

Figure gbm – Gradient Boosting Machine (gbm) Learning technique

```
> fitModel2<-trainControl(method="cv", number=5, verbose=T)
> gbmfit<-train(classe~.,data=dataTrain, method="gbm", trControl=fitModel2, v
erbose=F)
```

Loading required package: gbm

Loading required package: survival

Attaching package: 'survival'

The following object is masked from 'package:caret':

cluster

Loading required package: splines

Loading required package: parallel

Loaded gbm 2.1.1

Loading required package: plyr

```
+ Fold1: shrinkage=0.1, interaction.depth=1, n.minobsinnode=10, n.trees=150
- Fold1: shrinkage=0.1, interaction.depth=1, n.minobsinnode=10, n.trees=150
+ Fold1: shrinkage=0.1, interaction.depth=2, n.minobsinnode=10, n.trees=150
- Fold1: shrinkage=0.1, interaction.depth=2, n.minobsinnode=10, n.trees=150
+ Fold1: shrinkage=0.1, interaction.depth=3, n.minobsinnode=10, n.trees=150
- Fold1: shrinkage=0.1, interaction.depth=3, n.minobsinnode=10, n.trees=150
+ Fold2: shrinkage=0.1, interaction.depth=1, n.minobsinnode=10, n.trees=150
- Fold2: shrinkage=0.1, interaction.depth=1, n.minobsinnode=10, n.trees=150
+ Fold2: shrinkage=0.1, interaction.depth=2, n.minobsinnode=10, n.trees=150
- Fold2: shrinkage=0.1, interaction.depth=2, n.minobsinnode=10, n.trees=150
+ Fold2: shrinkage=0.1, interaction.depth=3, n.minobsinnode=10, n.trees=150
- Fold2: shrinkage=0.1, interaction.depth=3, n.minobsinnode=10, n.trees=150
+ Fold3: shrinkage=0.1, interaction.depth=1, n.minobsinnode=10, n.trees=150
- Fold3: shrinkage=0.1, interaction.depth=1, n.minobsinnode=10, n.trees=150
+ Fold3: shrinkage=0.1, interaction.depth=2, n.minobsinnode=10, n.trees=150
- Fold3: shrinkage=0.1, interaction.depth=2, n.minobsinnode=10, n.trees=150
+ Fold3: shrinkage=0.1, interaction.depth=3, n.minobsinnode=10, n.trees=150
- Fold3: shrinkage=0.1, interaction.depth=3, n.minobsinnode=10, n.trees=150
+ Fold4: shrinkage=0.1, interaction.depth=1, n.minobsinnode=10, n.trees=150
- Fold4: shrinkage=0.1, interaction.depth=1, n.minobsinnode=10, n.trees=150
+ Fold4: shrinkage=0.1, interaction.depth=2, n.minobsinnode=10, n.trees=150
- Fold4: shrinkage=0.1, interaction.depth=2, n.minobsinnode=10, n.trees=150
+ Fold4: shrinkage=0.1, interaction.depth=3, n.minobsinnode=10, n.trees=150
- Fold4: shrinkage=0.1, interaction.depth=3, n.minobsinnode=10, n.trees=150
+ Fold5: shrinkage=0.1, interaction.depth=1, n.minobsinnode=10, n.trees=150
- Fold5: shrinkage=0.1, interaction.depth=1, n.minobsinnode=10, n.trees=150
+ Fold5: shrinkage=0.1, interaction.depth=2, n.minobsinnode=10, n.trees=150
- Fold5: shrinkage=0.1, interaction.depth=2, n.minobsinnode=10, n.trees=150
+ Fold5: shrinkage=0.1, interaction.depth=3, n.minobsinnode=10, n.trees=150
- Fold5: shrinkage=0.1, interaction.depth=3, n.minobsinnode=10, n.trees=150
```

Aggregating results

Selecting tuning parameters

Fitting n.trees = 150, interaction.depth = 3, shrinkage = 0.1, n.minobsinnode = 10 on full training set

```
> gbmfit$finalModel
```

A gradient boosted model with multinomial loss function.

150 iterations were performed.

