# Constant Information Among All Models:

|  |
| --- |
| Date: 6-12-2017  512 events, 5 tracks per event, 4 hits per track, 5 noisy hits per event  128 epochs, batch size = 32, validation split = 0.25  All models tested on the same training data and training target. |

# Model 1:

|  |  |
| --- | --- |
| Sequence of 3 LSTM of size 32 each.  Optimizer is:  SGD(lr=0.01,  decay=10^-6,  momentum=0.9,  nesterov=True). | Test accuracy: 0.766249995679 |

# Model 2:

|  |  |
| --- | --- |
| From Model 1:  Changed optimizer to ‘rmsprop’. | Test accuracy: 0.837421871722 |
| Thoughts:  Changing the optimizer was a good idea. I’ll have to try other optimizers in the future. |

# Model 3:

|  |  |
| --- | --- |
| From Model 2:  Added “dropout=0.2” argument to all 3 LSTM layers. | Test accuracy: 0.817265633494 |
| Thoughts:  The dropout didn’t seem to affect much. Perhaps the data wasn’t overfitting in model 1. |

# Model 4:

|  |  |
| --- | --- |
| From Model 2:  Added another LSTM layer of size 32 in front. | Test accuracy: 0.829765621573 |
| Thoughts:  I’m surprised that accuracy did not change much with the addition of an entirely new LSTM layer. Perhaps 3 LSTM layers is sufficient for this tracking problem. |

# Model 5:

|  |  |
| --- | --- |
| From Model 2:  Doubled the number of neurons for each LSTM layer. Now each LSTM layer has 64 neurons. | Test accuracy: 0.838749993593 |
| Thoughts:  That horizontal line at the end makes me suspect that the model is overfitting.  Also, a little of topic, but these LSTM models are starting to take a long time to fit. It might be a good idea to request a GPU from Fermilab. |

# Model 6:

|  |  |
| --- | --- |
| From Model 5:  Added the argument “recurrent\_dropout=0.2” to each of the LSTM layers. | Test accuracy: 0.839374996722 |
| Thoughts:  Unfortunately, these new arguments did not change the accuracy. |

# Model 7:

|  |  |
| --- | --- |
| From Model 6:  Added the argument “dropout=0.2” to each of the LSTM layers. | Test accuracy: 0.841093748808 |
| Thoughts:  Not a great deal of change. The graph does take a bit longer (epochs) to converge. |