZIP file

</br>

2. Create a virtual environment (venv) with:
   - Windows:
     ```
     python -m venv {venv name}
     ```
   - Linux:
     ```
     python3 -m venv {venv name}
     ```

3. Enter the venv with:
   - Windows:
     ```
     .\{venv name}\Scripts\activate
     ```
   - Linux:
     ```
     source ./{venv name}/bin/activate
     ```

4. Enter the project directory with:
   ```
   cd school-project/
   ```

```
```

````
5. For normal use, install the dependencies and the project to the venv with:
  - Windows:
    ```
    python setup.py install
    ```
  - Linux:
    ```
    python3 setup.py install
    ```

*Note: In order to use an Nvidia GPU for training the networks, the latest Nvdia drivers must be installed and the CUDA Toolkit must be installed from
<a href="https://developer.nvidia.com/cuda-downloads">here</a>.*

## Usage

Run with:
- Windows:
  ```
  python school_project
  ```
- Linux:
  ```
  python3 school_project
  ```

## Development

Install the dependencies and the project to the venv in developing mode with:
- Windows:
  ```
  python setup.py develop
  ```
- Linux:
  ```
  python3 setup.py develop
  ```

Run Tests with:
- Windows:
  ```
  python -m unittest discover .\school_project\test\
  ```
- Linux:
  ```
  python3 -m unittest discover ./school_project/test/
  ```

Use Docker with:
- Build the Docker Image with:
  ```
  sudo docker build -t mcttn22/school-project ./
  ```
- Run the Docker Image with:
  ```
  sudo apt-get install x11-xserver-utils
  xhost +
  sudo docker run -v /tmp/.X11-unix:/tmp/.X11-unix -e DISPLAY=unix$DISPLAY mcttn22/school-project
  ```

````

````
Compile Project Report PDF with:
```
make all
```
*Note: This requires the Latexmk, pdflatek and Pygments libraries*
````

– Which will generate the following:

Tests passing Python 3

A-level Computer Science NEA Programming Project

This project is an investigation into how Artificial Neural Networks (ANNs) work and their applications in Image Recognition, by documenting all theory behind the project and developing applications of the theory, that allow for experimentation via a GUI. The ANNs are created without the use of any 3rd party Machine Learning Libraries and I currently have been able to achieve a prediction accuracy of 99.6% on the MNIST dataset. The report for this project is also included in this repository.

Installation

1. Download the Repository with:
   - ```
     git clone https://github.com/mcttn22/school-project.git
     ```
   - Or by downloading as a ZIP file

2. Create a virtual environment (venv) with:
   - Windows:
     ```
     python -m venv {venv name}
     ```
     ```

- Linux:

  ```
  python3 -m venv {venv name}
  ```

3. Enter the venv with:

   - Windows:

     ```
     .\{venv name}\Scripts\activate
     ```

   - Linux:

     ```
     source ./{venv name}/bin/activate
     ```

4. Enter the project directory with:

   ```
   cd school-project/
   ```

5. For normal use, install the dependencies and the project to the venv with:

   - Windows:

     ```
     python setup.py install
     ```

   - Linux:

     ```
     python3 setup.py install
     ```

Note: In order to use an Nvidia GPU for training the networks, the latest Nvdia drivers must be installed and the CUDA Toolkit must be installed from here.

Usage

Run with:

- Windows:

  ```
  python school_project
  ```

- Linux:

  ```
  python3 school_project
  ```

Development

Install the dependencies and the project to the venv in developing mode with:

- Windows:

  ```
  python setup.py develop
  ```

- Linux:

  ```
  python3 setup.py develop
  ```

Run Tests with:

- Windows:

```
python -m unittest discover .\school_project\test\
```

- Linux:

```
python3 -m unittest discover ./school_project/test/
```

Use Docker with:

- Build the Docker Image with:

```
sudo docker build -t mcttn22/school-project ./
```

- Run the Docker Image with:

```
sudo apt-get install x11-xserver-utils
xhost +
sudo docker run -v /tmp/.X11-unix:/tmp/.X11-unix -e DISPLAY=unix$DISPLAY mcttn22/school-prc
```

Compile Project Report PDF with:

```
make all
```

Note: This requires the Latexmk, pdflatek and Pygments libraries

- Also included was a license file that describes how others can use my code.

4.1.4 Organisation

I also utilised a TODO.md file for keeping track of what features and/or bugs need to be worked on.

4.2 Models package

This package is a self-contained package for creating the trained Artificial Neural Networks and can either be used with a CPU or a GPU, as well as containing the test and training data for all three datasets in a datasets directory. Whilst both the cpu and gpu subpackage are similar in functionality, the cpu subpackage uses NumPy for matrices whereas the gpu subpackage utilises NumPy alongisde the library CuPy which requires a GPU to be utilised for operations with the matrices. I have only shown the code for the cpu subpackage - the GPU subpackage is identical apart from calling CuPy instead of NumPy.

Both the cpu and gpu subpackage contain a utils subpackage that provides the tools for creating Artificial Neural Networks, and three modules that are the implementation of Artificial Neural Networks for each dataset.

4.2.1 Utils subpackage

The utils subpackage consists of a tools.py module that provides a ModelInterface class and helper functions for the model.py module, that contains an AbstractModel class that implements every method from the ModelInterface except for the load_dataset method.

- tools.py module:

```
"""Helper functions and ModelInterface class for model module."""

from abc import ABC, abstractmethod

import numpy as np

class ModelInterface(ABC):
    """Interface for ANN models."""
    @abstractmethod
    def _setup_layers(setup_values: callable) -> None:
        """Decorator that sets up model layers and sets up values of each layer
           with the method given.

        Args:
            setup_values (callable): the method that sets up the values of each
            layer.
        Raises:
            NotImplementedError: if this method is not implemented.

        """
        raise NotImplementedError

    @abstractmethod
    def create_model_values(self) -> None:
        """Create weights and bias/biases

        Raises:
            NotImplementedError: if this method is not implemented.

        """
        raise NotImplementedError

    @abstractmethod
    def load_model_values(self, file_location: str) -> None:
        """Load weights and bias/biases from .npz file.

        Args:
            file_location (str): the location of the file to load from.
        Raises:
            NotImplementedError: if this method is not implemented.

        """
        raise NotImplementedError

    @abstractmethod
    def load_datasets(self, train_dataset_size: int) -> tuple[np.ndarray, np.ndarray,
                                                              np.ndarray, np.ndarray]:
        """Load input and output datasets. For the input dataset, each column
           should represent a piece of data and each row should store the values
           of the piece of data.

        Args:
            train_dataset_size (int): the number of train dataset inputs to use.
        Returns:
```

```python
            tuple of train_inputs, train_outputs,
            test_inputs and test_outputs.
        Raises:
            NotImplementedError: if this method is not implemented.

        """
        raise NotImplementedError

    @abstractmethod
    def back_propagation(self, prediction: np.ndarray) -> None:
        """Adjust the weights and bias/biases via gradient descent.

        Args:
            prediction (numpy.ndarray): the matrice of prediction values
        Raises:
            NotImplementedError: if this method is not implemented.

        """
        raise NotImplementedError

    @abstractmethod
    def forward_propagation(self) -> np.ndarray:
        """Generate a prediction with the weights and bias/biases.

        Returns:
            numpy.ndarray of prediction values.
        Raises:
            NotImplementedError: if this method is not implemented.

        """
        raise NotImplementedError

    @abstractmethod
    def test(self) -> None:
        """Test trained weights and bias/biases.

        Raises:
            NotImplementedError: if this method is not implemented.

        """
        raise NotImplementedError

    @abstractmethod
    def train(self, epochs: int) -> None:
        """Train weights and bias/biases.

        Args:
            epochs (int): the number of forward and back propagations to
            ↪  do.
        Raises:
            NotImplementedError: if this method is not implemented.

        """
        raise NotImplementedError

    @abstractmethod
    def save_model_values(self, file_location: str) -> None:
        """Save the model by saving the weights then biases of each layer
        ↪  to
           a .npz file with a given file location.

           Args:
               file_location (str): the file location to save the model to.

```

```python
        """
        raise NotImplementedError

def relu(z: np.ndarray | int | float) -> np.ndarray | float:
    """Transfer function, transform input to max number between 0 and z.

    Args:
        z (numpy.ndarray | int | float):
        the numpy.ndarray | int | float to be transferred.
    Returns:
        numpy.ndarray | float,
        with all values | the value transferred to max number between 0-z.
    Raises:
        TypeError: if z is not of type numpy.ndarray | int | float.

    """
    return np.maximum(0.1*z, 0)  # Divide by 10 to stop overflow errors

def relu_derivative(output: np.ndarray) -> np.ndarray:
    """Calculate derivative of ReLu Transfer function with respect to z.

    Args:
        output (numpy.ndarray):
        the numpy.ndarray output of the ReLu transfer function.
    Returns:
        numpy.ndarray,
        derivative of the ReLu transfer function with respect to z.
    Raises:
        TypeError: if output is not of type numpy.ndarray.

    """
    output[output <= 0] = 0
    output[output > 0] = 1

    return output

def sigmoid(z: np.ndarray | int | float) -> np.ndarray | float:
    """Transfer function, transform input to number between 0 and 1.

    Args:
        z (numpy.ndarray | int | float):
        the numpy.ndarray | int | float to be transferred.
    Returns:
        numpy.ndarray | float,
        with all values | the value transferred to a number between 0-1.
    Raises:
        TypeError: if z is not of type numpy.ndarray | int | float.

    """
    return 1 / (1 + np.exp(-z))

def sigmoid_derivative(output: np.ndarray | int | float) -> np.ndarray | float:
    """Calculate derivative of sigmoid Transfer function with respect to z.

    Args:
        output (numpy.ndarray | int | float):
        the numpy.ndarray | int | float output of the sigmoid transfer function.
    Returns:
        numpy.ndarray | float,
        derivative of the sigmoid transfer function with respect to z.
    Raises:
        TypeError: if output is not of type numpy.ndarray | int | float.
```

```python

    """
    return output * (1 - output)

def calculate_loss(input_count: int,
                   outputs: np.ndarray,
                   prediction: np.ndarray) -> float:
    """Calculate average loss/error of the prediction to the outputs.

    Args:
        input_count (int): the number of inputs.
        outputs (np.ndarray):
        the train/test outputs array to compare with the prediction.
        prediction (np.ndarray): the array of prediction values.
    Returns:
        float loss.
    Raises:
        ValueError:
        if outputs is not a suitable multiplier with the prediction
        (incorrect shapes)

    """
    return np.squeeze(- (1/input_count) * np.sum(outputs *
    ↪   np.log(prediction) + (1 - outputs) * np.log(1 - prediction)))

def calculate_prediction_accuracy(prediction: np.ndarray,
                                  outputs: np.ndarray) -> float:
    """Calculate the percentage accuracy of the predictions.

    Args:
        prediction (np.ndarray): the array of prediction values.
        outputs (np.ndarray):
        the train/test outputs array to compare with the prediction.
    Returns:
        float prediction accuracy

    """
    return 100 - np.mean(np.abs(prediction - outputs)) * 100
```

- model.py module:

```python
"""Provides an abstract class for Artificial Neural Network models."""

from collections.abc import Generator
import time

import numpy as np

from .tools import (
                    ModelInterface,
                    relu,
                    relu_derivative,
                    sigmoid,
                    sigmoid_derivative,
                    calculate_loss,
                    calculate_prediction_accuracy
                    )

class _FullyConnectedLayer():
    """Fully connected layer for Deep ANNs,
       represented as a node of a Doubly linked list."""
    def __init__(self, learning_rate: float, input_neuron_count: int,
                 output_neuron_count: int, transfer_type: str) -> None:
```

```python
        """Initialise layer values.

        Args:
            learning_rate (float): the learning rate of the model.
            input_neuron_count (int):
            the number of input neurons into the layer.
            output_neuron_count (int):
            the number of output neurons into the layer.
            transfer_type (str): the transfer function type
            ('sigmoid' or 'relu')

        """
        # Setup layer attributes
        self.previous_layer = None
        self.next_layer = None
        self.input_neuron_count = input_neuron_count
        self.output_neuron_count = output_neuron_count
        self.transfer_type = transfer_type
        self.input: np.ndarray
        self.output: np.ndarray

        # Setup weights and biases
        self.weights: np.ndarray
        self.biases: np.ndarray
        self.learning_rate = learning_rate

    def __repr__(self) -> str:
        """Read values of the layer.

        Returns:
            a string description of the layers's
            weights, bias and learning rate values.

        """
        return (f"Weights: {self.weights.tolist()}\n" +
                f"Biases: {self.biases.tolist()}\n")

    def init_layer_values_random(self) -> None:
        """Initialise weights to random values and biases to 0s"""
        np.random.seed(1)  # Sets up pseudo random values for layer weight arrays
        self.weights = np.random.rand(self.output_neuron_count, self.input_neuron_count) - 0.5
        self.biases = np.zeros(shape=(self.output_neuron_count, 1))

    def init_layer_values_zeros(self) -> None:
        """Initialise weights to 0s and biases to 0s"""
        self.weights = np.zeros(shape=(self.output_neuron_count, self.input_neuron_count))
        self.biases = np.zeros(shape=(self.output_neuron_count, 1))

    def back_propagation(self, dloss_doutput) -> np.ndarray:
        """Adjust the weights and biases via gradient descent.

        Args:
            dloss_doutput (numpy.ndarray): the derivative of the loss of the
            layer's output, with respect to the layer's output.
        Returns:
            a numpy.ndarray derivative of the loss of the layer's input,
            with respect to the layer's input.
        Raises:
            ValueError:
            if dloss_doutput
```

```
            is not a suitable multiplier with the weights
            (incorrect shape)

        """
        match self.transfer_type:
            case 'sigmoid':
                dloss_dz = dloss_doutput * sigmoid_derivative(output=self.output)
            case 'relu':
                dloss_dz = dloss_doutput * relu_derivative(output=self.output)

        dloss_dweights = np.dot(dloss_dz, self.input.T)
        dloss_dbiases = np.sum(dloss_dz)

        assert dloss_dweights.shape == self.weights.shape

        dloss_dinput = np.dot(self.weights.T, dloss_dz)

        # Update weights and biases
        self.weights -= self.learning_rate * dloss_dweights
        self.biases -= self.learning_rate * dloss_dbiases

        return dloss_dinput

    def forward_propagation(self, inputs) -> np.ndarray:
        """Generate a layer output with the weights and biases.

        Args:
            inputs (np.ndarray): the input values to the layer.
        Returns:
            a numpy.ndarray of the output values.

        """
        self.input = inputs
        z = np.dot(self.weights, self.input) + self.biases
        if self.transfer_type == 'sigmoid':
            self.output = sigmoid(z)
        elif self.transfer_type == 'relu':
            self.output = relu(z)
        return self.output

class _Layers():
    """Manages linked list of layers."""
    def __init__(self) -> None:
        """Initialise linked list."""
        self.head = None
        self.tail = None

    def __iter__(self) -> Generator[_FullyConnectedLayer, None, None]:
        """Iterate forward through the network."""
        current_layer = self.head
        while True:
            yield current_layer
            if current_layer.next_layer is not None:
                current_layer = current_layer.next_layer
            else:
                break

    def __reversed__(self) -> Generator[_FullyConnectedLayer, None, None]:
        """Iterate back through the network."""
        current_layer = self.tail
        while True:
            yield current_layer
```

```python
            if current_layer.previous_layer is not None:
                current_layer = current_layer.previous_layer
            else:
                break

class AbstractModel(ModelInterface):
    """ANN model with variable number of hidden layers"""
    def __init__(self,
                 hidden_layers_shape: list[int],
                 train_dataset_size: int,
                 learning_rate: float,
                 use_relu: bool) -> None:
        """Initialise model values.

        Args:
            hidden_layers_shape (list[int]):
            list of the number of neurons in each hidden layer.
            train_dataset_size (int): the number of train dataset inputs to use.
            learning_rate (float): the learning rate of the model.
            use_relu (bool): True or False whether the ReLu Transfer function
            should be used.

        """
        # Setup model data
        self.train_inputs, self.train_outputs,\
        self.test_inputs, self.test_outputs = self.load_datasets(
                                              train_dataset_size=train_dataset_size
                                              )
        self.train_losses: list[float]
        self.test_prediction: np.ndarray
        self.test_prediction_accuracy: float
        self.training_progress = ""
        self.training_time: float

        # Setup model attributes
        self.__running = True
        self.input_neuron_count: int = self.train_inputs.shape[0]
        self.input_count = self.train_inputs.shape[1]
        self.hidden_layers_shape = hidden_layers_shape
        self.output_neuron_count = self.train_outputs.shape[0]
        self.layers_shape = [f'{layer}' for layer in (
                             [self.input_neuron_count] +
                             self.hidden_layers_shape +
                             [self.output_neuron_count]
                             )]
        self.use_relu = use_relu

        # Setup model values
        self.layers = _Layers()
        self.learning_rate = learning_rate

    def __repr__(self) -> str:
        """Read current state of model.

        Returns:
            a string description of the model's shape,
            weights, bias and learning rate values.

        """
        return (f"Layers Shape: {','.join(self.layers_shape)}\n" +
                f"Learning Rate: {self.learning_rate}")
```

```python

    def set_running(self, value: bool) -> None:
        """Set the running attribute to the given value.

        Args:
            value (bool): the value to set the running attribute to.

        """
        self.__running = value

    def _setup_layers(setup_values: callable) -> None:
        """Decorator that sets up model layers and sets up values of each layer
           with the method given.

        Args:
            setup_values (callable): the method that sets up the values of each
            layer.

        """
        def decorator(self, *args, **kwargs) -> None:
            # Check if setting up Deep Network
            if len(self.hidden_layers_shape) > 0:
                if self.use_relu:

                    # Add input layer
                    self.layers.head = _FullyConnectedLayer(
                                                learning_rate=self.learning_rate,
                                                input_neuron_count=self.input_neuron_count,
                                                output_neuron_count=self.hidden_layers_shape[0],
                                                transfer_type='relu'
                                                )
                    current_layer = self.layers.head

                    # Add hidden layers
                    for layer in range(len(self.hidden_layers_shape) - 1):
                        current_layer.next_layer = _FullyConnectedLayer(
                                        learning_rate=self.learning_rate,
                                        input_neuron_count=self.hidden_layers_shape[layer],
                                        output_neuron_count=self.hidden_layers_shape[layer + 1],
                                        transfer_type='relu'
                                        )
                        current_layer.next_layer.previous_layer = current_layer
                        current_layer = current_layer.next_layer
                else:

                    # Add input layer
                    self.layers.head = _FullyConnectedLayer(
                                                learning_rate=self.learning_rate,
                                                input_neuron_count=self.input_neuron_count,
                                                output_neuron_count=self.hidden_layers_shape[0],
                                                transfer_type='sigmoid'
                                                )
```

```python
                    current_layer = self.layers.head

                    # Add hidden layers
                    for layer in range(len(self.hidden_layers_shape) - 1):
                        current_layer.next_layer = _FullyConnectedLayer(
                                    learning_rate=self.learning_rate,
                                    input_neuron_count=self.hidden_layers_shape[layer],
                                    output_neuron_count=self.hidden_layers_shape[layer + 1],
                                    transfer_type='sigmoid'
                                    )
                        current_layer.next_layer.previous_layer = current_layer
                        current_layer = current_layer.next_layer

                # Add output layer
                current_layer.next_layer = _FullyConnectedLayer(
                                        learning_rate=self.learning_rate,
                                        input_neuron_count=self.hidden_layers_shape[-1],
                                        output_neuron_count=self.output_neuron_count,
                                        transfer_type='sigmoid'
                                        )
                current_layer.next_layer.previous_layer = current_layer
                self.layers.tail = current_layer.next_layer

            # Setup Perceptron Network
            else:
                self.layers.head = _FullyConnectedLayer(
                                        learning_rate=self.learning_rate,
                                        input_neuron_count=self.input_neuron_count,
                                        output_neuron_count=self.output_neuron_count,
                                        transfer_type='sigmoid'
                                        )
                self.layers.tail = self.layers.head

            setup_values(self, *args, **kwargs)

        return decorator

    @_setup_layers
    def create_model_values(self) -> None:
        """Create weights and bias/biases"""
        # Check if setting up Deep Network
        if len(self.hidden_layers_shape) > 0:

            # Initialise Layer values to random values
            for layer in self.layers:
                layer.init_layer_values_random()

        # Setup Perceptron Network
        else:

            # Initialise Layer values to zeros
            for layer in self.layers:
                layer.init_layer_values_zeros()

    @_setup_layers
    def load_model_values(self, file_location: str) -> None:
```

```
        """Load weights and bias/biases from .npz file.

        Args:
            file_location (str): the location of the file to load from.

        """
        data: dict[str, np.ndarray] = np.load(file=file_location)

        # Initialise Layer values
        i = 0
        keys = list(data.keys())
        for layer in self.layers:
            layer.weights = data[keys[i]]
            layer.biases = data[keys[i + 1]]
            i += 2

    def back_propagation(self, dloss_doutput) -> None:
        """Train each layer's weights and biases.

        Args:
            dloss_doutput (np.ndarray): the derivative of the loss of the
            output layer's output, with respect to the output layer's
            ↪  output.

        """
        for layer in reversed(self.layers):
            dloss_doutput =
            ↪  layer.back_propagation(dloss_doutput=dloss_doutput)

    def forward_propagation(self) -> np.ndarray:
        """Generate a prediction with the layers.

        Returns:
            a numpy.ndarray of the prediction values.

