Transcript

1. Hi, I am Sherry. In this presentation, I’m going to introduce my kidney transplant outcome prediction project.
2. Here is the training dataset of my predictive model. There are 34 patients in this dataset, and 6047 RNA-sequencings for each patient. And the outcome of each patient could be either “Acute Cellular Rejection” or “Normal/Non Specific”.
3. As the sample size of this dataset is pretty small, random forest is chosen as the classifier. Random forest is an ensemble learning method. It constructs multiple decision trees, and then it takes the majority of the prediction results from decision trees as its final result. In such a way, it could reduce model instability for small dataset.
4. Here is the performance of random forest. The average accuracy of 25 repeats of 5-fold cross validation is around 0.87.
5. This project creates a shiny app to produce prediction of kidney transplant outcome to help doctors to make decision whether the patient should have kidney transplant surgery. It takes patient’s RNA-sequencing as its input, and produce prediction result of acute rejection or stable as its output.
6. Here is a demonstration of the shiny app.
7. User can choose a patient from database (choose 32). The prediction result tab shows its prediction outcome, for this one, it is stable after transplant. The data tab shows how many RNA-sequencings of this patient, and all of the RNA-sequencings and their values. The Predictive Model: Random Forest tab shows a general introduction of this model.
8. Alternatively, user can upload their own patient’s data in a csv file (choose \_1). Prediction result for this one is Acute cellular rejection.
9. (choose incomplete) If the patient does not have all of the 6047 RNA-sequencings used in the predictive model, the app will show an insufficient data message. We can see this input only has 6046 RNA-sequencings.
10. That’s all for my presentation. Thank you for watching.