



Universidad de Carabobo
Facultad Experimental de Ciencias y Tecnología
Departamento de Computación
Aprendizaje Maquina

Redes Neuronales

Integrantes

Victor Mendoza CI: 21.476.548

2019, Julio

Análisis del Dataset con WEKA

Se realizara el estudio del dataset haciendo uso del archivo *vowel.arff* (Fig. 1). Al visualizar el dataset con WEKA (Fig. 2) podemos observar que se poseen 990 instancias donde cada una consta de 14 atributos.

```
@relation vowel
@attribute 'Train or Test' { Train, Test}
@attribute 'Speaker Number' { Andrew, Bill, David, Mark, Jo, Kate, Penny, Rose, Mike, Nick, Rich, Tim, Sarah, Sue, Wendy}
@attribute 'Sex' { Male, Female}
@attribute 'Feature 0' real
@attribute 'Feature 1' real
@attribute 'Feature 2' real
@attribute 'Feature 3' real
@attribute 'Feature 4' real
@attribute 'Feature 5' real
@attribute 'Feature 6' real
@attribute 'Feature 7' real
@attribute 'Feature 8' real
@attribute 'Feature 9' real
@attribute 'Class' { hid, hId, hEd, hAd, hYd, had, hOd, hod, hUd, hud, hed}
@data
Train,Andrew,Male,-3.639,0.418,-0.67,1.779,-0.168,1.627,-0.388,0.529,-0.874,-0.814,hid
Train,Andrew,Male,-3.327,0.496,-0.694,1.365,-0.265,1.933,-0.363,0.51,-0.621,-0.488,hId
Train,Andrew,Male,-2.12,0.894,-1.576,0.147,-0.707,1.559,-0.579,0.676,-0.809,-0.049,hEd
Train,Andrew,Male,-2.287,1.809,-1.498,1.012,-1.053,1.06,-0.567,0.235,-0.091,-0.795,hAd
Train,Andrew,Male,-2.598,1.938,-0.846,1.062,-1.633,0.764,0.394,-0.15,0.277,-0.396,hYd
Train,Andrew,Male,-2.852,1.914,-0.755,0.825,-1.588,0.855,0.217,-0.246,0.238,-0.365,had
Train,Andrew,Male,-3.482,2.524,-0.433,1.048,-1.995,0.902,0.322,0.45,0.377,-0.366,hOd
Train,Andrew,Male,-3.941,2.305,0.124,1.771,-1.815,0.593,-0.435,0.992,0.575,-0.301,hod
Train,Andrew,Male,-3.86,2.116,-0.939,0.688,-0.675,1.679,-0.512,0.928,-0.167,-0.434,hUd
Train,Andrew,Male,-3.648,1.812,-1.378,1.578,0.065,1.577,-0.466,0.702,0.06,-0.836,hud
Train,Andrew,Male,-3.032,1.739,-1.141,0.737,-0.834,1.386,-0.575,0.679,-0.018,-0.823,hed
Train,Andrew,Male,-3.653,0.373,-0.6,1.705,-0.222,1.765,-0.353,0.537,-0.797,-0.813,hid
Train,Andrew,Male,-3.237,0.436,-0.86,1.363,-0.251,1.915,-0.395,0.751,-0.774,-0.327,hId
Train,Andrew,Male,-2.135,0.954,-1.632,0.121,-0.704,1.6,-0.628,0.713,-0.903,-0.027,hEd
Train,Andrew,Male,-2.304,1.784,-1.506,0.981,-0.961,0.806,-0.294,-0.002,0.119,-0.76,hAd
Train,Andrew,Male,-2.54,2.144,-1.024,0.933,-1.567,1.024,0.188,-0.047,0.309,-0.633,hYd
Train,Andrew,Male,-2.826,2.003,-0.738,0.801,-1.669,0.939,0.245,-0.257,0.256,-0.458,had
Train,Andrew,Male,-3.582,2.374,-0.358,1.162,-1.953,0.621,0.339,0.355,0.415,-0.259,hOd
Train,Andrew,Male,-3.951,2.25,0.127,1.772,-1.906,0.567,-0.432,1.045,0.598,-0.293,hod
Train,Andrew,Male,-3.783,1.974,-1.2,0.606,-0.65,1.504,-0.134,0.528,0.392,-0.58,hUd
Train,Andrew,Male,-3.673,1.811,-1.405,1.621,0.044,1.572,-0.453,0.745,-0.066,-0.733,hud
Train,Andrew,Male,-2.946,1.649,-1.167,0.788,-0.909,1.3,-0.562,0.902,-0.07,-0.842,hed
Train,Andrew,Male,-3.665,0.337,-0.641,1.791,-0.194,1.686,-0.359,0.57,-0.676,-0.841,hid
```

Fig. 1: Dataset vowel.arff

Cada atributo posee un conjunto de características que se pueden observar a continuaicon:

- *Train or Test* (Fig. 3)
- *Speaker Number* (Fig. 4)
- *Sex* (Fig. 5)
- *Feature 0* (Fig. 6)
- *Feature 1* (Fig. 7)
- *Feature 2* (Fig. 8)
- *Feature 3* (Fig. 9)
- *Feature 4* (Fig. 10)
- *Feature 5* (Fig. 11)
- *Feature 6* (Fig. 12)
- *Feature 7* (Fig. 13)

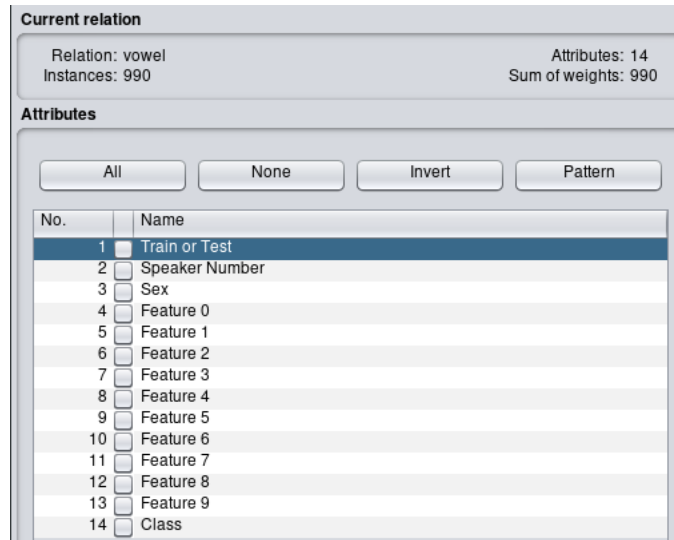


Fig. 2: Dataset vowel.arff en WEKA

Selected attribute

Name: Train or Test
Missing: 0 (0%)

Distinct: 2

Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	Train	528	528.0
2	Test	462	462.0

Fig. 3: Atributo Train or Test

- *Feature 8* (Fig. 14)
- *Feature 9* (Fig. 15)
- *Class* (Fig. 16)

También podemos observar la relación y distribución de los atributos con respecto al atributo *Train or Test*, donde el rojo representa *Train* y azul *Test* (Fig. 17), con respecto al atributo *Speaker Number* (Fig. 18), con respecto al atributo *Sex* (Fig. 19) y con respecto al atributo *Class* (Fig. 20).

Selected attribute

Name: Speaker Number
Missing: 0 (0%)

Distinct: 15

Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	Andrew	66	66.0
2	Bill	66	66.0
3	David	66	66.0
4	Mark	66	66.0
5	Jo	66	66.0
6	Kate	66	66.0
7	Penny	66	66.0
8	Rose	66	66.0
9	Mike	66	66.0
10	Mike	66	66.0

Fig. 4: Atributo Speaker Number

Selected attribute			
Name: Sex Missing: 0 (0%)		Distinct: 2	Type: Nominal Unique: 0 (0%)
No.	Label	Count	Weight
1	Male	528	528.0
2	Female	462	462.0

