1. Why do we scale our dataset after splitting it into training and validation/test sets?

1 point



The order of operations does not matter.



**During model design, we must hold out our validation/test data for all steps to preserve model integrity.**



T​he dimensionality of the data will be incorrect if we preprocess it before splitting it.

2.Question 2

True or false: PCA finds the most important dimensions in your dataset and drops the others.

1 point



False. PCA uses a neural network to find important components.



**False. PCA will find new directions in your dataset that contain the maximal variance, and then projects the data onto this new plane.**



True. The most important dimensions already exist as features and must be simply found by their variance.

3.Question 3

Is an autoencoder supervised or unsupervised?

1 point



Both. Autoencoders can be either supervised or unsupervised.



**Unsupervised. It is trained to predict its own input, hence it has no explicit labels.**



Supervised. Neural networks cannot learn without labels.

4.Question 4

Which of these snippets express a ReLU function?

1 point



*np.minimum(0, data@w+b)*



***np.maximum(0, data@w+b)***

5.Question 5

True or false: the hidden layers of an autoencoder are hyperparameters, and different architectures may lead to different performances.

1 point



**Yes. There is no set way of knowing what will work best.**



Performance is bottlenecked by latent space dimensionality, not layer sizes.

6.Question 6

What other uses do you think you an autoencoder could be useful?

1 point



Your answer cannot be more than 10000 characters.