Question 3

Consider the results in the following table, then answer the questions

Model	Train loss	Test loss	Train accuracy	Test accuracy
1	0.20	0.60	95%	80%
2	0.80	0.75	70%	68%
3	0.30	0.32	92%	90%

1. Which model has the best generalization? Justify your answer.

Best generalization implies lack of overfitting, overfitting can be seem in model 1, the perfomance drops considerably 15%, second model is not overfitting but is underfitting that can be implied by the magnitude of loss, thus the model with best generalization is the third with the least magnitude loss.

2. Which model is overfitting? Justify your answer.

As previously said and justified, the first one

3. Which model is underfitting? Justify your answer.

As previously said and justified, the second one.

4. Which strategies could you come up with to minimize overfitting?

We could apply data augmentation, by applying transformation to the input the model would learn features in distinct contexts and thus generalize better.

apply the dropout technique and randomly decrease the ammount of inputs per neurons thin each layer of an epoch

finally we could reduce the ammount of samples seen by the model.

5. Which strategies could you come up with to minimize underfitting?

improve the quality of the data so features are more easily distinguishable, add more data to compensate for a lousy dataset or improve hyperparameter tuning. Note that the last one could be applied in both overfitting and underfitting.

6. What happens if we apply the dropout technique directly to the weight matrix instead of after the activation function?