Fig. 5: Atributo Sex

Selected attribute	
Name: Feature 0 Missing: 0 (0%)	
Distinct: 853	Type: Numeric Unique: 728 (74%)
Statistic	Value
Minimum	-5.211
Maximum	-0.941
Mean	-3.204
StdDev	0.869

Fig. 6: Atributo Feature 0

Perceptron Multicapa

Para el entrenamiento se utilizara una red neuronal del tipo perceptron multicapa el cual utiliza la técnica *back propagation* para clasificar las instancias. Otra característica utilizada en la red que provee WEKA es que utiliza la función de activación *sigmoid*. Los diferentes parámetros (Fig. 21) a utilizar son:

- *seed*: Es utilizado para generar los pesos iniciales de las conexiones iniciales de los nodos, así como de aleatorizar la data de entrenamiento. Se utilizara con un valor de 5.
- *momentum*: Momento aplicado en la actualización de los pesos. Se utilizara con un valor de 0.9.
- *hiddenLayers*: Cantidad de capas ocultas en la red. Se utilizaran 50 capas escondidas.
- *normalizeAttributes*: Indica si los atributos serán normalizados. Se dejara el valor true por defecto.
- *numDecimalPlaces*: Cantidad de decimales a utilizar para la salida. Se tomaran en cuenta 5 decimales.
- *batchSize*: Cantidad de instancias a procesar. Se escogerá un batch de 64.
- *decay*: Indica si hay una disminución en la velocidad de aprendizaje. Se colocara en true la disminución del aprendizaje con el paso del tiempo.
- *normalizeNumericClass*: Indica si se normalizara las clases numéricas. Se dejara el valor true por defecto.

Selected attribute	
Name: Feature 1 Missing: 0 (0%)	
Distinct: 877	Type: Numeric Unique: 771 (78%)
Statistic	Value
Minimum	-1.274
Maximum	5.074
Mean	1.882
StdDev	1.175

Fig. 7: Atributo Feature 1

Selected attribute		
Name: Feature 2		Type: Numeric
Missing: 0 (0%)	Distinct: 815	Unique: 664 (67%)
Statistic	Value	
Minimum	-2.487	
Maximum	1.431	
Mean	-0.508	
StdDev	0.712	

Fig. 8: Atributo Feature 2

Selected attribute		
Name: Feature 3		Type: Numeric
Missing: 0 (0%)	Distinct: 836	Unique: 703 (71%)
Statistic	Value	
Minimum	-1.409	
Maximum	2.377	
Mean	0.515	
StdDev	0.759	

Fig. 9: Atributo Feature 3

- *learningRate*: Velocidad de aprendizaje al momento de actualizar los pesos. La velocidad del aprendizaje sera de 0.4.

Los parámetros para evaluar el modelo se observan en la Fig. 22.

Una vez entrenado el modelo y probado se clasificaron correctamente 890 instancias (Fig. 23) (Verdaderos Positivos) el cual representa un 89.899% de aciertos y se clasificaron incorrectamente 100 instancias (Verdaderos Negativos) el cual representa un 10.101% de fallos. Adicionalmente podemos observar en detalle la precisión con respecto a cada clase en la Fig. 24. En la Fig. 25 podemos observar la matriz de confusión de las clases, cabe recordar que existen 90 instancias para cada clase.

Curvas ROC

Una curva *ROC* (*Receiver Operating Characteristic* - Característica Operativa del Receptor) es la representación de la razón o ratio de verdaderos positivos (VPR = Razón de Verdaderos Positivos) frente a la razón o ratio de falsos positivos (FPR = Razón de Falsos Positivos) también según se varía el umbral de discriminación (valor a partir del cual decidimos que un caso es un positivo). También puede significar *Relative Operating Characteristic* (Característica Operativa Relativa) porque es una comparación de dos características operativas (VPR y FPR).

