AI HW4 Prompt

Implement a FFNN in Keras on the MNIST fashion dataset.

This implementation should use the techniques of stochastic gradient descent, batch normalization, momentum (value close to 1), he\_normal weight initialization, ELU activation function, and gradient clipping.

You should also be using sparse\_categorical\_crossentropy as your loss function and then compile your model with a learning rate of 0.01.

Your laers should be the input layer, a dense hidden layer with 300 nodes, another dense hidden layer with 100 nodes and a softmax output layer with 10 nodes.

Please also print out the model summary before training.

Write a report that answers only these specific questions:

1. What does momentum do?

2. What is a he\_normal weight initialization and why use it?

3. Why use the elu activation function? What benefit does it have?

4. What is sparse categorical cross entropy?

5. What is batch normalization accomplishing?

6 What was the accuracy and loss of your model?