        """
        output = self.train_inputs
        for layer in self.layers:
            output = layer.forward_propagation(inputs=output)
        return output

    def test(self) -> None:
        """Test the layers' trained weights and biases."""
        output = self.test_inputs
        for layer in self.layers:
            output = layer.forward_propagation(inputs=output)
        self.test_prediction = output

        # Calculate performance of model
        self.test_prediction_accuracy = calculate_prediction_accuracy(

                                                      ↪  prediction=self.test_prediction,
                                                      outputs=self.test_outputs
                                                      )

    def train(self, epoch_count: int) -> None:
        """Train layers' weights and biases.

           Args:
               epoch_count (int): the number of training epochs.

        """
        self.layers_shape = [f'{layer}' for layer in (
                            [self.input_neuron_count] +
```

```python
                              self.hidden_layers_shape +
                              [self.output_neuron_count]
                              )]
        self.train_losses = []
        training_start_time = time.time()
        for epoch in range(epoch_count):
            if not self.__running:
                break
            self.training_progress = f"Epoch {epoch} / {epoch_count}"
            prediction = self.forward_propagation()
            loss = calculate_loss(input_count=self.input_count,
                                  outputs=self.train_outputs,
                                  prediction=prediction)
            self.train_losses.append(loss)
            if not self.__running:
                break
            dloss_doutput = -(1/self.input_count) * ((self.train_outputs -
            ↪   prediction)/(prediction * (1 - prediction)))
            self.back_propagation(dloss_doutput=dloss_doutput)
        self.training_time = round(number=time.time() -
        ↪   training_start_time,
                                   ndigits=2)

    def save_model_values(self, file_location: str) -> None:
        """Save the model by saving the weights then biases of each layer
        ↪   to
           a .npz file with a given file location.

           Args:
               file_location (str): the file location to save the model to.

        """
        saved_model: list[np.ndarray] = []
        for layer in self.layers:
            saved_model.append(layer.weights)
            saved_model.append(layer.biases)
        np.savez(file_location, *saved_model)
```

4.2.2 Artificial Neural Network implementations

The following three modules implement the AbstractModel class from the above model.py module from the utils subpackage, on the three datasets.

- cat_recognition.py module:

```python
"""Implementation of Artificial Neural Network model on Cat Recognition
↪   dataset."""

import h5py
import numpy as np

from .utils.model import AbstractModel

class CatRecognitionModel(AbstractModel):
    """ANN model that trains to predict if an image is a cat or not a
    ↪   cat."""
    def __init__(self,
                 hidden_layers_shape: list[int],
                 train_dataset_size: int,
                 learning_rate: float,
                 use_relu: bool) -> None:
```

```python
        """Initialise Model's Base class.

        Args:
            hidden_layers_shape (list[int]):
            list of the number of neurons in each hidden layer.
            train_dataset_size (int): the number of train dataset inputs to use.
            learning_rate (float): the learning rate of the model.
            use_relu (bool): True or False whether the ReLu Transfer function
            should be used.

        """
        super().__init__(hidden_layers_shape=hidden_layers_shape,
                         train_dataset_size=train_dataset_size,
                         learning_rate=learning_rate,
                         use_relu=use_relu)

    def load_datasets(self, train_dataset_size: int) -> tuple[np.ndarray, np.ndarray,
                                                                np.ndarray, np.ndarray]:
        """Load image input and output datasets.

        Args:
            train_dataset_size (int): the number of train dataset inputs to use.
        Returns:
            tuple of image train_inputs, train_outputs,
            test_inputs and test_outputs numpy.ndarrys.

        Raises:
            FileNotFoundError: if file does not exist.

        """
        # Load datasets from h5 files
        # (h5 files stores large amount of data with quick access)
        train_dataset: h5py.File = h5py.File(
             r'school_project/models/datasets/train-cat.h5',
             'r'
             )
        test_dataset: h5py.File = h5py.File(
              r'school_project/models/datasets/test-cat.h5',
              'r'
              )

        # Load input arrays,
        # containing the RGB values for each pixel in each 64x64 pixel image,
        # for 209 images
        train_inputs: np.ndarray = np.array(train_dataset['train_set_x'][:])
        test_inputs: np.ndarray = np.array(test_dataset['test_set_x'][:])

        # Load output arrays of 1s for cat and 0s for not cat
        train_outputs: np.ndarray = np.array(train_dataset['train_set_y'][:])
        test_outputs: np.ndarray = np.array(test_dataset['test_set_y'][:])

        # Reshape input arrays into 1 dimension (flatten),
        # then divide by 255 (RGB)
        # to standardize them to a number between 0 and 1
        train_inputs = train_inputs.reshape((train_inputs.shape[0],
                                             -1)).T / 255
```

```python
        test_inputs = test_inputs.reshape((test_inputs.shape[0], -1)).T / 255

        # Reshape output arrays into a 1 dimensional list of outputs
        train_outputs = train_outputs.reshape((1, train_outputs.shape[0]))
        test_outputs = test_outputs.reshape((1, test_outputs.shape[0]))

        # Reduce train datasets' sizes to train_dataset_size
        train_inputs = (train_inputs.T[:train_dataset_size]).T
        train_outputs = (train_outputs.T[:train_dataset_size]).T

        return train_inputs, train_outputs, test_inputs, test_outputs
```

- mnist.py module:

```python
"""Implementation of Artificial Neural Network model on MNIST dataset."""

import pickle
import gzip

import numpy as np

from .utils.model import AbstractModel

class MNISTModel(AbstractModel):
    """ANN model that trains to predict Numbers from images."""
    def __init__(self, hidden_layers_shape: list[int],
                 train_dataset_size: int,
                 learning_rate: float,
                 use_relu: bool) -> None:
        """Initialise Model's Base class.

        Args:
            hidden_layers_shape (list[int]):
            list of the number of neurons in each hidden layer.
            train_dataset_size (int): the number of train dataset inputs to use.
            learning_rate (float): the learning rate of the model.
            use_relu (bool): True or False whether the ReLu Transfer function
            should be used.

        """
        super().__init__(hidden_layers_shape=hidden_layers_shape,
                         train_dataset_size=train_dataset_size,
                         learning_rate=learning_rate,
                         use_relu=use_relu)

    def load_datasets(self, train_dataset_size: int) -> tuple[np.ndarray, np.ndarray,
                                                              np.ndarray, np.ndarray]:
        """Load image input and output datasets.
        Args:
            train_dataset_size (int): the number of dataset inputs to use.
        Returns:
            tuple of image train_inputs, train_outputs,
            test_inputs and test_outputs numpy.ndarrys.

        Raises:
            FileNotFoundError: if file does not exist.

        """
```

```python
        # Load datasets from pkl.gz file
        with gzip.open(
              'school_project/models/datasets/mnist.pkl.gz',
              'rb'
              ) as mnist:
            (train_inputs, train_outputs),\
            (test_inputs, test_outputs) = pickle.load(mnist, encoding='bytes')

        # Reshape input arrays into 1 dimension (flatten),
        # then divide by 255 (RGB)
        # to standardize them to a number between 0 and 1
        train_inputs = np.array(train_inputs.reshape((train_inputs.shape[0],
                                                      -1)).T / 255)
        test_inputs = np.array(test_inputs.reshape(test_inputs.shape[0], -1).T / 255)

        # Represent number values
        # with a one at the matching index of an array of zeros
        train_outputs = np.eye(np.max(train_outputs) + 1)[train_outputs].T
        test_outputs = np.eye(np.max(test_outputs) + 1)[test_outputs].T

        # Reduce train datasets' sizes to train_dataset_size
        train_inputs = (train_inputs.T[:train_dataset_size]).T
        train_outputs = (train_outputs.T[:train_dataset_size]).T

        return train_inputs, train_outputs, test_inputs, test_outputs
```

- xor.py module

```python
"""Implementation of Artificial Neural Network model on XOR dataset."""

import numpy as np

from .utils.model import AbstractModel

class XORModel(AbstractModel):
    """ANN model that trains to predict the output of a XOR gate with two
       inputs."""
    def __init__(self,
                 hidden_layers_shape: list[int],
                 train_dataset_size: int,
                 learning_rate: float,
                 use_relu: bool) -> None:
        """Initialise Model's Base class.

        Args:
            hidden_layers_shape (list[int]):
            list of the number of neurons in each hidden layer.
            train_dataset_size (int): the number of train dataset inputs to use.
            learning_rate (float): the learning rate of the model.
            use_relu (bool): True or False whether the ReLu Transfer function
            should be used.

        """
        super().__init__(hidden_layers_shape=hidden_layers_shape,
                         train_dataset_size=train_dataset_size,
                         learning_rate=learning_rate,
                         use_relu=use_relu)

```

```python
    def load_datasets(self, train_dataset_size: int) -> tuple[np.ndarray, np.ndarray,
                                                              np.ndarray, np.ndarray]:
        """Load XOR input and output datasets.

        Args:
            train_dataset_size (int): the number of dataset inputs to use.
        Returns:
            tuple of XOR train_inputs, train_outputs,
            test_inputs and test_outputs numpy.ndarrys.

        """
        inputs: np.ndarray = np.array([[0, 0, 1, 1],
                                       [0, 1, 0, 1]])
        outputs: np.ndarray = np.array([[0, 1, 1, 0]])

        # Reduce train datasets' sizes to train_dataset_size
        inputs = (inputs.T[:train_dataset_size]).T
        outputs = (outputs.T[:train_dataset_size]).T

        return inputs, outputs, inputs, outputs
```

4.3 Frames package

I have used tkinter for the User Interface and the frames package which consists of tkinter frames to be loaded onto the main window when needed. The package also includes a hyper-parameter-defaults.json file, which stores optimum default values for the hyper-parameters to be set to.

- hyper-parameter-defaults.json file contents:

```json
{
    "MNIST": {
        "description": "An Image model trained on recognising numbers from images.",
        "epochCount": 150,
        "hiddenLayersShape": [1000, 1000],
        "minTrainDatasetSize": 1,
        "maxTrainDatasetSize": 60000,
        "maxLearningRate": 1
    },
    "Cat Recognition": {
        "description": "An Image model trained on recognising if an image is a cat or not.",
        "epochCount": 3500,
        "hiddenLayersShape": [100, 100],
        "minTrainDatasetSize": 1,
        "maxTrainDatasetSize": 209,
        "maxLearningRate": 0.3
    },
    "XOR": {
        "description": "For experimenting with Artificial Neural Networks, a XOR gate model has been used for its lesser computation time.",
        "epochCount": 4700,
        "hiddenLayersShape": [100, 100],
        "minTrainDatasetSize":  2,
        "maxTrainDatasetSize": 4,
        "maxLearningRate": 1
```

```
    }
}
```

- create_model.py module:

```
"""Tkinter frames for creating an Artificial Neural Network model."""

import json
import threading
import tkinter as tk
import tkinter.font as tkf

from matplotlib.figure import Figure
from matplotlib.backends.backend_tkagg import FigureCanvasTkAgg
import numpy as np

class HyperParameterFrame(tk.Frame):
    """Frame for hyper-parameter page."""
    def __init__(self, root: tk.Tk, width: int,
                 height: int, bg: str, dataset: str) -> None:
        """Initialise hyper-parameter frame widgets.

        Args:
            root (tk.Tk): the widget object that contains this widget.
            width (int): the pixel width of the frame.
            height (int): the pixel height of the frame.
            bg (str): the hex value or name of the frame's background
            ↪  colour.
            dataset (str): the name of the dataset to use
            ('MNIST', 'Cat Recognition' or 'XOR')
        Raises:
            TypeError: if root, width or height are not of the correct
            ↪  type.

        """
        super().__init__(master=root, width=width, height=height, bg=bg)
        self.root = root
        self.WIDTH = width
        self.HEIGHT = height
        self.BG = bg

        # Setup hyper-parameter frame variables
        self.dataset = dataset
        self.use_gpu: bool
        self.default_hyper_parameters = self.load_default_hyper_parameters(

                                                       ↪  dataset=dataset
                                                       )

        # Setup widgets
        self.title_label = tk.Label(master=self,
                                    bg=self.BG,
                                    font=('Arial', 20),
                                    text=dataset)
        self.about_label = tk.Label(
                                master=self,
                                bg=self.BG,
                                font=('Arial', 14),

                                ↪  text=self.default_hyper_parameters['description']
                                )
        self.learning_rate_scale = tk.Scale(
                               master=self,
```

```
                          bg=self.BG,
                          orient='horizontal',
                          label="Learning Rate",
                          length=185,
                          from_=0,
                          to=self.default_hyper_parameters['maxLearningRate'],
                          resolution=0.01
                          )
        self.learning_rate_scale.set(value=0.1)
        self.epoch_count_scale = tk.Scale(master=self,
                                          bg=self.BG,
                                          orient='horizontal',
                                          label="Epoch Count",
                                          length=185,
                                          from_=0,
                                          to=10_000,
                                          resolution=100)
        self.epoch_count_scale.set(
                             value=self.default_hyper_parameters['epochCount']
                             )
        self.train_dataset_size_scale = tk.Scale(
                   master=self,
                   bg=self.BG,
                   orient='horizontal',
                   label="Train Dataset Size",
                   length=185,
                   from_=self.default_hyper_parameters['minTrainDatasetSize'],
                   to=self.default_hyper_parameters['maxTrainDatasetSize'],
                   resolution=1
                   )
        self.train_dataset_size_scale.set(
                    value=self.default_hyper_parameters['maxTrainDatasetSize']
                    )
        self.hidden_layers_shape_label = tk.Label(
                                master=self,
                                bg=self.BG,
                                font=('Arial', 12),
                                text="Enter the number of neurons in each\n" +
                                        "hidden layer, separated by commas:"
                                )
        self.hidden_layers_shape_entry = tk.Entry(master=self)
        self.hidden_layers_shape_entry.insert(0, ",".join(
            f"{neuron_count}" for neuron_count in self.default_hyper_parameters['hiddenLayersShape']
            ))
        self.use_relu_check_button_var = tk.BooleanVar(value=True)
        self.use_relu_check_button = tk.Checkbutton(
                                        master=self,
                                        width=13, height=1,
                                        font=tkf.Font(size=12),
                                        text="Use ReLu",
                                        variable=self.use_relu_check_button_var
                                                        )
        self.use_gpu_check_button_var = tk.BooleanVar()
        self.use_gpu_check_button = tk.Checkbutton(
                                        master=self,
                                        width=13, height=1,
```

```python
                                        font=tkf.Font(size=12),
                                        text="Use GPU",
                                        variable=self.use_gpu_check_button_var
                                                       )
        self.model_status_label = tk.Label(master=self,
                                           bg=self.BG,
                                           font=('Arial', 15))

        # Pack widgets
        self.title_label.grid(row=0, column=0, columnspan=3)
        self.about_label.grid(row=1, column=0, columnspan=3)
        self.learning_rate_scale.grid(row=2, column=0, pady=(50,0))
        self.epoch_count_scale.grid(row=3, column=0, pady=(30,0))
        self.train_dataset_size_scale.grid(row=4, column=0, pady=(30,0))
        self.hidden_layers_shape_label.grid(row=2, column=1,
                                            padx=30, pady=(50,0))
        self.hidden_layers_shape_entry.grid(row=3, column=1, padx=30)
        self.use_relu_check_button.grid(row=2, column=2, pady=(30, 0))
        self.use_gpu_check_button.grid(row=3, column=2, pady=(30, 0))
        self.model_status_label.grid(row=5, column=0,
                                     columnspan=3, pady=50)

    def load_default_hyper_parameters(self, dataset: str) -> dict[
                                                 str,
                                                 str | int | list[int] | float
                                                 ]:
        """Load the dataset's default hyper-parameters from the json file.

           Args:
               dataset (str): the name of the dataset to load hyper-parameters
               for. ('MNIST', 'Cat Recognition' or 'XOR')
            Returns:
                a dictionary of default hyper-parameter values.
        """
        with open('school_project/frames/hyper-parameter-defaults.json') as f:
            return json.load(f)[dataset]

    def create_model(self) -> object:
        """Create and return a Model using the hyper-parameters set.

           Returns:
               a Model object.
        """
        self.use_gpu = self.use_gpu_check_button_var.get()

        # Validate hidden layers shape input
        hidden_layers_shape_input = [layer for layer in self.hidden_layers_shape_entry.get().replace(' ', '').split(',')]
        for layer in hidden_layers_shape_input:
            if not layer.isdigit():
                self.model_status_label.configure(
                                        text="Invalid hidden layers shape",
                                        fg='red'
                                        )
                raise ValueError

        # Create Model
        if not self.use_gpu:
            if self.dataset == "MNIST":
```

```python
                from school_project.models.cpu.mnist import MNISTModel as Model
            elif self.dataset == "Cat Recognition":
                from school_project.models.cpu.cat_recognition import CatRecognitionModel as Model
            elif self.dataset == "XOR":
                from school_project.models.cpu.xor import XORModel as Model
            model = Model(
                hidden_layers_shape = [int(neuron_count) for neuron_count in hidden_layers_shape_input],
                train_dataset_size = self.train_dataset_size_scale.get(),
                learning_rate = self.learning_rate_scale.get(),
                use_relu = self.use_relu_check_button_var.get()
                )
            model.create_model_values()

        else:
            try:
                if self.dataset == "MNIST":
                    from school_project.models.gpu.mnist import MNISTModel as Model
                elif self.dataset == "Cat Recognition":
                    from school_project.models.gpu.cat_recognition import CatRecognitionModel as Model
                elif self.dataset == "XOR":
                    from school_project.models.gpu.xor import XORModel as Model
                model = Model(hidden_layers_shape = [int(neuron_count) for neuron_count in hidden_layers_shape_input],
                              train_dataset_size = self.train_dataset_size_scale.get(),
                              learning_rate = self.learning_rate_scale.get(),
                              use_relu = self.use_relu_check_button_var.get())
                model.create_model_values()
            except ImportError as ie:
                self.model_status_label.configure(
                                        text="Failed to initialise GPU",
                                        fg='red'
                                        )
                raise ImportError
        return model

class TrainingFrame(tk.Frame):
    """Frame for training page."""
    def __init__(self, root: tk.Tk, width: int,
                 height: int, bg: str,
                 model: object, epoch_count: int) -> None:
        """Initialise training frame widgets.

        Args:
            root (tk.Tk): the widget object that contains this widget.
            width (int): the pixel width of the frame.
            height (int): the pixel height of the frame.
            bg (str): the hex value or name of the frame's background colour.
            model (object): the Model object to be trained.
            epoch_count (int): the number of training epochs.
        Raises:
            TypeError: if root, width or height are not of the correct type.

        """
```

```python
        super().__init__(master=root, width=width, height=height, bg=bg)
        self.root = root
        self.WIDTH = width
        self.HEIGHT = height
        self.BG = bg

        # Setup widgets
        self.model_status_label = tk.Label(master=self,
                                           bg=self.BG,
                                           font=('Arial', 15))
        self.training_progress_label = tk.Label(master=self,
                                                bg=self.BG,
                                                font=('Arial', 15))
        self.loss_figure: Figure = Figure()
        self.loss_canvas: FigureCanvasTkAgg = FigureCanvasTkAgg(
                                                     figure=self.loss_figure,
                                                     master=self
                                                     )


        # Pack widgets
        self.model_status_label.pack(pady=(30,0))
        self.training_progress_label.pack(pady=30)

        # Start training thread
        self.model_status_label.configure(
                                  text="Training weights and biases...",
                                  fg='red'
                                  )
        self.train_thread: threading.Thread = threading.Thread(
                                                          target=model.train,
                                                          args=(epoch_count,)
                                                          )
        self.train_thread.start()

    def plot_losses(self, model: object) -> None:
        """Plot losses of Model training.

           Args:
               model (object): the Model object thats been trained.

