

APRENDIZAJE ESTADÍSTICO

Boletín 4: Redes Neuronales

ANDRÉS CAMPOS CUIÑA

FECHA DE ENTREGA: 01/12/2021

ÍNDICE

1	Ejercicio 1.....	1
2	Ejercicio 2.....	3
3	Ejercicio 3.....	15

1 EJERCICIO 1

Se tiene una red neuronal con una capa oculta, $s_l = \{2, 3, 1\}$, $W(1) = (-2, 1; 1, -1; 3, -1)$, $W(2) = (2, 3, 1)$, $b(1) = (0; -1; 1)$, y $b(2) = (0)$. La red utiliza la función de activación sigmoide en todas las neuronas salvo la de salida, cuya función de activación es la identidad. Además, la función de coste es el error cuadrático. Dado el ejemplo (1, 1, 1):

1. Calcula $z_i(l)$ y $a_i(l)$ para todas las neuronas.

```

a1 = [[1.]
      [1.]]
z2 = [[-1.]
      [-1.]
      [ 3.]]
a2 = [[0.26894142]
      [0.26894142]
      [0.95257413]]
a3 = [[2.29728123]]
z3 = [[2.29728123]]

```

2. Usando el algoritmo de retro-propagación del error, calcula $\delta_i(l)$ para todas las neuronas.

```

delta3 = [[1.29728123]]
delta2 = [[0.51012194]
          [0.76518291]
          [0.05860683]]

```

3. Usando el algoritmo de retro-propagación del error, determina los valores finales de cada peso ($W_{ij}(l)$) y bias ($b_i(l)$) de la red neuronal tras finalizar la primera iteración del algoritmo, asumiendo un valor de $\lambda = 1$, y una tasa de aprendizaje $\alpha = 0,5$.

```

W1 = [[-1.25506097  0.24493903]
      [ 0.11740854 -0.88259146]
      [ 1.47069658 -0.52930342]]

```

b1 = [[-0.25506097]

[-1.38259146]

[0.97069658]]

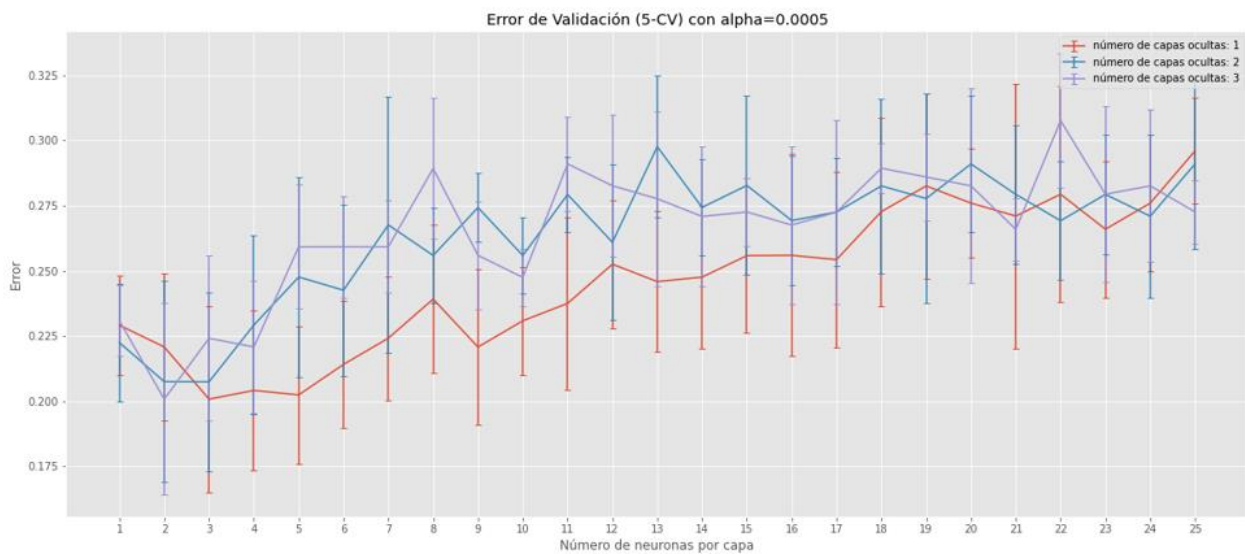
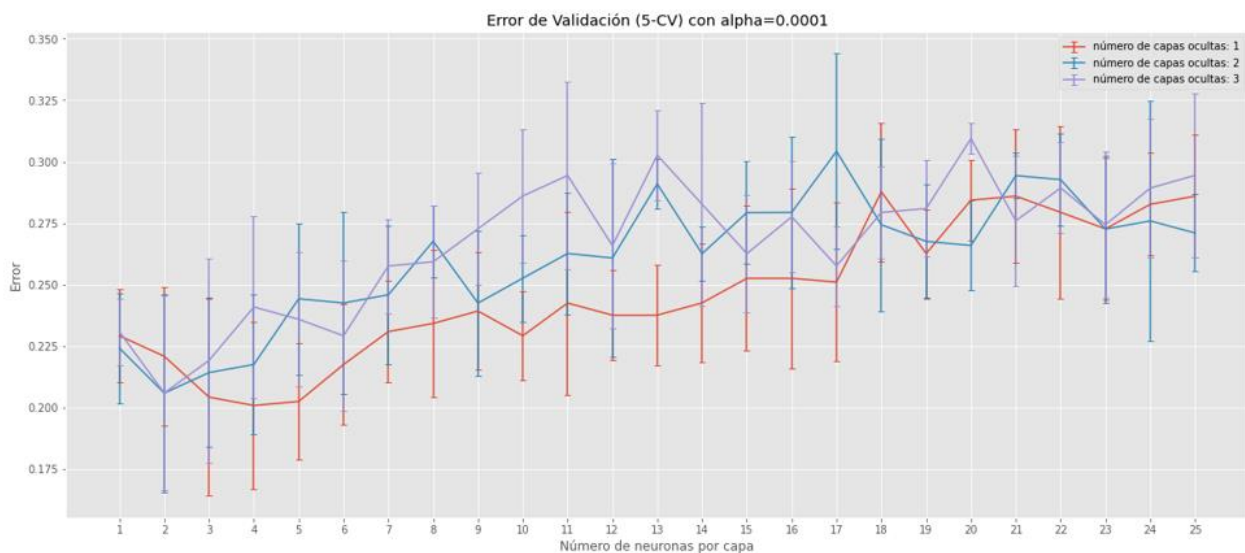
W2 = [[0.82555367 1.32555367 -0.11787827]]

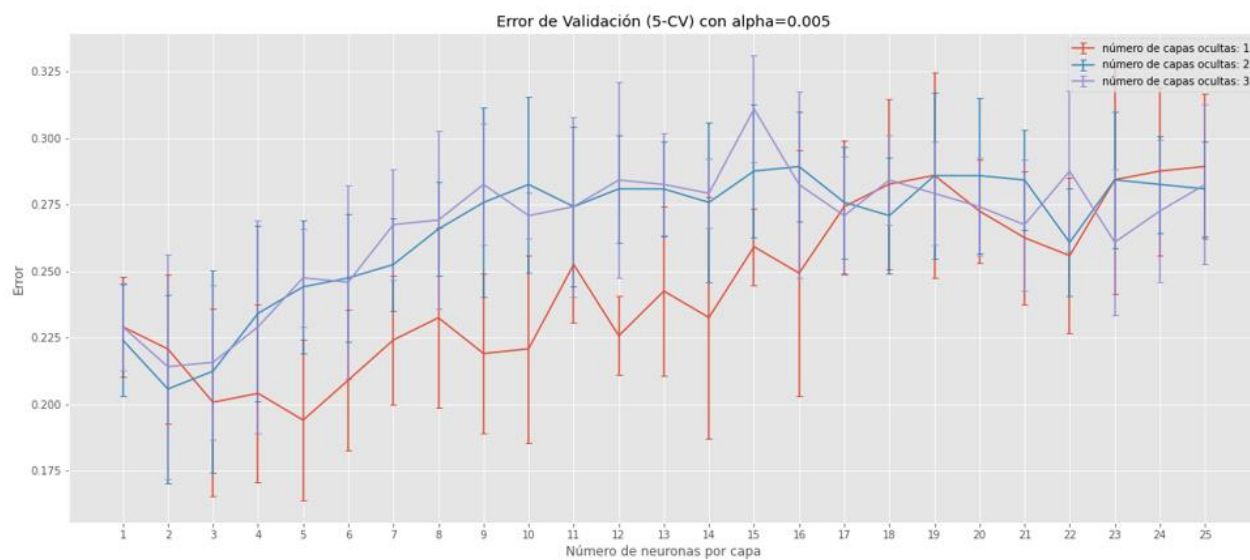
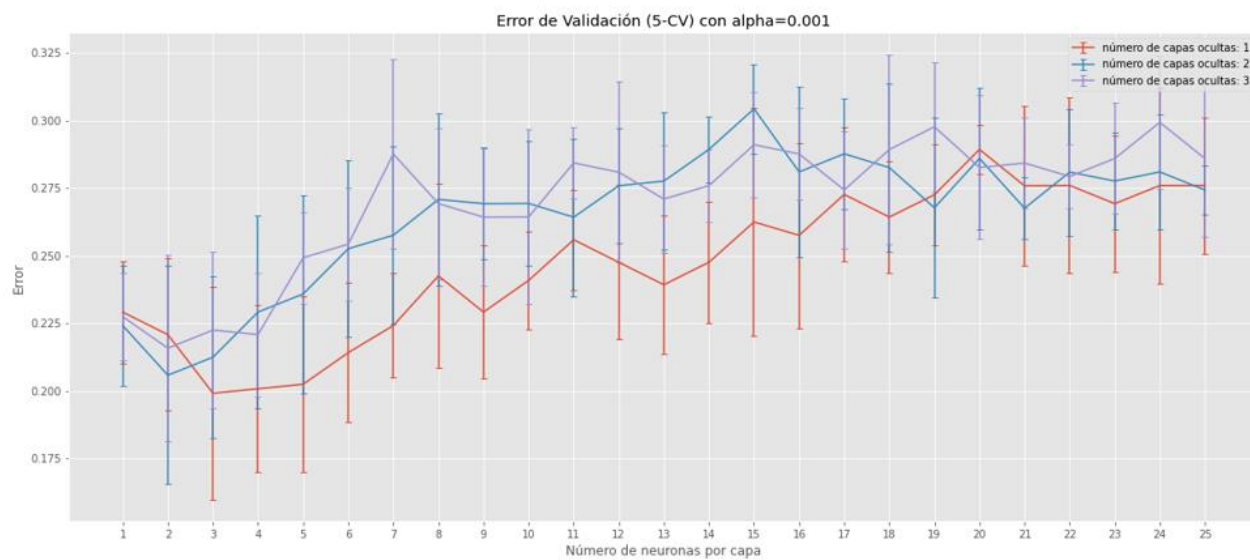
b2 = [[-0.64864062]]

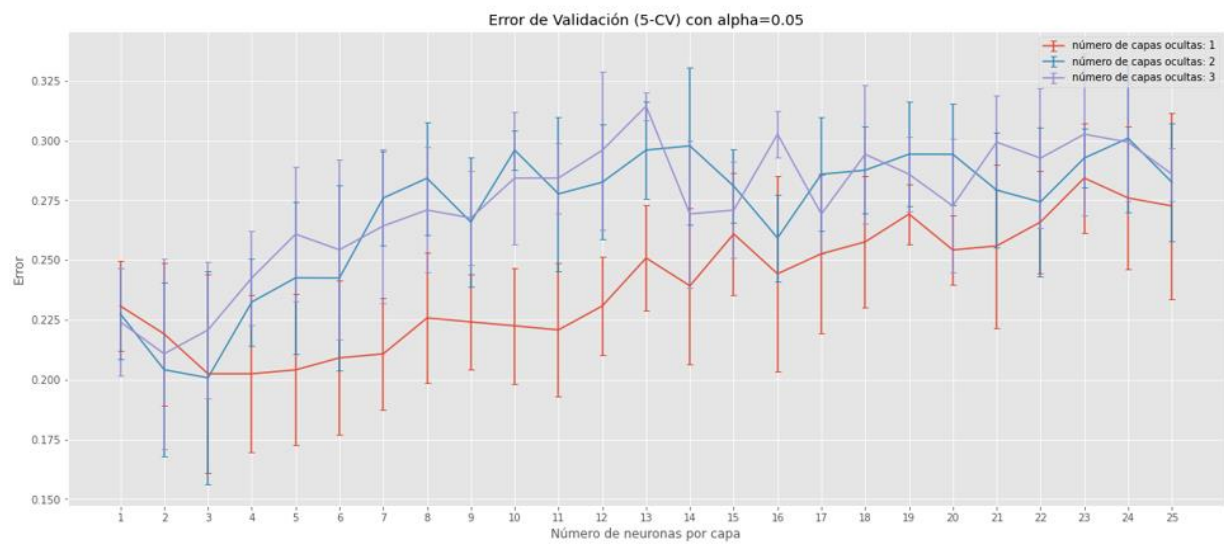
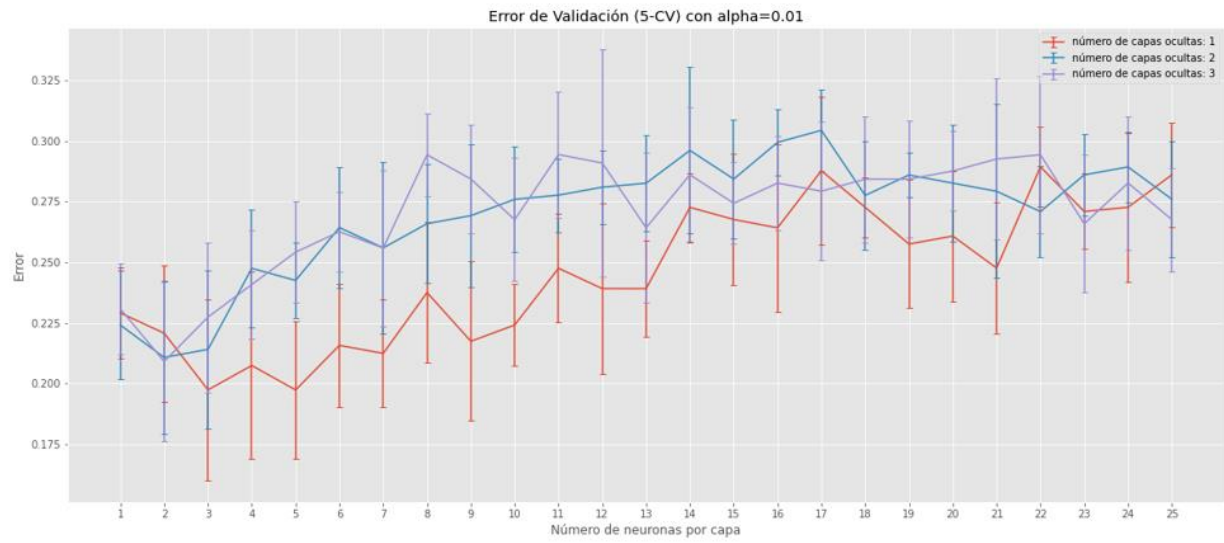
2 EJERCICIO 2

