

INTELIGENCIA ARTIFICIAL

Práctica 5

2018/2019

Jesús Aguas Acín -- 736935@unizar.es

Perceptrón

P5_demo.py

Carga datos `x_train` e `y_train` (60.000 muestras de entrenamiento y validación) y datos `x_test` e `y_test` (10.000 muestras para comprobar el resultado final de la red entrenada).

Define y entrena una única capa de perceptrones.

La entrada es de 28*28 pixeles = 784

Función de coste: *categorical_crossentropy*

		% Training	% Test	Time
Sigmoidal	SGD	90 %	90,68 %	23,2 seg.
	Adam	93,27 %	92,74 %	15 seg.
	RMSprop	93,22 %	92,73 %	14,5 seg.
Softmax	SGD	90,16 %	90,82 %	15 seg.
	Adam	93,27 %	92,78 %	15,8 seg.
	RMSprop	93,22 %	92,70 %	15,6 seg.

Función de coste: *mean_squared_error*

		% Training	% Test	Time
Sigmoidal	SGD	77,15 %	78,20 %	12,7 seg.
	Adam	92,34 %	92,34 %	14 seg.
	RMSprop	92,38 %	92,38 %	14,6 seg.
Softmax	SGD	76,42 %	77,82 %	13,5 seg.
	Adam	94,04 %	93,16 %	17 seg.
	RMSprop	93,78 %	93,20 %	13,7 seg.

Perceptrón multi-nivel con una capa oculta

Capa oculta de 10 neuronas

Función de coste: *categorical_crossentropy*

		% Training	% Test	Time
Sigmoidal - Sigmoidal	SGD	82,58 %	83,36 %	17 seg.
	Adam	92,94 %	92,54 %	17 seg.
	RMSprop	92,70 %	92,38 %	16,6 seg.
RELU - Sigmoidal	SGD	91,26 %	91,47 %	15 seg.
	Adam	93,86 %	93,16 %	15,8 seg.
	RMSprop	93,88 %	93,36 %	15,6 seg.
Sigmoidal - Softmax	SGD	85,42 %	86,16 %	16,8 seg.
	Adam	93,51 %	92,86 %	17,3 seg.
	RMSprop	93,42 %	93,00 %	17,5 seg.
RELU - Softmax	SGD	91,57 %	91,74 %	17,6 seg.
	Adam	94,23 %	93,51 %	17,8 seg.
	RMSprop	94,18 %	93,67 %	17,4 seg.

Función de coste: *mean_squared_error*

		% Training	% Test	Time
Sigmoidal - Sigmoidal	SGD	22,60 %	22,82 %	15,1 seg.
	Adam	91,28 %	91,26 %	22,6 seg.
	RMSprop	92,09 %	91,86 %	22,8 seg.
RELU - Sigmoidal	SGD	52,49 %	53,26 %	15 seg.
	Adam	92,11 %	91,94 %	15,8 seg.
	RMSprop	92,62 %	92,57 %	15,6 seg.
Sigmoidal - Softmax	SGD	27,15 %	27,27 %	16,2 seg.
	Adam	93,67 %	92,91 %	16,7 seg.
	RMSprop	93,56 %	93,18 %	16,5 seg.
RELU - Softmax	SGD	55,62 %	55,98 %	16 seg.
	Adam	93,54 %	93,19 %	16,6 seg.
	RMSprop	94,67 %	93,73 %	18,5 seg.

Capa oculta de 500 neuronas (500-10)

Función de coste: *categorical_crossentropy*

		% Training	% Test	Time
Sigmoidal - Sigmoidal	SGD	89,25 %	89,72 %	101 seg.
	Adam	99,70 %	97,77 %	154 seg.
	RMSprop	99,59 %	97,92 %	142 seg.
RELU - Sigmoidal	SGD	93,63 %	93,63 %	113 seg.
	Adam	99,82 %	98,23 %	107 seg.
	RMSprop	99,81 %	98,17 %	130 seg.
Sigmoidal - Softmax	SGD	89,34 %	89,79 %	117 seg.
	Adam	99,71 %	98,02 %	136 seg.
	RMSprop	99,62 %	98,00 %	123 seg.
RELU - Softmax	SGD	93,97 %	94,02 %	115 seg.
	Adam	99,83 %	98,29 %	109 seg.
	RMSprop	99,82 %	98,13 %	83 seg.

Función de coste: *mean_squared_error*

		% Training	% Test	Time
Sigmoidal - Sigmoidal	SGD	43,95 %	44,96 %	76 seg.
	Adam	99,15 %	97,86 %	120 seg.
	RMSprop	99,22 %	98,12 %	94 seg.
RELU - Sigmoidal	SGD	79,43 %	79,43 %	82 seg.
	Adam	99,61 %	98,21 %	110 seg.
	RMSprop	99,54 %	98,32 %	112 seg.
Sigmoidal - Softmax	SGD	41,03 %	41,82 %	81 seg.
	Adam	99,43 %	97,98 %	115 seg.
	RMSprop	99,34 %	97,90 %	98 seg.
RELU - Softmax	SGD	79,11 %	79,78 %	77 seg.
	Adam	99,63 %	98,06 %	105 seg.
	RMSprop	99,69 %	98,18 %	103 seg.

