**RECONOCIMIENTO DE FORMAS**

**MEMORIA PERCEPTRÓN Y ENSEMBLES**



**Autores:**

González Rodríguez, Enrique; Matrícula: y160329

Jiménez Martín, Carlos; Matrícula: y160190

**FACULTAD DE INGENIERÍA INFORMÁTICA UNIVERSIDAD POLITÉCNICA DE MADRID**

**CURSO 2019-2020**

**Índice**

[1. Perceptrón 3](#_Toc28358718)

[1.1 Iris Dataset 3](#_Toc28358719)

[1.2 Wine Dataset 4](#_Toc28358720)

[1.3 Cancer Dataset 5](#_Toc28358721)

[1.4 Isolet Dataset 5](#_Toc28358722)

[1.5 MNIST Dataset 6](#_Toc28358723)

[2. Ensembles 7](#_Toc28358724)

[2.1 Implementación 7](#_Toc28358725)

[2.2 Iris Dataset 8](#_Toc28358726)

[2.3 Wine Dataset 8](#_Toc28358727)

[2.4 Cancer Dataset 8](#_Toc28358728)

[2.5 Isolet Dataset 9](#_Toc28358729)

[2.6 MNIST Dataset 9](#_Toc28358730)

# Perceptrón

Para llevar a cabo la parte correspondiente del algoritmo Perceptrón en esta práctica, se ha utilizado el algoritmo ya implementado de la librería “Scikit Learn”.

Como parámetros se han utilizado:

* **eta0:** es una constante por la que se multiplican los “updates” que elabora el algoritmo.Se ha utilizado un valor igual a 1/k, donde k es el número de iteraciones del proceso de aprendizaje efectivo.
* **n\_jobs:** es el número de CPUs que se van a utilizar durante el algoritmo. Se ha puesto con valor igual a -1 para que se utilicen el mayor número de CPUs disponibles.
* **early\_stopping:** es un parámetro que sirve para detener el aprendizaje del clasificador cuando el porcentaje de mejora no está mejorando.
* **validation\_fraction:** es un valor que establece que porcentaje de datos se reservan para validar y poder detener el aprendizaje.

A continuación, se muestran los distintos resultados obtenidos para los dataset Iris, Wine, Cancer, Isolet y MNIST.

Se puede observar que con el parámetro early\_stoping=False se obtiene todo el rato el mismo resultado, pero sin embargo con early\_stoping=True se obtienen resultados diferentes debido a que el aprendizaje se detiene cuando la validación muestra un porcentaje de acierto menor que en la iteración anterior.

Por otro lado, estos resultados son orientativos, ya que el Perceptrón tiene un parámetro llamado “**random\_state”** el cual baraja los datos de una forma distinta para cada ejecución basándose en un número aleatorio, por lo que, si se ejecuta el código repetidamente, veremos cómo los resultados varían.

## 1.1 Iris Dataset

**early\_stopping**=True:

Precisión con validation\_fraction: 0.025 : 0.70 (+/- 0.31)

Precisión con validation\_fraction: 0.05 : 0.81 (+/- 0.24)

Precisión con validation\_fraction: 0.075 : 0.67 (+/- 0.31)

Precisión con validation\_fraction: 0.1 : 0.62 (+/- 0.38)

Precisión con validation\_fraction: 0.2 : 0.69 (+/- 0.36)

Precisión con validation\_fraction: 0.3 : 0.71 (+/- 0.31)

Precisión con validation\_fraction: 0.4 : 0.68 (+/- 0.41)

Precisión con validation\_fraction: 0.5 : 0.82 (+/- 0.29)

Precisión con validation\_fraction: 0.6 : 0.47 (+/- 0.29)

Precisión con validation\_fraction: 0.7 : 0.67 (+/- 0.28)

Precisión con validation\_fraction: 0.8 : 0.71 (+/- 0.30)

Precisión con validation\_fraction: 0.9 : 0.53 (+/- 0.44)

Precisión con validation\_fraction: 0.95 : 0.40 (+/- 0.27)

**early\_stopping**=False:

Precisión con validation\_fraction: 0.025 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.05 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.075 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.1 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.2 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.3 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.4 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.5 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.6 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.7 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.8 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.9 : 0.71 (+/- 0.25)

Precisión con validation\_fraction: 0.95 : 0.71 (+/- 0.25)

## 1.2 Wine Dataset

**early\_stopping**=True:

Precisión con validation\_fraction: 0.025 : 0.55 (+/- 0.14)

