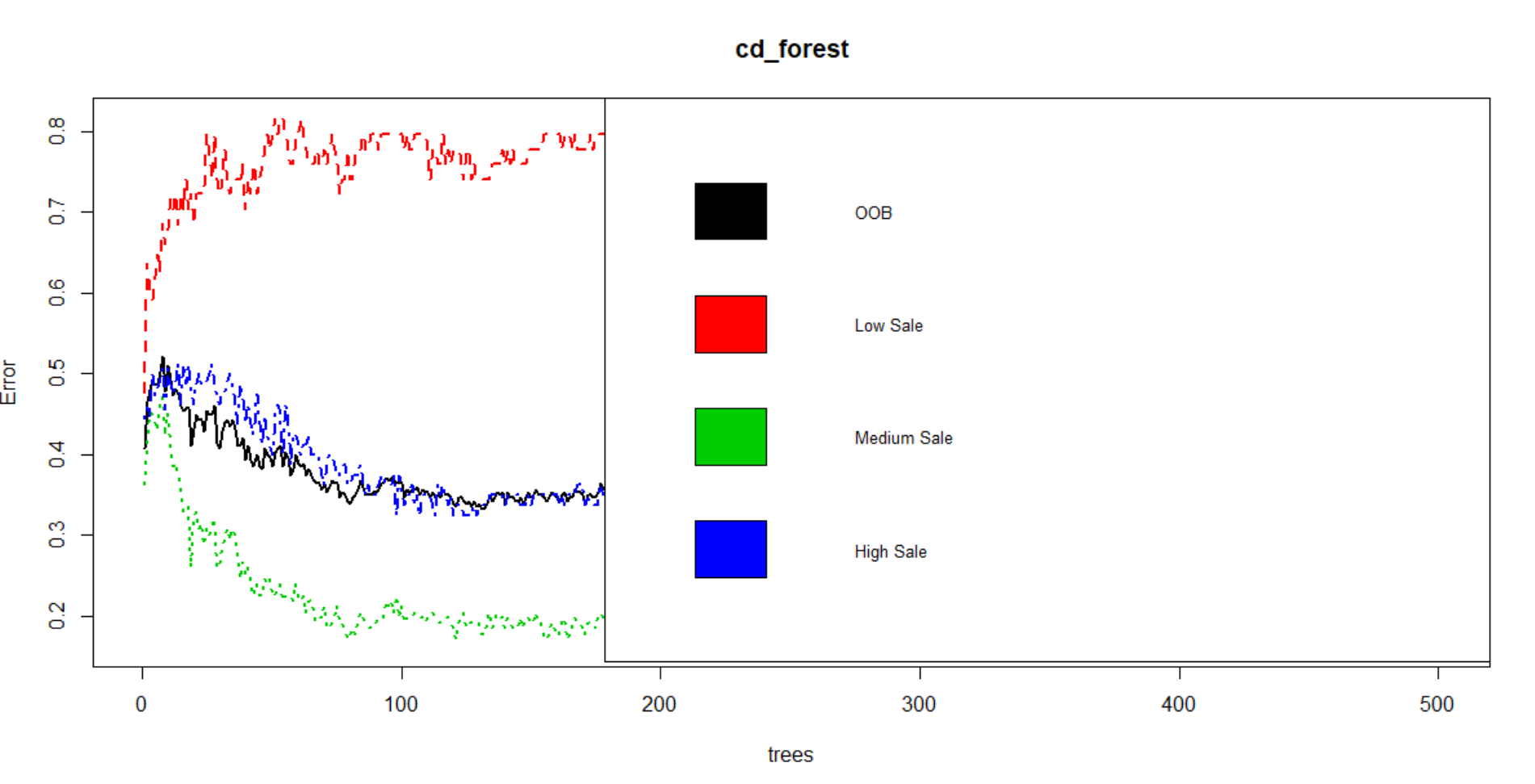
**Random Forest**

**Company Data Solution**

* The output data here is provided in numerical format, continuous data. Before moving further we need to categorize it. By understanding the business moments, we can categorize the data in following manner:

1. Sales:

* We can divide sales in 3 categories:
* Less than 5000$ = Low Sale
* 5000 – 9000$ = Medium Sale
* Above 9000$ = High Sale
* After trying different no. of trees, I settled on 500 trees and got a 66.67 % accuracy. Following it the graph for error rate with no. of trees:



**Fraud Check Solution**

* The output data here is provided in numerical format, continuous data. Before moving further we need to categorize it. By the given boundary condition we can categorize them:
* Taxable\_income:
* Less than equal to 30000 = Risky
* Above 30000 = Good
* For 500 no. of trees I got a accuracy of 80.87% and the following graph:

