DL23 Higgs Classification: adversarial network fundamentals

Hello, this is the first survey to lecture2, on adversarial neural network fundamentals.
There are 10 questions in this survey.

Which software packages are used to construct the adversarial neural network?

Please choose **all** that apply:

- [ ] Pytorch
- [ ] Keras
- [ ] Tensorflow
- [ ] Theano

How many hidden layers does the classifier network have? In the comment, please add the number of nodes in each layer.

Please choose all that apply and provide a comment:

- [ ] 1
- [ ] 2
- [ ] 3
- [ ] 4
- [ ] 5

How many hidden layers does the adversary network have? In the comment, please add the number of nodes in each layer.

Please choose all that apply and provide a comment:

- [ ] 1
- [ ] 2
- [ ] 3
- [ ] 4
- [ ] 5
- [ ] 6

How many passes through the data did we perform during the training?

Please choose **all** that apply:

- [ ] 10
- [ ] 50
- [ ] 100
- [ ] 200
- [ ] 5000

Pass through the data is called epoch.

In the file ann_classification.py, there is a hyperparameter 'batch', set to batch=5000. How would you explain to your colleague, what this hyperparameter does?

Please write your answer here:

Which activation function is used in the classifier output layer?

Please choose **all** that apply:

- [ ] sigmoid
- [ ] relu
- [ ] tanh
- [ ] linear

Images of frequently used activation functions: https://machine-learning.paperspace.com/wiki/activation-function

Softmax is a mathematical function that converts a vector of numbers into a vector of probabilities, where the probabilities of each value are proportional to the relative scale of each value in the vector. For which Gaussian Mixture Model layer would you use the softmax activation?

Please choose **all** that apply:

- [ ] means of the Gaussians
- [ ] widths of the Gaussians
- [ ] normalization coefficients of the Gaussians

If you load 200k data events, how many events will the code ann_classification.py use for training (as opposed to testing and validation) of the network?

Please choose **all** that apply:

- [ ] 200k*0.2
- [ ] 200k*0.3
- [ ] 200k*(1-0.2)
- [ ] 200k*(1-0.3)
- [ ] 200k*(1-0.2)*(1-0.3)

The regularization parameter, balancing the losses of the classifier and adversary, is set to lambda_reg = 3 . Which action would you take, if you found that the background myy distribution, predicted by your combined model, exhibits sculpting?

Please choose **all** that apply:

- [ ] retrain with lower lambda_reg, eg lambda_reg = 1
- [ ] retrain with higher lambda_reg, eg lambda_reg = 10

The results are stored in a files called clf_standalone_results.csv and ANN_results.csv. Each of these files has three columns: "myy,y_pred,y_test". What do you expect to find in the column labelled "y_test"?

Please choose **all** that apply:

- [ ] Predicted probabilities (values between 0 and 1), that the event is signal-like.
- [ ] True labels; either of 0 or 1, that the event is signal-like.

End.

Submit your survey.
Thank you for completing this survey.

DL23 Higgs classification: advanced adversarial neural network concepts

Hello, this is the second survey to Lecture2, which contains questions on advanced adversarial neural network concepts.
There are 5 questions in this survey.

Which events are used in training the adversary network?

Please choose **all** that apply:

- [ ] Only signal (label=1) events.
- [ ] Only background (label=0) events.
- [ ] Both signal and background events.

Re- question 1 (Which events are used in training the adversary network?): how would you explain to your fellow student, why this is done?

Please write your answer here:

In layers.py , you will see a setting: "disable_eager_execution()" What is tensorflow eager execution, and why do you think we disabled it?

Please write your answer here:

In the file ann_classification.py, you will see a parameter:

```
lr = 1e-5 # Relative learning rates for classifier and adversary
```

What is the role of the learning rate ratio?

Please choose **all** that apply:

- [ ] It sets the learning rate of the adversary to be 1e-5 that of the classifier.
- [ ] It sets the learning rate of the classifier to be 1e-5 that of the adversary.

What might be the benefit of using the learning rate lr = 1e-5 setting in the previous question?

Please write your answer here:

See section III. NEURAL NETWORKS of Ref. Chase Shimmin et al, Decorrelated Jet Substructure Tagging using Adversarial Neural Networks, https://arxiv.org/abs/1703.03507

That's all, thank you!

Submit your survey.
Thank you for completing this survey.