        """
        self.model_status_label.configure(
                 text=f"Weights and biases trained in {model.training_time}s",
                 fg='green'
                 )
        graph: Figure.axes = self.loss_figure.add_subplot(111)
        graph.set_title("Learning rate: " +
                        f"{model.learning_rate}")
        graph.set_xlabel("Epochs")
        graph.set_ylabel("Loss Value")
        graph.plot(np.squeeze(model.train_losses))
        self.loss_canvas.get_tk_widget().pack()
```

This outputs the following for the hyper-parameter frame shown in figure 17:

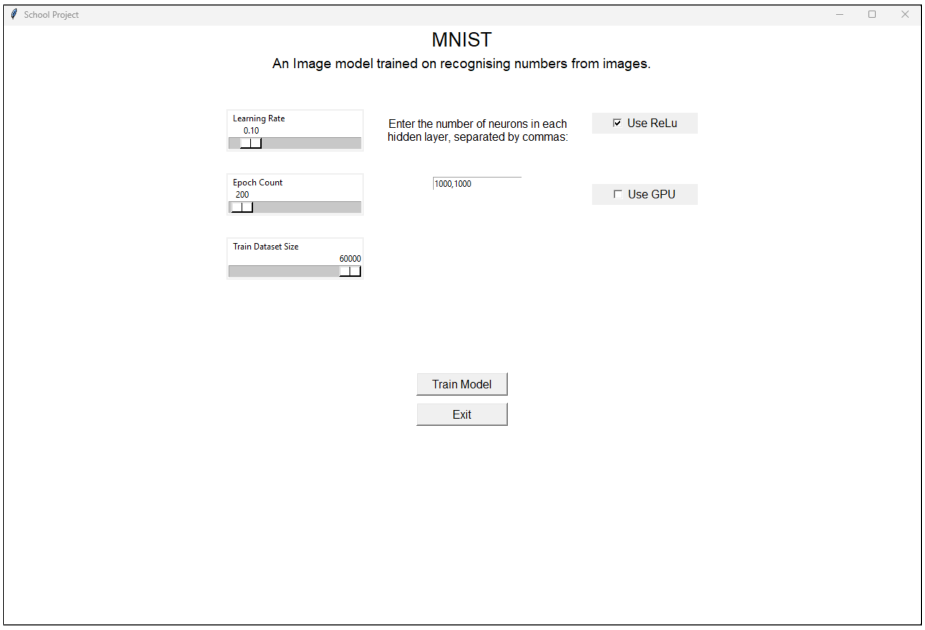


Figure 17: Hyper parameter frame - showing MNIST parameters

And outputs the following for the training frame, shown in figure 18, during training:

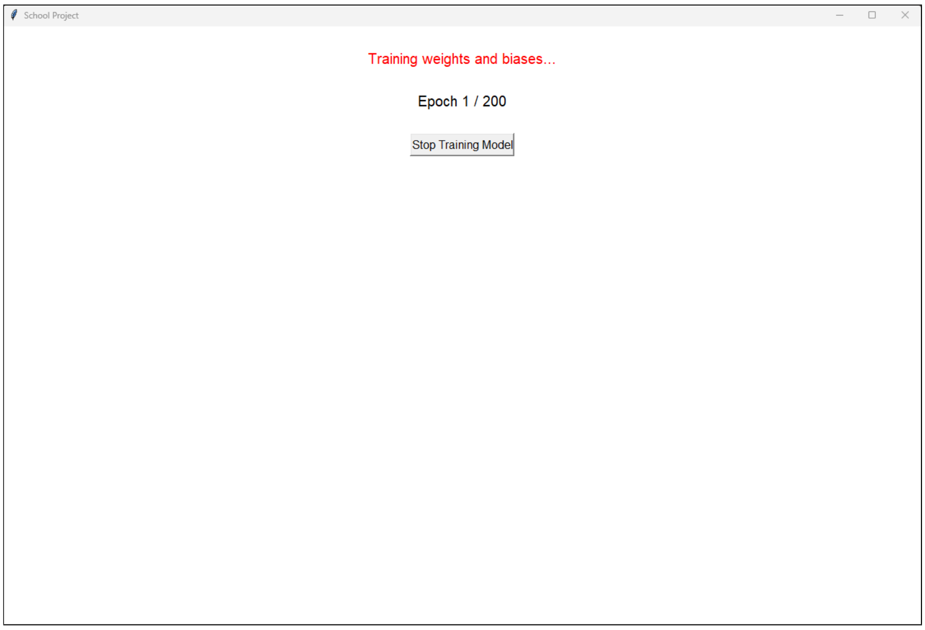


Figure 18: Training frame showing epoch count

And outputs the following for the training frame once training has completed as shown in figure 19:

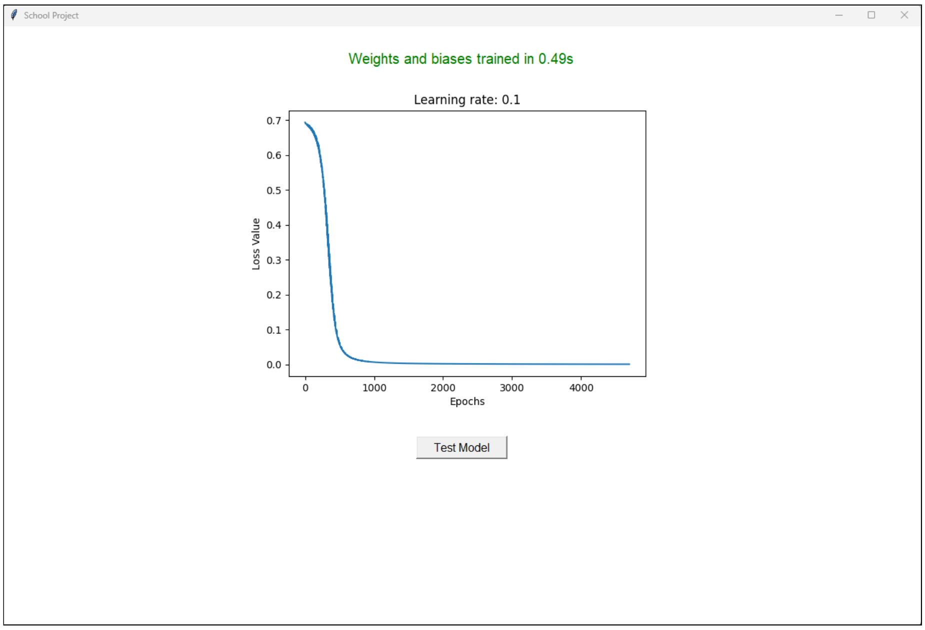


Figure 19: Training frame showing loss value against epochs

- load_model.py module:

```python
"""Tkinter frames for loading a saved Artificial Neural Network Model."""

import sqlite3
import tkinter as tk
import tkinter.font as tkf

class LoadModelFrame(tk.Frame):
    """Frame for load model page."""
    def __init__(self, root: tk.Tk,
                 width: int, height: int,
                 bg: str, connection: sqlite3.Connection,
                 cursor: sqlite3.Cursor, dataset: str) -> None:
        """Initialise load model frame widgets.

        Args:
            root (tk.Tk): the widget object that contains this widget.
            width (int): the pixel width of the frame.
            height (int): the pixel height of the frame.
            bg (str): the hex value or name of the frame's background
            ↪ colour.
            connection (sqlite3.Connection): the database connection
            ↪ object.
            cursor (sqlite3.Cursor): the database cursor object.
            dataset (str): the name of the dataset to use
            ('MNIST', 'Cat Recognition' or 'XOR')
        Raises:
```

```python
            TypeError: if root, width or height are not of the correct type.

        """
        super().__init__(master=root, width=width, height=height, bg=bg)
        self.root = root
        self.WIDTH = width
        self.HEIGHT = height
        self.BG = bg

        # Setup load model frame variables
        self.connection = connection
        self.cursor = cursor
        self.dataset = dataset
        self.use_gpu: bool
        self.model_options = self.load_model_options()

        # Setup widgets
        self.title_label = tk.Label(master=self,
                                    bg=self.BG,
                                    font=('Arial', 20),
                                    text=dataset)
        self.about_label = tk.Label(
                    master=self,
                    bg=self.BG,
                    font=('Arial', 14),
                    text=f"Load a pretrained model for the {dataset} dataset."
                    )
        self.model_status_label = tk.Label(master=self,
                                           bg=self.BG,
                                           font=('Arial', 15))

        # Don't give loaded model options if no models have been saved for the
        # dataset.
        if len(self.model_options) > 0:
            self.model_option_menu_label = tk.Label(
                                                master=self,
                                                bg=self.BG,
                                                font=('Arial', 14),
                                                text="Select a model to load or delete:"
                                                )
            self.model_option_menu_var = tk.StringVar(
                                                   master=self,
                                                   value=self.model_options[0]
                                                   )
            self.model_option_menu = tk.OptionMenu(
                                                    self,
                                                    self.model_option_menu_var,
                                                    *self.model_options
                                                    )
            self.use_gpu_check_button_var = tk.BooleanVar()
            self.use_gpu_check_button = tk.Checkbutton(
                                        master=self,
                                        width=7, height=1,
                                        font=tkf.Font(size=12),
                                        text="Use GPU",
                                        variable=self.use_gpu_check_button_var
                                        )
```

```
        else:
            self.model_status_label.configure(
                                   text='No saved models for this dataset.',
                                   fg='red'
                                   )

        # Pack widgets
        self.title_label.grid(row=0, column=0, columnspan=3)
        self.about_label.grid(row=1, column=0, columnspan=3)
        if len(self.model_options) > 0:  # Check if options should be given
            self.model_option_menu_label.grid(row=2, column=0, padx=(0,30), pady=(30,0))
            self.use_gpu_check_button.grid(row=2, column=2, rowspan=2, pady=(30,0))
            self.model_option_menu.grid(row=3, column=0, padx=(0,30), pady=(10,0))
        self.model_status_label.grid(row=4, column=0,
                                     columnspan=3, pady=50)

    def load_model_options(self) -> list[str]:
        """Load the model options from the database.

           Returns:
               a list of the model options.
        """
        sql = f"""
        SELECT Name FROM Models WHERE Dataset=?
        """
        parameters = (self.dataset.replace(" ", "_"),)
        self.cursor.execute(sql, parameters)

        # Save the string value contained within the tuple of each row
        model_options = []
        for model_option in self.cursor.fetchall():
            model_options.append(model_option[0])

        return model_options

    def load_model(self) -> object:
        """Create model using saved weights and biases.

           Returns:
              a Model object.

        """
        self.use_gpu = self.use_gpu_check_button_var.get()

        # Query data of selected saved model from database
        sql = """
        SELECT * FROM Models WHERE Dataset=? AND Name=?
        """
        parameters = (self.dataset.replace(" ", "_"), self.model_option_menu_var.get())
        self.cursor.execute(sql, parameters)
        data = self.cursor.fetchone()
        hidden_layers_shape_input = [layer for layer in data[3].replace(' ', '').split(',') if layer != '']

        # Create Model
        if not self.use_gpu:
            if self.dataset == "MNIST":
                from school_project.models.cpu.mnist import MNISTModel as Model
```

```python
            elif self.dataset == "Cat Recognition":
                from school_project.models.cpu.cat_recognition import CatRecognitionModel as Model
            elif self.dataset == "XOR":
                from school_project.models.cpu.xor import XORModel as Model
            model = Model(
                hidden_layers_shape=[int(neuron_count) for neuron_count in hidden_layers_shape_input],
                train_dataset_size=data[6],
                learning_rate=data[4],
                use_relu=data[7]
                )
            model.load_model_values(file_location=data[2])

        else:
            try:
                if self.dataset == "MNIST":
                    from school_project.models.gpu.mnist import MNISTModel as Model
                elif self.dataset == "Cat Recognition":
                    from school_project.models.gpu.cat_recognition import CatRecognitionModel as Model
                elif self.dataset == "XOR":
                    from school_project.models.gpu.xor import XORModel as Model
                model = Model(
                    hidden_layers_shape=[int(neuron_count) for neuron_count in hidden_layers_shape_input],
                    train_dataset_size=data[6],
                    learning_rate=data[4],
                    use_relu=data[7]
                    )
                model.load_model_values(file_location=data[2])
            except ImportError as ie:
                self.model_status_label.configure(
                                              text="Failed to initialise GPU",
                                              fg='red'
                                              )
                raise ImportError
        return model
```

This outputs the following for the load model frame when models have been saved for the dataset as shown in figure 20:

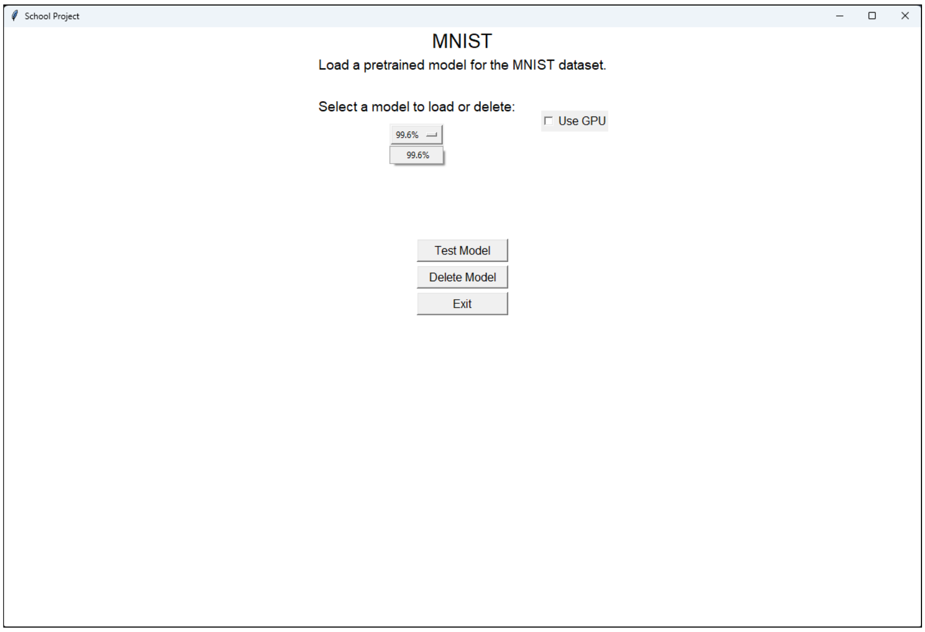


Figure 20: Load model frame

And outputs the following for the load model frame when no models have been saved for the dataset as shown in figure 21:

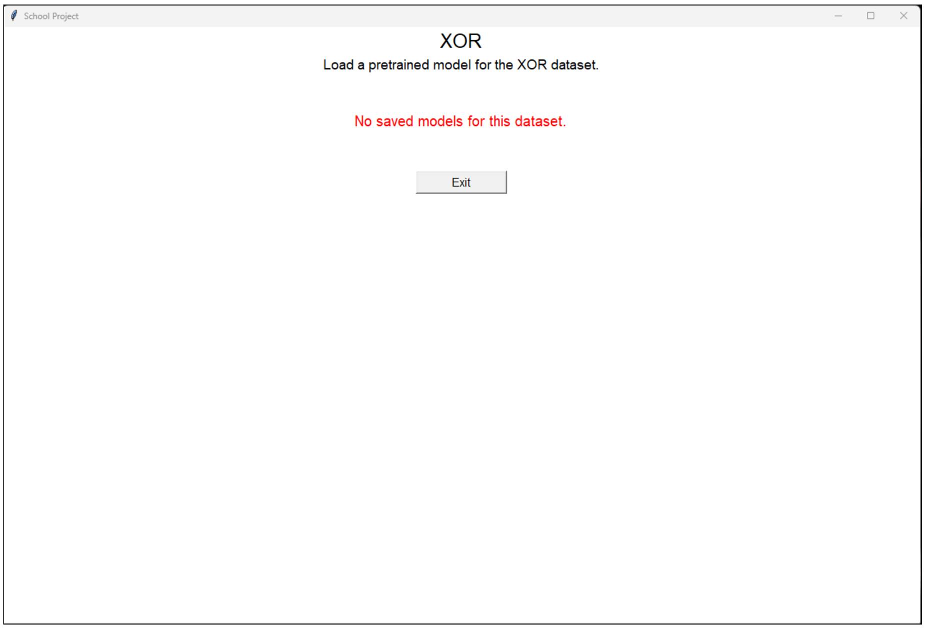


Figure 21: Load model frame showing error condition for an attempted load of a non-existent model

4.4 Project Entrypoint - __main__.py module

This module is the entrypoint to the project and loads the main window of the User Interface:

```python
"""The entrypoint of A-level Computer Science NEA Programming Project."""

import os
import sqlite3
import threading
import tkinter as tk
import tkinter.font as tkf
import uuid

from school_project.frames import (HyperParameterFrame, TrainingFrame,
                                   LoadModelFrame, TestMNISTFrame,
                                   TestCatRecognitionFrame, TestXORFrame)

class SchoolProjectFrame(tk.Frame):
    """Main frame of school project."""
    def __init__(self, root: tk.Tk, width: int, height: int, bg: str) -> None:
        """Initialise school project pages.

        Args:
            root (tk.Tk): the widget object that contains this widget.
            width (int): the pixel width of the frame.
            height (int): the pixel height of the frame.
            bg (str): the hex value or name of the frame's background colour.
        Raises:
            TypeError: if root, width or height are not of the correct type.

        """
        super().__init__(master=root, width=width, height=height, bg=bg)
        self.root = root.title("School Project")
        self.WIDTH = width
        self.HEIGHT = height
        self.BG = bg

        # Setup school project frame variables
        self.hyper_parameter_frame: HyperParameterFrame
        self.training_frame: TrainingFrame
        self.load_model_frame: LoadModelFrame
        self.test_frame: TestMNISTFrame | TestCatRecognitionFrame | TestXORFrame
        self.connection, self.cursor = self.setup_database()
        self.model = None

        # Record if the model should be saved after testing,
        # as only newly created models should be given the option to be saved.
        self.saving_model: bool

        # Setup school project frame widgets
        self.exit_hyper_parameter_frame_button = tk.Button(
                                    master=self,
                                    width=13,
                                    height=1,
                                    font=tkf.Font(size=12),
                                    text="Exit",
                                    command=self.exit_hyper_parameter_frame
                                    )
        self.exit_load_model_frame_button = tk.Button(
                                         master=self,
                                         width=13,
                                         height=1,
```

```python
                                    font=tkf.Font(size=12),
                                    text="Exit",
                                    command=self.exit_load_model_frame
                                    )
        self.train_button = tk.Button(master=self,
                                      width=13,
                                      height=1,
                                      font=tkf.Font(size=12),
                                      text="Train Model",
                                      command=self.enter_training_frame)
        self.stop_training_button = tk.Button(
                                        master=self,
                                        width=15, height=1,
                                        font=tkf.Font(size=12),
                                        text="Stop Training Model",
                                        command=lambda: self.model.set_running(
                                                    value=False
                                                    )
                                        )
        self.test_created_model_button = tk.Button(
                                            master=self,
                                            width=13, height=1,
                                            font=tkf.Font(size=12),
                                            text="Test Model",
                                            command=self.test_created_model
                                            )
        self.test_loaded_model_button = tk.Button(
                                             master=self,
                                             width=13, height=1,
                                             font=tkf.Font(size=12),
                                             text="Test Model",
                                             command=self.test_loaded_model
                                             )
        self.delete_loaded_model_button = tk.Button(
                                             master=self,
                                             width=13, height=1,
                                             font=tkf.Font(size=12),
                                             text="Delete Model",
                                             command=self.delete_loaded_model
                                             )
        self.save_model_label = tk.Label(
                                    master=self,
                                    text="Enter a name for your trained model:",
                                    bg=self.BG,
                                    font=('Arial', 15)
                                    )
        self.save_model_name_entry = tk.Entry(master=self, width=13)
        self.save_model_button = tk.Button(master=self,
                                           width=13,
                                           height=1,
                                           font=tkf.Font(size=12),
                                           text="Save Model",
                                           command=self.save_model)
        self.exit_button = tk.Button(master=self,
                                     width=13, height=1,
                                     font=tkf.Font(size=12),
                                     text="Exit",
                                     command=self.enter_home_frame)

        # Setup home frame
        self.home_frame = tk.Frame(master=self,
                                   width=self.WIDTH,
                                   height=self.HEIGHT,
                                   bg=self.BG)
```

```
        self.title_label = tk.Label(
                        master=self.home_frame,
                        bg=self.BG,
                        font=('Arial', 20),
                        text="A-level Computer Science NEA Programming Project"
                        )
        self.about_label = tk.Label(
            master=self.home_frame,
            bg=self.BG,
            font=('Arial', 14),
            text="An investigation into how Artificial Neural Networks work, " +
            "the effects of their hyper-parameters and their applications " +
            "in Image Recognition.\n\n" +
            " - Max Cotton"
            )
        self.model_menu_label = tk.Label(master=self.home_frame,
                                         bg=self.BG,
                                         font=('Arial', 14),
                                         text="Create a new model " +
                                         "or load a pre-trained model "
                                         "for one of the following datasets:")
        self.dataset_option_menu_var = tk.StringVar(master=self.home_frame,
                                                    value="MNIST")
        self.dataset_option_menu = tk.OptionMenu(self.home_frame,
                                                 self.dataset_option_menu_var,
                                                 "MNIST",
                                                 "Cat Recognition",
                                                 "XOR")
        self.create_model_button = tk.Button(
                                    master=self.home_frame,
                                    width=13, height=1,
                                    font=tkf.Font(size=12),
                                    text="Create Model",
                                    command=self.enter_hyper_parameter_frame
                                    )
        self.load_model_button = tk.Button(master=self.home_frame,
                                           width=13, height=1,
                                           font=tkf.Font(size=12),
                                           text="Load Model",
                                           command=self.enter_load_model_frame)

        # Grid home frame widgets
        self.title_label.grid(row=0, column=0, columnspan=4, pady=(10,0))
        self.about_label.grid(row=1, column=0, columnspan=4, pady=(10,50))
        self.model_menu_label.grid(row=2, column=0, columnspan=4)
        self.dataset_option_menu.grid(row=3, column=0, columnspan=4, pady=30)
        self.create_model_button.grid(row=4, column=1)
        self.load_model_button.grid(row=4, column=2)

        self.home_frame.pack()

        # Setup frame attributes
        self.grid_propagate(flag=False)
        self.pack_propagate(flag=False)

    @staticmethod
    def setup_database() -> tuple[sqlite3.Connection, sqlite3.Cursor]:
        """Create a connection to the pretrained_models database file and
           setup base table if needed.

           Returns:
               a tuple of the database connection and the cursor for it.

