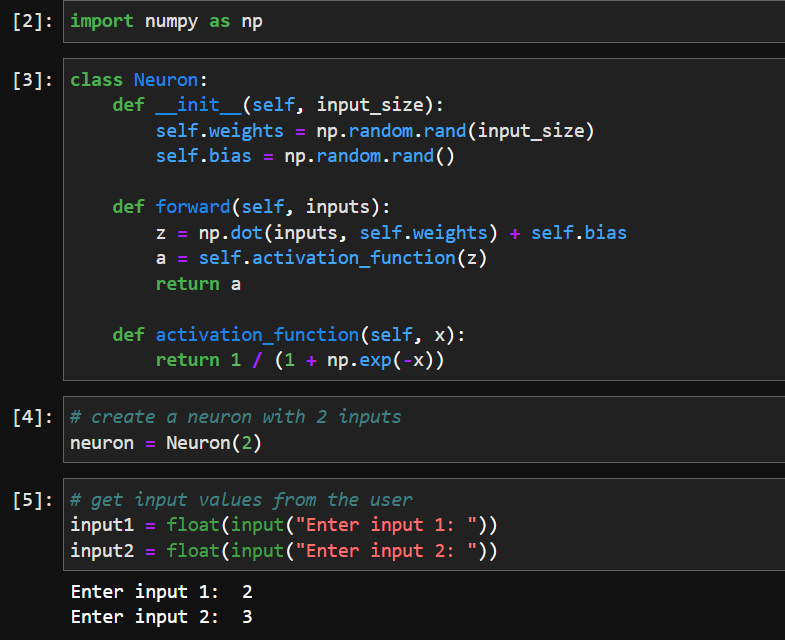
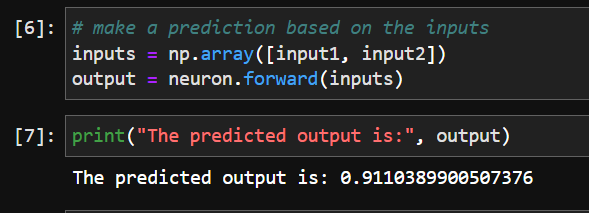
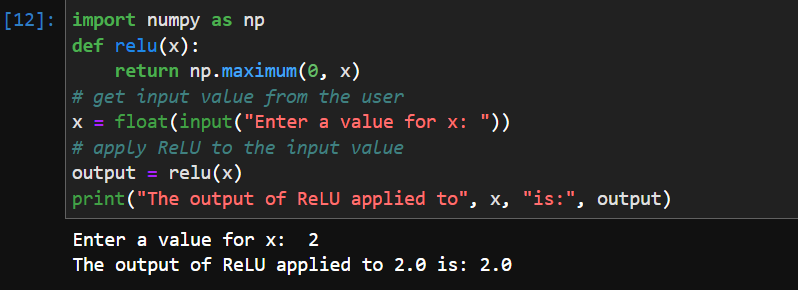
1. Write the Python code to implement a single neuron.