Con las curvas *ROC* permiten estudiar distintos parámetros y su relación entre si, estos pueden ser:

Selected attribute		
Name: Feature 4		Type: Numeric
Missing: 0 (0%)	Distinct: 803	Unique: 653 (66%)
Statistic	Value	
Minimum	-2.127	
Maximum	1.831	
Mean	-0.306	
StdDev	0.665	

Fig. 10: Atributo Feature 4

Selected attribute		
Name: Feature 5		Type: Numeric
Missing: 0 (0%)	Distinct: 798	Unique: 635 (64%)
Statistic	Value	
Minimum	-0.836	
Maximum	2.327	
Mean	0.63	
StdDev	0.604	

Fig. 11: Atributo Feature 5

Selected attribute		
Name: Feature 6		Type: Numeric
Missing: 0 (0%)	Distinct: 748	Unique: 562 (57%)
Statistic	Value	
Minimum	-1.537	
Maximum	1.403	
Mean	-0.004	
StdDev	0.462	

Fig. 12: Atributo Feature 6

- Verdaderos Positivos o éxitos (VP)
- Verdaderos Negativos o rechazos correctos (VN)
- Falsos Positivos o Error tipo I (FP)
- Falsos Negativos o Error tipo II (FN)
- Sensibilidad o Razón de verdaderos positivos (VPR), también se le conoce como razón de éxitos. $VPR = VP/(VP + FN)$
- Ratio o Razón de falsos positivos (FPR), también se le conoce como razón de falsas alarmas. $FPR = FP/(FP + VN)$
- Exactitud o Accuracy (ACC). $ACC = (VP + VN)/(P + N)$
- Especificidad o Razón de verdaderos negativos (SPC). $SPC = 1 - FPR$
- Valor predictivo positivo (PPV), tambien se le conoce como precisión en recuperación de información. $PPV = VP/(VP + FP)$
- Valor predictivo negativo (NPV). $NPV = VN/(VN + FN)$
- Ratio o Razón de falsos descubrimientos (FDR). $FDR = FP/(FP + VP)$

Selected attribute		
Name: Feature 7		Type: Numeric
Missing: 0 (0%)	Distinct: 794	Unique: 630 (64%)
Statistic	Value	
Minimum	-1.293	
Maximum	2.039	
Mean	0.337	
StdDev	0.573	

Fig. 13: Atributo Feature 7

Selected attribute		
Name: Feature 8 Missing: 0 (0%)		Type: Numeric Unique: 625 (63%)
Distinct: 788		
Statistic	Value	
Minimum	-1.613	
Maximum	1.309	
Mean	-0.303	
StdDev	0.57	

Fig. 14: Atributo Feature 8

Selected attribute		
Name: Feature 9 Missing: 0 (0%)		Type: Numeric Unique: 591 (60%)
Distinct: 775		
Statistic	Value	
Minimum	-1.68	
Maximum	1.396	
Mean	-0.071	
StdDev	0.604	

Fig. 15: Atributo Feature 9

Selected attribute				
Name: Class Missing: 0 (0%)			Type: Nominal Unique: 0 (0%)	
Distinct: 11				
No.	Label	Count	Weight	
3	hEd	90	90.0	▲ ▼
4	hAd	90	90.0	
5	hYd	90	90.0	
6	had	90	90.0	
7	hOd	90	90.0	
8	hod	90	90.0	
9	hUd	90	90.0	
10	hud	90	90.0	
11	hed	90	90.0	