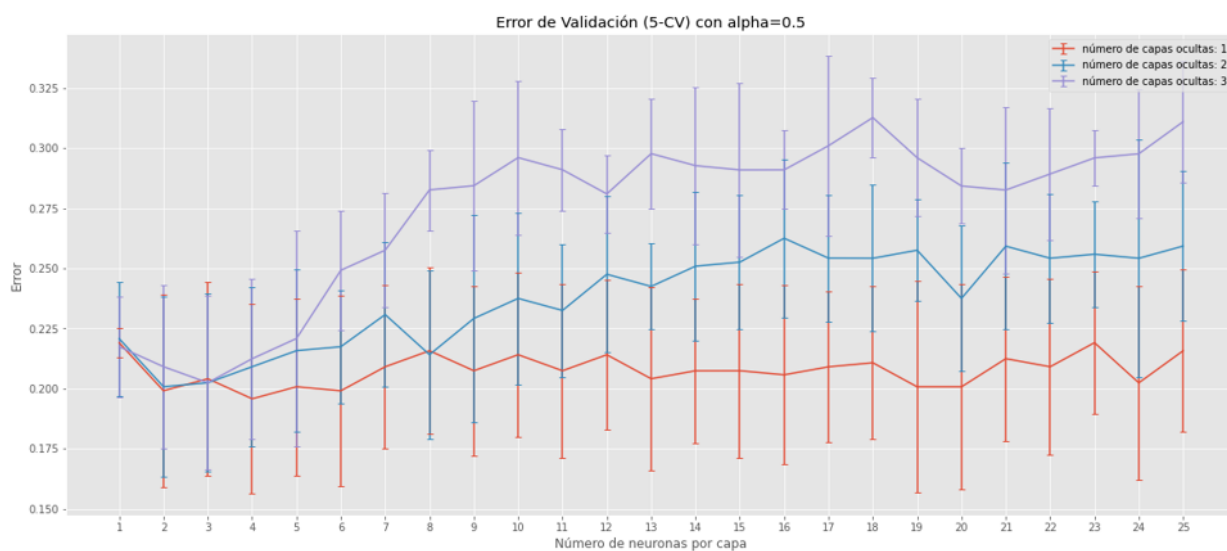
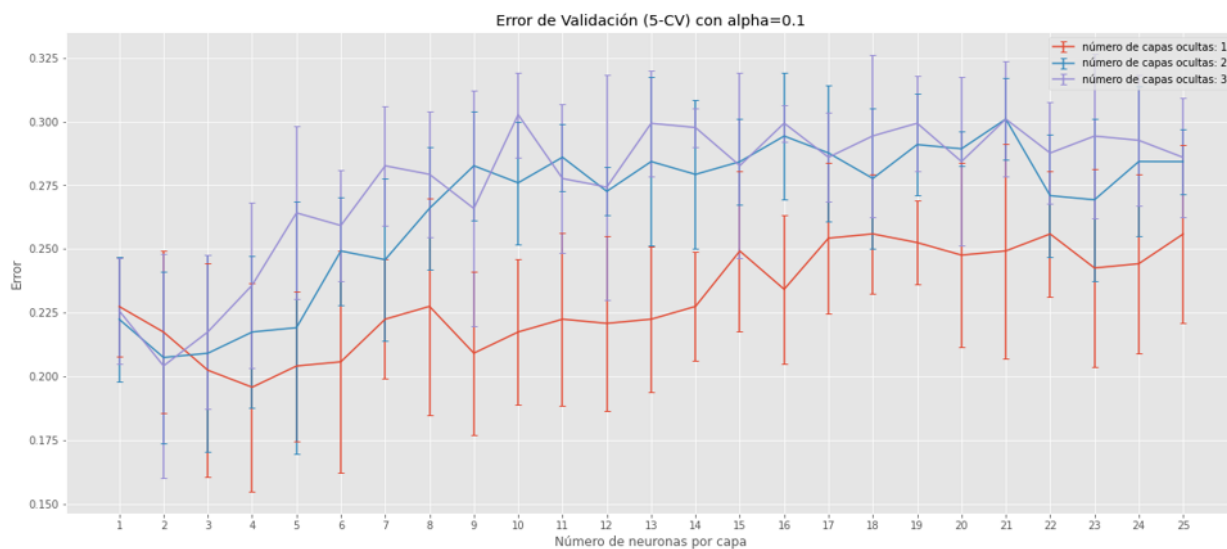
Realiza la experimentación con MLPClassifier usando los valores por defecto de los parámetros, excepto para `activation='tanh'` y `solver='lbfgs'`, y `max_iter=2000`.

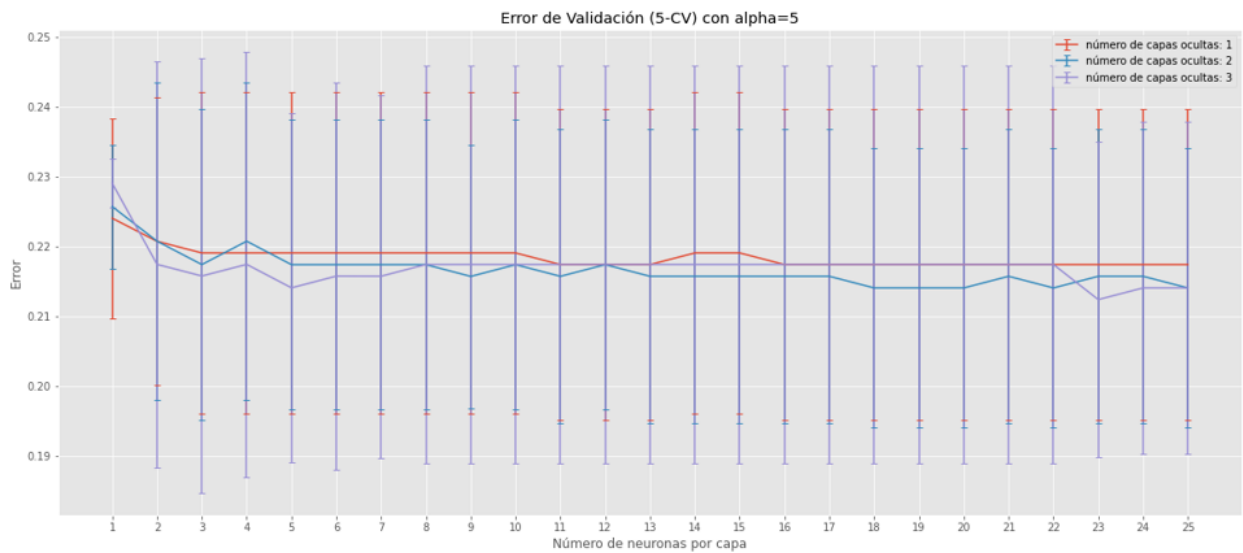
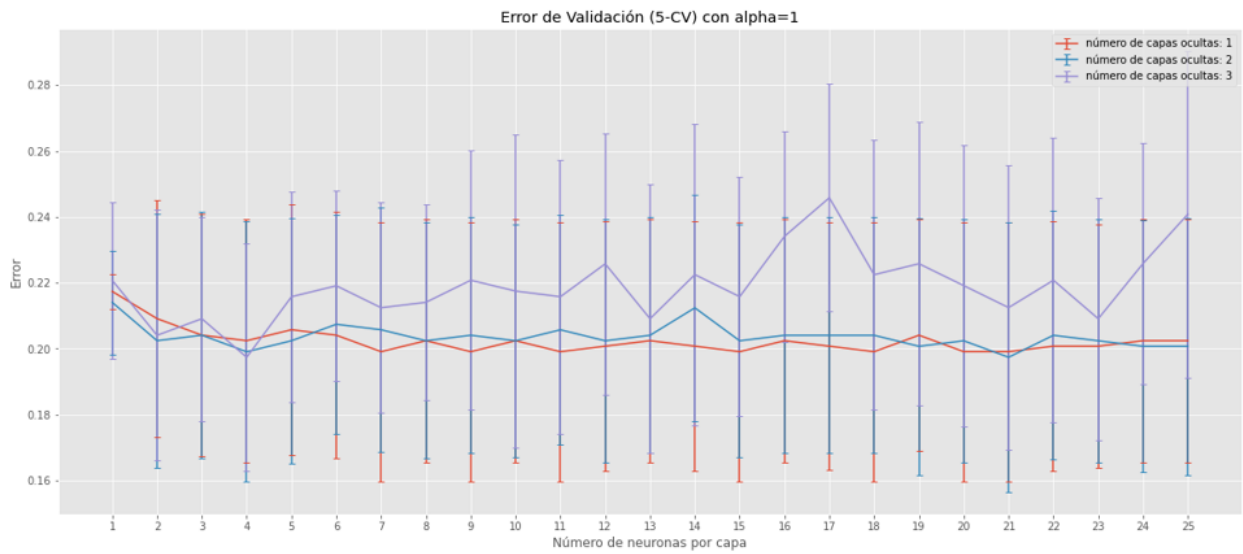
Muestra las gráficas del error de entrenamiento con validación cruzada (5-CV) frente al valor de los hiper-parámetros, y justifica la elección de los valores más apropiados. Para cada combinación de valor del número de capas ocultas y parámetro de regularización se debe generar una gráfica donde se represente en el eje horizontal el número de neuronas en la capa oculta.

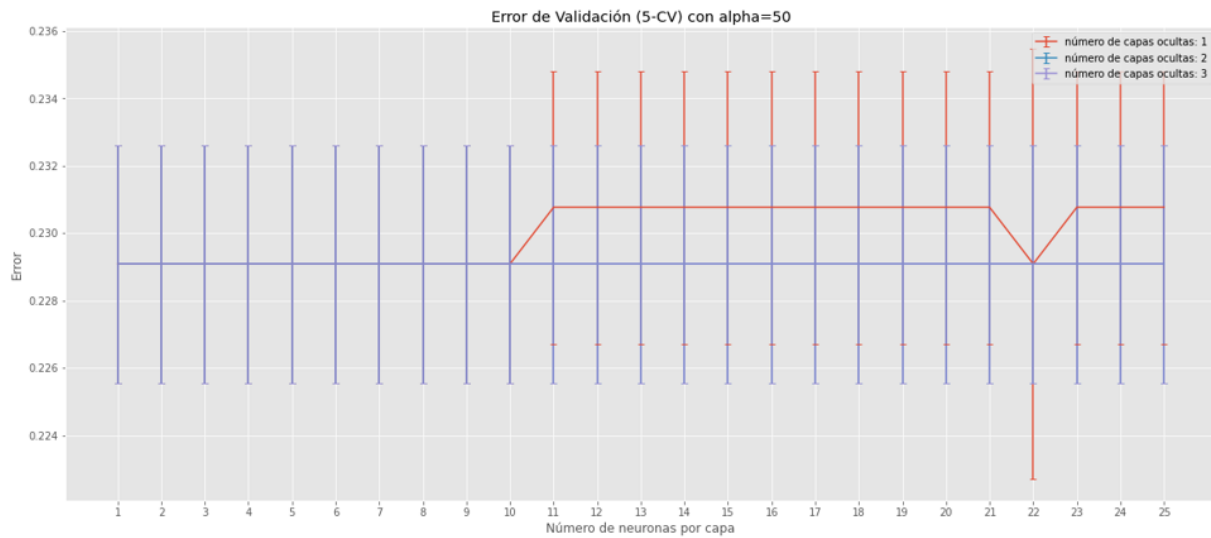
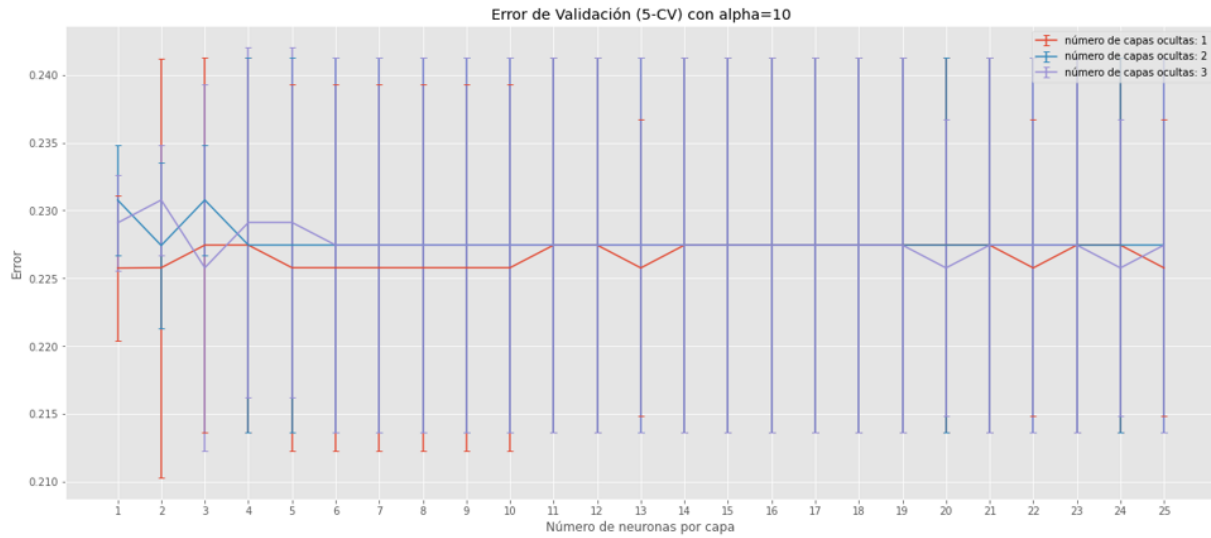








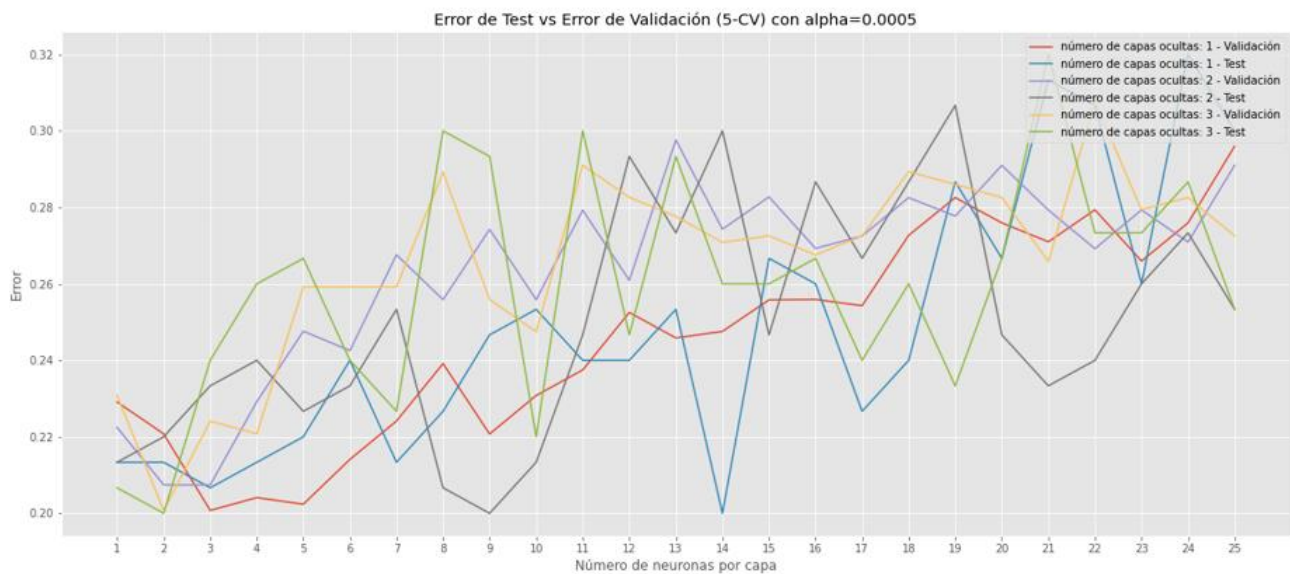
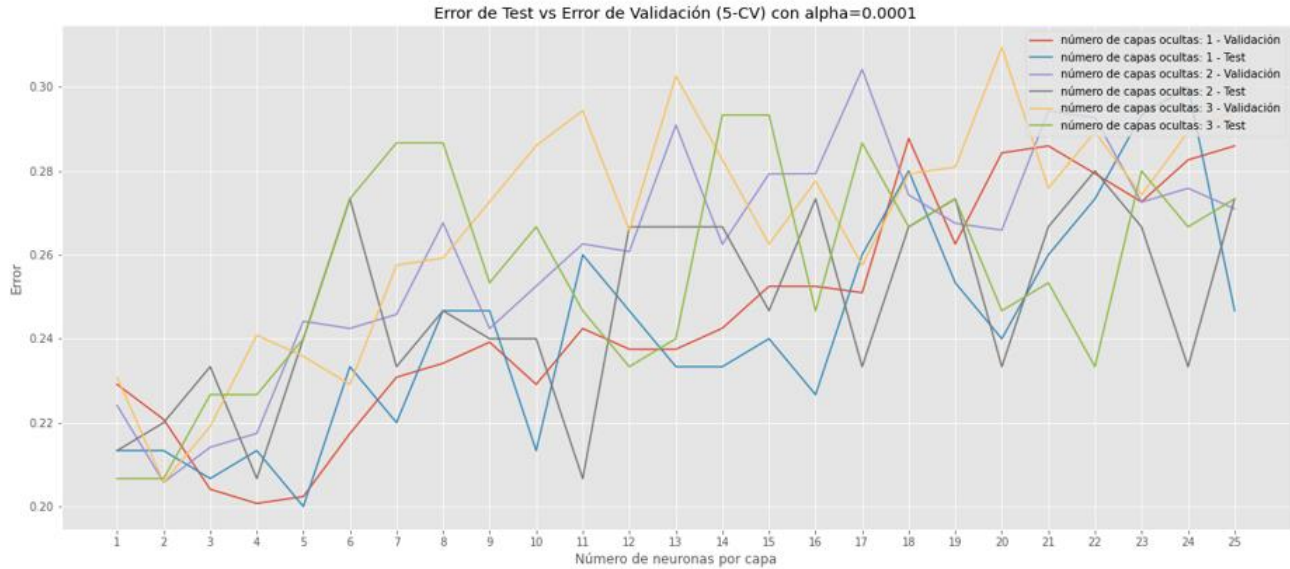


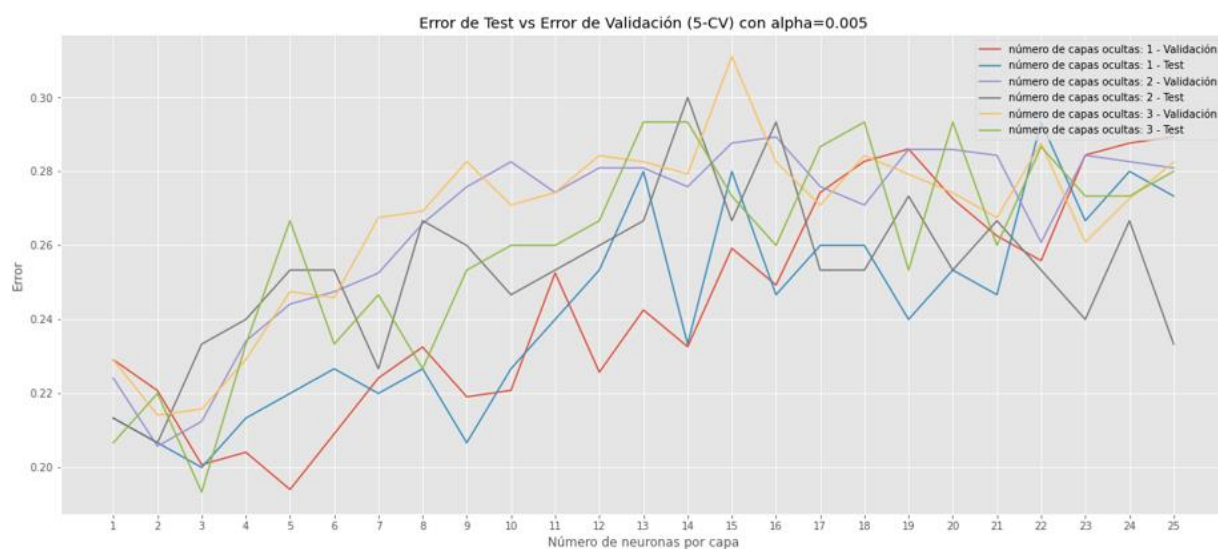
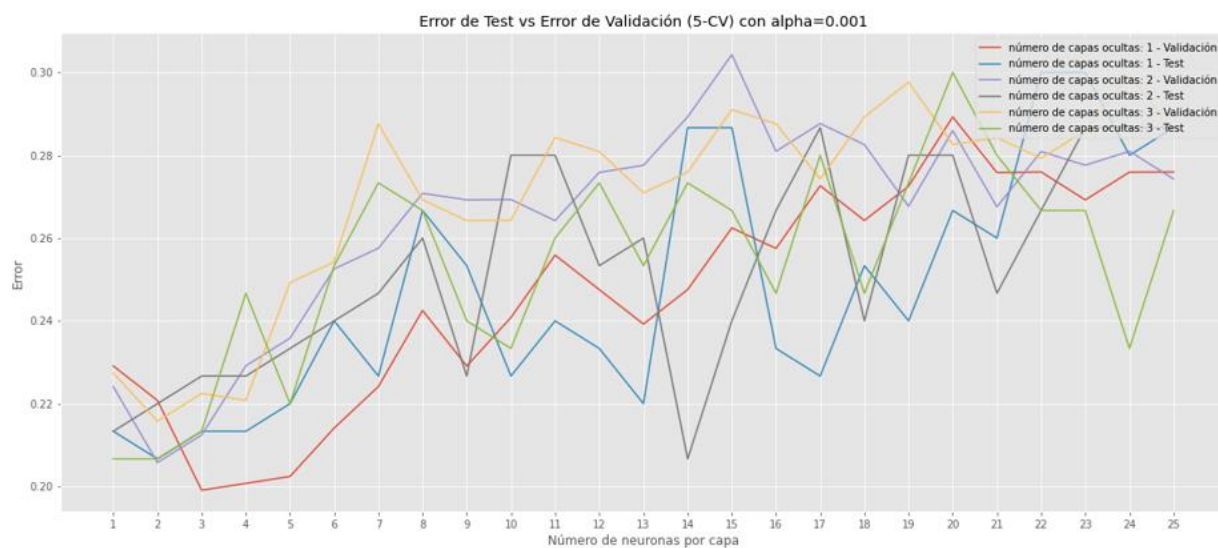


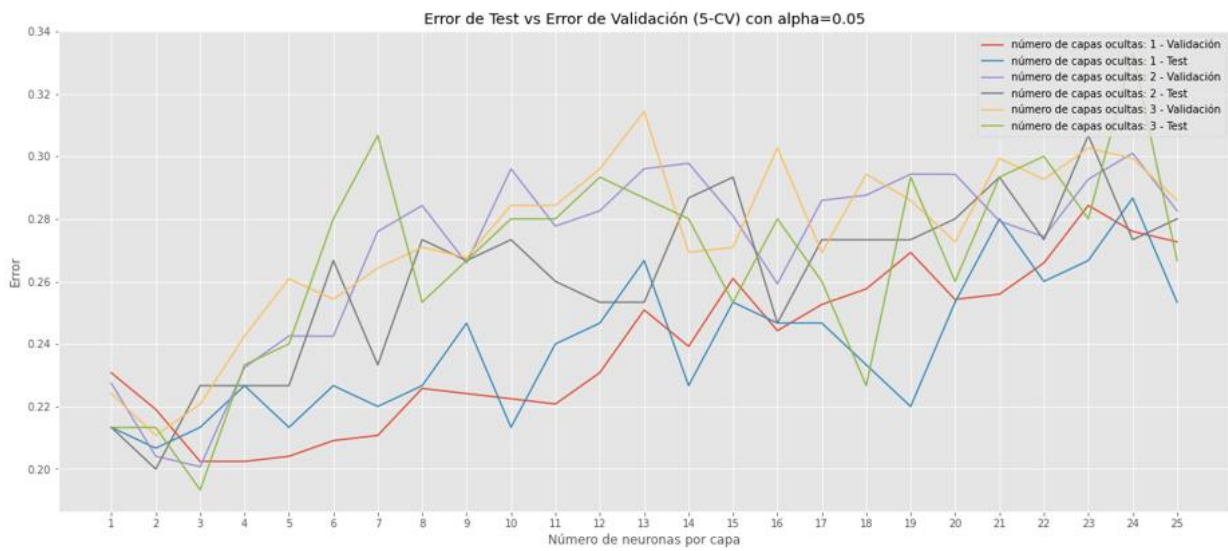
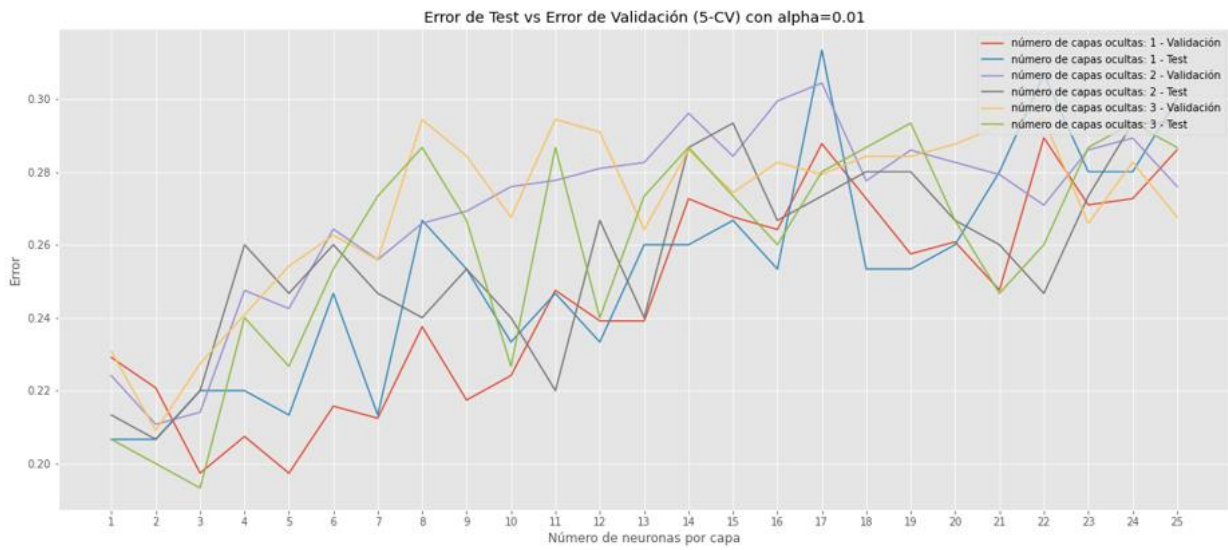
Menor error de validación cruzada, su desviación estándar y el valor del hiper-parámetro para el que se consigue:

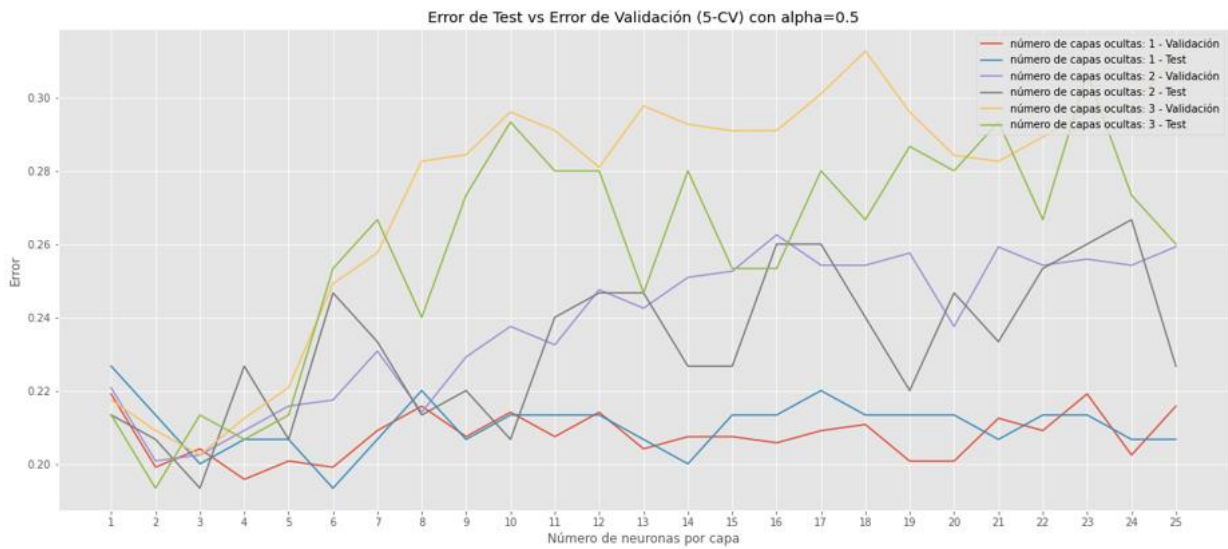
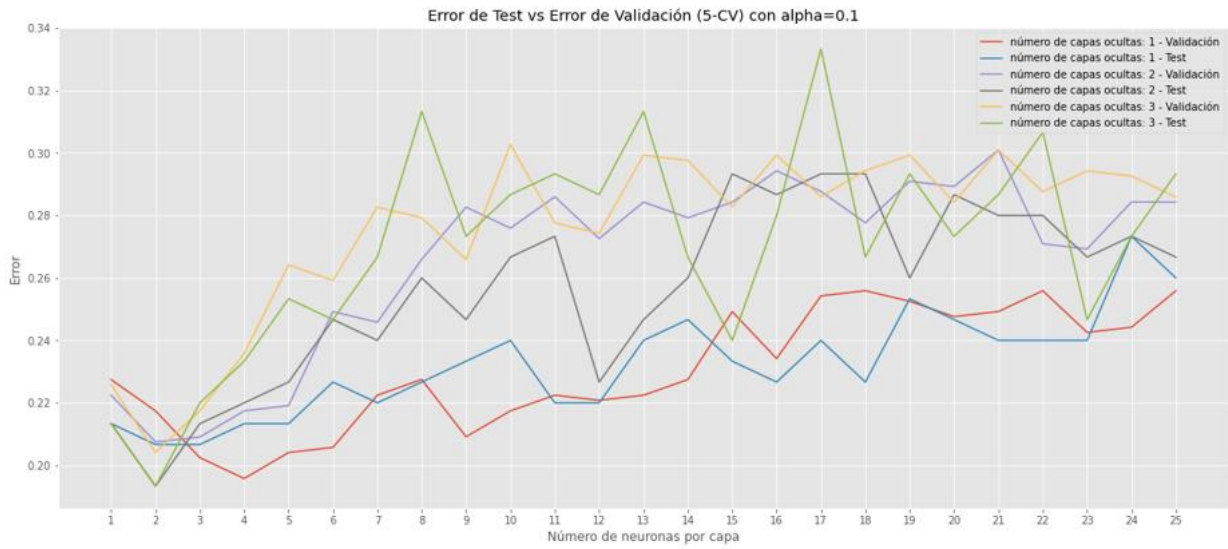
param_hidden_layer_sizes	param_alpha	n_hidden_layers	n_neurons_per_layer	mean_test_score	std_test_score	rank_test_score	
229	(5,)	0.005	1	5	0.194048	0.030102	1

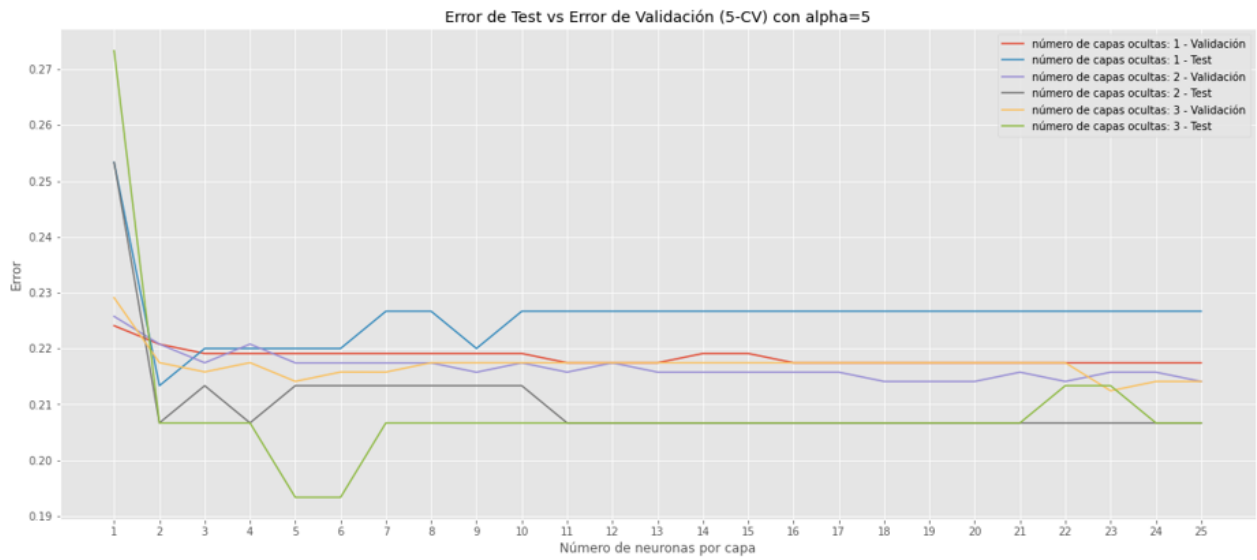
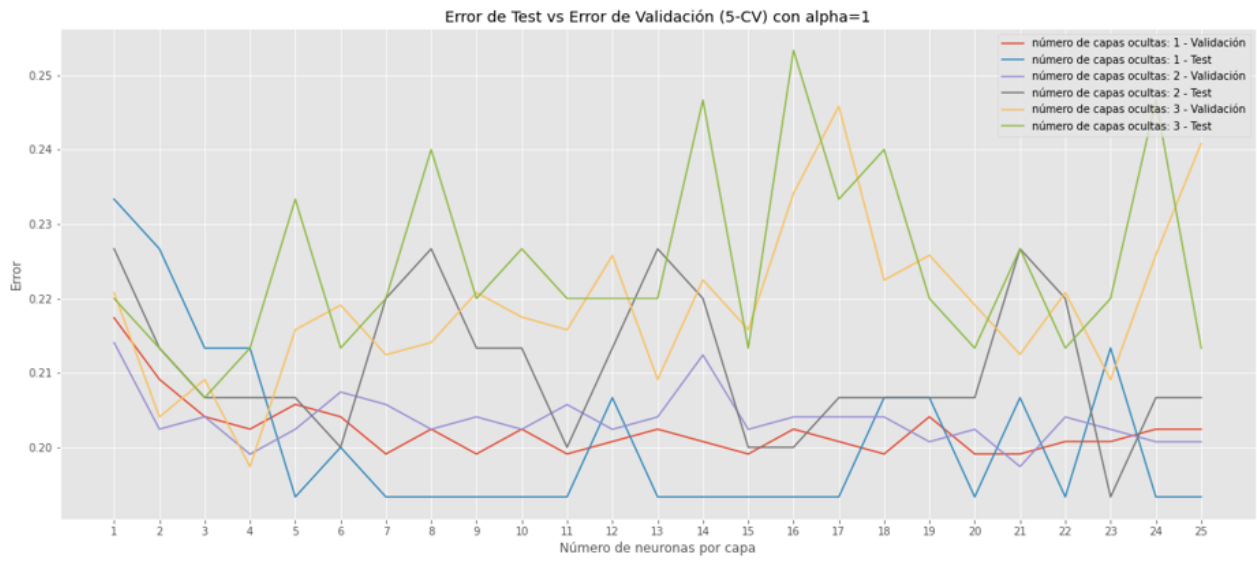
Muestra la gráfica del error de test frente al valor de los hiper-parámetros (siguiendo el mismo esquema que en el apartado anterior), y valora si la gráfica del error de entrenamiento con validación cruzada ha hecho una buena estimación del error de test.

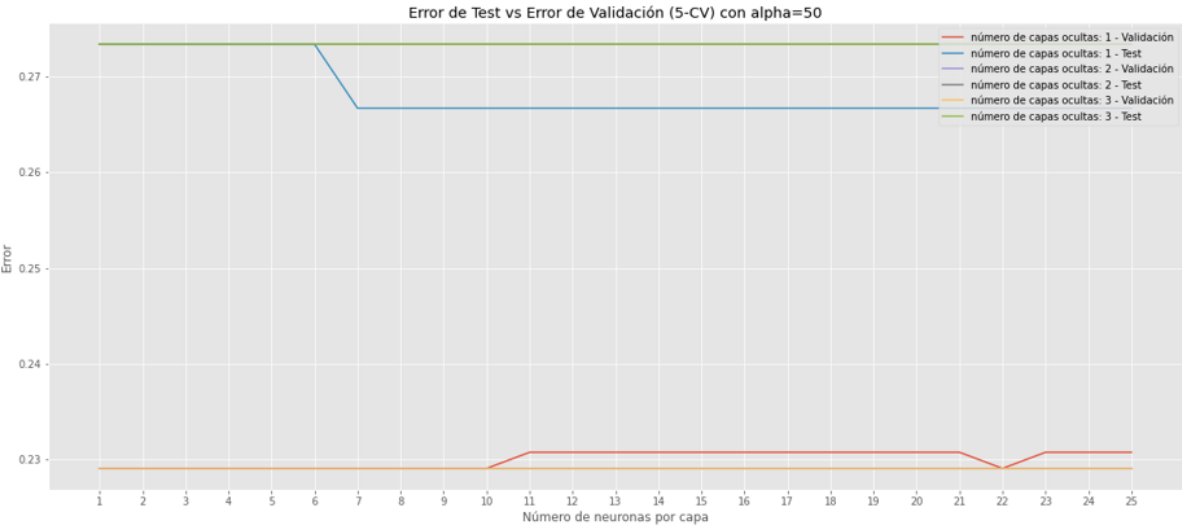
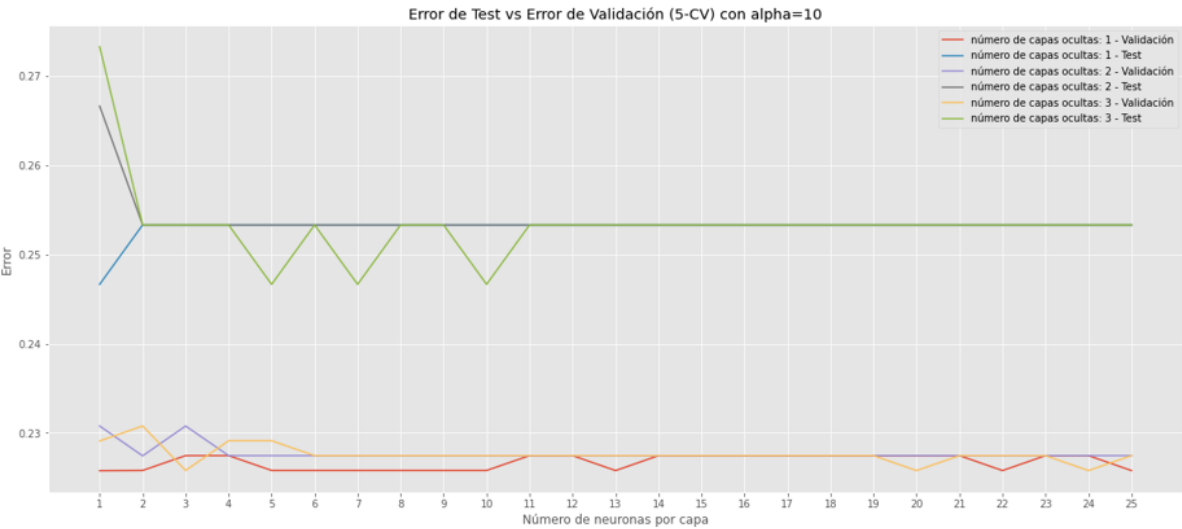












El menor error de test y el valor del hiper-parámetro para el que se consigue:

604	(5,)	1	1	5	0.193333	0.0
-----	------	---	---	---	----------	-----

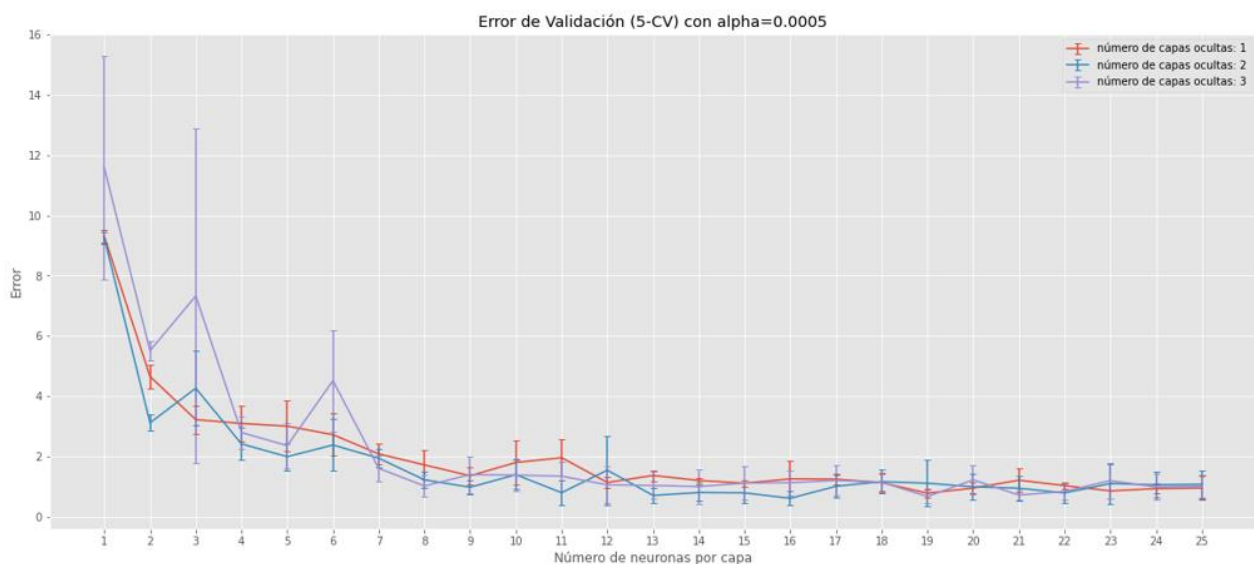
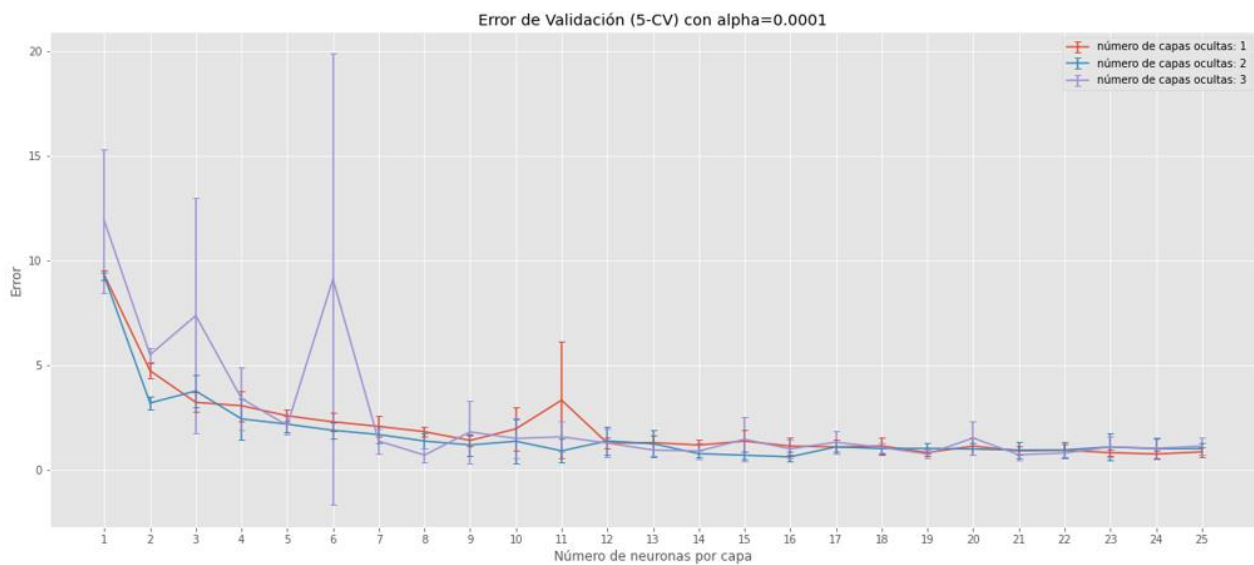
El error de test para el valor del hiper-parámetro seleccionado por la validación cruzada:

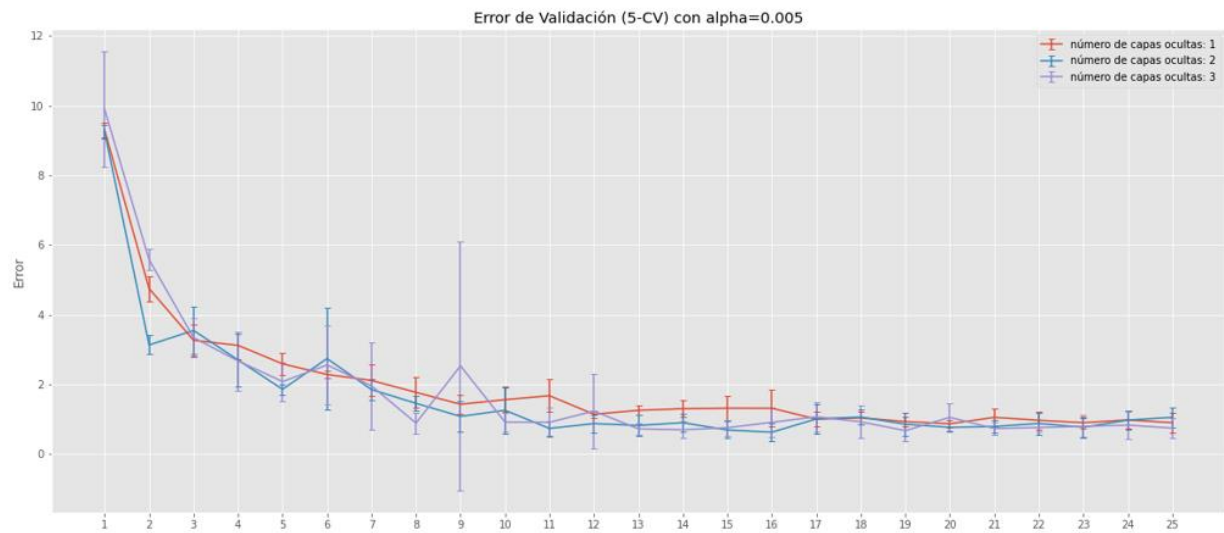
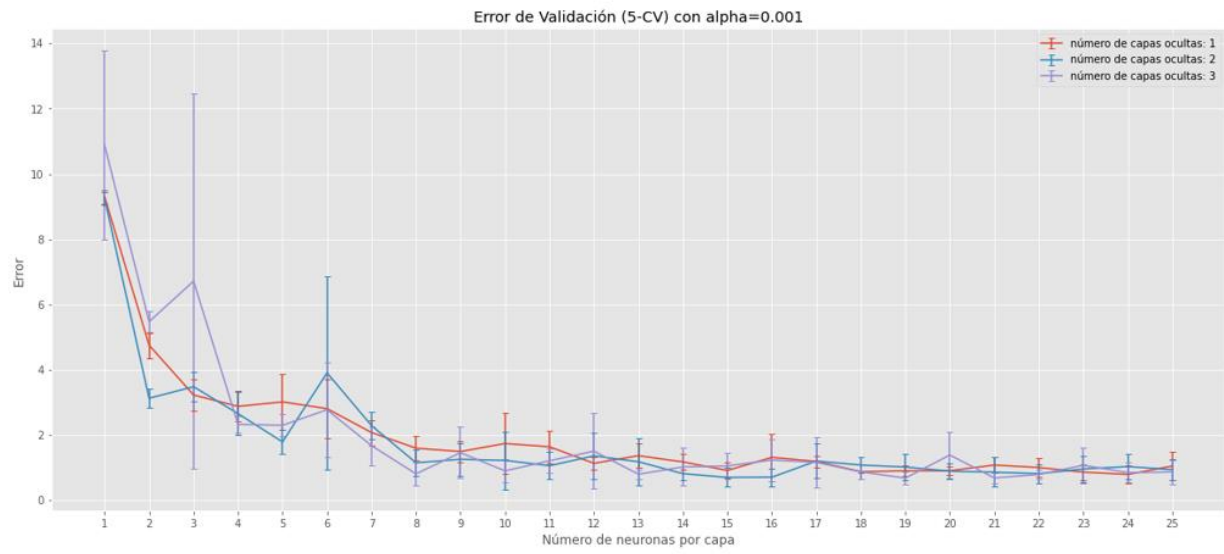
0.21999999999999997

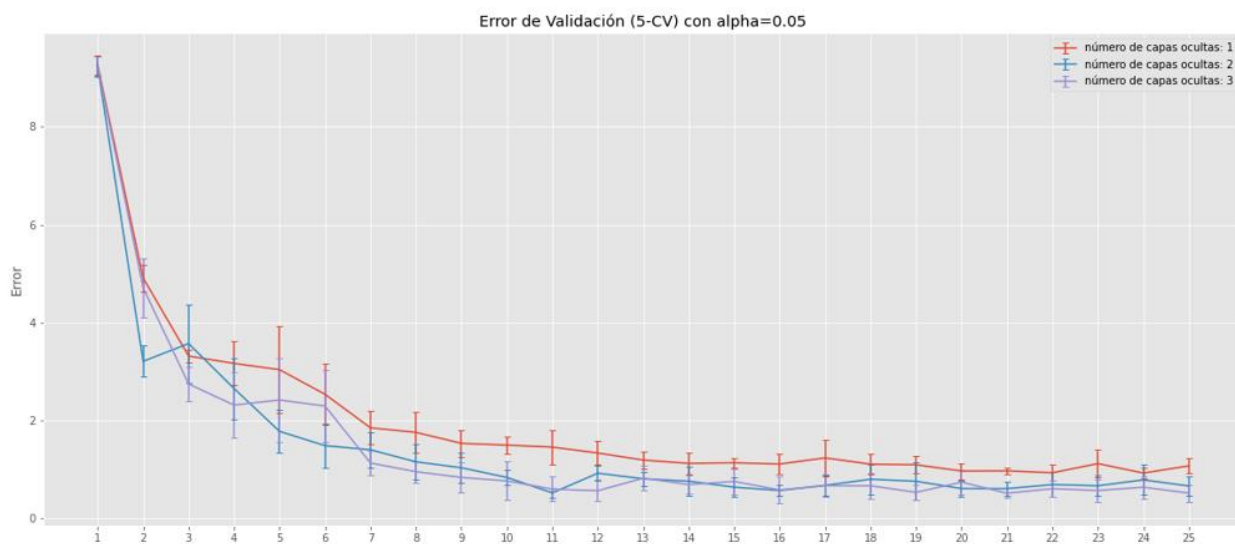
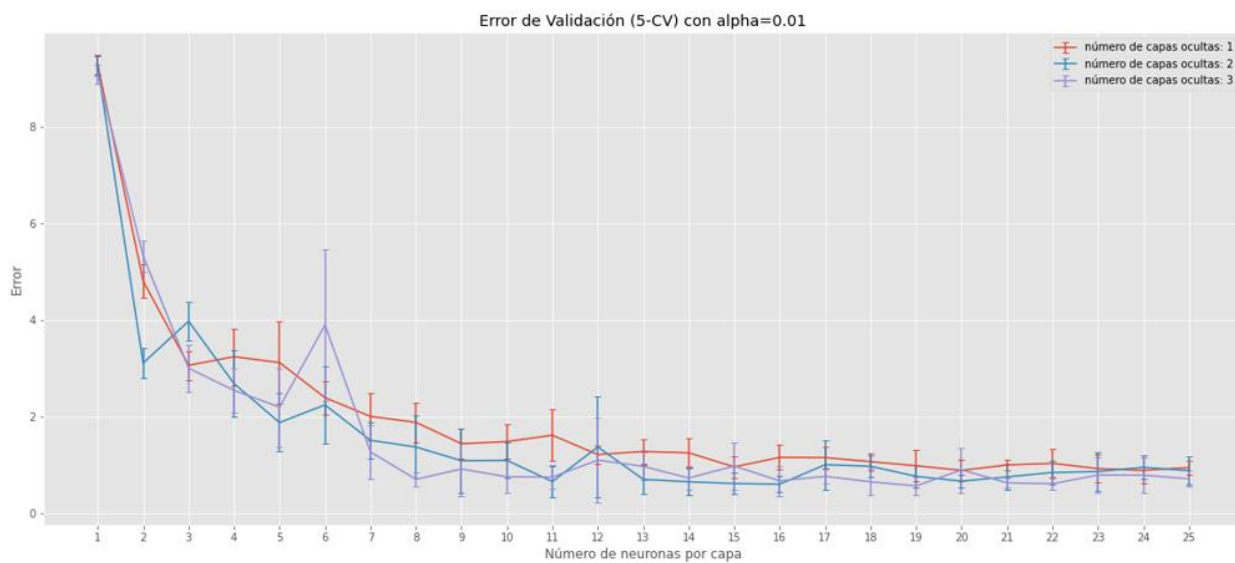
3 EJERCICIO 3