Precisión con validation\_fraction: 0.05 : 0.52 (+/- 0.20)

Precisión con validation\_fraction: 0.075 : 0.53 (+/- 0.22)

Precisión con validation\_fraction: 0.1 : 0.52 (+/- 0.28)

Precisión con validation\_fraction: 0.2 : 0.47 (+/- 0.23)

Precisión con validation\_fraction: 0.3 : 0.50 (+/- 0.27)

Precisión con validation\_fraction: 0.4 : 0.51 (+/- 0.23)

Precisión con validation\_fraction: 0.5 : 0.46 (+/- 0.29)

Precisión con validation\_fraction: 0.6 : 0.40 (+/- 0.24)

Precisión con validation\_fraction: 0.7 : 0.43 (+/- 0.23)

Precisión con validation\_fraction: 0.8 : 0.47 (+/- 0.24)

Precisión con validation\_fraction: 0.9 : 0.40 (+/- 0.22)

Precisión con validation\_fraction: 0.95 : 0.34 (+/- 0.08)

**early\_stopping**=False:

Precisión con validation\_fraction: 0.025 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.05 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.075 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.1 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.2 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.3 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.4 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.5 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.6 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.7 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.8 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.9 : 0.56 (+/- 0.24)

Precisión con validation\_fraction: 0.95 : 0.56 (+/- 0.24)

## 1.3 Cancer Dataset

**early\_stopping**=True:

Precisión con validation\_fraction: 0.025 : 0.83 (+/- 0.27)

Precisión con validation\_fraction: 0.05 : 0.89 (+/- 0.08)

Precisión con validation\_fraction: 0.075 : 0.83 (+/- 0.26)

Precisión con validation\_fraction: 0.1 : 0.84 (+/- 0.26)

Precisión con validation\_fraction: 0.2 : 0.83 (+/- 0.22)

Precisión con validation\_fraction: 0.3 : 0.72 (+/- 0.38)

Precisión con validation\_fraction: 0.4 : 0.75 (+/- 0.42)

Precisión con validation\_fraction: 0.5 : 0.77 (+/- 0.35)

Precisión con validation\_fraction: 0.6 : 0.77 (+/- 0.39)

Precisión con validation\_fraction: 0.7 : 0.81 (+/- 0.21)

Precisión con validation\_fraction: 0.8 : 0.69 (+/- 0.43)

Precisión con validation\_fraction: 0.9 : 0.62 (+/- 0.42)

Precisión con validation\_fraction: 0.95 : 0.68 (+/- 0.39)

**early\_stopping**=False:

Precisión con validation\_fraction: 0.025 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.05 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.075 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.1 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.2 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.3 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.4 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.5 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.6 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.7 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.8 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.9 : 0.86 (+/- 0.13)

Precisión con validation\_fraction: 0.95 : 0.86 (+/- 0.13)

## 1.4 Isolet Dataset

**early\_stopping**=True:

Precisión con validation\_fraction: 0.025 : 0.93 (+/- 0.02)

Precisión con validation\_fraction: 0.05 : 0.93 (+/- 0.03)

Precisión con validation\_fraction: 0.075 : 0.93 (+/- 0.03)

Precisión con validation\_fraction: 0.1 : 0.93 (+/- 0.03)

Precisión con validation\_fraction: 0.2 : 0.92 (+/- 0.02)

Precisión con validation\_fraction: 0.3 : 0.92 (+/- 0.03)

Precisión con validation\_fraction: 0.4 : 0.92 (+/- 0.02)

Precisión con validation\_fraction: 0.5 : 0.92 (+/- 0.04)

Precisión con validation\_fraction: 0.6 : 0.90 (+/- 0.03)

Precisión con validation\_fraction: 0.7 : 0.90 (+/- 0.02)

Precisión con validation\_fraction: 0.8 : 0.86 (+/- 0.06)

Precisión con validation\_fraction: 0.9 : 0.83 (+/- 0.06)

Precisión con validation\_fraction: 0.95 : 0.75 (+/- 0.08)

**early\_stopping**=False:

Precisión con validation\_fraction: 0.025 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.05 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.075 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.1 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.2 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.3 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.4 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.5 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.6 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.7 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.8 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.9 : 0.94 (+/- 0.01)

Precisión con validation\_fraction: 0.95 : 0.94 (+/- 0.01)

## 1.5 MNIST Dataset

**early\_stopping**=True:

Precisión con validation\_fraction: 0.025 : 0.86 (+/- 0.03)