There were 51 predictors of which 42 had non-zero influence.

```
> class(gbmfit)
[1] "train"          "train.formula"
> predgbm<-predict(gbmfit, newdata=dataTest)
> confusionMatrix(predgbm, dataTest$classe)
Confusion Matrix and Statistics
```

	Reference				
Prediction	A	B	C	D	E
A	1651	52	0	0	0
B	14	1040	30	2	12
C	8	40	977	35	10
D	1	3	16	919	13
E	0	4	3	8	1047

Overall Statistics

```
Accuracy : 0.9573
95% CI : (0.9519, 0.9624)
No Information Rate : 0.2845
P-Value [Acc > NIR] : < 2.2e-16
```

```
Kappa : 0.946
McNemar's Test P-Value : NA
```

Statistics by Class:

	Class: A	Class: B	Class: C	Class: D	Class: E
Sensitivity	0.9863	0.9131	0.9522	0.9533	0.9677
Specificity	0.9877	0.9878	0.9809	0.9933	0.9969
Pos Pred Value	0.9695	0.9472	0.9131	0.9653	0.9859
Neg Pred Value	0.9945	0.9793	0.9898	0.9909	0.9927
Prevalence	0.2845	0.1935	0.1743	0.1638	0.1839
Detection Rate	0.2805	0.1767	0.1660	0.1562	0.1779
Detection Prevalence	0.2894	0.1866	0.1818	0.1618	0.1805
Balanced Accuracy	0.9870	0.9504	0.9666	0.9733	0.9823

```
>
> predtrain<-predict(gbmfit, newdata=dataTrain)
> confusionMatrix(predtrain, dataTrain$classe)
Confusion Matrix and Statistics
```

	Reference				
Prediction	A	B	C	D	E
A	3873	63	1	0	1
B	21	2547	57	6	15
C	9	41	2312	67	17
D	2	1	22	2168	19
E	1	6	4	11	2473

Overall Statistics

```
Accuracy : 0.9735
95% CI : (0.9707, 0.9761)
No Information Rate : 0.2843
P-Value [Acc > NIR] : < 2.2e-16
```

```
Kappa : 0.9665
McNemar's Test P-Value : 1.539e-11
```

Statistics by Class:

	Class: A	Class: B	Class: C	Class: D	Class: E
Sensitivity	0.9916	0.9582	0.9649	0.9627	0.9794
Specificity	0.9934	0.9911	0.9882	0.9962	0.9980
Pos Pred Value	0.9835	0.9626	0.9452	0.9801	0.9912

Neg Pred Value	0.9966	0.9900	0.9926	0.9927	0.9954
Prevalence	0.2843	0.1935	0.1744	0.1639	0.1838
Detection Rate	0.2819	0.1854	0.1683	0.1578	0.1800
Detection Prevalence	0.2867	0.1926	0.1781	0.1610	0.1816
Balanced Accuracy	0.9925	0.9747	0.9766	0.9794	0.9887

```
>
> predtrain<-predict(gbmfit, newdata= dataTrain)
> confusionMatrix(predtrain, dataTrain$classe)
Confusion Matrix and Statistics
```

	Reference				
Prediction	A	B	C	D	E
A	3873	63	1	0	1
B	21	2547	57	6	15
C	9	41	2312	67	17
D	2	1	22	2168	19
E	1	6	4	11	2473

Overall Statistics

```
Accuracy : 0.9735
95% CI : (0.9707, 0.9761)
No Information Rate : 0.2843
P-Value [Acc > NIR] : < 2.2e-16
```

```
Kappa : 0.9665
McNemar's Test P-Value : 1.539e-11
```

Statistics by Class:

	Class: A	Class: B	Class: C	Class: D	Class: E
Sensitivity	0.9916	0.9582	0.9649	0.9627	0.9794
Specificity	0.9934	0.9911	0.9882	0.9962	0.9980
Pos Pred Value	0.9835	0.9626	0.9452	0.9801	0.9912
Neg Pred Value	0.9966	0.9900	0.9926	0.9927	0.9954
Prevalence	0.2843	0.1935	0.1744	0.1639	0.1838
Detection Rate	0.2819	0.1854	0.1683	0.1578	0.1800
Detection Prevalence	0.2867	0.1926	0.1781	0.1610	0.1816
Balanced Accuracy	0.9925	0.9747	0.9766	0.9794	0.9887

```
>
```