

Ejercicio 1

Descripción: El dataset contiene la información de largo y ancho de pétalos y sépalos de tres tipos de flores:

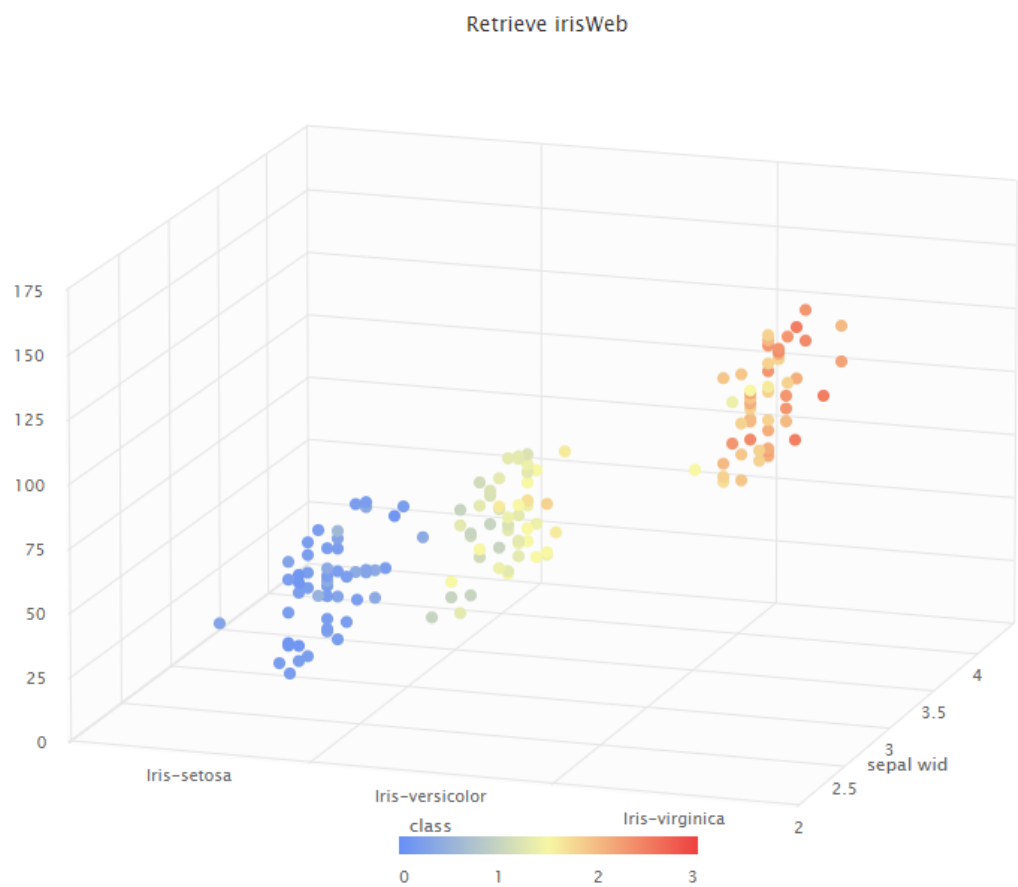
- Iris setosa
- Iris versicolor
- Iris virginica

El objetivo del dataset , es utilizar estos datos como entrenamiento para modelos de análisis en búsqueda de predecir el tipo de una flor dados los parámetros de largo y ancho de pétalos y sépalos como se mencionó anteriormente.

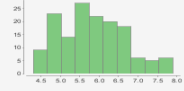
<i>sepal len</i>		<i>sepal wid</i>		<i>petal len</i>		<i>petal wid</i>	
Media	5,843 33333	Media	3,054	Media	3,758 66667	Media	1,198 66667
Error típico	0,067 61132	Error típico	0,035 40283	Error típico	0,144 06432	Error típico	0,062 31181
Mediana	5,8	Mediana	3	Mediana	4,35	Mediana	1,3
Moda	5	Moda	3	Moda	1,5	Moda	0,2
Desviación estándar	0,828 06613	Desviación estándar	0,433 59431	Desviación estándar	1,764 42042	Desviación estándar	0,763 16074
Varianza de la muestra	0,685 69351	Varianza de la muestra	0,188 00403	Varianza de la muestra	3,113 17942	Varianza de la muestra	0,582 41432
Curtosis	- 0,552 06404	Curtosis	0,290 78106	Curtosis	- 1,401 9208	Curtosis	- 1,339 75417
Coeficiente de asimetría	0,314 91096	Coeficiente de asimetría	0,334 05266	Coeficiente de asimetría	- 0,274 46425	Coeficiente de asimetría	- 0,104 99656
Rango	3,6	Rango	2,4	Rango	5,9	Rango	2,4
Mínimo	4,3	Mínimo	2	Mínimo	1	Mínimo	0,1
Máximo	7,9	Máximo	4,4	Máximo	6,9	Máximo	2,5
Suma	876,5	Suma	458,1	Suma	563,8	Suma	179,8
Cuenta	150	Cuenta	150	Cuenta	150	Cuenta	150

Datos estadísticos del dataset Iris.

Ejercicio 2 y 3 – Rapidminer



Visualización 3d de clases en función del ancho del sépalo con entonación de color del ancho del pétalo

sepal len	Real	0		Min 4.300	Max 7.900	Average 5.843	Deviation 0.828
sepal wid	Real	0		Min 2	Max 4.400	Average 3.054	Deviation 0.434
petal len	Real	0		Min 1	Max 6.900	Average 3.759	Deviation 1.764
petal wid	Real	0		Min 0.100	Max 2.500	Average 1.199	Deviation 0.763
class	Nominal	0		Least Iris-virginica (50)	Most Iris-setosa (50)	Values Iris-setosa (50), Iris-versicolor (50), Iris-virginica (50) Details...	

Estadísticas de rapidminer

Comparación:

Los valores para media, mediana y desviación estándar fueron exactamente iguales, por lo que se considera a ambas herramientas capaces de analizar este dataset.