Quiz 2 MANB 2153

Name: Ala'a Al-Shargabi

Matric Number:MAN108031

Questions/Instructions.

Answer all questions in a Jupyter Notebook (colab). Push your notebook to your github

under Quiz2-Yourfirstname.

1. Using the problem and data at https://www.kaggle.com/ronitf/heart-disease-uci,

download the data (3KB)

2. What type of problem are we trying to solve here? Classification or regression?

**Since the goal is to identify refers to the presence of heart disease in the patient, so it is a classification problem.**

3. You have decided you want to create a neural network model to perform the

classification/regression for that problem.

a. Use train\_test\_split function to split the data into training 70%, test 30%

b. Design a neural network model with any architecture. Try to be unique

here.

c. Use Keras codes to create the model

d. Compile and fit the model

e. What is the accuracy after 20 epochs?

**epochs 20 is :acc: 0.4717 means 47%**

4. What is underfitting and overfitting?

**Overfitting:**

**occurs when a statistical model or machine learning algorithm captures the noise of the data. Intuitively, overfitting occurs when the model or the algorithm fits the data too well. Specifically, overfitting occurs if the model or algorithm shows low bias but high variance. Overfitting is often a result of an excessively complicated model, and it can be prevented by fitting multiple models and using validation or cross-validation to compare their predictive accuracies on test data.**

**Underfitting :occurs when a statistical model or machine learning algorithm cannot capture the underlying trend of the data. Intuitively, underfitting occurs when the model or the algorithm does not fit the data well enough. Specifically, underfitting occurs if the model or algorithm shows low variance but high bias. Underfitting is often a result of an excessively simple model.**

**#Both overfitting and underfitting lead to poor predictions on new data sets.**

5. Explain how you can reduce the overfitting of the model.

**Overfitting is often a result of an excessively complicated model, and it can be prevented by fitting multiple models and using validation or cross-validation to compare their predictive accuracies on test data. It can be also done by:**

**1. Reduce the network’s capacity by removing layers or reducing the number of elements in the hidden layers**

**2.Apply regularization, which comes down to adding a cost to the loss function for large weights**

**3.Use Dropout layers, which will randomly remove certain features by setting them to zero**