

Ejercicio 1

Lognitud del sepalo, ancho del sepalo, longitud del petalo y ancho del petalo son atributos continuos, lo que quiere decir que los datos pueden cambiar a lo largo del tiempo y pueden o no ser numeros enteros. Las variables anteriormente mencionadas también comparten las características de: la unidad de medición es centímetros, no tienen valores perdidos y su rol es “Feature”, lo que significa que son variables que pueden ser definidas y observadas.

El atributo Clase es la variable objetivo. Comparte con los demás atributos, que no tiene valores perdidos. El tipo de esta variable es “Categorical”, quiere decir que representa un valor dentro de un conjunto de valores predefinidos, este caso los tipos de valores admitidos son: Iris Setosa, Iris Versicolour y Iris Virginica.

Lognitud del sepalo

Media: 5,83
Desvío estandar: 0,83
Varianza: 0,69

Ancho del sepalo

Media: 3,05
Desvío estandar: 0,43
Varianza: 0,19

Longitud del petalo

Media: 3,76
Desvío estandar: 1,76
Varianza: 3,11

Ancho del petalo

Media: 1,20
Desvio estandar: 0,76
Varianza: 0,58

Clase

Iris Setosa: 50
Iris Versicolour: 50
Iris Virginica: 50

Ejercicio 2

Ejercicio 3

Los datos proporcionados por RapidMiner, son parecidos a los anteriormente mencionados, las diferencias principales son: los nombres de los atributos a1, a2, a3 y a4, no detallan lo que representan, existe un atributo identificador adicional.

Lognitud del sepalo

Media: 5,84

Desvio estandar: 0,83

Ancho del sepalo

Media: 3,05

Desvio estandar: 0,43

Longitud del petalo

Media: 4,76

Desvio estandar: 1,76

Ancho del petalo

Media: 1,20

Desvio estandar: 0,76

Clase

Iris Setosa: 50

Iris Versicolour: 50

Iris Virginica: 50

La compración de ambos conjunto de datos muestra que son bastante parecidos, tal vez pueda deberse diferencia minima por la falta de precisión de los números decimales.