Fig. 16: Atributo Class

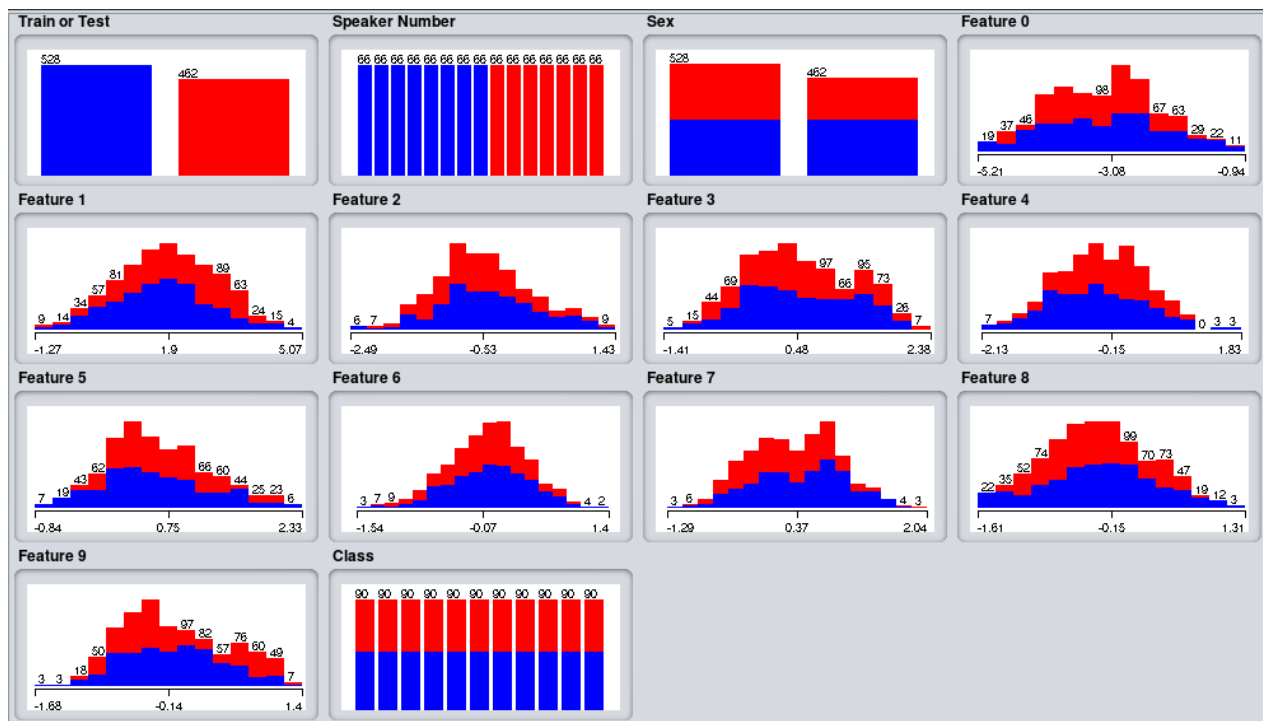


Fig. 17: Relacion Train or Test

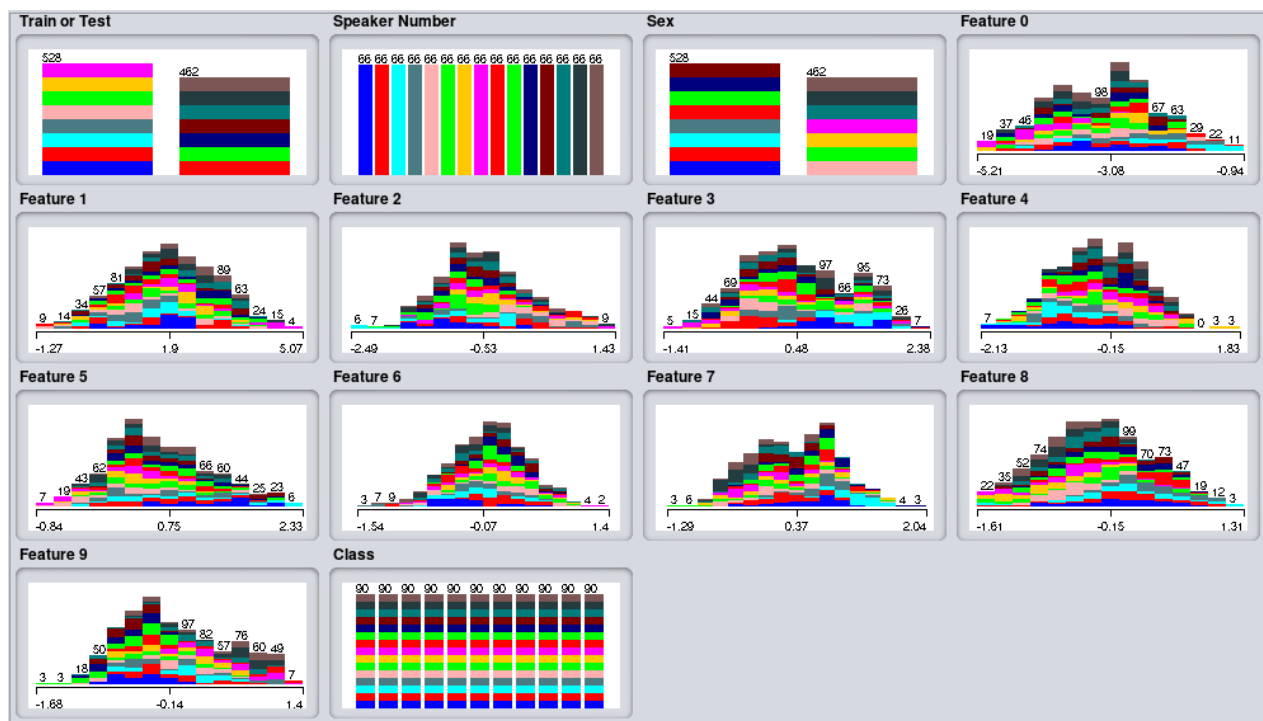