Precisión con validation\_fraction: 0.05 : 0.86 (+/- 0.03)

Precisión con validation\_fraction: 0.075 : 0.86 (+/- 0.02)

Precisión con validation\_fraction: 0.1 : 0.85 (+/- 0.05)

Precisión con validation\_fraction: 0.2 : 0.86 (+/- 0.03)

Precisión con validation\_fraction: 0.3 : 0.85 (+/- 0.02)

Precisión con validation\_fraction: 0.4 : 0.86 (+/- 0.02)

Precisión con validation\_fraction: 0.5 : 0.85 (+/- 0.02)

Precisión con validation\_fraction: 0.6 : 0.86 (+/- 0.03)

Precisión con validation\_fraction: 0.7 : 0.86 (+/- 0.02)

Precisión con validation\_fraction: 0.8 : 0.86 (+/- 0.03)

Precisión con validation\_fraction: 0.9 : 0.83 (+/- 0.02)

Precisión con validation\_fraction: 0.95 : 0.82 (+/- 0.03)

**early\_stopping**=False:

Precisión con validation\_fraction: 0.025 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.05 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.075 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.1 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.2 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.3 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.4 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.5 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.6 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.7 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.8 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.9 : 0.87 (+/- 0.02)

Precisión con validation\_fraction: 0.95 : 0.87 (+/- 0.02)

# Ensembles

## 2.1 Implementación

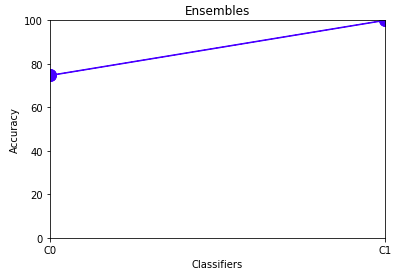
Para esta parte de la práctica tuvimos que programar el algoritmo exprofeso, utilizando el clasificador del Perceptrón de la librería “Scikit Learn” como base.

De manera genérica, lo que hace el algoritmo es coger un conjunto de datos e ir creando clasificadores de forma iterativa. Esta generación se produce de manera que, los datos mal predichos de un clasificador son los datos de entrada para el entrenamiento del siguiente clasificador. Este proceso se repite, de manera teórica, hasta alcanzar una mejora predefinida de la precisión inicial del clasificador a partir del conjunto de clasificadores producidos. Por ejemplo, si buscamos una mejora del 5% y la precisión inicial es del 75%, el algoritmo acaba cuando se han creado tantos clasificadores que den de resultado una precisión igual o mayor al 80%.

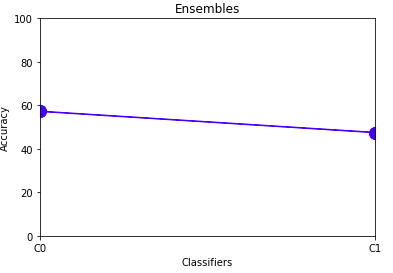
A la hora de implementarlo, a pesar del concepto teórico, nos hemos encontrado con dos casos excepcionales que nos llevan a cortar el algoritmo sin haber alcanzado la mejora buscada de la precisión:

* Cuando los datos de entrada de entrenamiento de un clasificador contienen datos pertenecientes a una sola clase, ya que no se clasificar un dataset si no hay un mínimo de dos clases.
* Cuando uno de los clasificadores no falla a la hora de clasificar los datos, de manera que el siguiente tiene datos de entrada para su entrenamiento.

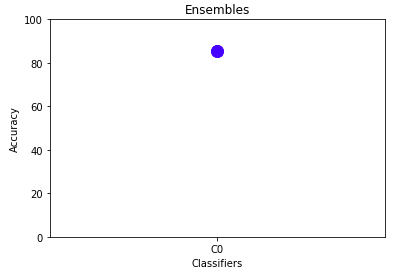
## 2.2 Iris Dataset



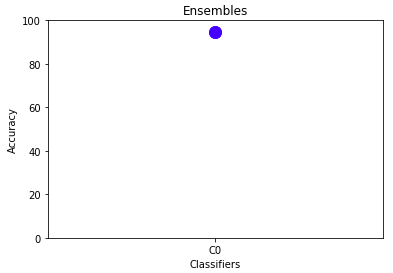
## 2.3 Wine Dataset



## 2.4 Cancer Dataset



## 2.5 Isolet Dataset



## 2.6 MNIST Dataset

