

Aprendizaje Automático Profundo (Deep Learning)

Practica 4 – Mutiperceptrón

Material de Lectura: Capítulos 2 y 3 del libro *Neural Networks and Deep Learning*.

- 1) El archivo **Vinos.csv** tiene información referida a 13 características químicas y/o visuales de varias muestras de vinos pertenecientes a 3 clases distintas.

Utilice el 80% de los ejemplos del archivo **Vinos.csv** para entrenar un mutiperceptrón que sea capaz que distinguir entre las 3 clases de vinos. Observe la tasa de acierto obtenida sobre el 20% restante.

Fuente de datos: *Wine Data Set* - <https://archive.ics.uci.edu/ml/datasets/wine>

- 2) El archivo **Balance.csv** tiene información sobre un experimento psicológico realizado para evaluar el aprendizaje en los niños. Cada fila de la tabla tiene las características de una balanza, referidas a la longitud de los brazos izquierdo y derecho de la balanza y al peso que hay en cada brazo, y un atributo que indica si la balanza se inclina al lado izquierdo (**L**), derecho (**R**), o está balanceada (**B**).

Utilice una parte de los ejemplos para entrenar un mutiperceptrón que sea capaz que predecir si la balanza está inclinada a derecha, a izquierda o si está balanceada. Analice la precisión de la red sobre los ejemplos de entrenamiento y sobre los de testeo.

Fuente: **Balance Scale Data Set** - <https://archive.ics.uci.edu/ml/datasets/Balance+Scale>

- 3) El archivo **ZOO.csv** contiene información de 101 animales caracterizados por los siguientes atributos

1. Nombre del animal	10. Vertebrado
2. Tiene Pelo	11. Branquias
3. Plumas	12. Venenoso
4. Huevos	13. Aletas
5. Leche	14. Patas
6. Vuela	15. Cola
7. Acuático	16. Domesticado
8. Depredador	17. Tamaño gato
9. Dentado	18. Clase

Salvo los atributos 1 y 18 que contienen texto y el 14 que contiene el número de patas del animal, el resto toma el valor 1 si el animal posee la característica y 0 si no. Hay 7 valores de clase posible (atributo 18): mamífero, ave, pez, invertebrado, insecto, reptil y anfibio.

Entrene un mutiperceptrón que sea capaz de clasificar un animal en una de las 7 clases.

Utilice el 70% de los ejemplos para entrenar y el 30% para realizar el testeo.

Realice al menos 10 ejecuciones independientes de la configuración seleccionada para respaldar sus afirmaciones referidas a la performance del modelo.

Fuente de Datos: **Zoo Data Set** - <https://archive.ics.uci.edu/ml/datasets/zoo>

- 4) Los archivos **Segment_Train.csv** y **Segment_Test.csv** contienen información referida a regiones de 3x3 pixels pertenecientes a 7 imágenes distintas. Cada una corresponde a uno de los siguientes tipos de superficie: ladrillo, cielo, follaje, cemento, ventana, camino y pasto.

Cada región de 3x3 ha sido caracterizada por 19 atributos numéricos:

1. region-centroid-col: la columna del pixel central de la región.
2. region-centroid-row: la fila del pixel central de la región.
3. region-pixel-count: el número de pixels de la región = 9.
4. short-line-density-5: el resultado de un algoritmo de extracción de líneas que cuenta la cantidad de líneas de bajo contraste que atraviesan la región.
5. short-line-density-2: ídem anterior para líneas de alto contraste.
6. vedge-mean: medida del contraste entre pixels adyacentes. Este atributo contiene el valor promedio y el siguiente la desviación. Estas medidas sirven para detectar la presencia de un eje vertical.
7. vegde-sd: (ver 6)
8. hedge-mean: ídem 6 para eje horizontal. Contiene el valor medio y el siguiente la desviación.
9. hedge-sd: (see 8).
10. intensity-mean: El promedio calculado sobre la región de la forma $(R + G + B)/3$
11. rawred-mean: el promedio sobre la región de los valores R.
12. rawblue-mean: el promedio sobre la región de los valores B.
13. rawgreen-mean: el promedio sobre la región de los valores G.
14. exred-mean: Medida de exceso de color rojo: $(2R - (G + B))$
15. exblue-mean: Medida de exceso de color azul: $(2B - (G + R))$
16. exgreen-mean: Medida de exceso de color verde: $(2G - (R + B))$
17. value-mean: Transformación no lineal 3D de RGB.
18. saturatoin-mean: (ver 17)
19. hue-mean: (ver 17)

El atributo 20 corresponde al número de imagen de la cual fue extraída la región de 3x3. Sus valores son: 1 (ladrillo), 2 (cemento), 3 (follaje), 4 (pasto), 5 (camino), 6 (cielo), 7 (ventana)

Entrene una red neuronal multiperceptrón para que dada una región de 3x3, representada a través de los 19 atributos indicados anteriormente, sea capaz de identificar a cuál de las 7 imágenes corresponde.

Utilice los ejemplos del archivo **Segment_Train.csv** para entrenar y los del archivo **Segment_Test.csv** para realizar el testeo.

Realice al menos 10 ejecuciones independientes de la configuración seleccionada para respaldar sus afirmaciones referidas a la performance del modelo.

Fuente: **Image Segmentation Data Set** - <https://archive.ics.uci.edu/ml/datasets/Image+Segmentation>