        """
```

```python
        connection = sqlite3.connect(
                                database='school_project/saved_models.db'
                                )
        cursor = connection.cursor()
        cursor.execute("""
        CREATE TABLE IF NOT EXISTS Models
        (Model_ID INTEGER PRIMARY KEY,
        Dataset TEXT NOT NULL,
        File_Location TEXT NOT NULL,
        Hidden_Layers_Shape TEXT NOT NULL,
        Learning_Rate FLOAT NOT NULL,
        Name TEXT NOT NULL,
        Train_Dataset_Size INTEGER NOT NULL,
        Use_ReLu INTEGER NOT NULL,
        UNIQUE (Dataset, Name))
        """)
        return (connection, cursor)

    def enter_hyper_parameter_frame(self) -> None:
        """Unpack home frame and pack hyper-parameter frame."""
        self.home_frame.pack_forget()
        self.hyper_parameter_frame = HyperParameterFrame(
                                    root=self,
                                    width=self.WIDTH,
                                    height=self.HEIGHT,
                                    bg=self.BG,
                                    dataset=self.dataset_option_menu_var.get()
                                    )
        self.hyper_parameter_frame.pack()
        self.train_button.pack()
        self.exit_hyper_parameter_frame_button.pack(pady=(10,0))

    def enter_load_model_frame(self) -> None:
        """Unpack home frame and pack load model frame."""
        self.home_frame.pack_forget()
        self.load_model_frame = LoadModelFrame(
                                    root=self,
                                    width=self.WIDTH,
                                    height=self.HEIGHT,
                                    bg=self.BG,
                                    connection=self.connection,
                                    cursor=self.cursor,
                                    dataset=self.dataset_option_menu_var.get()
                                    )
        self.load_model_frame.pack()

        # Don't give option to test loaded model if no models have been saved
        # for the dataset.
        if len(self.load_model_frame.model_options) > 0:
            self.test_loaded_model_button.pack()
            self.delete_loaded_model_button.pack(pady=(5,0))

        self.exit_load_model_frame_button.pack(pady=(5,0))

    def exit_hyper_parameter_frame(self) -> None:
        """Unpack hyper-parameter frame and pack home frame."""
        self.hyper_parameter_frame.pack_forget()
        self.train_button.pack_forget()
        self.exit_hyper_parameter_frame_button.pack_forget()
        self.home_frame.pack()

    def exit_load_model_frame(self) -> None:
        """Unpack load model frame and pack home frame."""
        self.load_model_frame.pack_forget()
```

```python
        self.test_loaded_model_button.pack_forget()
        self.delete_loaded_model_button.pack_forget()
        self.exit_load_model_frame_button.pack_forget()
        self.home_frame.pack()

    def enter_training_frame(self) -> None:
        """Load untrained model from hyper parameter frame,
           unpack hyper-parameter frame, pack training frame
           and begin managing the training thread.
        """
        try:
            self.model = self.hyper_parameter_frame.create_model()
        except (ValueError, ImportError) as e:
            return
        self.hyper_parameter_frame.pack_forget()
        self.train_button.pack_forget()
        self.exit_hyper_parameter_frame_button.pack_forget()
        self.training_frame = TrainingFrame(
                root=self,
                width=self.WIDTH,
                height=self.HEIGHT,
                bg=self.BG,
                model=self.model,
                epoch_count=self.hyper_parameter_frame.epoch_count_scale.get()
                )
        self.training_frame.pack()
        self.stop_training_button.pack()
        self.manage_training(train_thread=self.training_frame.train_thread)

    def manage_training(self, train_thread: threading.Thread) -> None:
        """Wait for model training thread to finish,
           then plot training losses on training frame.

        Args:
            train_thread (threading.Thread):
            the thread running the model's train() method.
        Raises:
            TypeError: if train_thread is not of type threading.Thread.

        """
        if not train_thread.is_alive():
            self.training_frame.training_progress_label.pack_forget()
            self.training_frame.plot_losses(model=self.model)
            self.stop_training_button.pack_forget()
            self.test_created_model_button.pack(pady=(30,0))
        else:
            self.training_frame.training_progress_label.configure(
                                            text=self.model.training_progress
                                            )
            self.after(100, self.manage_training, train_thread)

    def test_created_model(self) -> None:
        """Unpack training frame, pack test frame for the dataset
           and begin managing the test thread."""
        self.saving_model = True
        self.training_frame.pack_forget()
        self.test_created_model_button.pack_forget()
        if self.hyper_parameter_frame.dataset == "MNIST":
            self.test_frame = TestMNISTFrame(
                                    root=self,
                                    width=self.WIDTH,
                                    height=self.HEIGHT,
                                    bg=self.BG,
                                    use_gpu=self.hyper_parameter_frame.use_gpu,
```

```python
                                model=self.model
                                )
        elif self.hyper_parameter_frame.dataset == "Cat Recognition":
            self.test_frame = TestCatRecognitionFrame(
                                root=self,
                                width=self.WIDTH,
                                height=self.HEIGHT,
                                bg=self.BG,
                                use_gpu=self.hyper_parameter_frame.use_gpu,
                                model=self.model
                                )
        elif self.hyper_parameter_frame.dataset == "XOR":
            self.test_frame = TestXORFrame(root=self,
                                        width=self.WIDTH,
                                        height=self.HEIGHT,
                                        bg=self.BG,
                                        model=self.model)
        self.test_frame.pack()
        self.manage_testing(test_thread=self.test_frame.test_thread)

    def test_loaded_model(self) -> None:
        """Load saved model from load model frame, unpack load model frame,
           pack test frame for the dataset and begin managing the test thread."""
        self.saving_model = False
        try:
            self.model = self.load_model_frame.load_model()
        except (ValueError, ImportError) as e:
            return
        self.load_model_frame.pack_forget()
        self.test_loaded_model_button.pack_forget()
        self.delete_loaded_model_button.pack_forget()
        self.exit_load_model_frame_button.pack_forget()
        if self.load_model_frame.dataset == "MNIST":
            self.test_frame = TestMNISTFrame(
                                     root=self,
                                     width=self.WIDTH,
                                     height=self.HEIGHT,
                                     bg=self.BG,
                                     use_gpu=self.load_model_frame.use_gpu,
                                     model=self.model
                                     )
        elif self.load_model_frame.dataset == "Cat Recognition":
            self.test_frame = TestCatRecognitionFrame(
                                     root=self,
                                     width=self.WIDTH,
                                     height=self.HEIGHT,
                                     bg=self.BG,
                                     use_gpu=self.load_model_frame.use_gpu,
                                     model=self.model
                                     )
        elif self.load_model_frame.dataset == "XOR":
            self.test_frame = TestXORFrame(root=self,
                                        width=self.WIDTH,
                                        height=self.HEIGHT,
                                        bg=self.BG,
                                        model=self.model)
        self.test_frame.pack()
        self.manage_testing(test_thread=self.test_frame.test_thread)

    def manage_testing(self, test_thread: threading.Thread) -> None:
        """Wait for model test thread to finish,
           then plot results on test frame.

        Args:
```

```python
            test_thread (threading.Thread):
            the thread running the model's predict() method.
        Raises:
            TypeError: if test_thread is not of type threading.Thread.

        """
        if not test_thread.is_alive():
            self.test_frame.plot_results(model=self.model)
            if self.saving_model:
                self.save_model_label.pack(pady=(30,0))
                self.save_model_name_entry.pack(pady=10)
                self.save_model_button.pack()
            self.exit_button.pack(pady=(20,0))
        else:
            self.after(1_000, self.manage_testing, test_thread)

    def save_model(self) -> None:
        """Save the model, save the model information to the database, then
           enter the home frame."""
        model_name = self.save_model_name_entry.get().strip()

        # Check if model name is empty
        if len(model_name) == 0:
            self.test_frame.model_status_label.configure(
                                           text="Model name can not be blank",
                                           fg='red'
                                           )
            return

        # Check if model contains double spaces or greater
        elif '  ' in model_name:
            self.test_frame.model_status_label.configure(
                                           text="Only single spaces are allowed",
                                           fg='red'
                                           )
            return


        # Check if model name has already been taken
        dataset = self.dataset_option_menu_var.get().replace(" ", "_")
        sql = """
        SELECT Name FROM Models WHERE Dataset=?
        """
        parameters = (dataset,)
        self.cursor.execute(sql, parameters)
        for saved_model_name in self.cursor.fetchall():
            if saved_model_name[0] == model_name:
                self.test_frame.model_status_label.configure(
                                                       text="Model name taken",
                                                       fg='red'
                                                       )
                return

        # Save model to random hex file name
        file_location = f"school_project/saved-models/{uuid.uuid4().hex}.npz"
        self.model.save_model_values(file_location=file_location)

        # Save the model information to the database
        sql = """
        INSERT INTO Models
        (Dataset, File_Location, Hidden_Layers_Shape, Learning_Rate, Name,
        ↪  Train_Dataset_Size, Use_ReLu)
        VALUES (?, ?, ?, ?, ?, ?, ?)
        """
```

```python
        parameters = (
                    dataset,
                    file_location,
                    self.hyper_parameter_frame.hidden_layers_shape_entry.get(),
                    self.hyper_parameter_frame.learning_rate_scale.get(),
                    model_name,
                    self.hyper_parameter_frame.train_dataset_size_scale.get(),
                    self.hyper_parameter_frame.use_relu_check_button_var.get()
                    )
        self.cursor.execute(sql, parameters)
        self.connection.commit()

        self.enter_home_frame()

    def delete_loaded_model(self) -> None:
        """Delete saved model file and model data from the database."""
        dataset = self.dataset_option_menu_var.get().replace(" ", "_")
        model_name = self.load_model_frame.model_option_menu_var.get()

        # Delete saved model
        sql = f"""SELECT File_Location FROM Models WHERE Dataset=? AND Name=?"""
        parameters = (dataset, model_name)
        self.cursor.execute(sql, parameters)
        os.remove(self.cursor.fetchone()[0])

        # Remove model data from database
        sql = """DELETE FROM Models WHERE Dataset=? AND Name=?"""
        parameters = (dataset, model_name)
        self.cursor.execute(sql, parameters)
        self.connection.commit()

        # Reload load model frame with new options
        self.exit_load_model_frame()
        self.enter_load_model_frame()

    def enter_home_frame(self) -> None:
        """Unpack test frame and pack home frame."""
        self.model = None  # Free up trained Model from memory
        self.test_frame.pack_forget()
        if self.saving_model:
            self.save_model_label.pack_forget()
            self.save_model_name_entry.delete(0, tk.END)  # Clear entry's text
            self.save_model_name_entry.pack_forget()
            self.save_model_button.pack_forget()
        self.exit_button.pack_forget()
        self.home_frame.pack()

def main() -> None:
    """Entrypoint of project."""
    root = tk.Tk()
    school_project_frame = SchoolProjectFrame(root=root, width=1280,
                                    height=835, bg='white')
    school_project_frame.pack(side='top', fill='both', expand=True)
    root.mainloop()

    # Stop model training when GUI closes
    if school_project_frame.model is not None:
        school_project_frame.model.set_running(value=False)