A partir del mejor resultado se ha probado con 784 neuronas training: 99,66% test: 98,34%
Y 1000 neuronas training: 99,85% test 98,37%.
512 neuronas con 30 épocas training: 99,81% test 98,51%

Perceptrón multi-nivel con dos capas oculta

Capas ocultas de 10 neuronas (10-10-10)

Función de coste: *categorical_crossentropy*

		% Training	% Test
RELU - RELU - Sigmoidal	SGD	91,49 %	91,53 %
	Adam	93,81 %	93,00 %
	RMSprop	93,80 %	93,44 %
RELU - RELU - Softmax	SGD	92,01 %	92,04 %
	Adam	94,67 %	93,70 %
	RMSprop	94,63 %	93,78 %
Sigmoidal - RELU - Sigmoidal	SGD	83,56 %	83,48 %
	Adam	94,05 %	93,17 %
	RMSprop	93,87 %	93,12 %
Sigmoidal - RELU - Softmax	SGD	85,93 %	85,89 %
	Adam	94,08 %	93,16 %
	RMSprop	94,16 %	93,33 %

Función de coste: *mean_squared_error*

		% Training	% Test
RELU - RELU - Sigmoidal	SGD	43,49 %	43,33 %
	Adam	85,39 %	85,17 %
	RMSprop	93,92 %	93,62 %
RELU - RELU - Softmax	SGD	30,59 %	30,23 %
	Adam	95,08 %	93,83 %
	RMSprop	94,80 %	93,47 %
Sigmoidal - RELU - Sigmoidal	SGD	19,10 %	19,07 %
	Adam	92,41 %	91,51 %
	RMSprop	92,70 %	91,66 %
Sigmoidal - RELU - Softmax	SGD	12,85 %	13,11 %
	Adam	94,76 %	93,55 %
	RMSprop	94,28 %	93,32 %

Capas ocultas de 500 neuronas (500-500-10)

Función de coste: *mean_squared_error*

		% Training	% Test	Time
RELU - RELU - Softmax	Adam	99,51 %	98,06 %	182 seg.
	RMSprop	99,44 %	97,98 %	148 seg.
Sigmoidal - RELU - Sigmoidal	Adam	99,52 %	97,96 %	164 seg.
	RMSprop	99,54 %	98,12 %	164 seg.
Sigmoidal - RELU - Softmax	Adam	99,38 %	97,89 %	172 seg.
	RMSprop	99,48 %	97,99 %	133 seg.
Sigmoidal - Sigmoidal - Sigmoidal	Adam	99,54 %	98,07 %	159 seg.
	RMSprop	99,39 %	98,24 %	138 seg.

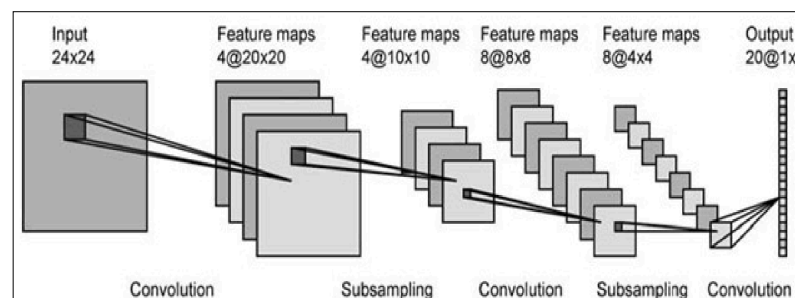
Red Convolucional

```
('x_train shape:', (60000, 28, 28, 1))
(60000, 'train samples')
(10000, 'test samples')
Epoch 1/12
2019-01-10 02:38:36.132193: I tensorflow/core/platform/cpu_feature_guard.cc:141] Your CPU supports
instructions that this TensorFlow binary was not compiled to use: SSE4.1 SSE4.2 AVX AVX2 FMA
2019-01-10 02:38:36.132820: I tensorflow/core/common_runtime/process_util.cc:69] Creating new thread
pool with default inter op setting: 4. Tune using inter_op_parallelism_threads for best performance.
60000/60000 [=====] - 144s 2ms/step - loss: 0.2683 - acc: 0.9174
Epoch 2/12
60000/60000 [=====] - 132s 2ms/step - loss: 0.0850 - acc: 0.9748
Epoch 3/12
60000/60000 [=====] - 129s 2ms/step - loss: 0.0665 - acc: 0.9799
Epoch 4/12
60000/60000 [=====] - 129s 2ms/step - loss: 0.0543 - acc: 0.9837
Epoch 5/12
60000/60000 [=====] - 125s 2ms/step - loss: 0.0460 - acc: 0.9861
Epoch 6/12
60000/60000 [=====] - 124s 2ms/step - loss: 0.0406 - acc: 0.9874
Epoch 7/12
60000/60000 [=====] - 143s 2ms/step - loss: 0.0363 - acc: 0.9889
Epoch 8/12
60000/60000 [=====] - 143s 2ms/step - loss: 0.0344 - acc: 0.9891
Epoch 9/12
60000/60000 [=====] - 145s 2ms/step - loss: 0.0315 - acc: 0.9904
Epoch 10/12
60000/60000 [=====] - 131s 2ms/step - loss: 0.0301 - acc: 0.9906
Epoch 11/12
60000/60000 [=====] - 136s 2ms/step - loss: 0.0291 - acc: 0.9914
Epoch 12/12
60000/60000 [=====] - 130s 2ms/step - loss: 0.0264 - acc: 0.9918
('Test loss:', 0.02712853874079865)
('Test accuracy:', 0.9912)
```

Podemos observar unos resultados excepcionales, la precisión obtenida en los test es del 99%, que comparada con la de entrenamiento, 99%, indica que no se ha producido sobreajuste, ya que no existe ninguna gran ventaja acertando los datos de entrenamiento, que los del test.

Se ha conseguido evitar el sobreajuste mediante *Dropout* (apagar aleatoriamente algunas neuronas al entrenar para mejorar la generalización de la red).

Ej. Red convolucional



cada pixel de salida es una combinación lineal de los píxeles de entrada

Dropout effect

