Big Data Technology and Application Institute

# STAGE SUMMARY AND OUTLOOK
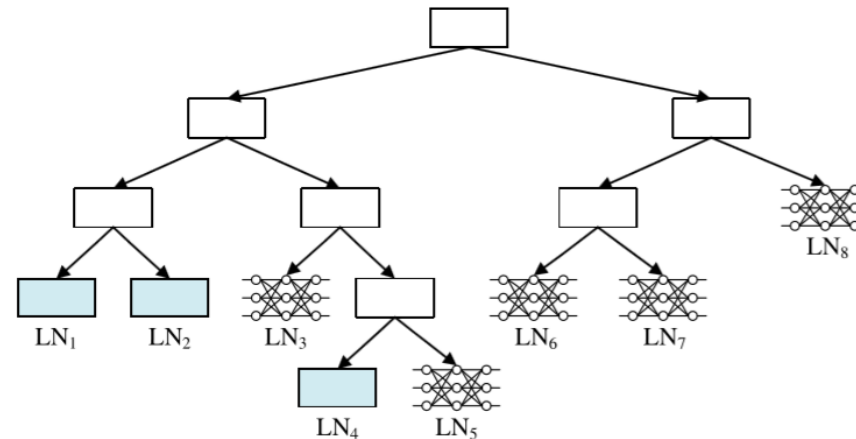# OF ELM-T PROJECT

Speaker : Yan Dasen

# Outline

▶ **Parallel performance**    ▶ **Accuracy**    ▶ **Outlook**

# Parallel performance

We have parallelized the ELMT model by using spark，hope to find some improvement of testing or training time



Experiments show that ELMT takes longer testing and training time than DT in parallel programs.
The reason is that the step of building the tree structure can be highly parallelized, but the calculation of the generalized inverse in the ELM node cannot be parallelized.

# Accuracy

In addition to the parameter, the test accuracy has a great relationship with the way of dividing data. Different way of dividing data, DT and ELMT comparison results are different.

Overall, the advantages of ELMT in terms of accuracy are not obvious in our experiment, and need to more in-depth research.

# Outlook

■ Parametric sensitivity analysis

■ In addition to tree size, training and testing time, and accuracy, find more model comparison indicators for comparison and research to explore the advantages of ELMT. Such as noise immunity，processing data with missing values

# THANKS FOR
## YOUR ATTENTION