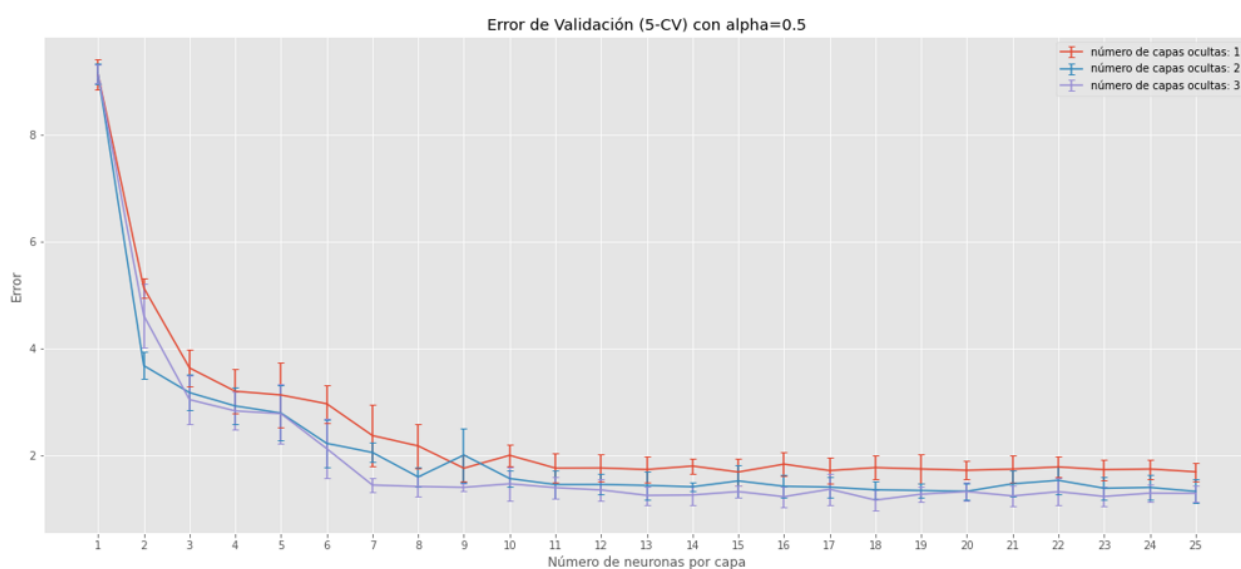
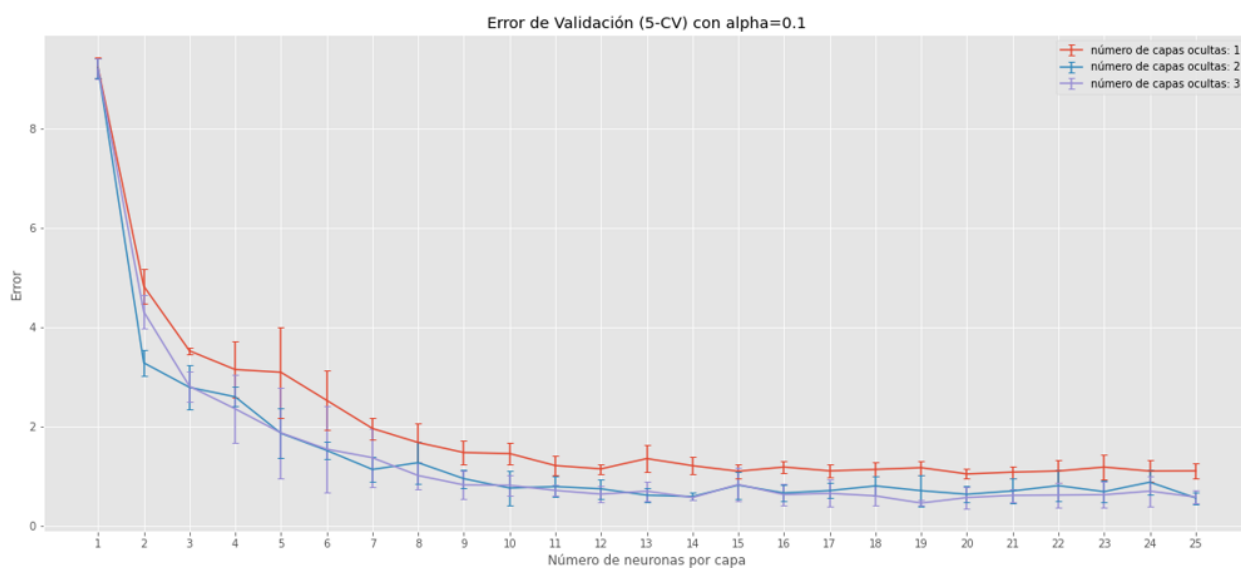
Realiza la experimentación con ***MLPRegressor*** usando los valores por defecto de los parámetros, excepto para **activation='tanh'** y **solver='lbfgs'**, y **max_iter=2000**.

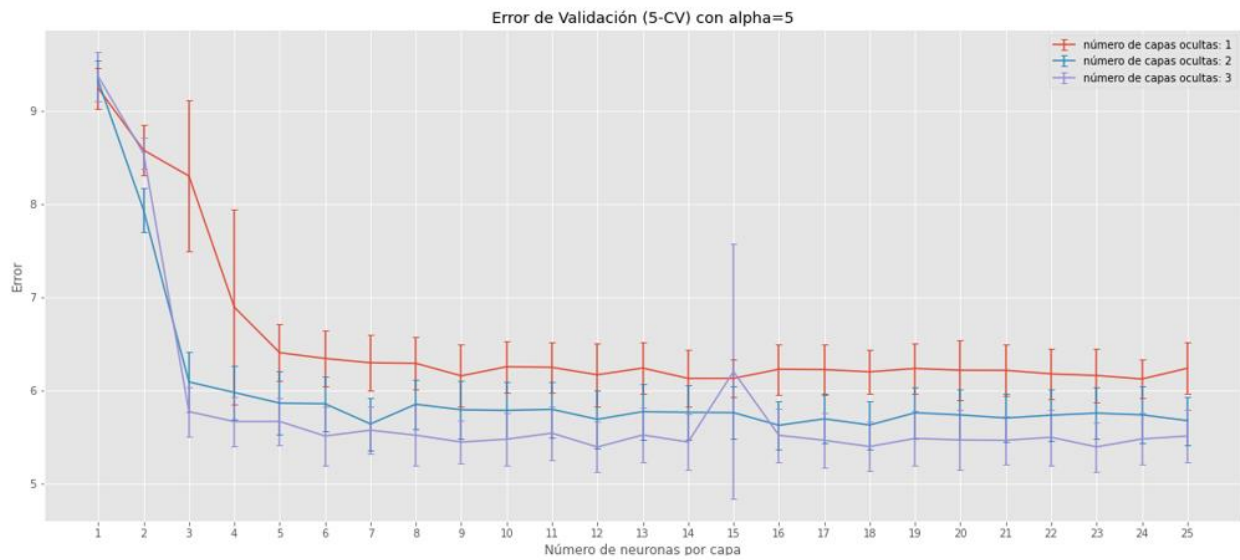
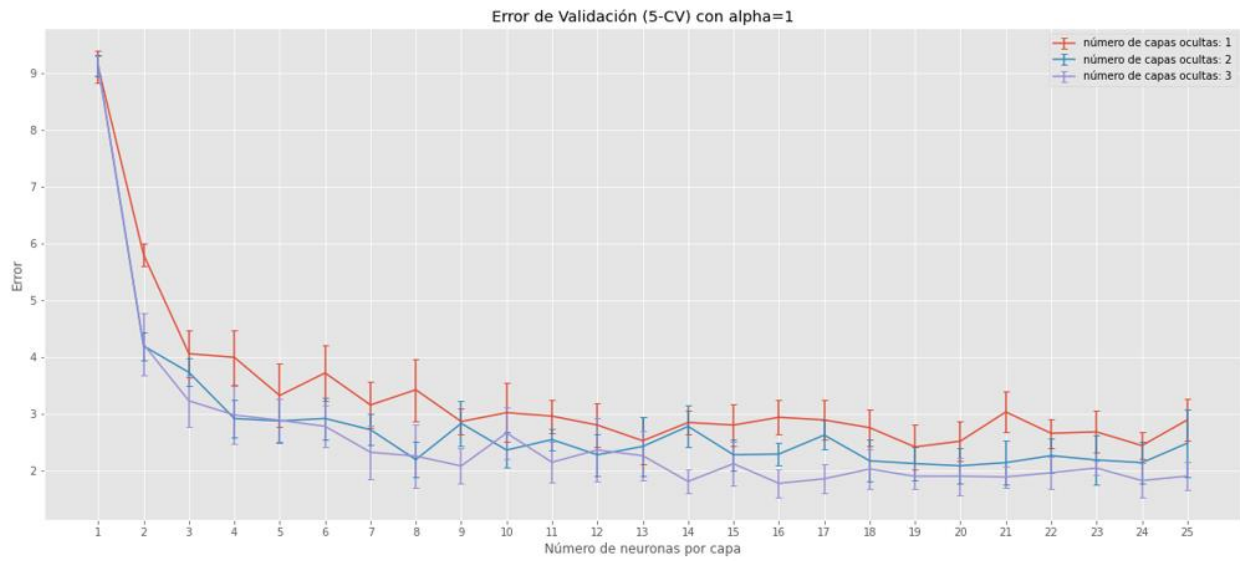
Muestra las gráficas del error de entrenamiento con validación cruzada (5-CV) frente al valor de los hiper-parámetros, y justifica la elección de los valores más apropiados. Para cada combinación de valor del número de capas ocultas y parámetro de regularización se debe generar una gráfica donde se represente en el eje horizontal el número de neuronas en la capa oculta.

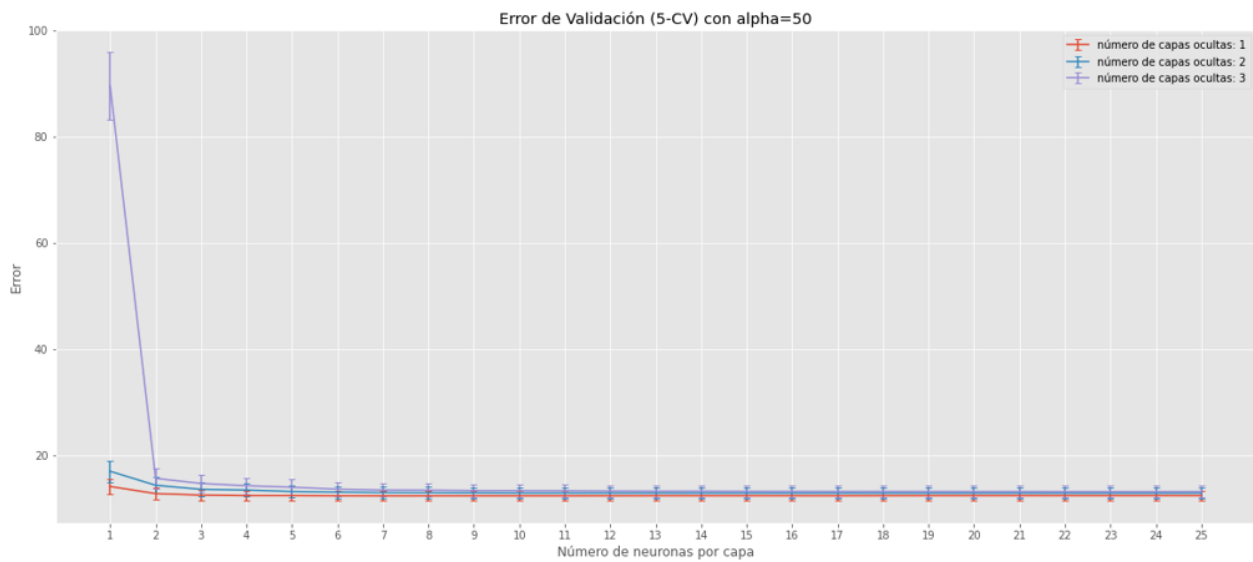
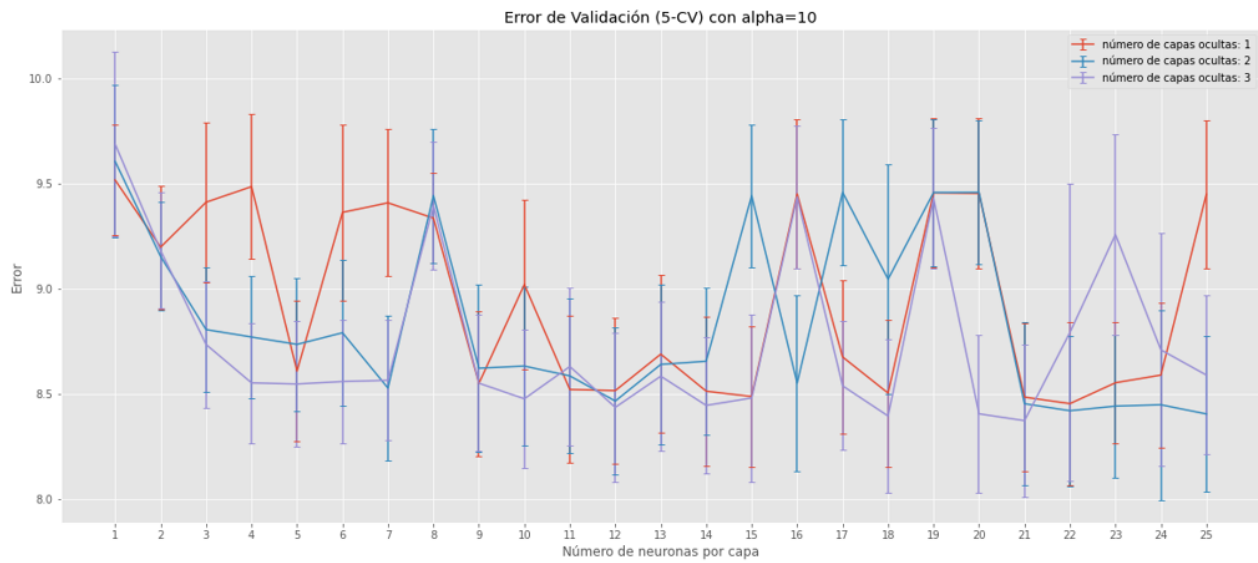








