Question 4

```
Validation Accuracy is: 0.7012418854078465
  Test Accuracy is: 0.6974315551792266
```

Process:

For this ensemble process, we select IRT model developed in q2 as base model.

Steps:

1. Randomly select samples with replacement from train data (size is the same as the original data size).
2. Train the model with selected data.
3. Make predictions for validation and test data.
4. Repeat step 1-3 for 3 times and record the predictions for each round.
5. Average the predictions by formula

$$y_{bagged} = 1(\sum_{i=1}^{m} \frac{y_i}{m} > 0.5)$$

   This is the same as taking a majority vote.
6. Calculate the accuracy using the final average predicted correctness.

Do you obtain better performance using the ensemble? Why or why not?

Ans: No, we didn't obtain better performance using the ensemble. The ensembled accuracy (around 0.69) is always close to non-ensembled (around 0.69) accuracy. This might be because there is adequate data to train all models, the validation/test data matches the training distribution. Under this condition, generalizing data is relatively useless.

Notice: The following graph indicates the training process of the not ensemble model, and we also can use the validation accuracy and test accuracy of non-ensemble model to compare with the ensemble model and we can see there is no big difference between the accuracies of these two models.

Out[4]:
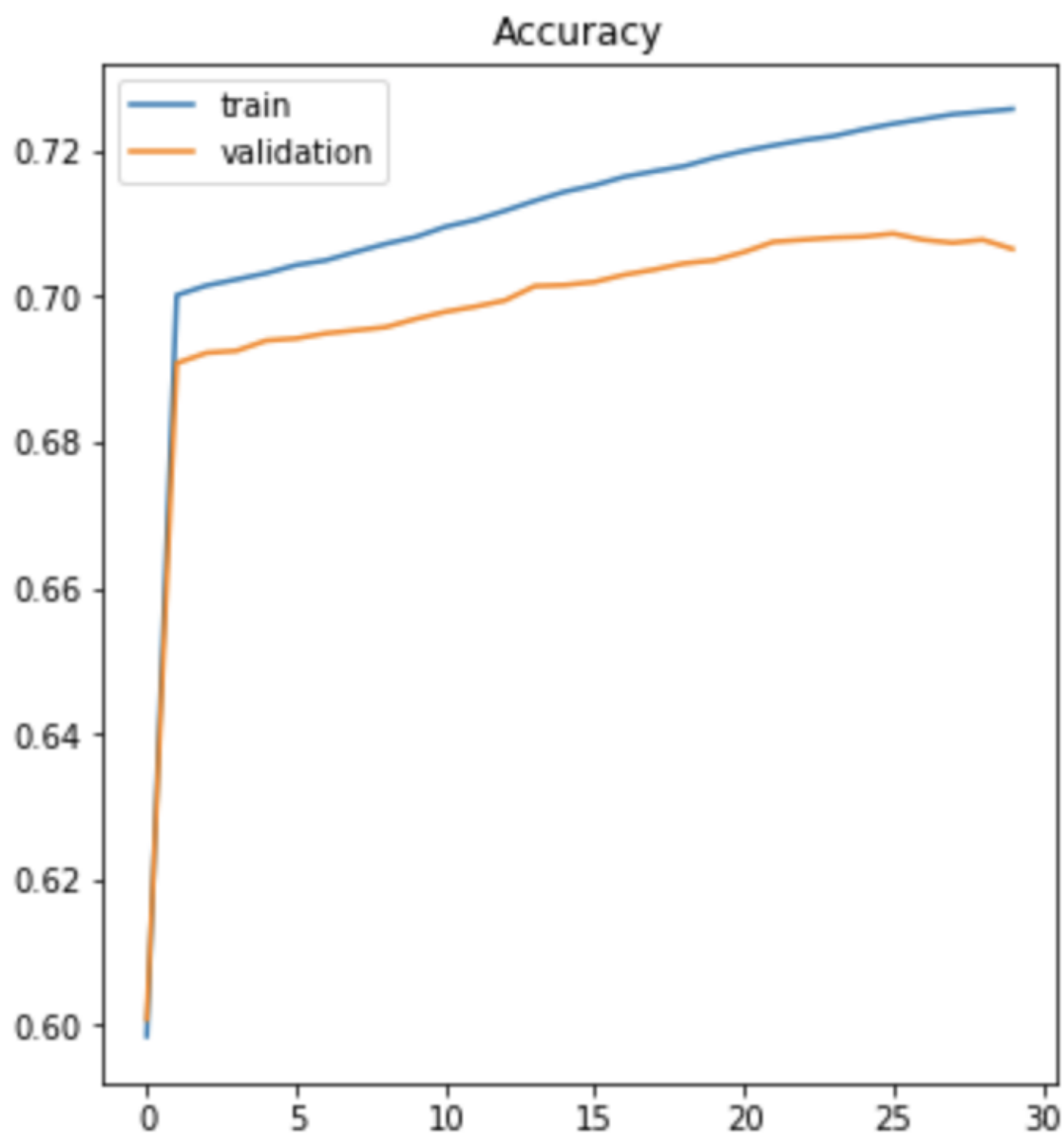
Validation Accuracy is: 0.7061812023708721
Test Accuracy is: 0.6996895286480384