Is there any difference in the performance when using the clean and noisy datasets?

Yes we observed a difference. The noisy data is achieved by an automated action unit (AU) recognition system and is not 100% accurate. However the label designated to each sample is always correct, the AUs detected may be wrongly allocated or missing. The clean dataset was obtained from human experts therefore it is 100% correct.

The invalid AUs found in the noisy dataset means that there are errors in the decision tree. Thus, classifying unseen data may result into a difference between the feasible choices and the clean dataset tree. This may mean a viable choice may be introduced or missed out when it wouldn't have been in the clean dataset having reduced performance compared to the clean set.

We noticed this difference in performance in our result. Making use of the data in the Emotion Recall classification table we can see there is a clear difference when using the clean data set over the noisy data set.

Emotion Recall Classification Table

88.39	70.13
75.53	81.07
67.89	52.44
88.94	78.26
47.93	35.79
70.88	63.21

We also observed a visual difference in the trees resulted after using the clean and noisy data sets. We noticed the from pruning examples. as we provided more of the noisy data, the optimal tree size was reduced. This would suggest that the noisy data had missing AUs compared to the clean data set resulting in less complex decision tree once pruned.