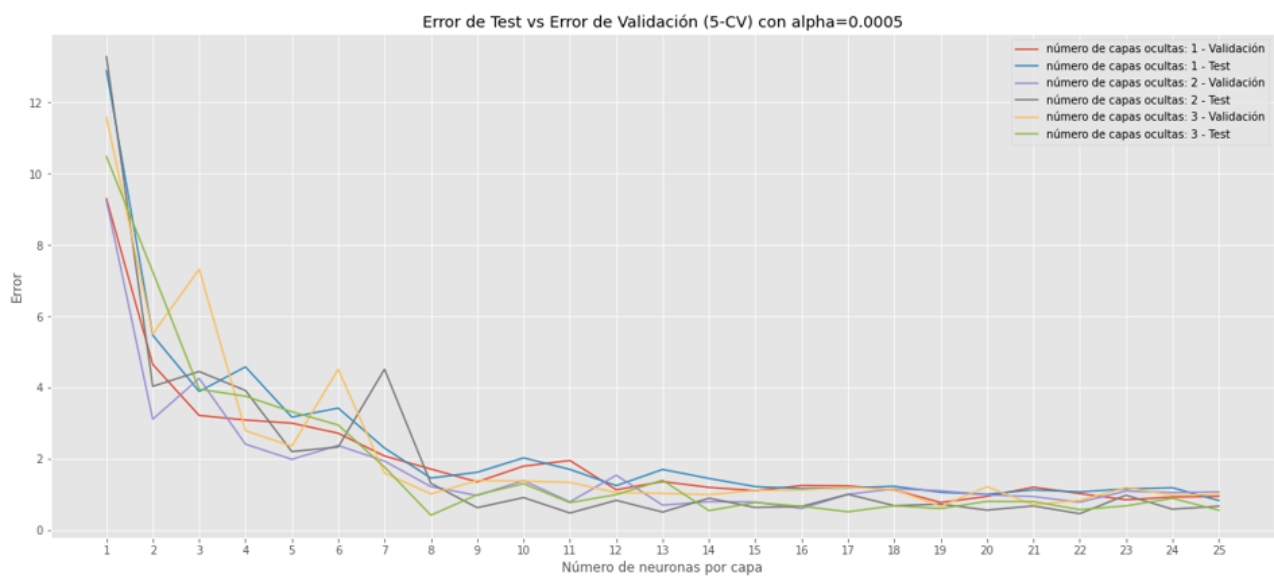
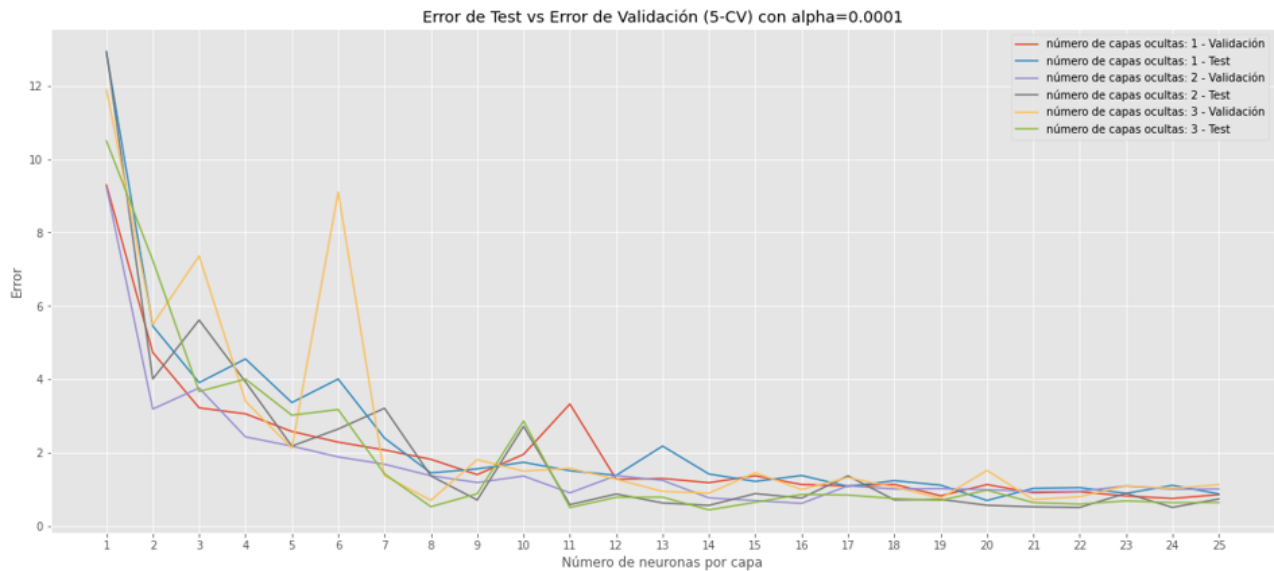


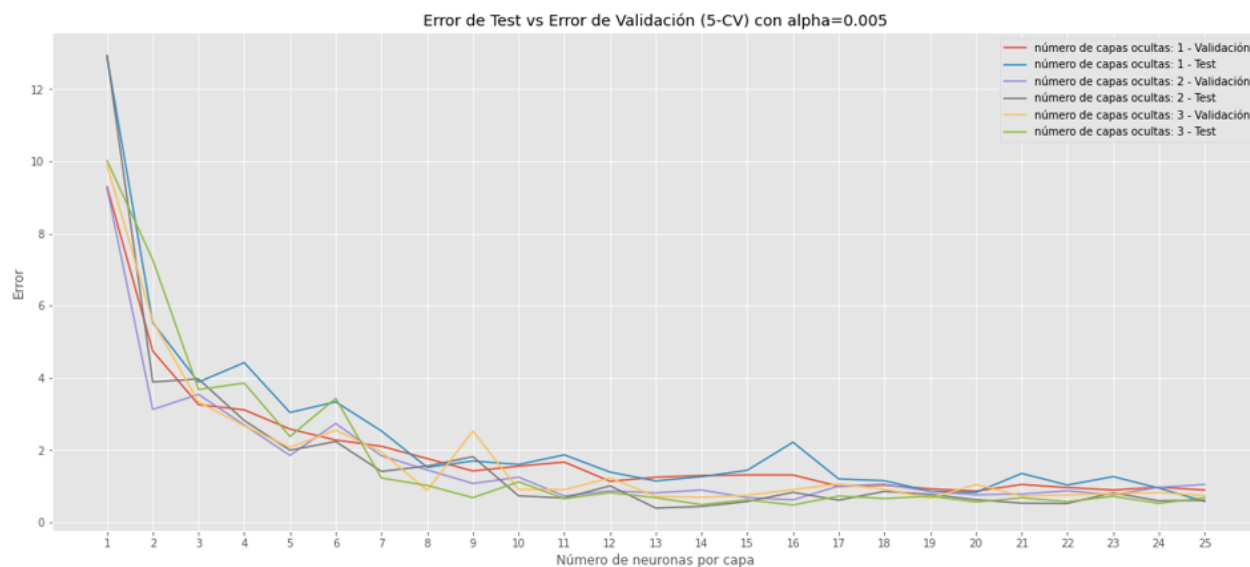
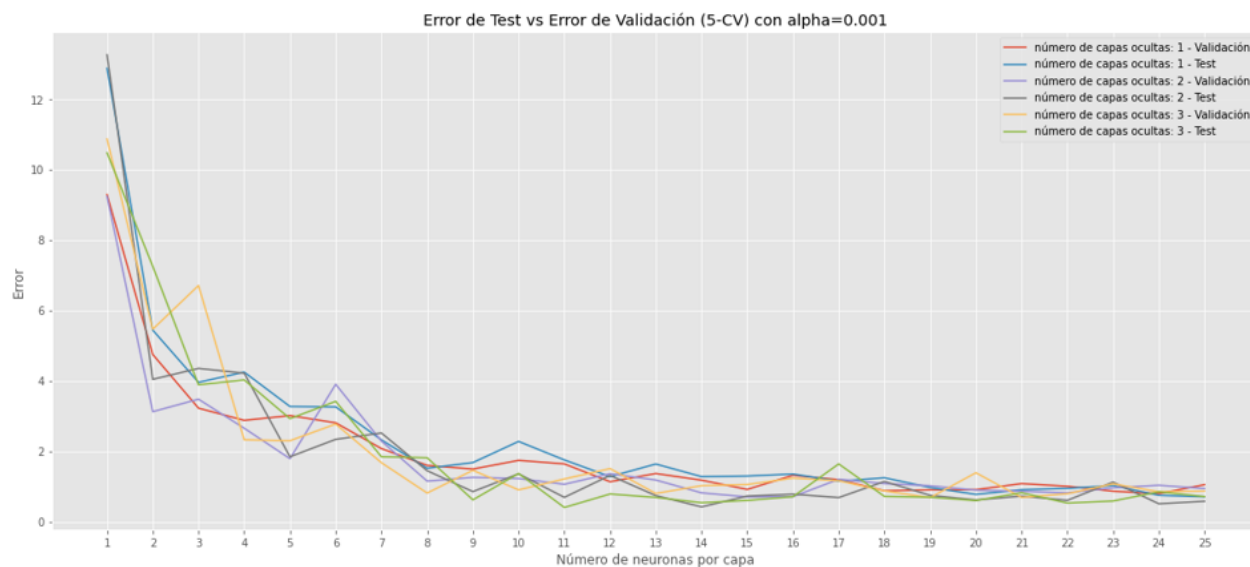


Menor error de validación cruzada, su desviación estándar y el valor del hiper-parámetro para el que se consigue:

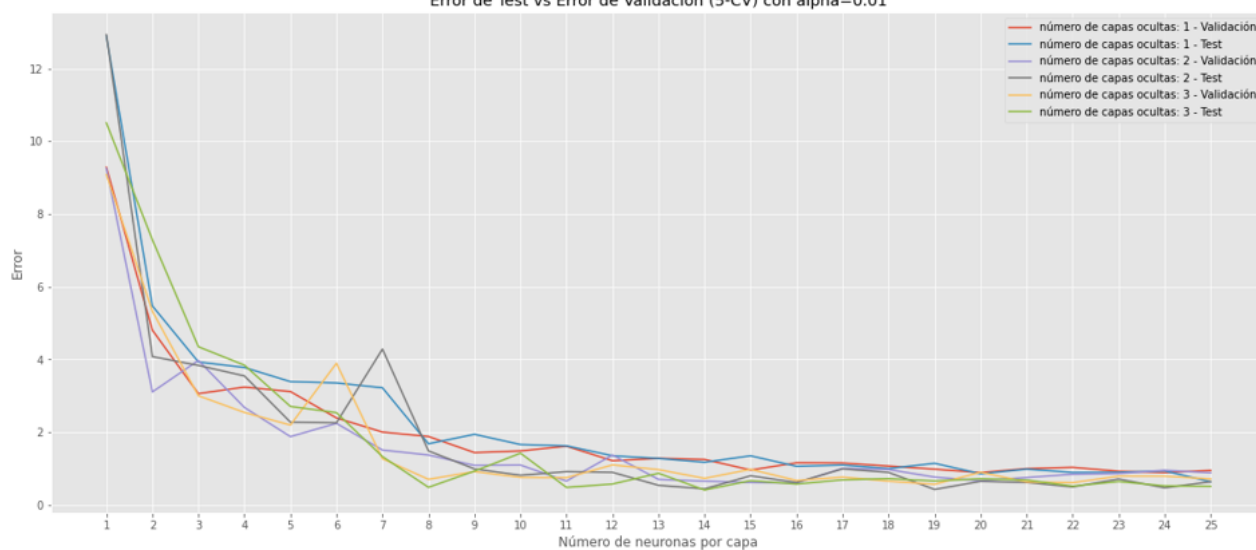
	param_hidden_layer_sizes	param_alpha	n_hidden_layers	n_neurons_per_layer	mean_test_score	std_test_score	rank_test_score
518	(19, 19, 19)	0.1	3	19	0.4667	0.056198	1

Muestra la gráfica del error de test frente al valor de los hiper-parámetros (siguiendo el mismo esquema que en el apartado anterior), y valora si la gráfica del error de entrenamiento con validación cruzada ha hecho una buena estimación del error de test.