if __name__ == "__main__":
    main()
```

Which outputs the following for the home frame:

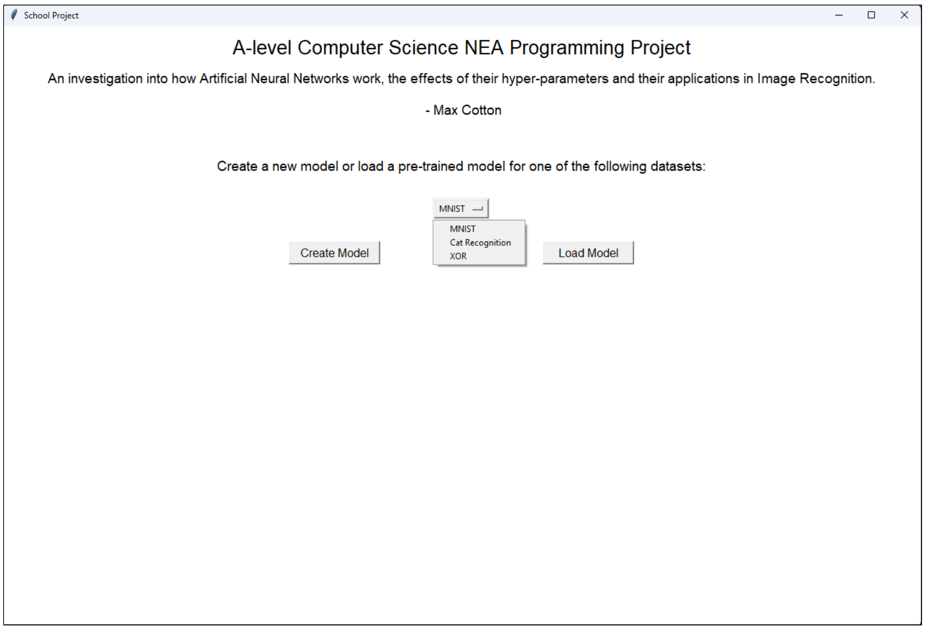


Figure 22: Home frame - the entry point to the program

5 Testing

Testing on the project source code consists of manual testing, where the inputs to frames were tested to ensure correct behaviour and error handling, and automated testing through the use of unit tests that run upon every commit to GitHub.

5.1 Summary of tests

Test Type	Description	Result
Manual	Hyper Parameter Frame - Use GPU Validation	Pass
Manual	Hyper Parameter Frame - Non-Numeric Hidden Layers Shape Validation	Pass
Manual	Hyper Parameter Frame - Negative Hidden Layers Shape Validation	Pass
Manual	Hyper Parameter Frame - Invalid Delimiter Hidden Layers Shape Validation	Pass
Manual	Load Model Frame - Use GPU Validation	Pass
Manual	Test Frames - Taken Trained Model Name Validation	Pass
Manual	Test Frames - Empty Trained Model Name Validation	Pass
Manual	Test Frames - Invalid Delimiter Trained Model Name Validation	Pass
Unit	test_database.py - test_database_structure	Pass
Unit	test_database.py - test_not_null_constraint	Pass
Unit	test_database.py - test_unique_constraint	Pass
Unit	test_database.py - test_save_load_consistency	Pass
Unit	test_model.py - test_train_dataset_size	Pass
Unit	test_model.py - test_network_shape	Pass
Unit	test_model.py - test_learning_rates	Pass
Unit	test_model.py - test_relu_model_transfer_types	Pass
Unit	test_model.py - test_sigmoid_model_transfer_types	Pass
Unit	test_model.py - test_weight_matrice_shapes	Pass
Unit	test_model.py - test_biase_matrice_shapes	Pass
Unit	test_model.py - test_layer_output_shapes	Pass
Unit	test_model.py - test_save_model	Pass
Unit	test_tools.py - test_relu	Pass
Unit	test_tools.py - test_sigmoid	Pass

5.2 Manual Testing - Input Validation Testing

The following tests check the input validation of each frames' inputs.

5.2.1 Hyper Parameter Frame

- Use GPU Validation:

Description	Select Use GPU checkbox without a GPU present.
Expected Result	The exception should be handled and a useful error message should be displayed.
Actual Result	Expected Result
Test Status	Pass

Evidence:

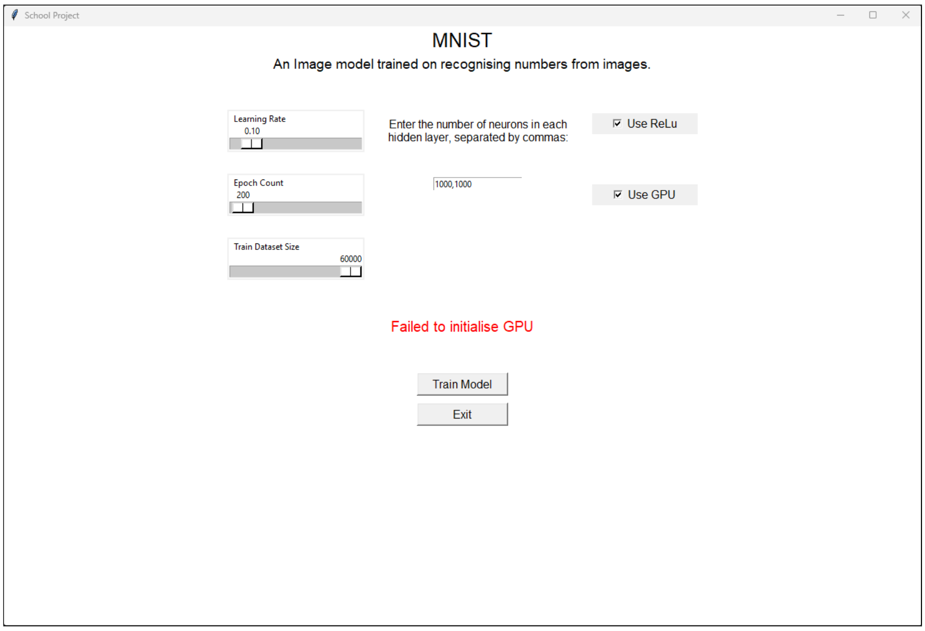


Figure 23: Use GPU Validation evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#use-gpu-validation`

- Non-Numeric Hidden Layers Shape Validation:

Description	Enter a non-numeric hidden layers shape.
Data Value	"test"
Data Type	Erroneous
Expected Result	The exception should be handled and a useful error message should be displayed.
Actual Result	Expected Result
Test Status	Pass

Evidence:

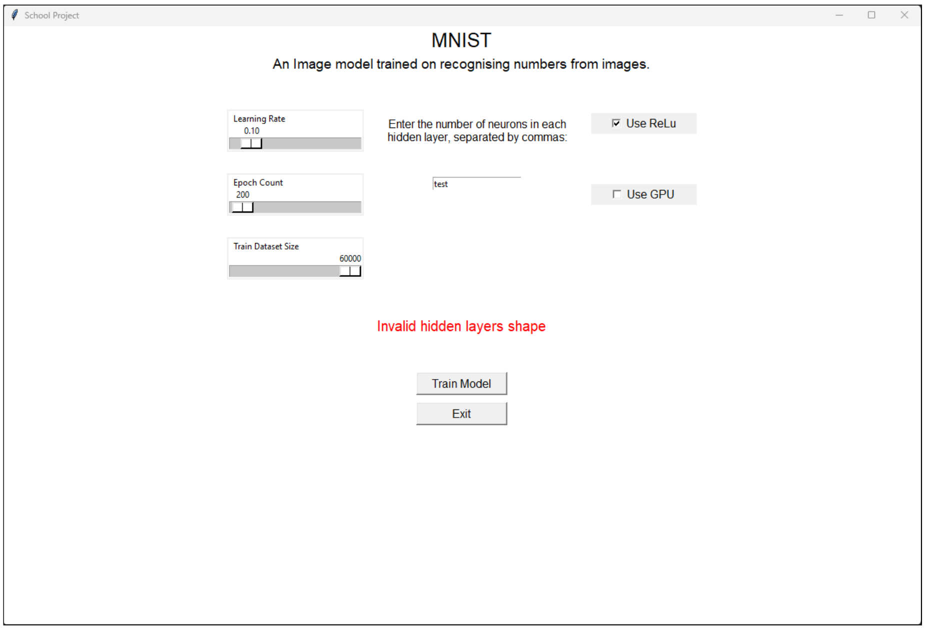


Figure 24: Non-Numeric Hidden Layers Shape Validation evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#non-numeric-hidden-layers-shape-validation`

- Negative Hidden Layers Shape Validation:

Description	Enter a negative hidden layers shape.
Data Value	"-100"
Data Type	Erroneous
Expected Result	The exception should be handled and a useful error message should be displayed.
Actual Result	Expected Result
Test Status	Pass

Evidence:

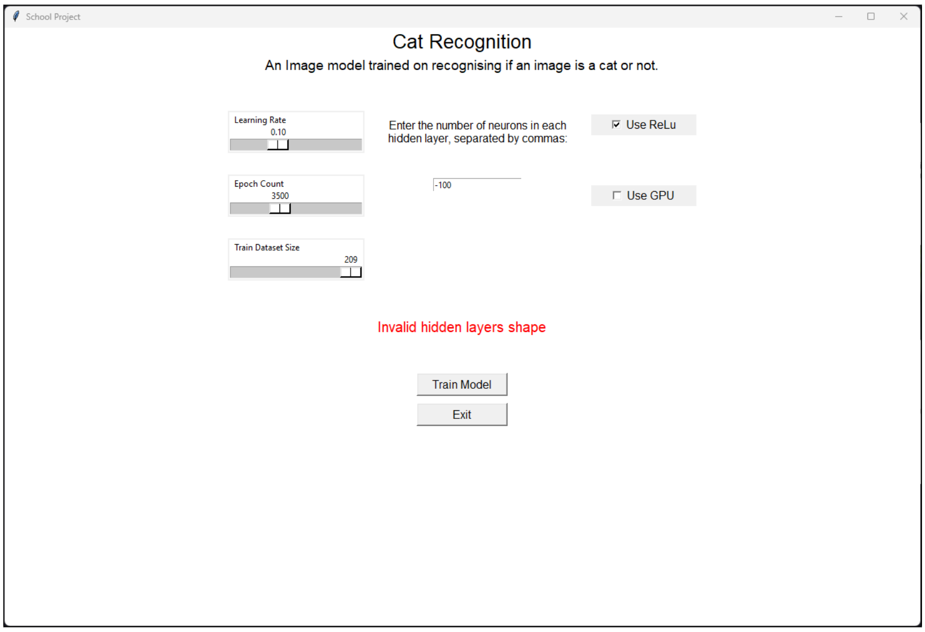


Figure 25: Negative Hidden Layers Shape Validation evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#negative-hidden-layers-shape-validation`

- Invalid Delimiter Hidden Layers Shape Validation:

Description	Enter a hidden layers shape with invalid delimiters.
Data Value	"100,,100"
Data Type	Erroneous
Expected Result	The exception should be handled and a useful error message should be displayed.
Actual Result	Expected Result
Test Status	Pass

Evidence:

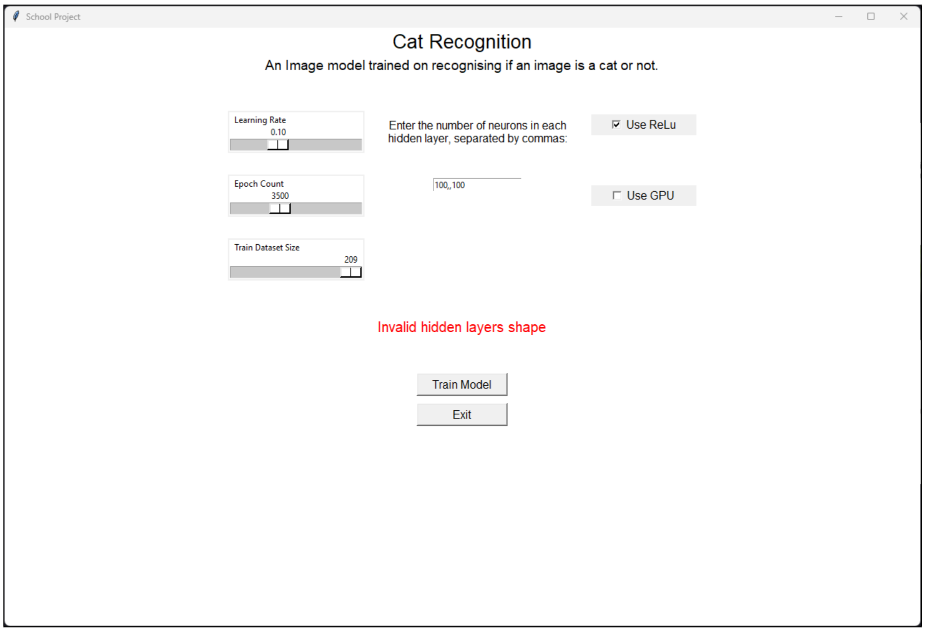


Figure 26: Invalid Delimiter Hidden Layers Shape Validation evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#invalid-delimiter-hidden-layers-shape-validation`

5.2.2 Load Model Frame

- Use GPU Validation:

Description	Select Use GPU checkbox without a GPU present.
Expected Result	The exception should be handled and a useful error message should be displayed.
Actual Result	Expected Result
Test Status	Pass

Evidence:

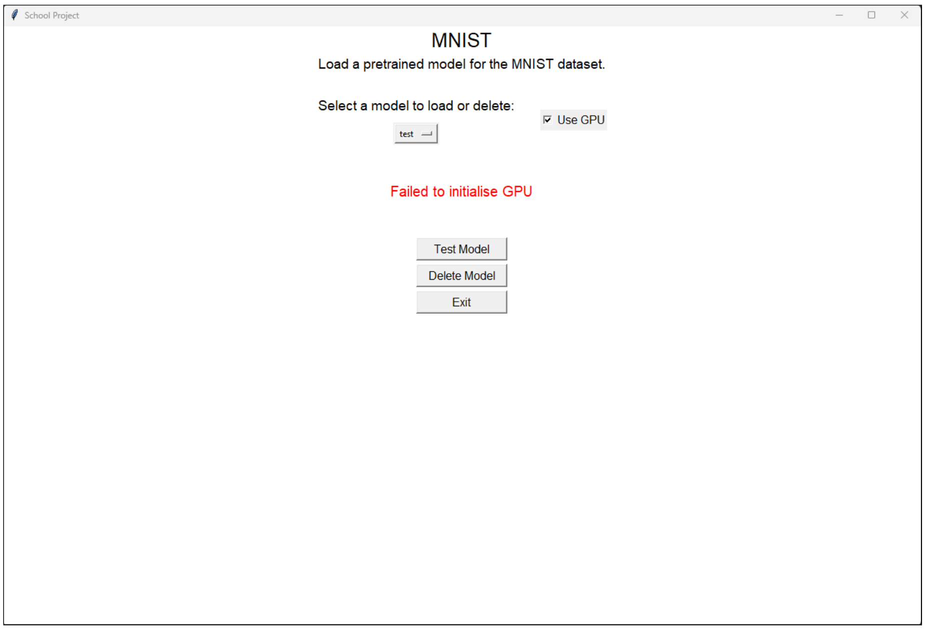


Figure 27: Use GPU Validation evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#use-gpu-validation-1`

5.2.3 Test Frames

- Taken Trained Model Name Validation:

Description	Try to save a trained model with an already taken name.
Data Value	"test"
Data Type	Erroneous
Expected Result	The exception should be handled and a useful error message should be displayed.
Actual Result	Expected Result
Test Status	Pass

Evidence:

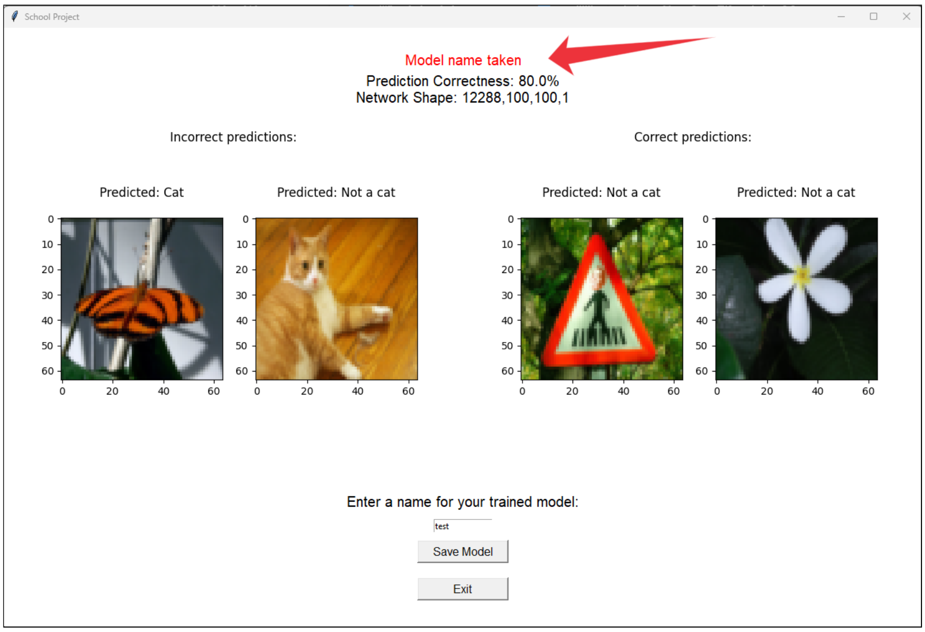


Figure 28: Taken Trained Model Name Validation evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#taken-trained-model-name-validation`

- Empty Trained Model Name Validation:

Description	Try to save a trained model with blank name.
Data Value	””
Expected Result	The exception should be handled and a useful error message should be displayed.
Actual Result	Expected Result
Test Status	Pass

Evidence:

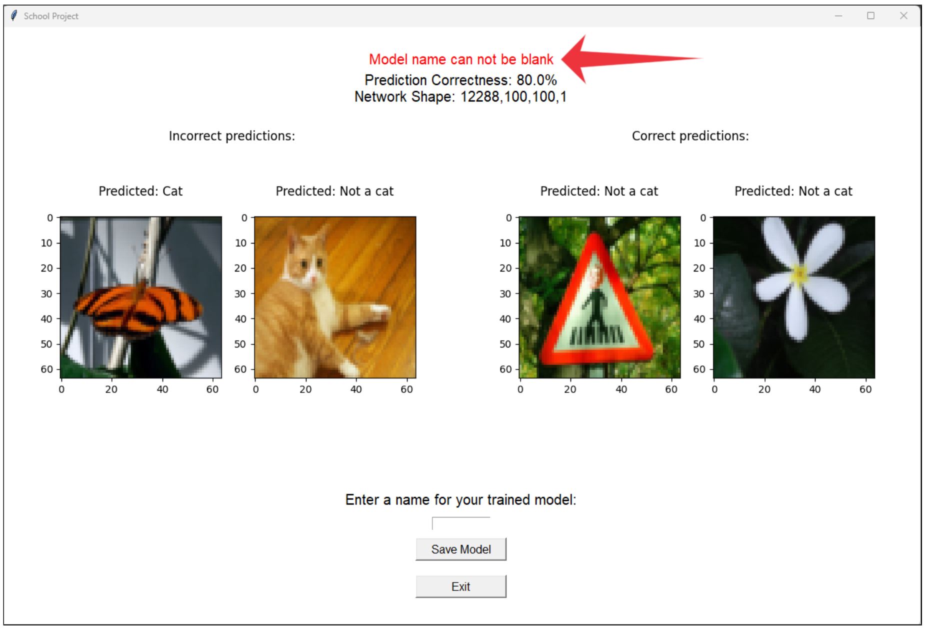


Figure 29: Empty Trained Model Name Validation evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#empty-trained-model-name-validation`

- Invalid Delimiter Trained Model Name Validation:

Description	Try to save a trained model with a name with incorrect delimiters.
Data Value	"test test"
Expected Result	The exception should be handled and a useful error message should be displayed.
Actual Result	Expected Result
Test Status	Pass

Evidence:

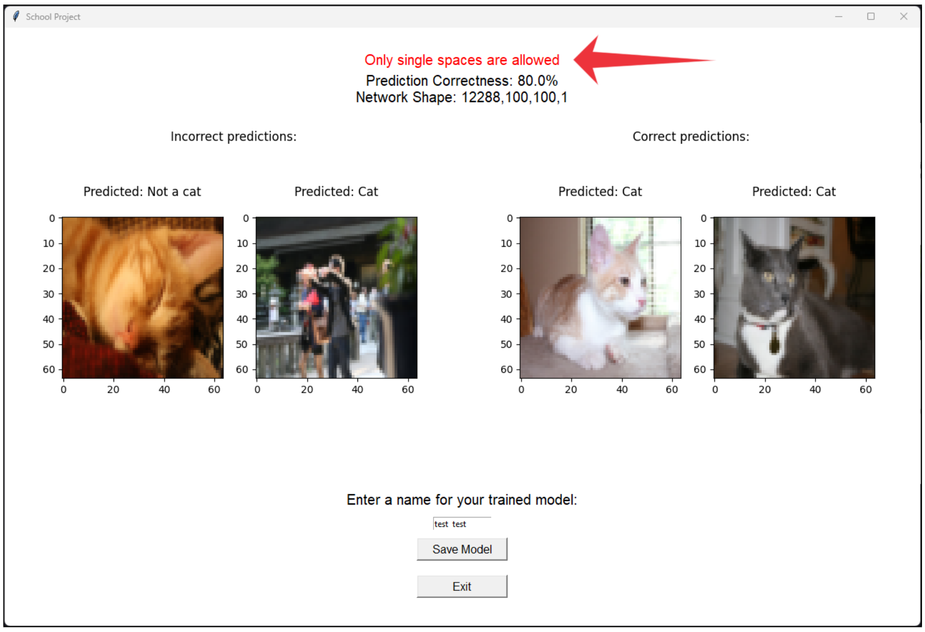


Figure 30: Invalid Delimiter Trained Model Name Validation evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#invalid-delimiter-trained-model-name-validation`

5.3 Automated Testing

5.3.1 Unit Tests

Within the test package, I have written the following unit tests:

- Unit tests for the database in a test_database.py module:
  - test_database_structure:

Description	Test that the database tables are set up correctly.
Expected Result	Check that the 'Models' table exists in the database and that the table's info matches the following format: [(0, 'Model_ID', 'INTEGER', 0, None, 1), (1, 'Dataset', 'TEXT', 1, None, 0), (2, 'File_Location', 'TEXT', 1, None, 0), (3, 'Hidden_Layers_Shape', 'TEXT', 1, None, 0), (4, 'Learning_Rate', 'FLOAT', 1, None, 0), (5, 'Name', 'TEXT', 1, None, 0), (6, 'Train_Dataset_Size', 'INTEGER', 1, None, 0), (7, 'Use_ReLu', 'INTEGER', 1, None, 0)]
Actual Result	Expected Result
Test Status	Pass

– test_not_null_constraint:

Description	Test that the NOT NULL constraint is setup.
Data Value	("Test_Dataset", f"school_project/saved-models/uuid.uuid4().hex.npz", "100, 100", 0.1, "Test_Name", 100)
Data Type	Erroneous
Expected Result	A sqlite3.IntegrityError should be raised.
Actual Result	Expected Result
Test Status	Pass

– test_unique_constraint:

Description	Test that the UNIQUE (Dataset, Name) constraint is setup.
Data Value	("Test_Dataset", f"school_project/saved-models/uuid.uuid4().hex.npz", "100, 100", 0.1, "Test_Name", 100, True)
Data Type	Erroneous
Expected Result	A sqlite3.IntegrityError should be raised.
Actual Result	Expected Result
Test Status	Pass

– test_save_load_consistency:

Description	Test that data is not changed between saving and loading.
Data Value	(”Test_Dataset”, f”school_project/saved-models/uuid.uuid4().hex.npz”, ”100, 100”, 0.1, ”Test_Name”, 100, True)
Data Type	Normal
Expected Result	Data is not changed between saving and loading.
Actual Result	Expected Result
Test Status	Pass

– Evidence:

```python
"""Unit tests for database."""

import sqlite3
import unittest
import uuid

class TestDatabase(unittest.TestCase):
    """Unit tests for database."""
    def __init__(self, *args, **kwargs) -> None:
        """Initialise unit tests."""
        super(TestDatabase, self).__init__(*args, **kwargs)

    def test_database_strucure(self) -> None:
        """Test that the database tables are set up correctly."""
        connection = sqlite3.connect(database='school_project/saved_models.db')
        cursor = connection.cursor()

        # Check that 'Models' table is in the database
        cursor.execute("SELECT name FROM sqlite_master WHERE type='table'")
        self.assertIn(member="Models", container=cursor.fetchall()[0])

        # Check that 'Models' table has the correct attributes
        expected_table_info = [(0, 'Model_ID', 'INTEGER', 0, None, 1),
                               (1, 'Dataset', 'TEXT', 1, None, 0),
                               (2, 'File_Location', 'TEXT', 1, None, 0),
                               (3, 'Hidden_Layers_Shape', 'TEXT', 1, None, 0),
                               (4, 'Learning_Rate', 'FLOAT', 1, None, 0),
                               (5, 'Name', 'TEXT', 1, None, 0),
                               (6, 'Train_Dataset_Size', 'INTEGER', 1, None, 0),
                               (7, 'Use_ReLu', 'INTEGER', 1, None, 0)]
        cursor.execute("PRAGMA table_info(Models)")
        table_info = cursor.fetchall()
        for expected_attribute, attribute in zip (expected_table_info,
                                                  table_info):
            for expected_info, info in zip(expected_attribute, attribute):
```

```python
                self.assertEqual(first=expected_info, second=info)

    def test_not_null_constraint(self) -> None:
        """Test that the NOT NULL constraint is setup."""
        connection = sqlite3.connect(database='school_project/saved_models.db')
        cursor = connection.cursor()

        # Try to insert record with the last attribute missing
        test_data = ("Test_Dataset",
                     f"school_project/saved-models/{uuid.uuid4().hex}.npz",
                     "100, 100",
                     0.1,
                     "Test_Name",
                     100)
        sql = """
        INSERT INTO Models
        (Dataset, File_Location, Hidden_Layers_Shape, Learning_Rate, Name, Train_Dataset_Size)
        VALUES (?, ?, ?, ?, ?, ?)
        """
        self.assertRaises(sqlite3.IntegrityError, cursor.execute, sql, test_data)

    def test_unique_constraint(self) -> None:
        """Test that the UNIQUE (Dataset, Name) constraint is setup."""
        connection = sqlite3.connect(database='school_project/saved_models.db')
        cursor = connection.cursor()

        # Save test data
        test_data = ("Test_Dataset",
                     f"school_project/saved-models/{uuid.uuid4().hex}.npz",
                     "100, 100",
                     0.1,
                     "Test_Name",
                     100,
                     True)
        sql = """
        INSERT INTO Models
        (Dataset, File_Location, Hidden_Layers_Shape, Learning_Rate, Name, Train_Dataset_Size, Use_ReLu)
        VALUES (?, ?, ?, ?, ?, ?, ?)
        """
        cursor.execute(sql, test_data)
        connection.commit()

        # Try to save the same data again
        test_data = ("Test_Dataset",
                     f"school_project/saved-models/{uuid.uuid4().hex}.npz",
                     "100, 100",
                     0.1,
                     "Test_Name",
                     100,
                     True)
        sql = """
        INSERT INTO Models
        (Dataset, File_Location, Hidden_Layers_Shape, Learning_Rate, Name, Train_Dataset_Size, Use_ReLu)
        VALUES (?, ?, ?, ?, ?, ?, ?)
```

```python
        """
        self.assertRaises(sqlite3.IntegrityError, cursor.execute, sql, test_data)

        # Remove test data from database
        sql = """
        DELETE FROM Models WHERE Dataset=? AND Name=?
        """
        parameters = (test_data[0], test_data[4])
        cursor.execute(sql, parameters)
        connection.commit()

    def test_save_load_consistency(self) -> None:
        """Test that data is not changed between saving and loading."""
        connection = sqlite3.connect(database='school_project/saved_models.db')
        cursor = connection.cursor()

        test_data = ("Test_Dataset",
                     f"school_project/saved-models/{uuid.uuid4().hex}.npz",
                     "100, 100",
                     0.1,
                     "Test_Name",
                     100,
                     True)

        # Save test data
        sql = """
        INSERT INTO Models
        (Dataset, File_Location, Hidden_Layers_Shape, Learning_Rate, Name, Train_Dataset_Size, Use_ReLu)
        VALUES (?, ?, ?, ?, ?, ?, ?)
        """
        cursor.execute(sql, test_data)
        connection.commit()

        # Load test data
        sql = """
        SELECT * FROM Models WHERE Dataset=? AND Name=?
        """
        cursor.execute(sql, (test_data[0], test_data[4]))
        loaded_data = cursor.fetchall()[0]

        # Delete test data from database
        sql = """
        DELETE FROM Models WHERE Dataset=? AND Name=?
        """
        parameters = (test_data[0], test_data[4])
        cursor.execute(sql, parameters)
        connection.commit()

        # Compare test data with loaded data
        for test_value, loaded_value in zip(test_data, loaded_data[1:]):
            self.assertEqual(first=test_value, second=loaded_value)

if __name__ == "__main__":
    unittest.main()
```

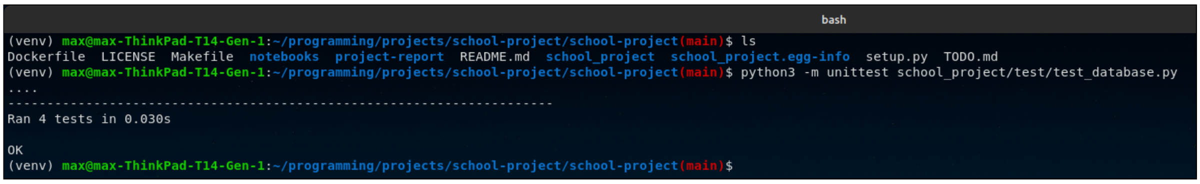


Figure 31: Unit tests for the database in a test_database.py module evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#test_databasepy`

- Unit tests for the utils subpackage of both the cpu and gpu subpackage of the models package. Similarly to the code for the cpu and gpu subpackage, it is not shown as they are identical part from the call to NumPy for the CPU and CuPy for the GPU.
  - test_model.py module:
    * test_train_dataset_size:

Description	Test the size of training dataset to be value chosen.
Data Value	hidden_layers_shape = [100, 100], train_dataset_size = 4, learning_rate = 0.1, use_relu = True
Data Type	Normal
Expected Result	The number of columns of the training input matrix should be equal to 4.
Actual Result	Expected Result
Test Status	Pass

    * test_network_shape:

Description	Test the neuron count of each layer to match the set shape of the network.
Data Value	hidden_layers_shape = [100, 100], train_dataset_size = 4, learning_rate = 0.1, use_relu = True
Data Type	Normal
Expected Result	The input neuron count of each layer should match [2, 100, 100, 1].
Actual Result	Expected Result
Test Status	Pass

    * test_learning_rates:

Description	Test learning rate of each layer to be the same.
Data Value	hidden_layers_shape = [100, 100], train_dataset_size = 4, learning_rate = 0.1, use_relu = True
Data Type	Normal
Expected Result	The learning rate of each layer should be 0.1.
Actual Result	Expected Result
Test Status	Pass

∗ test_relu_model_transfer_types:

Description	Test transfer type of each layer to match whats set.
Data Values	hidden_layers_shape = [100, 100], train_dataset_size = 4, learning_rate = 0.1, use_relu = True
Data Type	Normal
Expected Result	The transfer type of each layer should follow a pattern of ['relu', 'relu', 'sigmoid'].
Actual Result	Expected Result
Test Status	Pass

∗ test_sigmoid_model_transfer_types:

Description	Test transfer type of each layer to match whats set.
Data Values	hidden_layers_shape = [100, 100], train_dataset_size = 4, learning_rate = 0.1, use_relu = False
Data Type	Normal
Expected Result	The transfer type of each layer should follow a pattern of ['sigmoid', 'sigmoid', 'sigmoid']
Actual Result	Expected Result
Test Status	Pass

∗ test_weight_matrice_shapes:

Description	Test that each layer's weight matrix has the same number of columns as the layer's input matrix's number of rows, for the matrice multiplication.
Data Values	hidden_layers_shape = [100, 100], train_dataset_size = 4, learning_rate = 0.1, use_relu = True
Data Type	Normal
Expected Result	Each layer's weight matrix has the same number of columns as the layer's input matrix's number of rows.
Actual Result	Expected Result
Test Status	Pass

* test_bias_matrice_shapes:

Description	Test that each layer's bias matrix has the same number of rows as the result of the layer's weights and input multiplication, for element-wise addition of the biases.
Data Values	hidden_layers_shape = [100, 100], train_dataset_size = 4, learning_rate = 0.1, use_relu = True
Data Type	Normal
Expected Result	Each layer's bias matrix has the same number of rows as the result of the layer's weights and input multiplication.
Actual Result	Expected Result
Test Status	Pass

* test_layer_output_shapes:

Description	Test the shape of each layer's activation function's output.
Data Values	hidden_layers_shape = [100, 100], train_dataset_size = 4, learning_rate = 0.1, use_relu = True
Data Type	Normal
Expected Result	The shape of each layer's activation function's output should have the same number of rows as the layer's weight matrix and the same number of columns as the layer's input matrix.
Actual Result	Expected Result
Test Status	Pass

* test_save_model:

Description	Test that the weights and biases are saved correctly.
Data Values	hidden_layers_shape = [100, 100], train_dataset_size = 4, learning_rate = 0.1, use_relu = True, file_location = f"school_project/saved-models/uuid.uuid4().hex.npz"
Data Type	Normal
Expected Result	The weights and biases of each layer should not change between saving and loading.
Actual Result	Expected Result

* Evidence:

