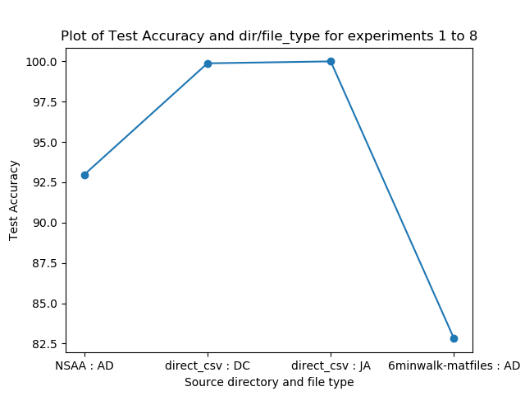
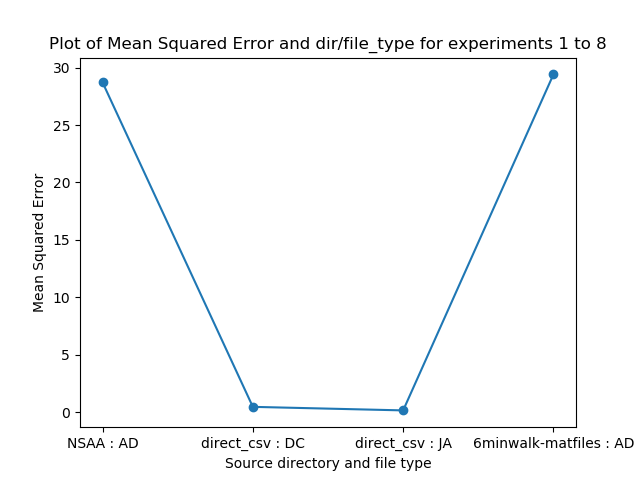
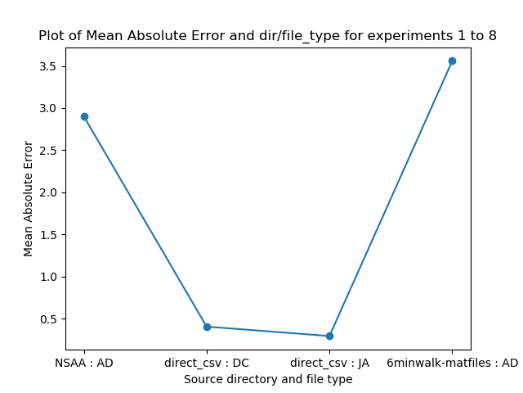
**Experiments and Results Discussion**

**Experiment Set 1: Performance of RNN on Different Source Data**

There are several things that can be noted straight away from looking at the results. The first is the difference between using raw joint angles for classification and NSAA overall score regression and using extracted statistical values from the same subject (i.e. the performance of the inner two values on x-axis vs outer two). We first compare row 1 and row 5, which are data from the same subjects that train a model to perform the same tasks with markedly different results. The only difference between the two is that, in the first case, the model is trained on raw joint angles (provided by the datacube file) for ‘6minwalk’ data, while the second uses extracted statistical values of many measurements for the same subjects. In using just the raw data values, we achieve a 99.88% accuracy (i.e. for each sequence of 60 rows of 66 joint angle values, the model can predict with 99.88% accuracy whether it came from a ‘D’ file or an ‘HC’ file); however, looking at more measurements (e.g. position, accelerometer values, etc.), performing manual feature extraction via computing of the statistical values and then reducing the dimensionality, and then training the model provides a much worse classification accuracy of 82.81%. This can also be seen when the same data sources are then used to train the RNN to perform regression for the overall NSAA score: the raw joint angle data gives a much smaller mean absolute error of 0.4037 (meaning that it predicts a score of between 0 and 34 which on average is 0.4037 away from the true value in either direction), compared with a much worse MAE of 3.56 from 6minwalk AD statistical value files.

This may seem counterintuitive at first glance, as the former is just using data direct from the provided ‘.mat’ files, while with the latter we actually process it further to ideally extract more important features from the data. However, a prominent benefit of using neural networks is that they are noted to perform better with raw data rather than manually extracted features. This most likely goes a fair way towards explaining this discrepancy in accuracies and mean errors. It should also be noted that training the RNN with raw joint angle data requires far fewer training epochs (~20) than for statistical values (~100). This will primarily be due to the larger amount of raw joint angle data that is fed through the network; in the case of training on all files within the data cube, the ‘x’ input shape is (8470, 60, 66) while for the corresponding AD files we only have (552, 10, 30) samples due to how computing statistical values dramatically reduces the raw amount of available data. This decrease in amount of available data also might help to account for the reduction in accuracy when using data sourced from AD files.

A further observation can be made about the experiments concerning the raw joint angle files (rows 1 to 4 in the table) in that they were performing much better than we were expecting them to be: by simply considering only the joint angle measurement of a subjects suit data, given 1 second’s worth of an input sequence to the RNN, it can correctly classify whether the frame comes from a healthy control subject or one with DMD to a very high accuracy of 99.88% and predict the overall NSAA score to within 0.4037 of the true value of between 0 and 34. This is extraordinarily high, much better than the ability of medical professionals and, most notably, this is only the first iteration of the experiments with raw measurement files. Hence, it is more likely that an oversight was made in the coding of the RNN or an incorrect assumption about its performance was somehow made. This investigation into the cause of this is an important next step in the project and further expanded upon later on.

**Overall Conclusions and Takeaways from Experiments**