Fig. 18: Relacion Speaker Number

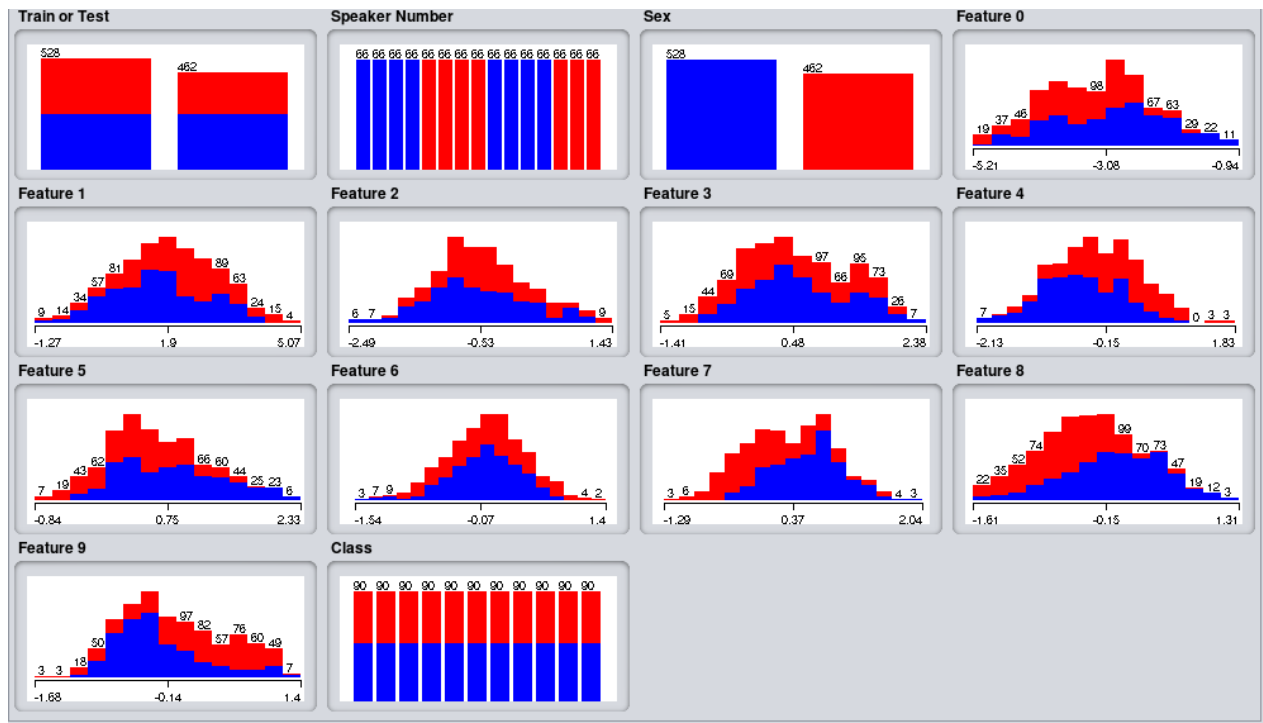


Fig. 19: Relacion Sex

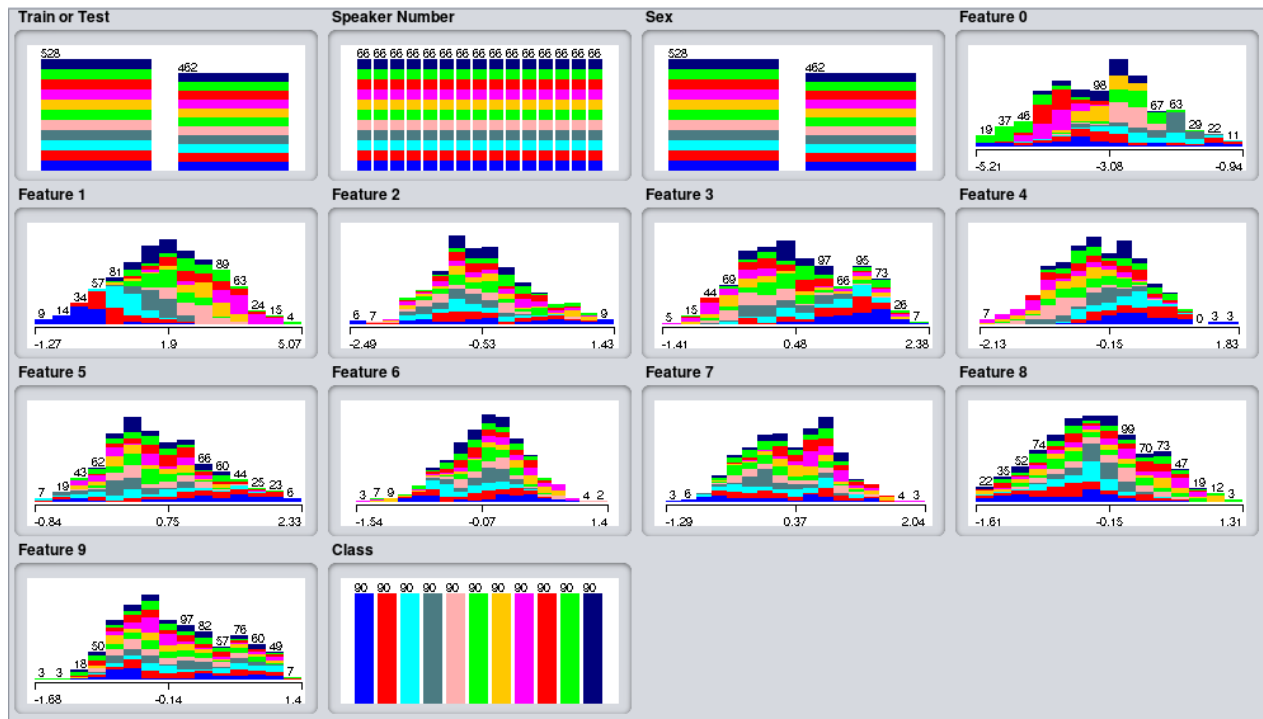


Fig. 20: Relacion Class

weka.classifiers.functions.MultilayerPerceptron

About

A classifier that uses backpropagation to learn a multi-layer perceptron to classify instances.