```python
"""Unit tests for model module."""

import os
import unittest
import uuid

import numpy as np

# Test XOR implementation of Model for its lesser computation time
from school_project.models.cpu.xor import XORModel

class TestModel(unittest.TestCase):
    """Unit tests for model module."""
    def __init__(self, *args, **kwargs) -> None:
        """Initialise unit tests and inputs."""
        super(TestModel, self).__init__(*args, **kwargs)

    def test_train_dataset_size(self) -> None:
        """Test the size of training dataset to be value
        ↪ chosen."""
        train_dataset_size = 4
        model = XORModel(hidden_layers_shape = [100, 100],
                         train_dataset_size = train_dataset_size,
                         learning_rate = 0.1,
                         use_relu = True)
        model.create_model_values()
        model.train(epoch_count=1)
        self.assertEqual(first=model.layers.head.input.shape[1],
                         second=train_dataset_size)

    def test_network_shape(self) -> None:
        """Test the neuron count of each layer to match the set
        ↪ shape of the
           network."""
        layers_shape = [2, 100, 100, 1]
        model = XORModel(hidden_layers_shape = [100, 100],
                         train_dataset_size = 4,
                         learning_rate = 0.1,
                         use_relu = True)
        model.create_model_values()
        model.train(epoch_count=1)
        for count, layer in enumerate(model.layers):
            self.assertEqual(first=layer.input_neuron_count,
                             second=layers_shape[count])

```

```python
    def test_learning_rates(self) -> None:
        """Test learning rate of each layer to be the same."""
        learning_rate = 0.1
        model = XORModel(hidden_layers_shape = [100, 100],
                         train_dataset_size = 4,
                         learning_rate = learning_rate,
                         use_relu = True)
        model.create_model_values()
        model.train(epoch_count=1)
        for layer in model.layers:
            self.assertEqual(first=layer.learning_rate, second=learning_rate)

    def test_relu_model_transfer_types(self) -> None:
        """Test transfer type of each layer to match whats set."""
        transfer_types = ['relu', 'relu', 'sigmoid']
        model = XORModel(hidden_layers_shape = [100, 100],
                              train_dataset_size = 4,
                              learning_rate = 0.1,
                              use_relu = True)
        model.create_model_values()
        model.train(epoch_count=1)
        for count, layer in enumerate(model.layers):
            self.assertEqual(first=layer.transfer_type,
                             second=transfer_types[count])

    def test_sigmoid_model_transfer_types(self) -> None:
        """Test transfer type of each layer to match whats set."""
        transfer_types = ['sigmoid', 'sigmoid', 'sigmoid']
        model = XORModel(hidden_layers_shape = [100, 100],
                                 train_dataset_size = 4,
                                 learning_rate = 0.1,
                                 use_relu = False)
        model.create_model_values()
        model.train(epoch_count=1)
        for count, layer in enumerate(model.layers):
            self.assertEqual(first=layer.transfer_type,
                             second=transfer_types[count])

    def test_weight_matrice_shapes(self) -> None:
        """Test that each layer's weight matrix has the same number of columns
        as the layer's input matrix's number of rows, for the matrice
        multiplication."""
        model = XORModel(hidden_layers_shape = [100, 100],
                         train_dataset_size = 4,
                         learning_rate = 0.1,
                         use_relu = True)
        model.create_model_values()
        model.train(epoch_count=1)
        for layer in model.layers:
            self.assertEqual(first=layer.weights.shape[1],
                             second=layer.input.shape[0])

    def test_bias_matrice_shapes(self) -> None:
        """Test that each layer's bias matrix has the same number of rows
        as the result of the layer's weights and input multiplication, for
        element-wise addition of the biases."""
        model = XORModel(hidden_layers_shape = [100, 100],
                         train_dataset_size = 4,
                         learning_rate = 0.1,
```

```
                         use_relu = True)
        model.create_model_values()
        model.train(epoch_count=1)
        for layer in model.layers:
            self.assertEqual(first=layer.biases.shape[0],
                             second=layer.weights.shape[0])

    def test_layer_output_shapes(self) -> None:
        """Test the shape of each layer's activation function's
        ↪  output."""
        model = XORModel(hidden_layers_shape = [100, 100],
                         train_dataset_size = 4,
                         learning_rate = 0.1,
                         use_relu = True)
        model.create_model_values()
        model.train(epoch_count=1)
        for layer in model.layers:
            self.assertEqual(
                         first=(layer.weights.shape[0],
                         ↪  layer.input.shape[1]),
                         second=layer.output.shape
                         )

    def test_save_model(self) -> None:
        """Test that the weights and biases are saved
        ↪  correctly."""
        initial_model = XORModel(hidden_layers_shape = [100,
        ↪  100],
                                 train_dataset_size = 4,
                                 learning_rate = 0.1,
                                 use_relu = True)
        initial_model.create_model_values()
        initial_model.train(epoch_count=1)

        # Save model values
        file_location =
        ↪  f"school_project/saved-models/{uuid.uuid4().hex}.npz"

        ↪  initial_model.save_model_values(file_location=file_location)

        # Create model from the saved values
        loaded_model = XORModel(hidden_layers_shape = [100, 100],
                                train_dataset_size = 4,
                                learning_rate = 0.1,
                                use_relu = True)

        ↪  loaded_model.load_model_values(file_location=file_location)

        # Remove the saved model values
        os.remove(path=file_location)

        # Compare initial and loaded model values
        for layer1, layer2 in zip(initial_model.layers,
        ↪  loaded_model.layers):
            self.assertTrue(np.array_equal(a1=layer1.weights,
                                           a2=layer2.weights))
            self.assertTrue(np.array_equal(a1=layer1.biases,
                                           a2=layer2.biases))

if __name__ == '__main__':
    unittest.main()
```

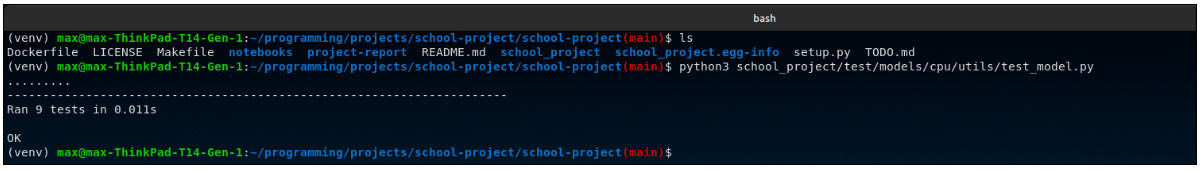


Figure 32: Unit tests for model module evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#test_modelpy`

- test_tools.py module:
  * test_relu:

Description	Test ReLu output range to be greater than or equal to zero.
Data Values	[-100, 0, 100]
Data Type	Boundary
Expected Result	The output of the ReLu transfer function should be greater than or equal to zero.
Actual Result	Expected Result
Test Status	Pass

  * test_sigmoid:

Description	Test sigmoid output range to be within 0-1.
Data Values	[-100, 0, 100]
Data Type	Boundary
Expected Result	The output of Sigmoid transfer function should be between zero and one.
Actual Result	Expected Result
Test Status	Pass

  * Evidence:

```python
"""Unit tests for tools module."""

import unittest

from school_project.models.cpu.utils import tools

class TestTools(unittest.TestCase):
    """Unit tests for the tools module."""
    def __init__(self, *args, **kwargs) -> None:
        """Initialise unit tests."""
        super(TestTools, self).__init__(*args, **kwargs)

    def test_relu(self) -> None:
        """Test ReLu output range to be >=0."""
        test_inputs = [-100, 0, 100]
        for test_input in test_inputs:
            output = tools.relu(z=test_input)
            self.assertGreaterEqual(a=output, b=0)
```

```python

    def test_sigmoid(self) -> None:
        """Test sigmoid output range to be within 0-1."""
        test_inputs = [-100, 0, 100]
        for test_input in test_inputs:
            output = tools.sigmoid(z=test_input)
            self.assertTrue(expr=output >= 0 and output <= 1)

if __name__ == '__main__':
    unittest.main()
```

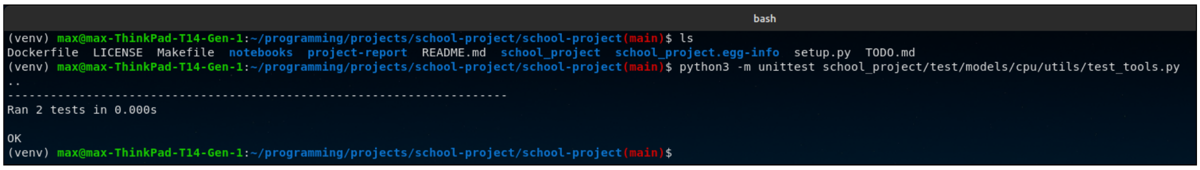


Figure 33: Unit tests for tools module evidence

Link to video evidence: `https://github.com/mcttn22/school-project/blob/main/project-report/testing-videos.md/#test_toolspy`

5.3.2 GitHub Automated Testing

With the following configuration entered in the .github/workflows/tests.yml file, the unit tests are run automatically on GitHub servers after each commit that is pushed to GitHub, and the status of the tests (either passing or failing) can be viewed on the repository's page. This automatic testing allows for a faster workflow and allows me to identify which changes (commits) cause issues within the code, allowing for easier maintenance of the project.

```yaml
name: Tests

on:
  push:
    branches: [ "main" ]
  pull_request:
    branches: [ "main" ]

permissions:
  contents: read

jobs:
  build:

    runs-on: ubuntu-latest

    steps:
    - uses: actions/checkout@v3
    - name: Set up Python 3.10
      uses: actions/setup-python@v3
      with:
        python-version: "3.10"
    - name: Install dependencies
      run: |
        python -m pip install --upgrade pip
        pip install numpy
    - name: Test
      run: |
        python -m unittest discover ./school_project/test/models/cpu
```

5.3.3 Docker

Basic Dockerfile instructions are included for its use in the README.md file, this allows the project to be quickly run and tested in Docker containers. Below shows the contents of the basic Dockerfile:

```dockerfile
FROM python:3.11

# Set a directory for the app
WORKDIR /usr/src/app

# Copy all the files to the container
COPY . .

# Install dependencies
RUN python setup.py install

# Run the project
CMD ["python", "./school_project"]
```

6 Investigation

This section outlines the code utilised to test the performance of a trained network. It then utilises this functionality to explore the effects of Hyper-Parameters on the Artificial Neural Network performance.

6.1 test_model module

The test_model module is contained within the frames package, and contains tkinter frames for investigating the behaviour of trained Artificial Neural Network models for each dataset. For each training dataset that an Artificial Neural Network is trained on, there is a corresponding test dataset with completely new images to be tested on to judge the performance of the trained model. As fewer images are needed for testing than for training, the Cat dataset only has 50 test images (compared to the 209 images for training) and the MNIST [9] dataset only has 10,000 test images (compared to the 60,000 images for training).Each frame displays the results of the testing along with a random selection of incorrect and correct predictions.

```python
"""Tkinter frames for testing a saved Artificial Neural Network model."""

import random
import threading
import tkinter as tk

from matplotlib.figure import Figure
from matplotlib.backends.backend_tkagg import FigureCanvasTkAgg
import numpy as np

class TestMNISTFrame(tk.Frame):
    """Frame for Testing MNIST page."""
    def __init__(self, root: tk.Tk, width: int,
                 height: int, bg: str,
                 use_gpu: bool, model: object) -> None:
        """Initialise test MNIST frame widgets.

        Args:
            root (tk.Tk): the widget object that contains this widget.
            width (int): the pixel width of the frame.
            height (int): the pixel height of the frame.
            bg (str): the hex value or name of the frame's background colour.
            use_gpu (bool): True or False whether the GPU should be used.
            model (object): The Model object to be tested.
        Raises:
            TypeError: if root, width or height are not of the correct type.

        """
        super().__init__(master=root, width=width, height=height, bg=bg)
        self.root = root
        self.WIDTH = width
        self.HEIGHT = height
        self.BG = bg

        # Setup test MNIST frame variables
        self.use_gpu = use_gpu

         # Setup widgets
        self.model_status_label = tk.Label(master=self,
                                           bg=self.BG,
```

```
                                           font=('Arial', 15))
        self.results_label = tk.Label(master=self,
                                      bg=self.BG,
                                      font=('Arial', 15))
        self.correct_prediction_figure = Figure()
        self.correct_prediction_canvas = FigureCanvasTkAgg(
                                        figure=self.correct_prediction_figure,
                                        master=self
                                        )
        self.incorrect_prediction_figure = Figure()
        self.incorrect_prediction_canvas = FigureCanvasTkAgg(
                                      figure=self.incorrect_prediction_figure,
                                      master=self
                                      )

        # Grid widgets
        self.model_status_label.grid(row=0, columnspan=3, pady=(30,0))
        self.results_label.grid(row=1, columnspan=3)
        self.incorrect_prediction_canvas.get_tk_widget().grid(row=2, column=0)
        self.correct_prediction_canvas.get_tk_widget().grid(row=2, column=2)

        # Start test thread
        self.model_status_label.configure(text="Testing trained model",
                                          fg='red')
        self.test_thread = threading.Thread(target=model.test)
        self.test_thread.start()

    def plot_results(self, model: object) -> None:
        """Plot results of Model test.

           Args:
               model (object): the Model object thats been tested.

        """
        self.model_status_label.configure(text="Testing Results:", fg='green')
        if not self.use_gpu:
            self.results_label.configure(
             text="Prediction Correctness: " +
             f"{round(number=100 - np.mean(np.abs(model.test_prediction.round() -
             ↪   model.test_outputs)) * 100, ndigits=1)}%\n" +
             f"Network Shape: " +
             f"{','.join(model.layers_shape)}\n"
             )

            test_inputs = np.squeeze(model.test_inputs).T
            test_outputs = np.squeeze(model.test_outputs).T.tolist()
            test_prediction = np.squeeze(model.test_prediction).T.tolist()

            # Randomly shuffle order of test_inputs, test_outputs and
            ↪   test_prediciton
            # whilst maintaining order between them
            test_data = list(zip(test_inputs,
                                 test_outputs,
                                 test_prediction))
            random.shuffle(test_data)
            test_inputs, test_outputs, test_prediction = zip(*test_data)

        elif self.use_gpu:

            import cupy as cp

            self.results_label.configure(
             text="Prediction Correctness: " +
```

```
             f"{round(number=100 - np.mean(np.abs(cp.asnumpy(model.test_prediction).round() - cp.asnumpy(model.test_outputs))) * 100, ndigits=1)}%\n" +
             f"Network Shape: " +
             f"{','.join(model.layers_shape)}\n"
             )

            test_inputs = cp.asnumpy(cp.squeeze(model.test_inputs)).T
            test_outputs = cp.asnumpy(cp.squeeze(model.test_outputs)).T.tolist()
            test_prediction = cp.squeeze(model.test_prediction).T.tolist()

            # Randomly shuffle order of test_inputs, test_outputs and test_prediciton
            # whilst maintaining order between them
            test_data = list(zip(test_inputs,
                                 test_outputs,
                                 test_prediction))
            random.shuffle(test_data)
            test_inputs, test_outputs, test_prediction = zip(*test_data)

        # Setup incorrect prediction figure
        self.incorrect_prediction_figure.suptitle("Incorrect predictions:")
        image_count = 0
        for i in range(len(test_prediction)):
            if test_prediction[i].index(max(test_prediction[i])) != test_outputs[i].index(max(test_outputs[i])):
                if image_count == 2:
                    break
                elif image_count == 0:
                    image = self.incorrect_prediction_figure.add_subplot(121)
                elif image_count == 1:
                    image = self.incorrect_prediction_figure.add_subplot(122)
                image.set_title(f"Predicted: {test_prediction[i].index(max(test_prediction[i]))}\n" +
                                f"Should have predicted: {test_outputs[i].index(max(test_outputs[i]))}")
                image.imshow(test_inputs[i].reshape((28,28)))
                image_count += 1

        # Setup correct prediction figure
        self.correct_prediction_figure.suptitle("Correct predictions:")
        image_count = 0
        for i in range(len(test_prediction)):
            if test_prediction[i].index(max(test_prediction[i])) == test_outputs[i].index(max(test_outputs[i])):
                if image_count == 2:
                    break
                elif image_count == 0:
                    image = self.correct_prediction_figure.add_subplot(121)
                elif image_count == 1:
                    image = self.correct_prediction_figure.add_subplot(122)
                image.set_title(f"Predicted: {test_prediction[i].index(max(test_prediction[i]))}")
                image.imshow(test_inputs[i].reshape((28,28)))
                image_count += 1

class TestCatRecognitionFrame(tk.Frame):
    """Frame for Testing Cat Recognition page."""
    def __init__(self, root: tk.Tk, width: int,
                 height: int, bg: str,
                 use_gpu: bool, model: object) -> None:
        """Initialise test cat recognition frame widgets.

        Args:
```

```python
            root (tk.Tk): the widget object that contains this widget.
            width (int): the pixel width of the frame.
            height (int): the pixel height of the frame.
            bg (str): the hex value or name of the frame's background colour.
            use_gpu (bool): True or False whether the GPU should be used.
            model (object): the Model object to be tested.
        Raises:
            TypeError: if root, width or height are not of the correct type.

        """
        super().__init__(master=root, width=width, height=height, bg=bg)
        self.root = root
        self.WIDTH = width
        self.HEIGHT = height
        self.BG = bg

        # Setup image recognition frame variables
        self.use_gpu = use_gpu

        # Setup widgets
        self.model_status_label = tk.Label(master=self,
                                           bg=self.BG,
                                           font=('Arial', 15))
        self.results_label = tk.Label(master=self,
                                      bg=self.BG,
                                      font=('Arial', 15))
        self.correct_prediction_figure = Figure()
        self.correct_prediction_canvas = FigureCanvasTkAgg(
                                        figure=self.correct_prediction_figure,
                                        master=self
                                        )
        self.incorrect_prediction_figure = Figure()
        self.incorrect_prediction_canvas = FigureCanvasTkAgg(
                                      figure=self.incorrect_prediction_figure,
                                      master=self
                                      )

        # Grid widgets
        self.model_status_label.grid(row=0, columnspan=3, pady=(30,0))
        self.results_label.grid(row=1, columnspan=3)
        self.incorrect_prediction_canvas.get_tk_widget().grid(row=2, column=0)
        self.correct_prediction_canvas.get_tk_widget().grid(row=2, column=2)

        # Start test thread
        self.model_status_label.configure(text="Testing trained model...",
                                          fg='red')
        self.test_thread = threading.Thread(target=model.test)
        self.test_thread.start()

    def plot_results(self, model: object) -> None:
        """Plot results of Model test

           Args:
               model (object): the Model object thats been tested.

        """
        self.model_status_label.configure(text="Testing Results:", fg='green')
        if not self.use_gpu:
            self.results_label.configure(
             text="Prediction Correctness: " +
             f"{round(number=100 - np.mean(np.abs(model.test_prediction.round() -
             ↪  model.test_outputs)) * 100, ndigits=1)}%\n" +
             f"Network Shape: " +
             f"{','.join(model.layers_shape)}\n"
```

```
             )

            # Randomly shuffle order of test_inputs, test_outputs and test_prediciton
            # whilst maintaining order between them
            test_data = list(zip(model.test_inputs.T,
                                 np.squeeze(model.test_outputs).T.tolist(),
                                 np.squeeze(model.test_prediction.round()).T.tolist()))
            random.shuffle(test_data)
            (test_inputs,
             test_outputs,
             test_prediction) = map(lambda arr: np.array(arr).T,
                                    zip(*test_data))

        elif self.use_gpu:

            import cupy as cp

            self.results_label.configure(
             text="Prediction Correctness: " +
             f"{round(number=100 - np.mean(np.abs(cp.asnumpy(model.test_prediction).round() - cp.asnumpy(model.test_outputs))) * 100, ndigits=1)}%\n" +
             f"Network Shape: " +
             f"{','.join(model.layers_shape)}\n"
             )

            # Randomly shuffle order of test_inputs, test_outputs and test_prediciton
            # whilst maintaining order between them
            test_data = list(zip(cp.asnumpy(model.test_inputs).T,
                                 cp.asnumpy(cp.squeeze(model.test_outputs)).T.tolist(),
                                 cp.asnumpy(cp.squeeze(model.test_prediction)).round().T.tolist()))
            random.shuffle(test_data)
            (test_inputs,
             test_outputs,
             test_prediction) = map(lambda arr: np.array(arr).T,
                                    zip(*test_data))

        # Setup incorrect prediction figure
        self.incorrect_prediction_figure.suptitle("Incorrect predictions:")
        image_count = 0
        for i in range(len(test_prediction)):
            if test_prediction[i] != test_outputs[i]:
                if image_count == 2:
                    break
                elif image_count == 0:
                    image = self.incorrect_prediction_figure.add_subplot(121)
                elif image_count == 1:
                    image = self.incorrect_prediction_figure.add_subplot(122)
                image.set_title(f"Predicted: {'Cat' if test_prediction[i] == 1 else 'Not a cat'}\n")
                image.imshow(test_inputs[:,i].reshape((64,64,3)))
                image_count += 1

        # Setup correct prediction figure
        self.correct_prediction_figure.suptitle("Correct predictions:")
        image_count = 0
        for i in range(len(test_prediction)):
            if test_prediction[i] == test_outputs[i]:
                if image_count == 2:
```

```
                    break
                elif image_count == 0:
                    image = self.correct_prediction_figure.add_subplot(121)
                elif image_count == 1:
                    image = self.correct_prediction_figure.add_subplot(122)
                image.set_title(f"Predicted: {'Cat' if test_prediction[i] == 1 else 'Not a cat'}\n")
                image.imshow(test_inputs[:,i].reshape((64,64,3)))
                image_count += 1

class TestXORFrame(tk.Frame):
    """Frame for Testing XOR page."""
    def __init__(self, root: tk.Tk, width: int,
                 height: int, bg: str, model: object) -> None:
        """Initialise test XOR frame widgets.

