The *ninth* Master of Computing and innovation Group Project weekly meeting will be held in **Room 464** at **3.00pm on Tuesday 1 May 2018**.

Meeting minutes on 1st May

Attendance: Zheng Xu, Junjie Guo

1 Apologies

None

2 Progress

We successfully implement the Xgboost algorithm in our system base on the Hadoop platform. It is similar to the C4.5 system which could train the model and produce a XML file to record the recursive tree. The testing result is slightly under our expecting.

2.1 Findings

1. Theoretically, using xgboost algorithm could have a higher accuracy with a deeper depth as one parameter (not modify other parameters). However, we find the accuracy of tree with depth 3 is higher than tree with depth 5.

   Answer: Check up the code carefully, maybe the our algorithm is wrong. We have found that the algorithm in the relevant paper has some problems. We have modified it to get the correct value to calculate the Gain.

2. We do not develop the system with Spark any more because we use Java more frequently.

   Answer: Have successfully implement the system on Hadoop with xgboost algorithm.

2.2 About the next steps

1. Improve the xgboost system (modify the code).

2. Testing the system with different parameters.

Note: Next meeting to be held on 8 May 2018.