🡪

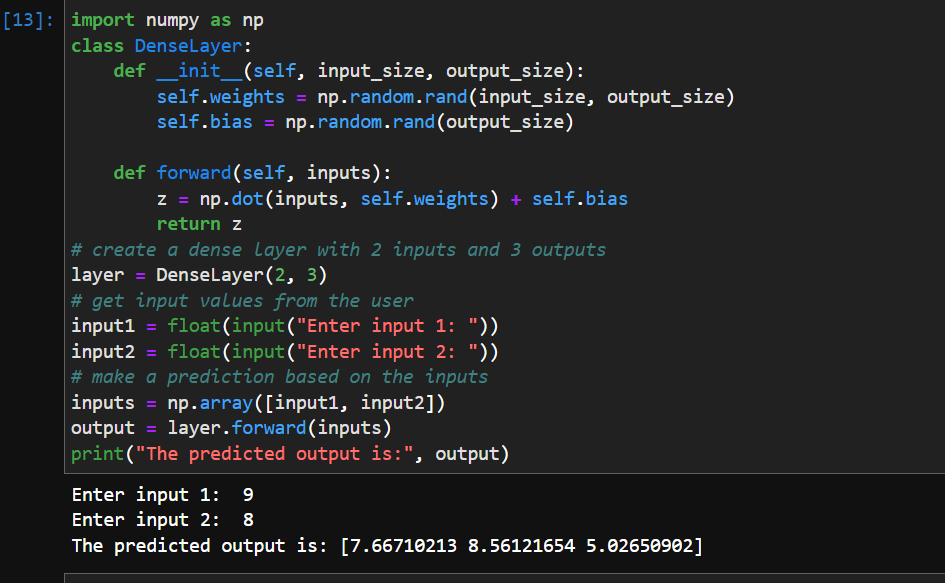




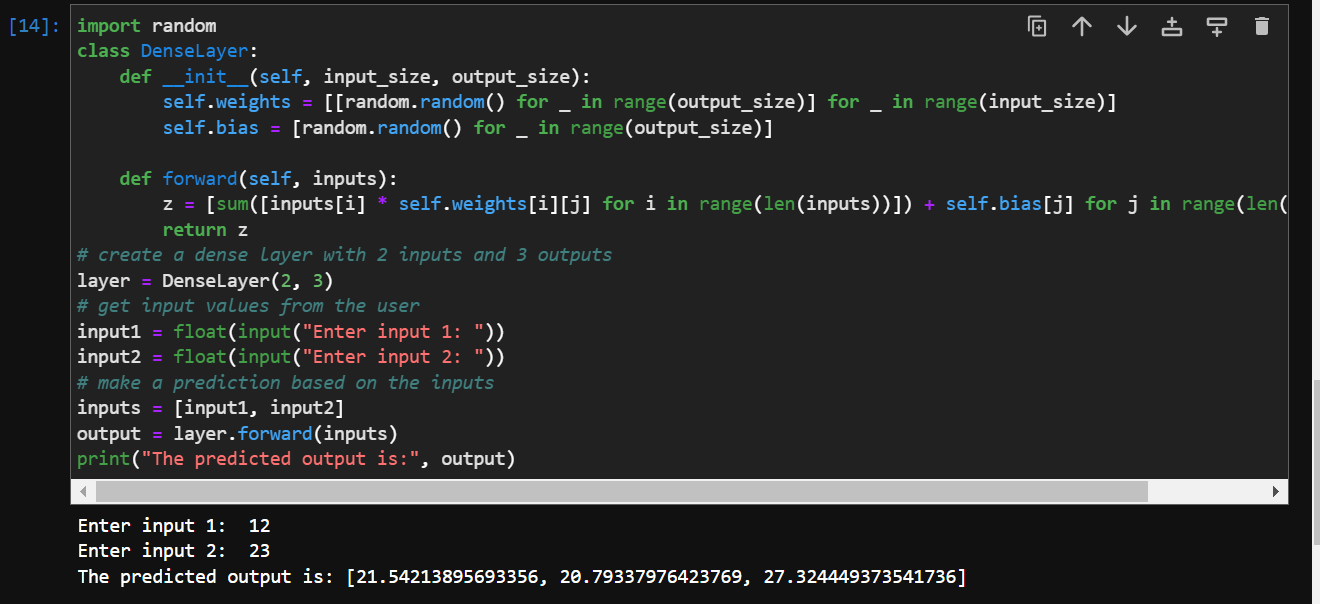
1. Write the Python code to implement ReLU.

🡪

1. Write the Python code for a dense layer in terms of matrix multiplication.

🡪

1. Write the Python code for a dense layer in plain Python (that is, with list comprehensions and functionality built into Python).

🡪

1. What is the “hidden size” of a layer?

🡪The "hidden size" of a layer refers to the number of neurons in that layer. In other words, it is the number of nodes in the layer that are not directly connected to the input or output layers.

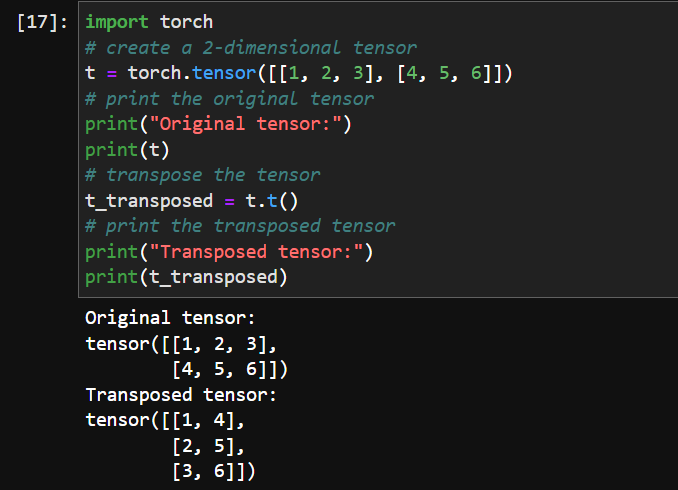
For example, in a neural network with an input layer, an output layer, and a single hidden layer, the hidden size of that layer would refer to the number of neurons in the hidden layer. The input and output layers have a fixed size determined by the input and output of the task being performed, but the number of neurons in the hidden layer can vary and is typically chosen by the user based on the complexity of the task.