        Args:
            root (tk.Tk): the widget object that contains this widget.
            width (int): the pixel width of the frame.
            height (int): the pixel height of the frame.
            bg (str): the hex value or name of the frame's background colour.
            model (object): the Model object to be tested.
        Raises:
            TypeError: if root, width or height are not of the correct type.

        """
        super().__init__(master=root, width=width, height=height, bg=bg)
        self.root = root
        self.WIDTH = width
        self.HEIGHT = height
        self.BG = bg

        # Setup widgets
        self.model_status_label = tk.Label(master=self,
                                           bg=self.BG,
                                           font=('Arial', 15))
        self.results_label = tk.Label(master=self,
                                      bg=self.BG,
                                      font=('Arial', 20))

        # Pack widgets
        self.model_status_label.pack(pady=(30,0))

        # Start test thread
        self.model_status_label.configure(text="Testing trained model...",
                                          fg='red')
        self.test_thread = threading.Thread(target=model.test)
        self.test_thread.start()

    def plot_results(self, model: object):
        """Plot results of Model test.

           Args:
               model (object): the Model object thats been tested.

        """
        self.model_status_label.configure(text="Testing Results:", fg='green')
        results = (
             f"Prediction Accuracy: " +
             f"{round(number=model.test_prediction_accuracy, ndigits=1)}%\n" +
             f"Network Shape: " +
             f"{','.join(model.layers_shape)}\n"
             )
        for i in range(model.test_inputs.shape[1]):
```

```
            results += f"{model.test_inputs[0][i]},"
            results += f"{model.test_inputs[1][i]} = "
            if np.squeeze(model.test_prediction)[i] >= 0.5:
                results += "1\n"
            else:
                results += "0\n"
        self.results_label.configure(text=results)
        self.results_label.pack()
```

This code outputs the following results, as shown in figures 34 to 36 depending on the dataset used:

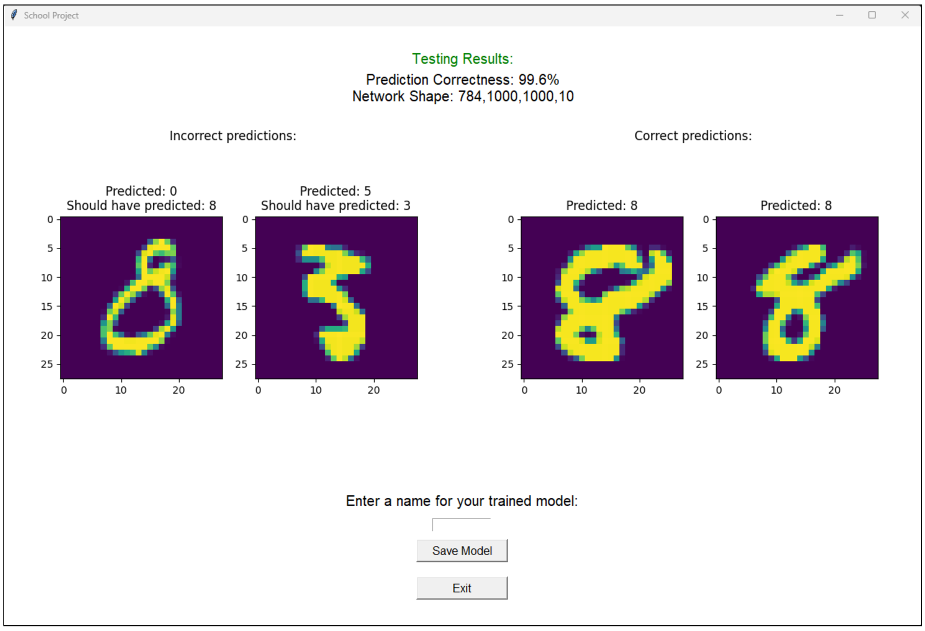


Figure 34: Model test results for MNIST database

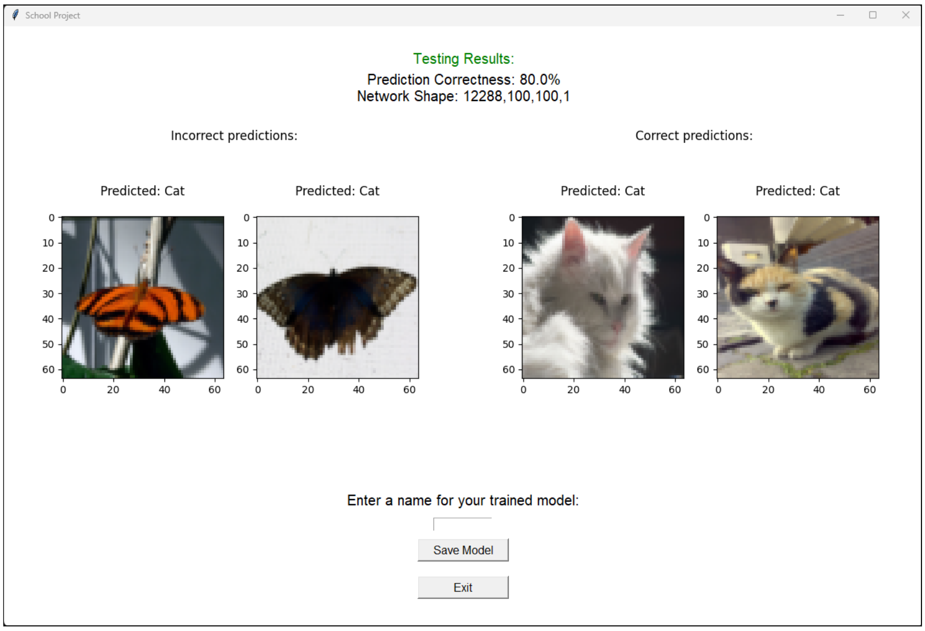


Figure 35: Model test results for Cat recognition database

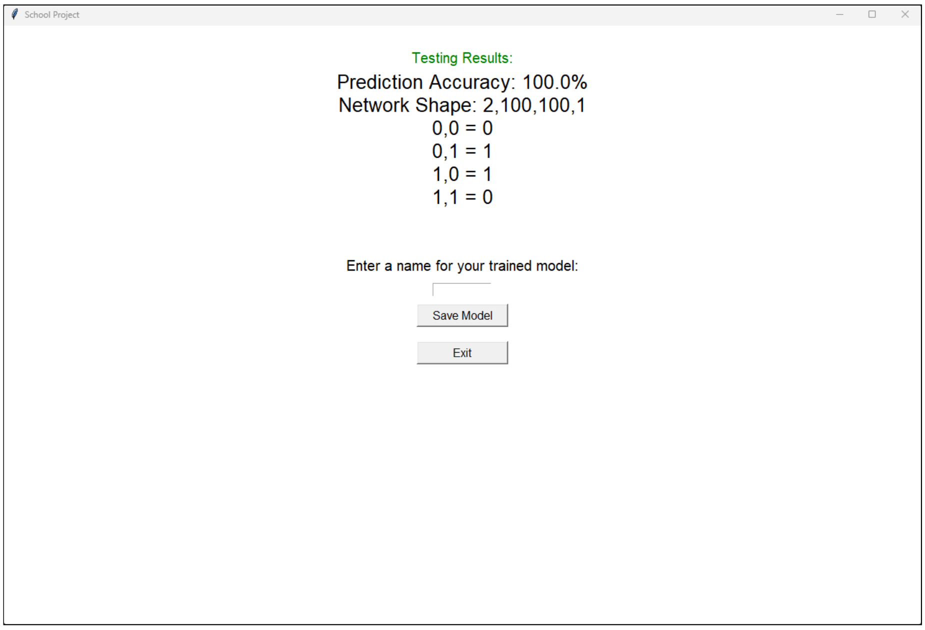


Figure 36: Model test results for XOR dataset

6.2 Exploration into the effects of Hyper-Parameters

As discussed in section 2.1 Artificial Neural Networks have a number of critical hyper-parameters which describe the shape of the network and the nature in which it learns. To explore this, I have conducted a series of experiments utilising the program to explore the fundamental impact of these hyper-parameters. In particular I explored the impact of the:

- Learning rate
- Number of Epochs undertaken during training
- Training dataset size
- Number of hidden layers
- Number of neurons in each layer
- Use of the ReLu vs Sigmoid transfer function
- Use of GPU vs CPU

For these investigations, I utilised Jupyter Notebook to run blocks of code and display the results. The output of each Jupyter Notebook is shown in the analysis below.

Analysis of the impact of the learning rate on the reduction of the loss value

The following code trains and tests models on the XOR dataset with varying learning rates, and then plots graphs of Loss Value against Epoch Count.

[17]:
```
import os

import matplotlib.pyplot as plt
import numpy as np

from school_project.models.cpu.xor import XORModel as Model

# Change to root directory of project
os.chdir(os.getcwd())

# Set width and height of figure
plt.rcParams["figure.figsize"] = [5, 10]

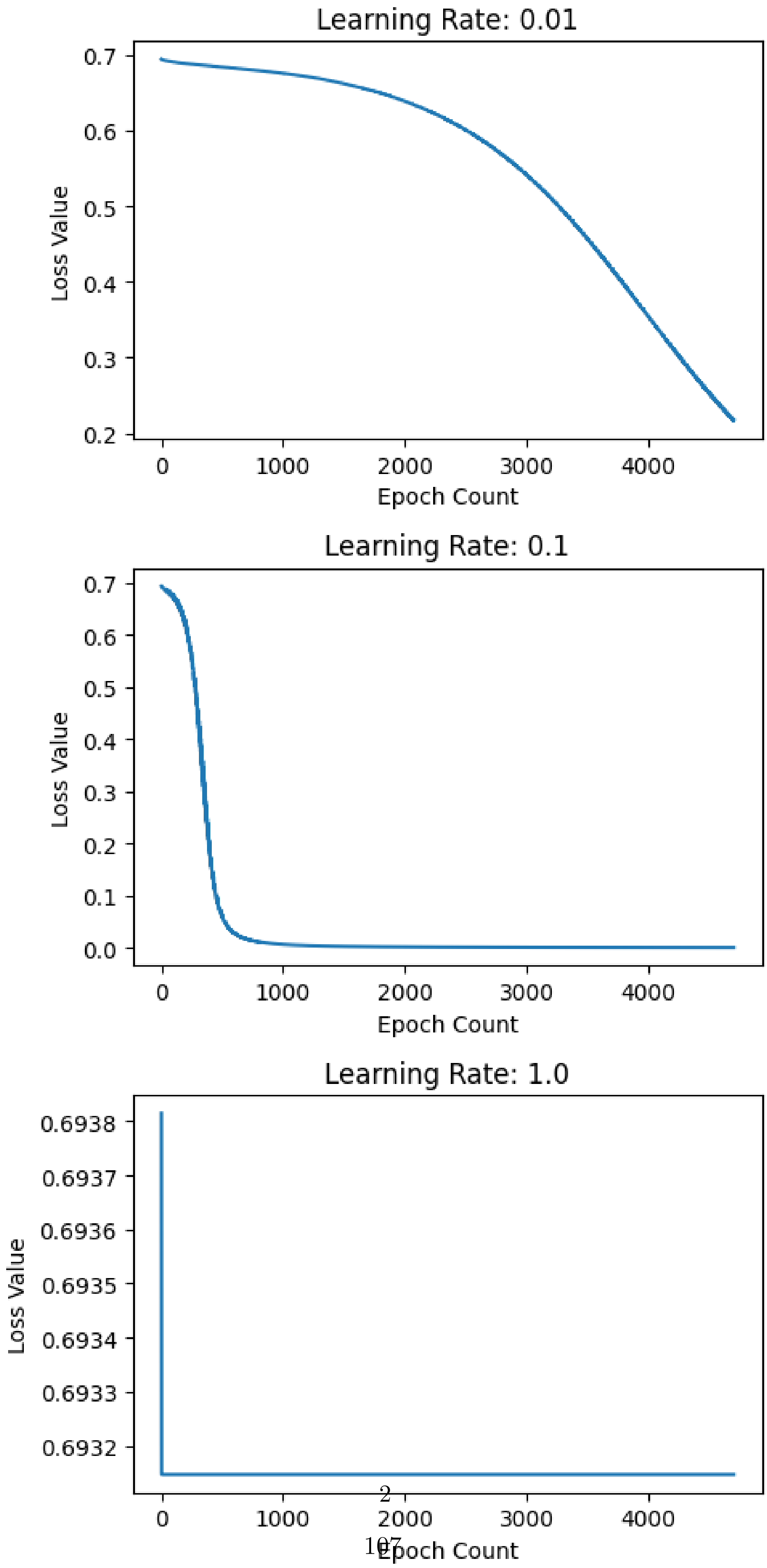
learning_rates = [0.01, 0.1, 1.0]

figure, axis = plt.subplots(nrows=len(learning_rates), ncols=1)

for count, learning_rate in enumerate(learning_rates):
    model = Model(hidden_layers_shape=[100, 100],
                  train_dataset_size=4,
                  learning_rate=learning_rate,
                  use_relu=True)
    model.create_model_values()
    model.train(epoch_count=4_700)
    model.test()

    axis[count].set_title(f"Learning Rate: {model.learning_rate}")
    axis[count].set_xlabel("Epoch Count")
    axis[count].set_ylabel("Loss Value")
    axis[count].plot(np.squeeze(model.train_losses))

plt.tight_layout()
plt.show()
```


Learning Rate: 0.01
Loss Value
Epoch Count
0.7
0.6
0.5
0.4
0.3
0.2
0
1000
2000
3000
4000
Learning Rate: 0.1
Loss Value
Epoch Count
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
1000
2000
3000
4000
Learning Rate: 1.0
Loss Value
Epoch Count
0.6938
0.6937
0.6936
0.6935
0.6934
0.6933
0.6932
0
1000
2000
3000
4000

As shown above, if the learning rate is set to too low of a value (0.01 in this case) the model will take more epochs to reduce the loss value, and may even get stuck in unwanted local minimums. If the learning rate is set to an optimal value (0.1 in this case) the model reduces the loss value efficiently and to a small enough value for predictions. On the other hand, if the learning rate is set to too high of a value (1.0 in this case) the model may learn too quickly and even 'jump over' minima, causing the loss value to stop reducing.

Analysis of the impact of training epoch count on network performance and training time taken

The following code trains models on the Cat Recognition dataset and tests the model at regular Epoch Count intervals, and then plots graphs of Test Prediction Accuracy against Epoch Count and Training Time against Epoch Count.

[6]:
```
from IPython.display import clear_output, display
import os

import matplotlib.pyplot as plt
import numpy as np

from school_project.models.gpu.cat_recognition import CatRecognitionModel as␣
 ↪Model

# Change to root directory of project
os.chdir(os.getcwd())

# Set width and height of figure
plt.rcParams["figure.figsize"] = [10, 5]

# Generate list of Epoch Counts from 1 to 5000, incremented by 500
epoch_count_interval = 500
epoch_counts = np.array(list(range(0, 5_000, epoch_count_interval)))

test_prediction_accuracies = np.array([])
training_times = np.array([])

# Create model object
model = Model(hidden_layers_shape=[100, 100],
              train_dataset_size=209,
              learning_rate=0.1,
              use_relu=True)
model.create_model_values()

for index, epoch_count in enumerate(epoch_counts):
    clear_output(wait=True)
```

```
    display(f"Progress: {round(number=index/len(epoch_counts) * 100,␣
↪ndigits=2)}%")

    model.train(epoch_count=epoch_count_interval)
    model.test()

    test_prediction_accuracies = np.append(test_prediction_accuracies,
                                           model.test_prediction_accuracy)

    # Add training times cumulatively
    if len(training_times) != 0:
        training_times = np.append(training_times,
                                   training_times[-1] + model.training_time)
    else:
        training_times = np.append(training_times,
                                   model.training_time)

clear_output(wait=True)
display("Progress: Complete")

figure, axis = plt.subplots(nrows=1, ncols=2)

axis[0].set_xlabel("Epoch Count")
axis[0].set_ylabel("Test Prediction Accuracy (%)")

# Plot regression line
axis[0].plot(epoch_counts, test_prediction_accuracies, marker='x')

# Determine gradient and y-intercept of training times regression line
m, c = np.polyfit(epoch_counts, training_times, deg=1)
print(f"Training Times Regression Line Gradient: {round(number=m, ndigits=2)}")

axis[1].set_xlabel("Epoch Count")
axis[1].set_ylabel("Training Time (s)")

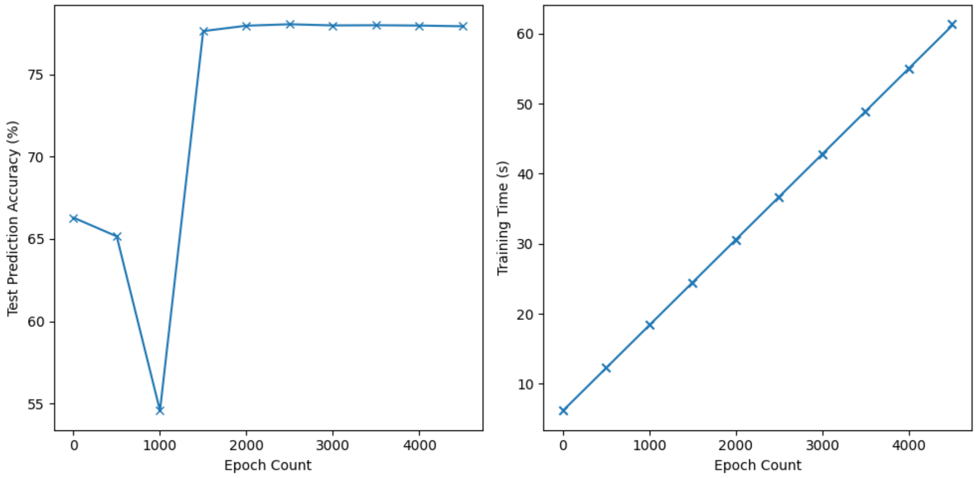
# Plot scatter graph of epoch counts and training times
axis[1].scatter(epoch_counts, training_times, marker='x')

# Plot regression line
axis[1].plot(epoch_counts, m * epoch_counts + c)

plt.tight_layout()
plt.show()
```

```
'Progress: Complete'

Training Times Regression Line Gradient: 0.01
```



As shown above, as the epoch count increases so does both the test prediction accuracy and the training time taken.

Analysis of the impact of training dataset size on network performance and training time taken

The following code trains and tests models on the Cat Recognition dataset with varying Train Dataset Sizes, and then plots graphs of Test Prediction Accuracy against Train Dataset Size and Training Time against Train Dataset Size.

[1]:
```
from IPython.display import clear_output, display
import os

import matplotlib.pyplot as plt
import numpy as np

from school_project.models.gpu.cat_recognition import CatRecognitionModel as␣
↪Model

# Change to root directory of project
os.chdir(os.getcwd())

# Set width and height of figure
plt.rcParams["figure.figsize"] = [10, 5]

# Generate list of train dataset sizes from 1 to 210, incremented by 13
train_dataset_sizes = np.array(list(range(1, 210, 13)))

test_prediction_accuracies = np.array([])
training_times = np.array([])

for index, train_dataset_size in enumerate(train_dataset_sizes):
    clear_output(wait=True)
    display(f"Progress: {round(number=index/len(train_dataset_sizes) * 100,␣
↪ndigits=2)}%")

    model = Model(hidden_layers_shape=[100, 100],
                  train_dataset_size=train_dataset_size,
                  learning_rate=0.1,
                  use_relu=True)
    model.create_model_values()
    model.train(epoch_count=2_000)
```

```
    model.test()

    test_prediction_accuracies = np.append(test_prediction_accuracies,
                                           model.test_prediction_accuracy)
    training_times = np.append(training_times,
                               model.training_time)

clear_output(wait=True)
display("Progress: Complete")

figure, axis = plt.subplots(nrows=1, ncols=2)

# Determine gradient and y-intercept of prediction accuracies regression line
m, c = np.polyfit(train_dataset_sizes, test_prediction_accuracies, deg=1)
print(f"Test Prediction Accuracies Regression Line Gradient: {round(number=m,␣
 ↪ndigits=2)}")

axis[0].set_xlabel("Train Dataset Size")
axis[0].set_ylabel("Test Prediction Accuracy (%)")

# Plot scatter graph of train dataset sizes and prediction accuracies
axis[0].scatter(train_dataset_sizes, test_prediction_accuracies, marker='x')

axis[0].plot(train_dataset_sizes, m * train_dataset_sizes + c)

# Determine gradient and y-intercept of training times regression line
m, c = np.polyfit(train_dataset_sizes, training_times, deg=1)
print(f"Training Times Regression Line Gradient: {round(number=m, ndigits=2)}")

axis[1].set_xlabel("Train Dataset Size")
axis[1].set_ylabel("Training Time (s)")

