

## ¿Cómo implementar *Gain Information* cuando existen *missing values*?

Cuando existen valores faltantes se puede asignar a ellos el valor más común entre todos los ejemplos del conjunto de entrenamiento. Otra manera sería clasificarlos asignándoles una probabilidad a cada posible valor para luego estimarlos basándose en las frecuencias observadas de los valores del ejemplo.

Otra forma es reemplazar los valores faltantes con la media de todos los valores para el atributo si éste es numérico o la moda si es nominal.

Por ejemplo, para el dataset “Tennis” si existieran valores faltantes para el atributo Outlook. Dado que es un atributo nominal, usaríamos la moda para asignar el valor correspondiente. En este caso corresponde a “Sunny”. Por lo tanto, todos los valores faltantes se reemplazarían con Sunny.