

The *tenth* Master of Computing and innovation Group Project weekly meeting will be held in **Room 464** at **3.00pm on Tuesday 8 May 2018**.

Meeting minutes on 8th May

Attendance: Zheng Xu, Junjie Guo

1 Apologies

None

2 Progress

We focused on improve the accuracy and response time. The result shows that the performance is better when the model is complicated. The number of nodes in the model seriously affect the response time. The response time will be long if the size of the model is large.

2.1 Findings

1. Although the accuracy is a little bit higher than C4.5 system, the false positive rate is also higher than C4.5.

Answer: Tutor will send some related documents to us. We will use one of false positive reduction techniques in our system.

2. The C4.5 system's training time is quite long (about 60 mins), maybe we could improve it. Compared to C4.5, xgboost system's training time is shorter (about 38 mins).

Answer: Maybe we could try to use multiple computers (multiple nodes) to implement our programs.

2.2 About the next steps

1. Finding the way to reduce the false positive rate.

Answer: Improving it by modify the algorithm or add the reduction techniques.

2. Testing system with other dataset.

Answer: We may use the DRAPA 98 dataset to testing our system. However, we just find the dataset's structures are different from KDDcup 99 that we need to modify the code.

Note: Next meeting to be held on 15 May 2018.