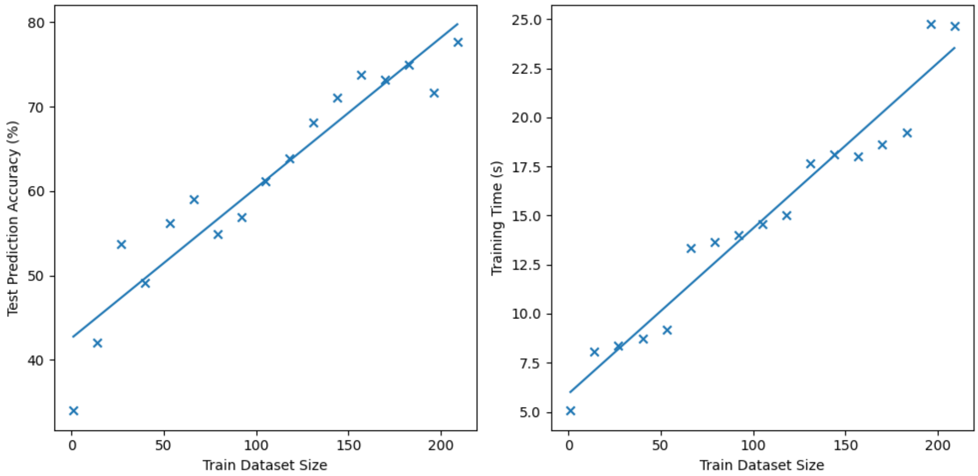
# Plot scatter graph of train dataset sizes and training times
axis[1].scatter(train_dataset_sizes, training_times, marker='x')

# Plot regression line
axis[1].plot(train_dataset_sizes, m * train_dataset_sizes + c)

plt.tight_layout()
plt.show()
```

```
'Progress: Complete'

Test Prediction Accuracies Regression Line Gradient: 0.18
Training Times Regression Line Gradient: 0.08
```



As shown above, as the train dataset size increases so does both the prediction accuracy and the training time taken. Therefore, I can predict that if I increase the size of the Cat Recognition dataset, I could improve the accuracy of the model trained on the dataset.

Analysis of the impact of layer count on network performance and training time taken

The following code trains and tests models on the Cat Recognition dataset with a varying number of layers, and then plots graphs of Test Prediction Accuracy against Layer Count and Training Time against Layer Count.

[1]:
```
from IPython.display import clear_output, display
import os

import matplotlib.pyplot as plt
import numpy as np

from school_project.models.gpu.cat_recognition import CatRecognitionModel as␣
↪Model

# Change to root directory of project
os.chdir(os.getcwd())

# Set width and height of figure
plt.rcParams["figure.figsize"] = [10, 5]

layer_counts = np.array(list(range(1, 5)))
neuron_count = 100
test_prediction_accuracies = np.array([])
training_times = np.array([])

for index, layer_count in enumerate(layer_counts):
    clear_output(wait=True)
    display(f"Progress: {round(number=index/len(layer_counts) * 100,␣
↪ndigits=2)}%")

    model = Model(
           hidden_layers_shape=[neuron_count for layer in range(layer_count)],
           train_dataset_size=209,
           learning_rate=0.1,
           use_relu=True
           )
    model.create_model_values()
```

```python
    model.train(epoch_count=3_500)
    model.test()

    test_prediction_accuracies = np.append(test_prediction_accuracies,
                                           model.test_prediction_accuracy)
    training_times = np.append(training_times,
                               model.training_time)

clear_output(wait=True)
display("Progress: Complete")

figure, axis = plt.subplots(nrows=1, ncols=2)

axis[0].set_xlabel("Layer Count")
axis[0].set_ylabel("Test Prediction Accuracy (%)")

axis[0].plot(layer_counts, test_prediction_accuracies, marker='x')

# Determine gradient and y-intercept of training times regression line
m, c = np.polyfit(layer_counts, training_times, deg=1)
print(f"Training Times Regression Line Gradient: {round(number=m, ndigits=2)}")

axis[1].set_xlabel("Layer Count")
axis[1].set_ylabel("Training Time (s)")

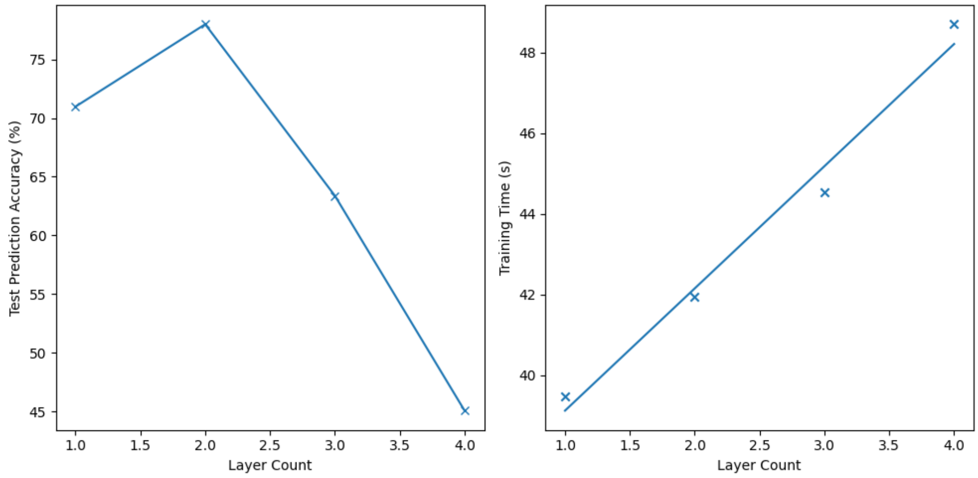
# Plot scatter graph of layer Counts and training times
axis[1].scatter(layer_counts, training_times, marker='x')

# Plot regression line
axis[1].plot(layer_counts, m * layer_counts + c)

plt.tight_layout()
plt.show()
```

```
'Progress: Complete'

Training Times Regression Line Gradient: 3.03
```



As shown above, as the layer count increases so does the training time taken and the test prediction accuracy at first. However, as the layer count continued to increase the prediction accuracy began to drop greatly (after 2 layers in this case). This is most likely due to the model overfitting and learning the training dataset too closely, causing it to fail on the new inputs of the test dataset.

Analysis of the impact of neuron count on network performance and training time taken

The following code trains and tests models on the Cat Recognition dataset with a varying number of neurons in each layer, and then plots graphs of Test Prediction Accuracy against Neuron Count and Training Time against Neuron Count.

[1]:
```
from IPython.display import clear_output, display
import os

import matplotlib.pyplot as plt
import numpy as np

from school_project.models.gpu.cat_recognition import CatRecognitionModel as␣
↪Model

# Change to root directory of project
os.chdir(os.getcwd())

# Set width and height of figure
plt.rcParams["figure.figsize"] = [10, 5]

# Generate list of neuron counts from 1 to 501, incremented by 100
neuron_counts = np.array(list(range(1, 501, 100)))

layer_count = 2
test_prediction_accuracies = np.array([])
training_times = np.array([])

for index, neuron_count in enumerate(neuron_counts):
    clear_output(wait=True)
    display(f"Progress: {round(number=index/len(neuron_counts) * 100,␣
↪ndigits=2)}%")

    model = Model(
           hidden_layers_shape=[neuron_count for layer in range(layer_count)],
           train_dataset_size=209,
           learning_rate=0.1,
           use_relu=True
```

```
            )
    model.create_model_values()
    model.train(epoch_count=3_500)
    model.test()

    test_prediction_accuracies = np.append(test_prediction_accuracies,
                                           model.test_prediction_accuracy)
    training_times = np.append(training_times,
                               model.training_time)

clear_output(wait=True)
display("Progress: Complete")

figure, axis = plt.subplots(nrows=1, ncols=2)

axis[0].set_xlabel("Neuron Count")
axis[0].set_ylabel("Test Prediction Accuracy (%)")

axis[0].plot(neuron_counts, test_prediction_accuracies, marker='x')

# Determine gradient and y-intercept of training times regression line
m, c = np.polyfit(neuron_counts, training_times, deg=1)
print(f"Training Times Regression Line Gradient: {round(number=m, ndigits=2)}")

axis[1].set_xlabel("Neuron Count")
axis[1].set_ylabel("Training Time (s)")

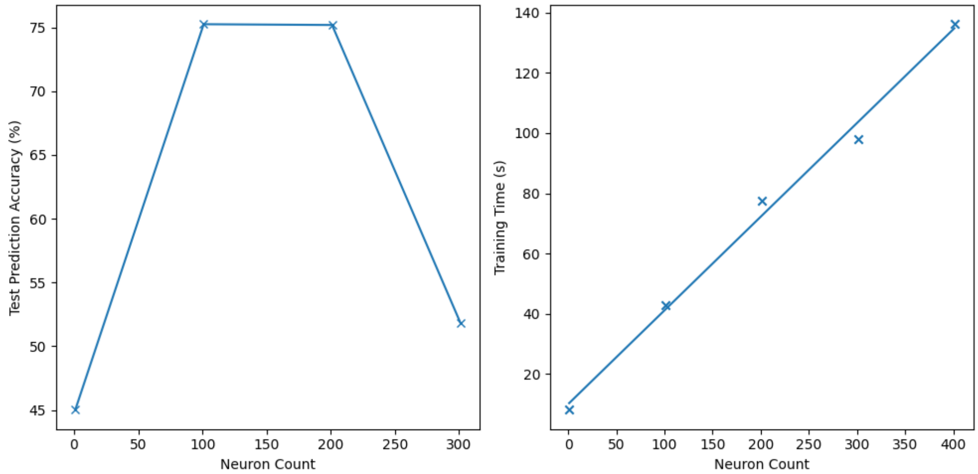
# Plot scatter graph of neuron counts and training times
axis[1].scatter(neuron_counts, training_times, marker='x')

# Plot regression line
axis[1].plot(neuron_counts, m * neuron_counts + c)

plt.tight_layout()
plt.show()
```

```
'Progress: Complete'

Training Times Regression Line Gradient: 0.31
```



As shown above, as the neuron count of each layer increases so does the training time taken and the test prediction accuracy at first. However, as the neuron count continued to increase the prediction accuracy began to drop greatly (after 200 neurons in this case). This is most likely due to the model overfitting and learning the training dataset too closely, causing it to fail on the new inputs of the test dataset.

Analysis of the impact of the transfer function on the reduction of the loss value

The following code trains and tests models on the XOR dataset using ReLu and then not using ReLu, and then plots graphs of Loss Value against Epoch Count.

[1]:
```
import os

import matplotlib.pyplot as plt
import numpy as np

from school_project.models.cpu.xor import XORModel as Model

# Change to root directory of project
os.chdir(os.getcwd())

# Set width and height of figure
plt.rcParams["figure.figsize"] = [10, 5]

figure, axis = plt.subplots(nrows=1, ncols=2)

model = Model(hidden_layers_shape=[100, 100],
              train_dataset_size=4,
              learning_rate=0.1,
              use_relu=True)
model.create_model_values()
model.train(epoch_count=4_700)
model.test()

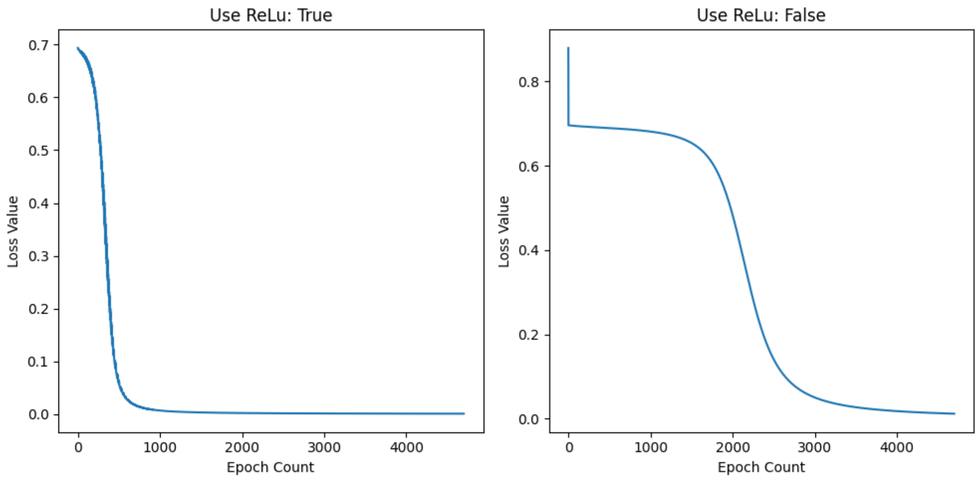
axis[0].set_title("Use ReLu: True")
axis[0].set_xlabel("Epoch Count")
axis[0].set_ylabel("Loss Value")
axis[0].plot(np.squeeze(model.train_losses))

model = Model(hidden_layers_shape=[100, 100],
              train_dataset_size=4,
              learning_rate=0.1,
              use_relu=False)
model.create_model_values()
```

```
model.train(epoch_count=4_700)
model.test()

axis[1].set_title("Use ReLu: False")
axis[1].set_xlabel("Epoch Count")
axis[1].set_ylabel("Loss Value")
axis[1].plot(np.squeeze(model.train_losses))

plt.tight_layout()
plt.show()
```



As shown above, when using the ReLu transfer function along with the Sigmoid transfer function, the loss value decreases at a much faster rate than without. The model without the ReLu transfer function does reach the same accuracy but takes far more training epochs to do so.

Analysis of the impact of using a CPU vs GPU on training time taken

The following code trains a model on the XOR dataset using the CPU and then using the GPU to train, and then outputs the training time taken.

[5]:
```python
import os

from school_project.models.cpu.cat_recognition import CatRecognitionModel as␣
↪CPUModel
from school_project.models.gpu.cat_recognition import CatRecognitionModel as␣
↪GPUModel

# Change to root directory of project
os.chdir(os.getcwd())

model = CPUModel(hidden_layers_shape=[100, 100],
                 train_dataset_size=209,
                 learning_rate=0.1,
                 use_relu=True)
model.create_model_values()
model.train(epoch_count=3_500)

print(f"CPU Training Time: {model.training_time}s")

model = GPUModel(hidden_layers_shape=[100, 100],
                 train_dataset_size=209,
                 learning_rate=0.1,
                 use_relu=True)
model.create_model_values()
model.train(epoch_count=3_500)

print(f"GPU Training Time: {model.training_time}s")
```

```
CPU Training Time: 160.33s
GPU Training Time: 43.24s
```

As shown above, the GPU is almost four times faster at training the model than the CPU, showing how beneficial it is to utilise the parallel computations of the GPU

6.3 Conclusions

The principle conclusion from this analysis is that both the shape of the network and selection of other hyper-parameters is critical to develop an optimum Artificial Neural Network.

Firstly, when considering the shape of the network, higher numbers of neurons per layer and higher layer counts do not necessarily equate to an increase in network performance. As can be seen on page 118, increasing neuron count passes through an ideal value before prediction accuracy begins to drop. In a similar manner, as can be seen on page 115, increasing the number of layers passes through an optimum value before performance decreases. Both of these behaviours are likely due to overfitting of the network to the training dataset and was an outcome predicted during the interview I undertook at the beginning of the project. In fact, it was described to me that finding the optimum network shape is part analysis and part trial and error.

Secondly, key parameters within the forward and backward propagation methods have an impact on network performance. Most importantly, is the learning rate where selecting too small a value results in an increase in required epochs and training time - the model is also at risk of converging on a local minimum resulting in a suboptimal conclusion. Alternatively, selecting too high a learning rate can lead to an oscillation around the optimal solution. The nature of the transfer function similarly has a significant effect, with the ReLu transfer function reaching a lower loss value faster - it should be noted that the Sigmoid function did also converge but required more epochs.

With regard to the size of the training dataset, a larger dataset does appear to generate a better solution. I would expect the effect of this to reduce with a significant training dataset size but the size of the training dataset for Cat recognition did not reach this level.

Lastly the use of a GPU decreased training time by approximately a factor of 4, which is the result of the GPU architecture being optimized for matrix calculations.

7 Evaluation

7.1 Third Party Feedback

I demonstrated the final version of my program to the same third party that I interviewed in the analysis, and their response is shown below:

"In my opinion, Max has definitely met the primary and secondary goals of this project. Firstly, and most importantly, he has researched and implemented, from first principles, an Artificial Neural Network that is flexible and abstracted to the point that it can tackle a range of problems. Max started the analysis for this project from a very theoretical and mathematical point of view before implementing code which has allowed him to extend its implementation to a range of datasets from the XOR problem to image analysis.

I was particularly impressed at the level of analysis he undertook into how Artificial Neural Networks work and the impact that different kinds of design decisions can have on implementation. He took on board suggestions to explore different types of transfer functions such as the ReLu function to increase the speed of training and it was nice to see comparative studies of this and other techniques. It was also great to see the ability to save and load trained models which has allowed him to train models on a desktop PC equipped with a graphics card to be utilized on a lower power laptop.

The analysis section exploring the impact on both learning rates and epoch count was very nice to see, as well as the identification of an optimal learning rate suitable for the image dataset he was working with.

In summary, it was fantastic to see a true maths-to-code example of implementing Artificial Neural Networks that didn't rely on the use of external AI libraries. I am certain he has learned a great deal regarding the fundamental properties and limitations of Artificial Neural Networks. I was also impressed by the usage of software engineering tools such as GitHub and Jupyter Notebook throughout the project."

7.2 Project Objectives Evaluation

7.2.1 Project Objectives

For the reader's convenience, I have restated the project objectives below:

Objective ID	Description
1	Learn how Artificial Neural Networks work and develop them from first principles
2	Implement the Artificial Neural Networks by creating trained models on image datasets
2.1	Allow use of Graphics Cards for faster training
2.2	Allow for the saving and loading of trained models
3	Develop a Graphical User Interface
3.1	Provide controls for hyper-parameters of models
3.2	Display and compare the results each model's predictions

7.2.2 Project Objective Evaluations

Objective ID	Evaluation	Status	3rd Party Evaluation
1	I have learnt how Artificial Neural Networks work from online resources, reports and interviewing a subject matter expert. I have proven the key mathematical principles from first principles and implemented these structures within Python code.	Fully met	Fully met
2	I have implemented trainable Artificial Neural Networks with configurable numbers of layers, number of neurons in each layer and the nature of the Transfer Functions. The Artificial Neural Networks have been trained and tested on a variety of datasets and operates at an accuracy level comparable with the resources learnt from.	Fully met	Fully met
2.1	The Artificial Neural Networks allow the use of a graphics card where applicable.	Fully met	Fully met
2.2	The trained Artificial Neural Networks' weights and biases can be saved to a data file and the features of the corresponding Artificial Neural Networks are saved to a database. These saved Artificial Neural Networks can be loaded independently.	Fully met	Fully met
3	A Graphical User Interface allowing configuration of all hyper-parameters, loading and saving of trained models and testing has been developed.	Fully met	Fully met
3.1	The Graphical User Interface allows user configuration of all utilised model hyper-parameters.	Fully met	Fully met
3.2	The model predictions can be compared in terms of both learning rate and overall accuracy.	Fully met	Fully met

7.3 Requirements Evaluation

ID	Description	Status	3rd Party Evaluation
1	Learn how Artificial Neural Networks work	Fully met	Fully met
2	Develop Artificial Neural Networks from first principles		
2.1	Provide utilities for creating Artificial Neural Networks	Fully met	Fully met
2.2	Allow for the saving and loading of trained models' weights and biases	Fully met	Fully met
2.3	Allow use of Graphics Cards for faster training	Fully met	Fully met
3	Implement the Artificial Neural Networks on image datasets		
3.1	Allow input of unique hyper-parameters	Fully met	Fully met
3.2	Allow unique datasets and train dataset size to be loaded	Fully met	Fully met
4	Use a database to store a model's features and the location of its weights and biases	Fully met	Fully met
5	Develop a Graphical User Interface		
5.1	Provide controls for hyper-parameters of models	Fully met	Fully met
5.2	Display details of models' training	Fully met	Fully met
5.3	Display the results of each model's predictions	Fully met	Fully met
5.4	Allow for the saving of trained models	Fully met	Fully met
5.5	Allow for the loading of saved trained models	Fully met	Fully met

7.4 Future Improvements

By taking into consideration my evaluation of the objectives and feedback from the third party, I believe that the following future improvements could be made for the project:

- For the analysis of the effects of hyper-parameters, repeated tests seeded with different parts of the training dataset could provide a more accurate analysis of the average behaviour of the Artificial Neural Networks.
- Performing data augmentation to expand the size of the training datasets, such as by shifting, rotating, cropping and zooming into training images or by adding noise to training images to produce more.
- Exploring Convolutional Neural Networks.
- Utilising a standardized file format for storing trained Artificial Neural Networks, so that they can be integrated with other machine learning libraries.

References

[1] Datacamp. Datacamp loss functions in machine learning explained. `https://www.datacamp.com/tutorial/loss-function-in-machine-learning`, 2023.

[2] Lenny Delligatti. *SysML distilled: A brief guide to the systems modeling language*. Addison-Wesley, 2013.

[3] Earl B Hunt. *Artificial intelligence*. Academic Press, 2014.

[4] IBM. Ibm machine learning resources. `https://developer.ibm.com/technologies/machine-learning/`, 2023.

[5] Marcopeix. Marcopeix. `https://github.com/marcopeix`, 2023.

[6] Hala Nelson. *Essential Math for AI*. "O'Reilly Media, Inc.", 2023.

[7] Researchgate. Researchgate. `https://www.researchgate.net`, 2023.

[8] W3 schools. Json - introduction. `https://www.w3schools.com/js/js_json_intro.asp`, 2023.

[9] Wikipedia. Mnist database. `https://en.wikipedia.org/wiki/MNIST_database`, 2023.