More

Capabilities

GUI

autoBuild

batchSize

debug

decay

doNotCheckCapabilities

hiddenLayers

learningRate

momentum

nominalToBinaryFilter

normalizeAttributes

normalizeNumericClass

numDecimalPlaces

reset

seed

trainingTime

Fig. 21: Parámetros de la red neuronal

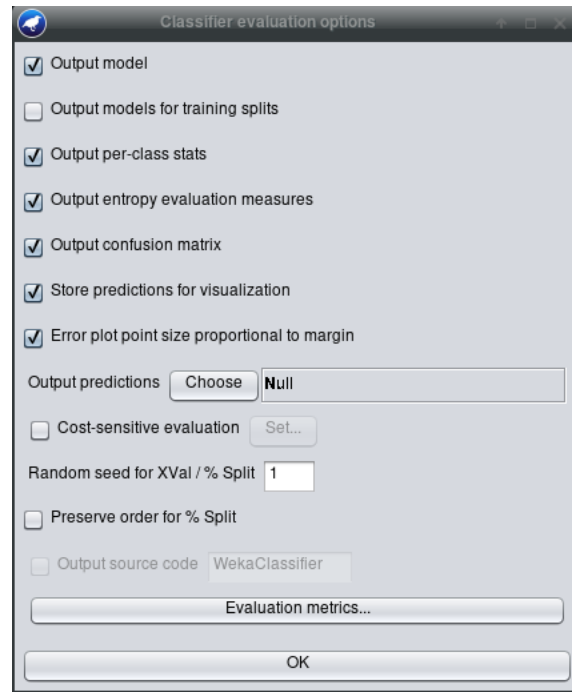


Fig. 22: Parámetros de evaluación

Correctly Classified Instances	890	89.899	%
Incorrectly Classified Instances	100	10.101	%
Kappa statistic	0.8889		
K&B Relative Info Score	83.6553	%	
K&B Information Score	2865.0577 bits	2.894	bits/instance
Class complexity order 0	3424.8373 bits	3.4594	bits/instance
Class complexity scheme	891.8767 bits	0.9009	bits/instance
Complexity improvement (Sf)	2532.9606 bits	2.5585	bits/instance
Mean absolute error	0.046		
Root mean squared error	0.1329		
Relative absolute error	27.8521	%	
Root relative squared error	46.2142	%	
Total Number of Instances	990		

Fig. 23: Resultados del modelo

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.889	0.002	0.976	0.889	0.930	0.925	0.982	0.942	hid
	0.944	0.012	0.885	0.944	0.914	0.906	0.991	0.966	hId
	0.933	0.016	0.857	0.933	0.894	0.883	0.996	0.946	hEd
	0.922	0.002	0.976	0.922	0.949	0.944	0.978	0.945	hAd
	0.878	0.006	0.940	0.878	0.908	0.900	0.993	0.959	hYd
	0.778	0.016	0.833	0.778	0.805	0.786	0.961	0.844	had
	0.833	0.007	0.926	0.833	0.877	0.867	0.909	0.858	hOd
	0.922	0.006	0.943	0.922	0.933	0.926	0.975	0.899	hod
	0.900	0.013	0.871	0.900	0.885	0.874	0.982	0.947	hUd
	0.956	0.014	0.869	0.956	0.910	0.902	0.991	0.970	hud
	0.933	0.018	0.840	0.933	0.884	0.873	0.984	0.914	hed
Weighted Avg.	0.899	0.010	0.902	0.899	0.899	0.890	0.977	0.926	

Fig. 24: Precisión de cada clase

=== Confusion Matrix ===

a	b	c	d	e	f	g	h	i	j	k	<-- classified as
80	2	1	0	0	0	0	0	1	6	0	a = hid
1	85	4	0	0	0	0	0	0	0	0	b = hId
0	3	84	1	0	0	0	0	1	0	1	c = hEd
0	2	2	83	0	2	0	0	0	0	1	d = hAd
0	0	0	0	79	7	3	0	0	0	1	e = hYd
0	0	3	1	3	70	0	0	1	0	12	f = had
1	3	4	0	2	2	75	0	2	0	1	g = hOd
0	0	0	0	0	0	1	83	2	4	0	h = hod
0	1	0	0	0	1	2	2	81	3	0	i = hUd
0	0	0	0	0	0	0	3	1	86	0	j = hud
0	0	0	0	0	2	0	0	4	0	84	k = hed

Fig. 25: Matriz de confusión