The hidden size of a layer is an important hyperparameter that can affect the performance of a neural network. A larger hidden size may allow the network to learn more complex features, but may also increase the risk of overfitting. On the other hand, a smaller hidden size may simplify the network and reduce the risk of overfitting, but may not be able to learn as complex of features.

1. What does the t method do in PyTorch?

🡪In PyTorch, the t() method is used to transpose a tensor.

A tensor in PyTorch is a multi-dimensional array, similar to a NumPy array. Transposing a tensor means to swap its rows and columns, effectively rotating the tensor. For example, if we have a 2-dimensional tensor t with shape (3, 4), transposing it would result in a tensor with shape (4, 3).



1. Why is matrix multiplication written in plain Python very slow?

🡪Matrix multiplication involves a large number of arithmetic operations, and performing those operations using the basic Python operations can be slow because Python is an interpreted language, which means that it has to interpret and execute each line of code at runtime. This can lead to a significant overhead when working with large matrices.

Moreover, the built-in data structures in Python, such as lists or tuples, are not optimized for numerical operations and lack the efficient memory access patterns that are required for efficient matrix multiplication.

In contrast, specialized numerical libraries such as NumPy or TensorFlow are written in optimized C or Fortran code, which allows for much faster execution of numerical operations. They also offer various optimizations such as parallelism, vectorization, and caching, which further improve the performance of matrix multiplication.

Therefore, if you need to perform matrix multiplication, it's recommended to use a specialized library such as NumPy, which provides optimized functions for matrix operations and can execute them much faster than plain Python.

1. In matmul, why is ac==br?

🡪In matrix multiplication, two matrices A and B can be multiplied only if the number of columns of A (denoted by 'c') is equal to the number of rows of B (denoted by 'r').

When we write matmul(A,B), A is assumed to be a matrix of shape (a, c) and B is assumed to be a matrix of shape (c, b). The resulting matrix C from the matrix multiplication of A and B would have the shape (a, b).

The element at row 'i' and column 'j' in C can be calculated by taking the dot product of the i-th row of A and the j-th column of B. This requires that the length of the row in A (i.e., 'c') must be equal to the length of the column in B (also 'c'). Therefore, we have the requirement that ac == br in order to perform matrix multiplication using matmul.

1. In Jupyter Notebook, how do you measure the time taken for a single cell to execute?

🡪To measure the time taken for a single cell to execute in a Jupyter Notebook, you can use the %%timeit magic command. This command runs the cell multiple times and provides the average time taken for the cell to execute.

To use the %%timeit command, you simply need to add it at the beginning of the cell you want to time. For example:

%%timeit

# Your code here

Once you run the cell with the %%timeit command, Jupyter Notebook will execute the cell multiple times with different inputs and record the time taken for each execution. It will then calculate the average time taken and display it along with other statistics such as standard deviation and confidence intervals.

Alternatively, you can also use the %time command to measure the time taken for a single execution of a cell. This command provides more detailed information than %%timeit, but it only runs the cell once. To use %time, simply add it at the beginning of the cell you want to time, like this:

%time

# Your code here

After running the cell with %time, Jupyter Notebook will display the time taken for the cell to execute, along with other detailed information about the execution, such as CPU time and memory usage.

1. What is elementwise arithmetic?

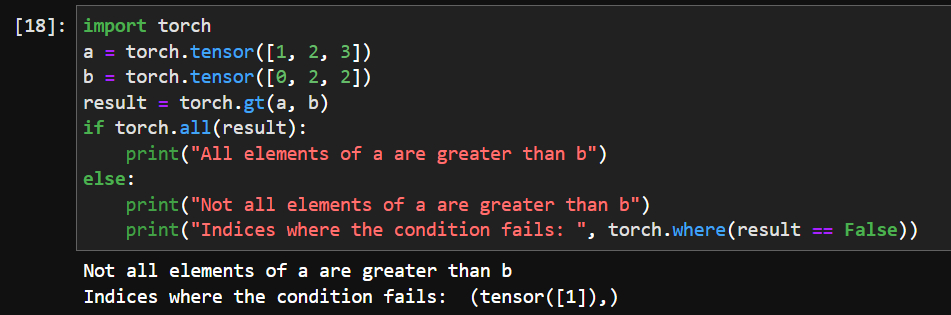
🡪Elementwise arithmetic refers to performing arithmetic operations on the corresponding elements of two or more arrays or matrices. In elementwise arithmetic, each element of one array is paired with the corresponding element of another array, and the arithmetic operation is performed on those pairs of elements.

