**Random Forest Regression**

Starting with the initial value: 20 trees and depth of 4, with maximum features set to auto, the initial RMSE is 0.0436615344541.

After tuning the number of trees and depth, the best RMSE we could get is 0.00980717675171 with number of trees 34 and depth 11.

The importance of each data is shown in the table below:(the higher value, the more important)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Week#** | **Day of week** | **Backup start time** | **Work flow id** | **File name** | **Backup time**  **(hour)** |
| 0.00172559 | 0.3023244 | 0.08029848 | 0.04701604 | 0.07674994 | 0.49188555 |

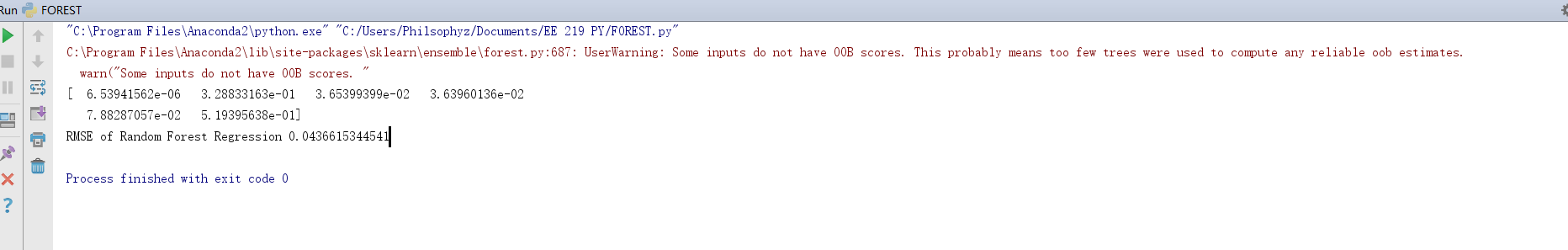


Figure initial setting with 20 trees and depth of 4

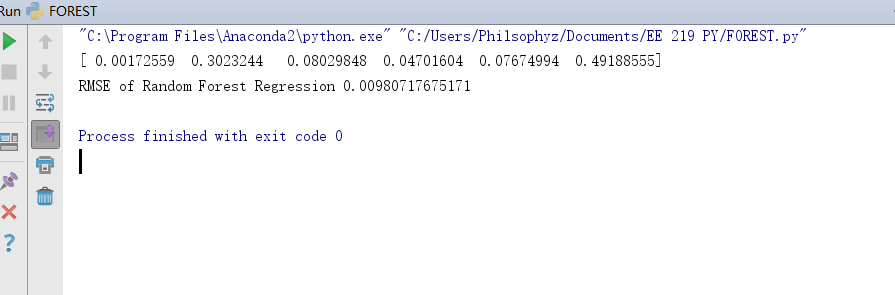


Figure best RMSE with 34 trees and depth of 11