

- 1) Utiliza el conjunto de datos *airline.arff* para pronosticar el número de pasajeros que volarán con la aerolínea los próximos 5 o 10 meses (a elegir). Selecciona un intervalo de confianza a elegir entre el 90 y 99%. Haz uso de los últimos 12 o 24 registros del histórico de datos (a elegir) para entrenar un modelo de perceptrón multicapa. Evalúa un conjunto de prueba compuesto por el 20% o el 30% del histórico de datos (a elegir). Muestra por pantalla los resultados cuantitativos obtenidos en el test para alguna de las siguientes métricas (a elegir únicamente una de ellas): MAE o RMS.

Variables a elegir	Valores fijados
<i>Meses a predecir</i>	5
<i>Intervalo de confianza</i>	95%
<i>Número de registros previos</i>	24
<i>Porcentaje para el test</i>	30%
<i>Métrica de error utilizada</i>	MAE

Captura de pantalla con los resultados cuantitativos de la métrica elegida:

```
=== Evaluation on training data ===
Target          1-step-ahead  2-steps-ahead  3-steps-ahead  4-steps-ahead  5-steps-ahead
=====
passenger_numbers
N                77            76            75            74            73
Mean absolute error  6.4801        7.547        8.4671        9.2469        10.1355

Total number of instances: 101

=== Evaluation on test data ===
Target          1-step-ahead  2-steps-ahead  3-steps-ahead  4-steps-ahead  5-steps-ahead
=====
passenger_numbers
N                43            42            41            40            39
Mean absolute error  27.4467       35.4344      40.2979      42.2294      41.1039

Total number of instances: 43
```

- 2) Repite el ejercicio anterior con los mismos valores, pero cambiando el modelo de perceptrón multicapa por uno de regresión lineal. Muestra por pantalla los resultados cuantitativos obtenidos del nuevo modelo y determina cuál de ellos reporta mejores resultados de predicción.

Captura de pantalla con los resultados obtenidos:

```
=== Evaluation on training data ===
Target          1-step-ahead  2-steps-ahead  3-steps-ahead  4-steps-ahead  5-steps-ahead
=====
passenger_numbers
N                77            76            75            74            73
Mean absolute error  4.5167        4.8309       5.0354       5.1539       5.5342

Total number of instances: 101

=== Evaluation on test data ===
Target          1-step-ahead  2-steps-ahead  3-steps-ahead  4-steps-ahead  5-steps-ahead
=====
passenger_numbers
N                43            42            41            40            39
Mean absolute error  44.2192       44.7665      47.0006      47.9422      49.0339

Total number of instances: 43
```

Funciona mejor el modelo de perceptrón multicapa . (teniendo solo en cuenta el 1 y el 2)

- 3) Utiliza el conjunto de datos *airline.arff* para pronosticar el número de pasajeros que volarán con la aerolínea las próximas 8-12 semanas (a elegir). Selecciona un intervalo de confianza a elegir entre el 90 y 99%. Haz uso de las últimas 6 o 12 semanas del histórico de datos (a elegir) para entrenar un modelo de árboles de decisión. Evalúa un conjunto de test compuesto por el 10% o el 20% del histórico de datos (a elegir). Haz una captura de pantalla de los resultados de predicción obtenidos en el test para alguna de las siguientes métricas (a elegir únicamente una de ellas): MSE, MAPE.

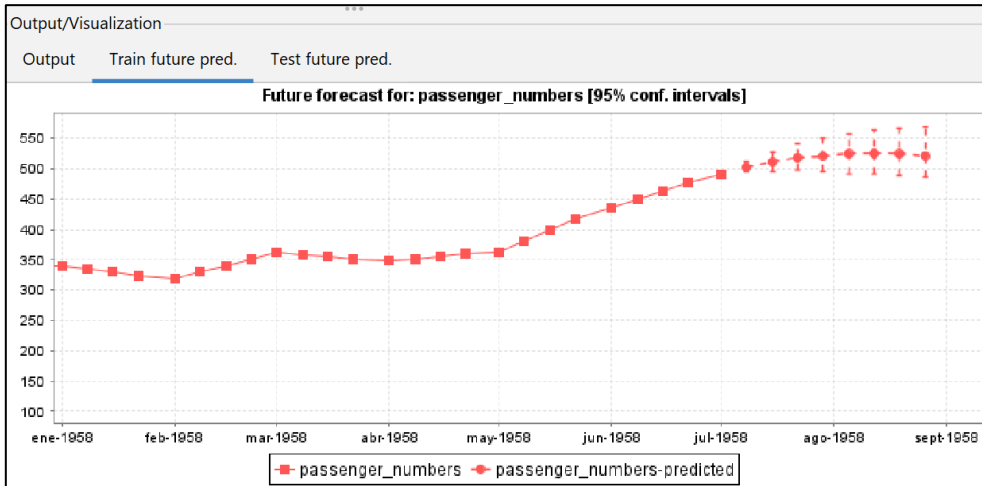
Variables a elegir	Valores fijados
<i>Semanas a predecir</i>	8
<i>Intervalo de confianza</i>	95%
<i>Número de registros previos</i>	12
<i>Porcentaje para el test</i>	20%
<i>Métrica de error utilizada</i>	MAPE

Captura de pantalla con los resultados cuantitativos de la métrica elegida:

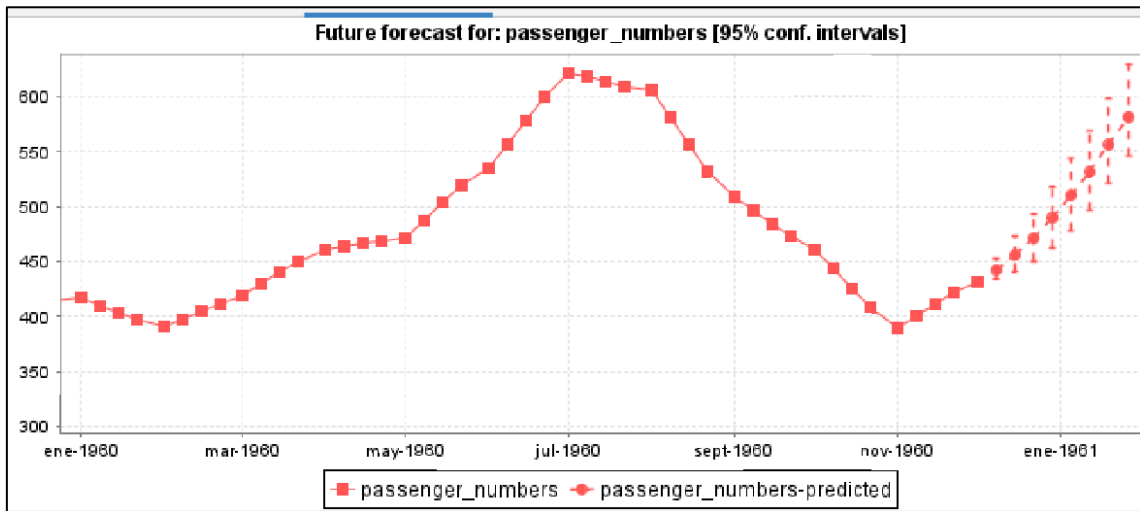
=== Evaluation on training data ===								
Target	1-step-ahead	2-steps-ahead	3-steps-ahead	4-steps-ahead	5-steps-ahead	6-steps-ahead	7-steps-ahead	8-steps-ahead
=====								
passenger_numbers								
N	445	444	443	442	441	440	439	438
Mean absolute percentage error	1.5618	2.8054	4.0233	4.9759	5.7437	6.3078	6.6904	6.9473
Total number of instances: 457								
=== Evaluation on test data ===								
Target	1-step-ahead	2-steps-ahead	3-steps-ahead	4-steps-ahead	5-steps-ahead	6-steps-ahead	7-steps-ahead	8-steps-ahead
=====								
passenger_numbers								
N	113	112	111	110	109	108	107	106
Mean absolute percentage error	1.433	2.9532	4.9158	6.7982	8.6285	10.3016	11.9638	13.945
Total number of instances: 113								

- 4) A partir de las especificaciones fijadas en el ejercicio anterior, realiza una captura de pantalla de las gráficas que se obtienen al predecir los datos tanto de entrenamiento como de test. En concreto, para el entrenamiento, muestra la línea temporal que abarca (aprox.) desde enero de 1958 hasta el último valor predicho. Para el test, muestra la gráfica desde enero de 1960 hasta el final.

Captura de pantalla del gráfico de entrenamiento:



**Captura de pantalla del gráfico de test:**



- 5) Utiliza el conjunto de datos *abalone.arff*. Selecciona un número de bolsas K=5 o K=10 (a elegir) y aplica un modelo de regresión lineal para determinar el error cometido al predecir la edad de las abulones durante el cross-validation. Reporta únicamente una de las siguientes métricas: MAE o RMSE.

Variables a elegir	Valores fijados
Número de bolsas	10
Métrica de error utilizada	RMSE
ERROR COMETIDO	2.2218

- 6) En base a las especificaciones fijadas en el ejercicio anterior, realiza una captura de pantalla mostrando los coeficientes que acompañan a cada una de las características de las abulones en el modelo de regresión lineal propuesto.

**Captura de pantalla de los coeficientes de las características**

```
Rings =
    0.3895 * Sex +
    11.0544 * Diameter +
    11.1838 * Height +
    9.0743 * Whole weight +
    -20.1358 * Shucked weight +
    -10.2093 * Viscera weight +
    8.7171 * Shell weight +
    2.9353
```

- 7) Utiliza el conjunto de datos *winequality-white.arff*. Selecciona un número de bolsas K=5 o K=10 (a elegir) y aplica un modelo de árboles de decisión para determinar el error cometido al predecir la calidad del vino. Reporta únicamente una de las siguientes métricas: MAE o RMSE.

VARIABLES A ELEGIR	VALORES
Número de bolsas	10
Métrica de error utilizada	RMSE
ERROR COMETIDO	0.7188

- 8) En base a las especificaciones fijadas en el ejercicio anterior, haz una captura de pantalla mostrando el árbol de decisión creado. ¿Qué característica es la más importante?

#### Captura de pantalla del árbol de decisión

```

alcohol <= 10.85 :
| volatile acidity <= 0.282 :
| | volatile acidity <= 0.207 :
| | | residual sugar <= 12.575 :
| | | | free sulfur dioxide <= 25.5 :
| | | | | residual sugar <= 2.95 : LM1 (127/77.651%)
| | | | | residual sugar > 2.95 :
| | | | | total sulfur dioxide <= 92 : LM2 (17/0%)
| | | | | total sulfur dioxide > 92 :
| | | | | alcohol <= 10.3 : LM3 (29/47.578%)
| | | | | alcohol > 10.3 : LM4 (15/66.397%)
| | | | free sulfur dioxide > 25.5 : LM5 (393/74.992%)
| | | residual sugar > 12.575 :
| | | | alcohol <= 9.05 :
| | | | | free sulfur dioxide <= 30.5 : LM6 (21/0%)
| | | | | free sulfur dioxide > 30.5 :
| | | | | density <= 0.998 : LM7 (8/0%)
| | | | | density > 0.998 :
| | | | | density <= 0.998 : LM8 (7/0%)
| | | | | density > 0.998 :
| | | | | fixed acidity <= 7.15 : LM9 (8/37.346%)
| | | | | fixed acidity > 7.15 :
| | | | | chlorides <= 0.056 : LM10 (9/81.289%)
| | | | | chlorides > 0.056 : LM11 (6/0%)
| | | alcohol > 9.05 : LM12 (91/58.694%)
| | volatile acidity > 0.207 :
| | | alcohol <= 9.85 :
| | | | residual sugar <= 12.65 :
| | | | | chlorides <= 0.044 :
| | | | | pH <= 3.275 : LM13 (116/60.469%)
| | | | | pH > 3.275 : LM14 (33/80.362%)
| | | | | chlorides > 0.044 : LM15 (296/62.13%)
| | | | residual sugar > 12.65 :
| | | | | residual sugar <= 15.05 :
| | | | | density <= 0.999 :
| | | | | chlorides <= 0.056 :
| | | | | sulphates <= 0.445 : LM16 (13/53.547%)
| | | | | sulphates > 0.445 :
| | | | | volatile acidity <= 0.265 :
| | | | | residual sugar <= 13.25 :
| | | | | total sulfur dioxide <= 167.25 : LM17 (4/29.624%)
| | | | | total sulfur dioxide > 167.25 : LM18 (5/0%)
| | | | | residual sugar > 13.25 :
| | | | | chlorides <= 0.052 : LM19 (10/0%)
| | | | | chlorides > 0.052 :
| | | | | fixed acidity <= 7.45 : LM20 (2/0%)
| | | | | fixed acidity > 7.45 : LM21 (2/0%)
| | | volatile acidity > 0.265 : LM22 (12/0%)

```

```

| | | | | | | | | | volatile acidity > 0.265 : LM22 (12/0%)
| | | | | | | | | | chlorides > 0.056 :
| | | | | | | | | | chlorides <= 0.061 :
| | | | | | | | | | sulphates <= 0.47 :
| | | | | | | | | | | fixed acidity <= 7.3 : LM23 (7/0%)
| | | | | | | | | | | fixed acidity > 7.3 : LM24 (2/0%)
| | | | | | | | | | sulphates > 0.47 : LM25 (4/0%)
| | | | | | | | | | chlorides > 0.061 : LM26 (13/0%)
| | | | | | | | density > 0.999 :
| | | | | | | | density <= 0.999 :
| | | | | | | | | residual sugar <= 14.2 :
| | | | | | | | | | citric acid <= 0.35 : LM27 (2/0%)
| | | | | | | | | | citric acid > 0.35 : LM28 (4/0%)
| | | | | | | | | residual sugar > 14.2 : LM29 (8/0%)
| | | | | | | | density > 0.999 :
| | | | | | | | | pH <= 3.085 : LM30 (8/0%)
| | | | | | | | | pH > 3.085 :
| | | | | | | | | | total sulfur dioxide <= 182.5 :
| | | | | | | | | | sulphates <= 0.68 :
| | | | | | | | | | | fixed acidity <= 8.5 :
| | | | | | | | | | | free sulfur dioxide <= 43.25 : LM31 (5/0%)
| | | | | | | | | | | free sulfur dioxide > 43.25 :
| | | | | | | | | | | | fixed acidity <= 7.15 : LM32 (4/0%)
| | | | | | | | | | | | fixed acidity > 7.15 : LM33 (2/0%)
| | | | | | | | | | | | fixed acidity > 8.5 : LM34 (3/0%)
| | | | | | | | | | sulphates > 0.68 : LM35 (7/0%)
| | | | | | | | | total sulfur dioxide > 182.5 :
| | | | | | | | | | citric acid <= 0.375 : LM36 (6/0%)
| | | | | | | | | | citric acid > 0.375 : LM37 (3/0%)
| | | | | | residual sugar > 15.05 :
| | | | | | total sulfur dioxide <= 131.5 :
| | | | | | | sulphates <= 0.475 : LM38 (26/0%)
| | | | | | | sulphates > 0.475 :
| | | | | | | | fixed acidity <= 6.25 : LM39 (4/0%)
| | | | | | | | fixed acidity > 6.25 : LM40 (7/0%)
| | | | | | total sulfur dioxide > 131.5 :
| | | | | | | total sulfur dioxide <= 150 : LM41 (32/0%)
| | | | | | | total sulfur dioxide > 150 : LM42 (83/69.564%)
| | | alcohol > 9.85 :
| | | | pH <= 3.325 :
| | | | | free sulfur dioxide <= 27.5 : LM43 (130/73.807%)
| | | | | free sulfur dioxide > 27.5 :
| | | | | sulphates <= 0.525 : LM44 (167/77.208%)
| | | | | sulphates > 0.525 :
| | | | | | total sulfur dioxide <= 172.5 : LM45 (65/62.966%)
| | | | | | total sulfur dioxide > 172.5 : LM46 (26/63.239%)
| | | | pH > 3.325 :
| | | | | density <= 0.994 : LM47 (45/76.232%)
| | | | | density > 0.994 : LM48 (77/83.245%)

```

```

| | volatile acidity > 0.282 :
| | | free sulfur dioxide <= 20.5 : LM49 (243/71.159%)
| | | free sulfur dioxide > 20.5 :
| | | | alcohol <= 9.85 :
| | | | | citric acid <= 0.235 : LM50 (165/53.83%)
| | | | | citric acid > 0.235 : LM51 (462/60.182%)
| | | | alcohol > 9.85 : LM52 (251/66.801%)
alcohol > 10.85 :
| | free sulfur dioxide <= 21.5 :
| | | free sulfur dioxide <= 10.5 : LM53 (98/107.925%)
| | | free sulfur dioxide > 10.5 : LM54 (366/88.527%)
| | free sulfur dioxide > 21.5 : LM55 (1349/83.105%)

```

La característica más importante **es el alcohol**.

- 9) Utiliza el conjunto de datos *insurance.arff*. Selecciona un número de bolsas  $K=5$  o  $K=10$  (a elegir) y aplica un modelo de perceptrón multicapa para determinar el error cometido al predecir el coste del seguro. Reporta el error cometido mostrando las métricas relativas (RAE y RRSE).

Variables a elegir	Valores
Número de bolsas	10
Métrica de error utilizada	RAE y RRSE
Relative absolute error	0.35487 (35.487%)
Root relative squared error	0.445184 (44.5184%)

- 10) En base a las especificaciones fijadas en el ejercicio anterior, haz una captura de pantalla mostrando una gráfica donde se represente el error cometido entre la etiqueta real y la etiqueta predicha.

Captura de pantalla del error cometido entre  $y$  e  $\hat{y}$

