

# PROYECTO 4

## TEORIA

CON LOS VALORES MAXDEPTH = 30 Y CP = 0

- **¿Cuál es el comportamiento de la precisión (% de aciertos) durante el entrenamiento y el test a medida que el parámetro “minsplit” aumenta? ¿Por qué?**

Según el parámetro “minsplit” aumenta, el porcentaje de aciertos baja ya que “minsplit” establece el número mínimo de observaciones que tiene que haber para que se haga el Split.

- **¿La precisión crece o decrece?**

La precisión decrece.

- **¿Hay un valor máximo o mínimo para la precisión?**

No hay máximo y mínimo ya que los valores de prueba van cambiando, pero mediante una función de la librería “caret” llamada `confusionMatrix()`, donde cruzamos la predicción con los datos de prueba para generar una matriz de confusión. Podemos obtener la precisión para cada una de las predicciones además de otros datos como Kappa.

Listening on http://127.0.0.1:6047  
Confusion Matrix and Statistics

	Reference	
Prediction	False	True
False	282	16
True	3	32

Accuracy : 0.9429  
95% CI : (0.9123, 0.9653)  
No Information Rate : 0.8559  
P-Value [Acc > NIR] : 3.897e-07

Kappa : 0.7394

Mcnemar's Test P-Value : 0.005905

Sensitivity : 0.9895  
Specificity : 0.6667  
Pos Pred Value : 0.9463  
Neg Pred Value : 0.9143  
Prevalence : 0.8559  
Detection Rate : 0.8468  
Detection Prevalence : 0.8949  
Balanced Accuracy : 0.8281

'Positive' Class : False

Confusion Matrix and Statistics

	Reference	
Prediction	False	True
False	279	18
True	6	30

Accuracy : 0.9279  
95% CI : (0.8947, 0.9533)  
No Information Rate : 0.8559  
P-Value [Acc > NIR] : 3.591e-05

Kappa : 0.674

Mcnemar's Test P-Value : 0.02474

Sensitivity : 0.9789

Specificity : 0.6250

Pos Pred Value : 0.9394

Neg Pred Value : 0.8333

Prevalence : 0.8559

Detection Rate : 0.8378

Detection Prevalence : 0.8919

Balanced Accuracy : 0.8020

'Positive' Class : False

#### Confusion Matrix and Statistics

	Reference	
Prediction	False	True
False	280	15
True	5	34

Accuracy : 0.9401

95% CI : (0.909, 0.963)

No Information Rate : 0.8533

P-Value [Acc > NIR] : 5.293e-07

Kappa : 0.7388

Mcnemar's Test P-Value : 0.04417

Sensitivity : 0.9825

Specificity : 0.6939

Pos Pred Value : 0.9492

Neg Pred Value : 0.8718

Prevalence : 0.8533

Detection Rate : 0.8383

Detection Prevalence : 0.8832

Balanced Accuracy : 0.8382

'Positive' Class : False

---

## Confusion Matrix and Statistics

Prediction	Reference	
	False	True
False	274	16
True	11	32

Accuracy : 0.9189  
 95% CI : (0.8842, 0.9459)  
 No Information Rate : 0.8559  
 P-Value [Acc > NIR] : 0.0003191

Kappa : 0.6565

Mcnemar's Test P-Value : 0.4414183

Sensitivity : 0.9614  
 Specificity : 0.6667  
 Pos Pred Value : 0.9448  
 Neg Pred Value : 0.7442  
 Prevalence : 0.8559  
 Detection Rate : 0.8228  
 Detection Prevalence : 0.8709  
 Balanced Accuracy : 0.8140

'Positive' Class : False

## Confusion Matrix and Statistics

Prediction	Reference	
	False	True
False	276	18
True	9	31

Accuracy : 0.9192  
 95% CI : (0.8846, 0.9461)  
 No Information Rate : 0.8533  
 P-Value [Acc > NIR] : 0.0001863

Kappa : 0.6505

Mcnemar's Test P-Value : 0.1236577

Sensitivity : 0.9684  
 Specificity : 0.6327  
 Pos Pred Value : 0.9388  
 Neg Pred Value : 0.7750  
 Prevalence : 0.8533  
 Detection Rate : 0.8263  
 Detection Prevalence : 0.8802  
 Balanced Accuracy : 0.8005

'Positive' Class : False

- ¿Cómo afecta el incremento del “minsplit” a la representación gráfica del árbol de decisión?

Pues según aumentamos el “minsplit”, el árbol va teniendo menos ramas.