

REALIZAR PREDICCIONES

Actividad 1

El archivo MAT `datosletras.mat` contiene `datos_entrena`, la tabla de datos utilizada para entrenar el modelo `modelo_knn`. También contiene `datos_prueba`, que es una tabla de datos (con las mismas características que `datos_entrena`) que el modelo nunca ha visto antes.

Recuerde que puede usar la función `predict` para obtener las predicciones de un modelo para datos nuevos.

```
preds = predict(model,newdata)
```

Tarea: Cargue los datos y ajústelos al modelo `modelo_knn`:

```
load datosletras.mat
modelo_knn = fitcknn(datos_entrena,"Character",...
    "NumNeighbors",5,"Standardize",true,...
    "DistanceWeight","squaredinverse");
```

Utilice la función `predict` para obtener las predicciones del modelo para los datos de prueba. Almacene el resultado en una variable llamada `letra_pred`.

Actividad 2

En este caso, se conocen las clases correctas para los datos de prueba. Se almacenan en la variable `Character` de la tabla `datos_prueba`.

Tarea: Utilice el operador `~=` para determinar la tasa de clasificación errónea (el número de predicciones incorrectas dividido entre el número total de predicciones). Almacene el resultado en una variable llamada `tasa_clas_err`.

Actividad 3

Las clases de respuesta no siempre se distribuyen por igual en los datos de entrenamiento o de prueba. La *pérdida* es una medida más justa de la clasificación errónea que incorpora la probabilidad de cada clase (basada en la distribución en los datos).

```
loss(model,testdata)
```

Tarea: Utilice la función `loss` para determinar la pérdida de datos de prueba para el modelo kNN `modelo_knn`. Almacene el resultado en una variable llamada `perdida`.

Tarea adicional

Puede calcular la pérdida en cualquier conjunto de datos donde se conozca la clase correcta. Intente determinar la pérdida de los datos de entrenamiento originales (`datos_entrena`). Esto se conoce

como *pérdida por resustitución* (la pérdida cuando los datos de entrenamiento se “resustituyen” en el modelo). Puede calcular la pérdida por resustitución directamente con:

```
resubLoss(modelo_knn)
```

Archivos requeridos:

```
datosletras.mat
```