For example, given two arrays A and B, elementwise addition would involve adding each element of A to the corresponding element of B, resulting in a new array C where each element is the sum of the corresponding elements in A and B.

Elementwise arithmetic is commonly used in numerical computations, particularly in linear algebra and machine learning. In Python, libraries such as NumPy provide efficient support for elementwise arithmetic operations on arrays and matrices.

Elementwise arithmetic is different from matrix multiplication, where each element in the resulting matrix is the sum of the products of the corresponding elements of two input matrices. In matrix multiplication, the input matrices must have compatible shapes, whereas in elementwise arithmetic, the input arrays must have the same shape.

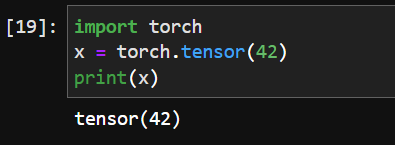
1. Write the PyTorch code to test whether every element of a is greater than the corresponding element of b.

🡪

1. What is a rank-0 tensor? How do you convert it to a plain Python data type?

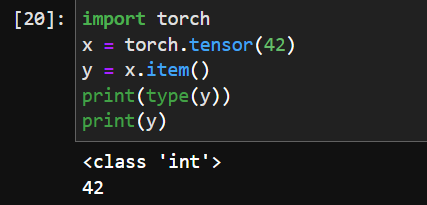
🡪A rank-0 tensor is a tensor with zero dimensions, also known as a scalar or a 0-d tensor. It represents a single numerical value, such as an integer or a floating-point number.

In PyTorch, a rank-0 tensor is created using the torch.tensor() function with a single argument, which is the scalar value:



This creates a rank-0 tensor x with the value of 42.

To convert a rank-0 tensor to a plain Python data type, you can use the .item() method. This method returns the scalar value of the tensor as a Python primitive type, such as an integer or a floating-point number:



1. How does elementwise arithmetic help us speed up matmul?

🡪Element-wise arithmetic can help speed up matrix multiplication by reducing the number of operations required to compute the result. Specifically, element-wise arithmetic can be used to implement a faster version of matrix multiplication known as “broadcasting matrix multiplication.”

In broadcasting matrix multiplication, the two matrices being multiplied are broadcasted to have the same shape. This allows us to perform element-wise multiplication between the two matrices, followed by a sum along one of the axes to compute the final result.

This approach can be much faster than traditional matrix multiplication because it takes advantage of the parallel processing capabilities of modern hardware. By performing many element-wise operations in parallel, we can significantly reduce the time it takes to compute the result of a matrix multiplication.

1. What are the broadcasting rules?

🡪Broadcasting is a powerful mechanism in NumPy and PyTorch that allows you to perform operations between tensors of different shapes. The broadcasting rules determine how tensors with different shapes are treated during arithmetic operations.

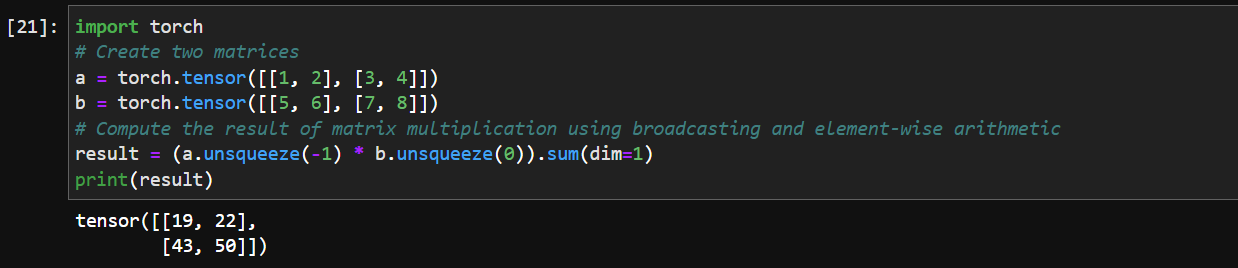
Here are the broadcasting rules:

I-f the two tensors have different numbers of dimensions, the shape of the tensor with fewer dimensions is padded with ones on its leading (left) side.

-If the shape of the two tensors does not match in any dimension, the tensor with shape equal to 1 in that dimension is stretched to match the other shape.

-If in any dimension the sizes disagree and neither is equal to 1, an error is raised.

These rules allow you to perform element-wise operations between tensors of different shapes in a flexible and intuitive manner.



1. What is expand\_as? Show an example of how it can be used to match the results of broadcasting.

🡪expand\_as is a PyTorch tensor method that allows you to expand a tensor to the size of another tensor. This can be useful when you want to perform operations between tensors of different sizes that are not broadcastable.

Here’s an example that shows how expand\_as can be used to match the results of broadcasting:

