**Hoja de Trabajo #8**

1. Cargar datos “iris.csv”
2. Graficar boxplots para cada variable independiente. Esto nos permite comparar los valores que toma cada variable agrupando de acuerdo al tipo de flor.

Hint: plot(x=datos$Species,y=datos$variable)

1. ¿Qué comportamiento observas en los datos?
2. ¿Hay algún tipo de flor que siempre tiene valores por debajo del resto de flores?
3. Ajusta un modelo para cada variable independiente por separado y grafica el modelo.
   1. ¿La gráfica es interpretable? ¿Tienen sentido las reglas que el modelo establece?
4. Ajusta un modelo incluyendo todas las variables y grafica el árbol.
   1. A pesar de incluir todas las variables, ¿el modelo incluye a todas al momento de generar reglas de decisión?
   2. ¿Cómo interpretas los resultados del modelo y las reglas de decisión?
5. Genera las predicciones que obtienes con el modelo que incluye todas las variables.
   1. Crea una nueva columna en el dataframe original llamada *Pred* con el vector de predicciones.
   2. Calcula el accuracy que obtienes con el modelo. Puede usar la función utilizada en la sesión de regresión logística para calcular la precisión.
6. Retomando lo aprendido en la sesión de la regresión logística, ajusta una regresión logística incluyendo todas las variables.
   1. Crea una nueva columna en el dataframe original llamada Pred\_logreg con el vector de predicciones.
   2. Calcula el accuracy que obtienes con la regresión logística.
   3. ¿Existe diferencia en la precisión entre un modelo y otro? ¿Qué modelo genera mejores predicciones? ¿Cuál sería tu recomendación si tuvieras que elegir entre un modelo y otro?