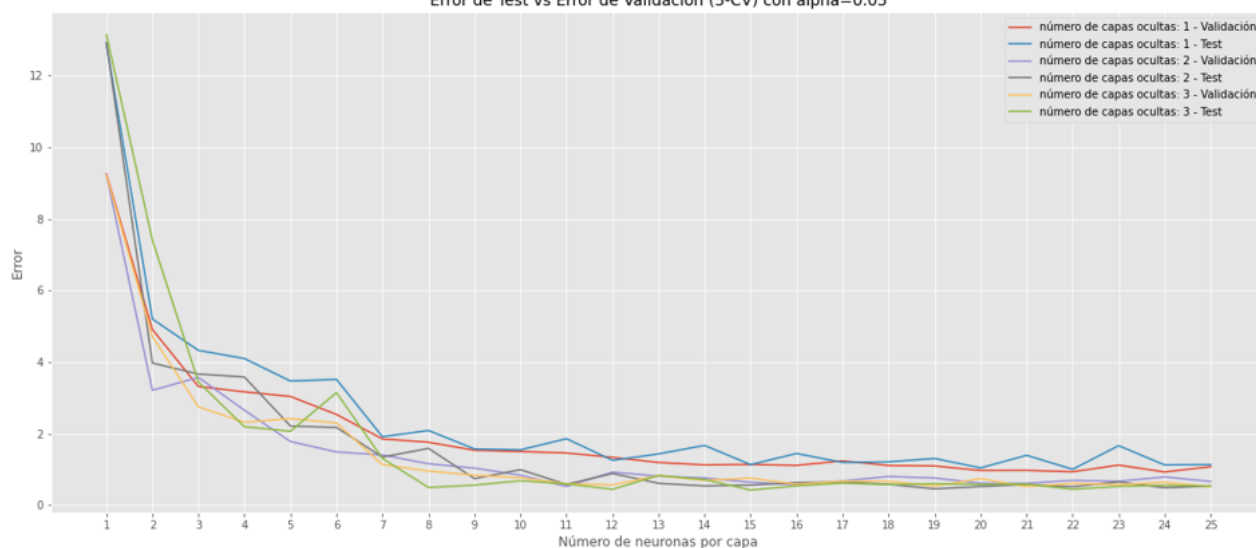


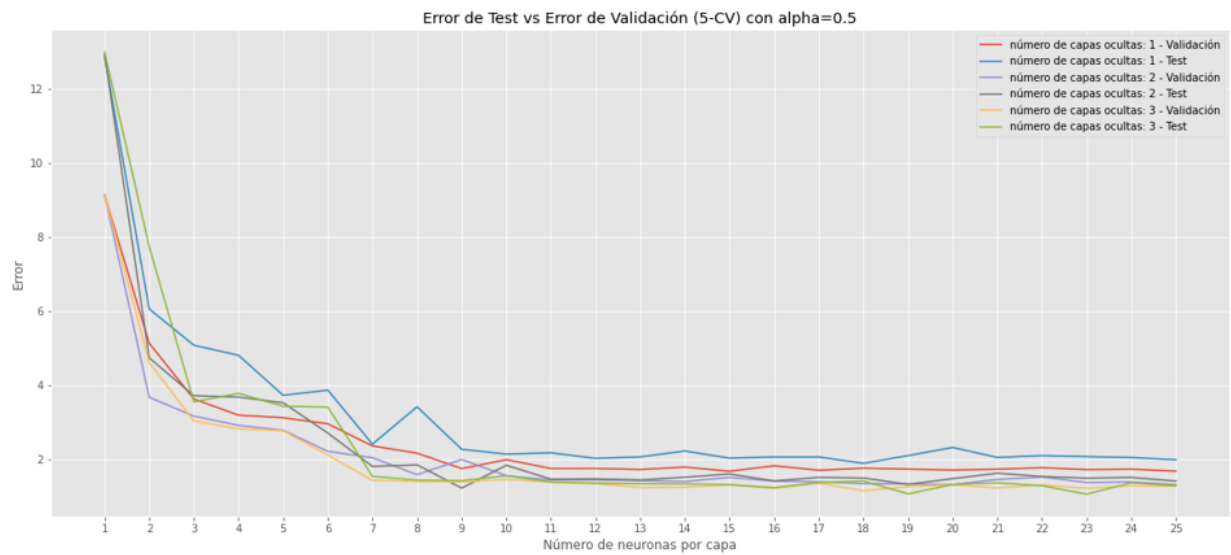
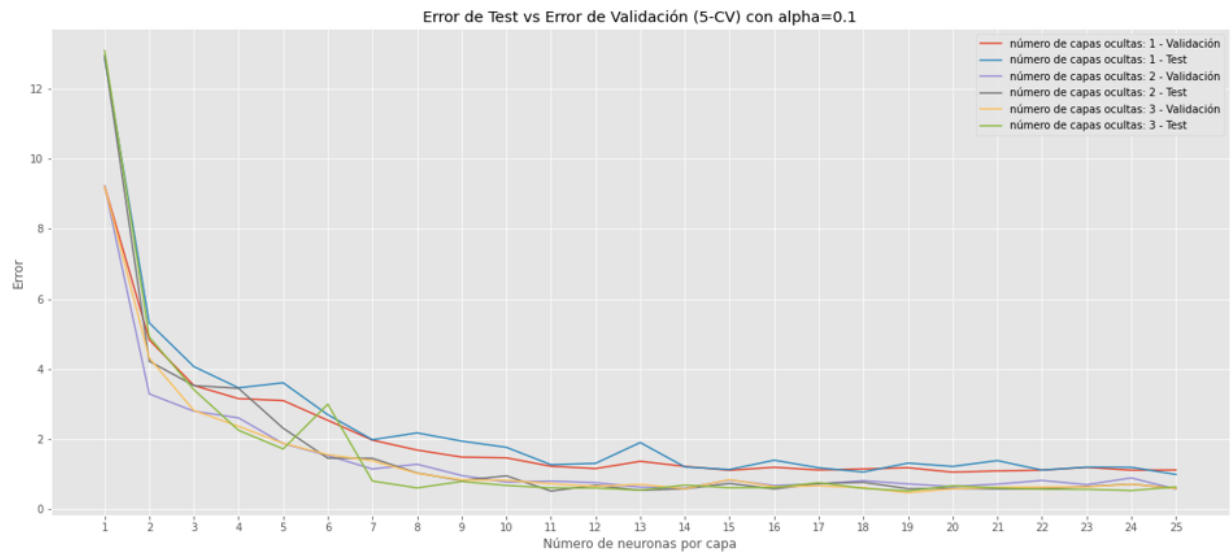


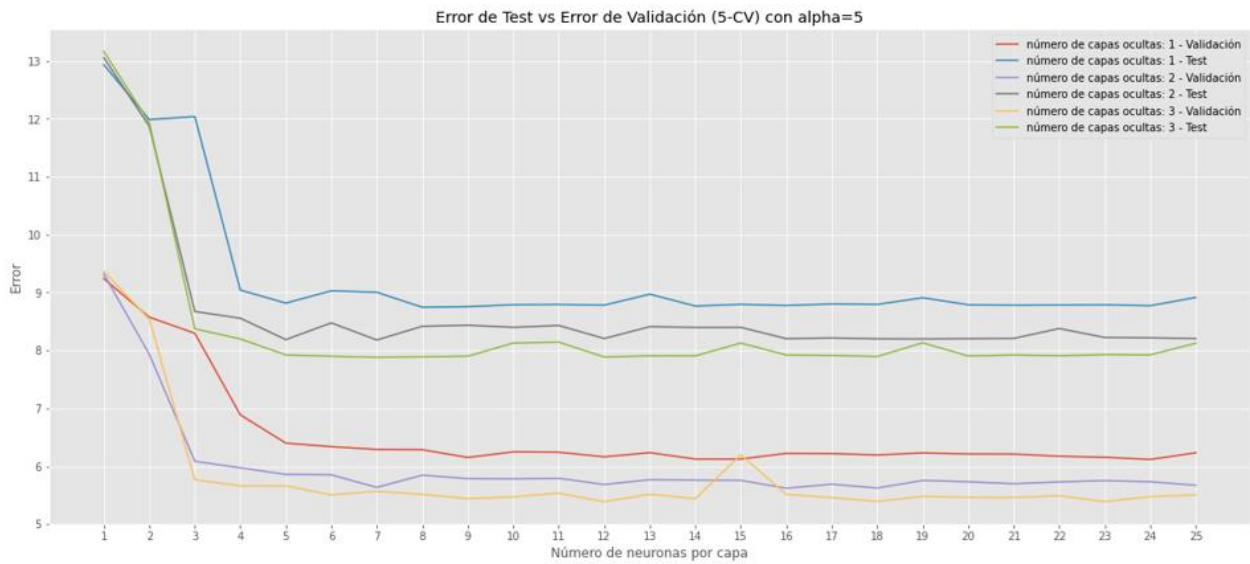
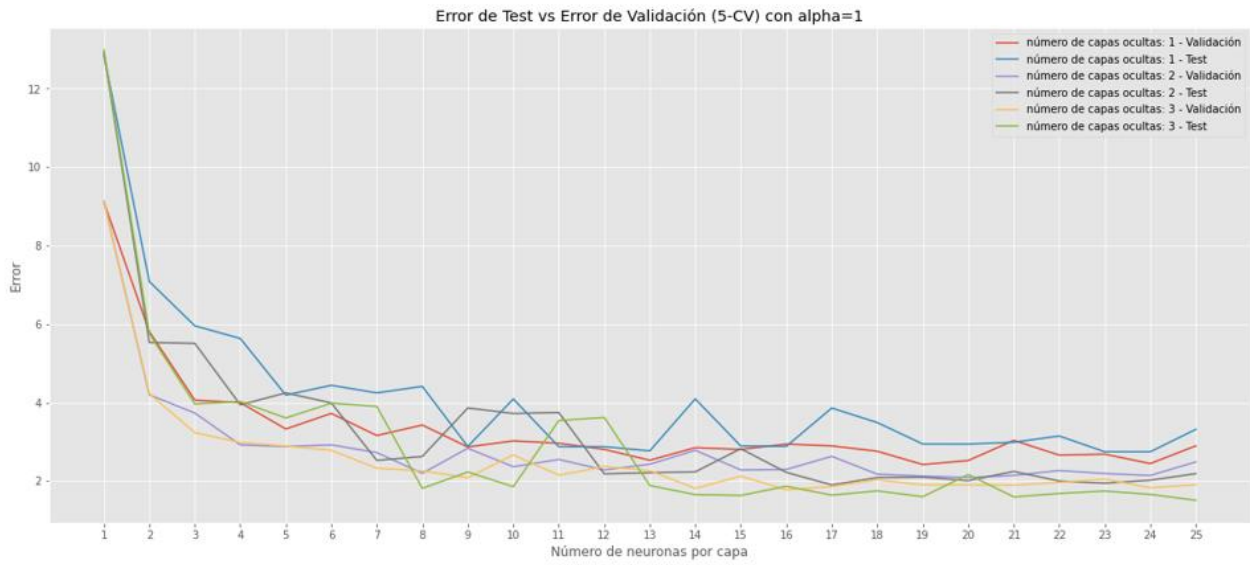
Error de Test vs Error de Validación (5-CV) con $\alpha=0.01$

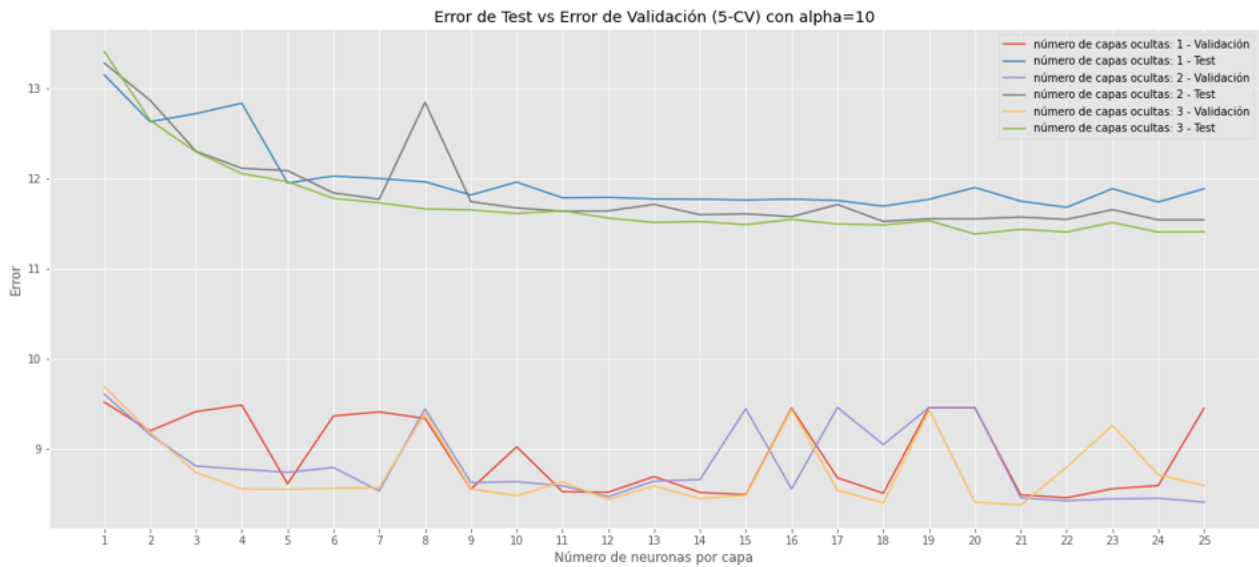


Error de Test vs Error de Validación (5-CV) con $\alpha=0.05$









El menor error de test y el valor del hiper-parámetro para el que se consigue:

	param_hidden_layer_sizes	param_alpha	n_hidden_layers	n_neurons_per_layer	mean_test_score	std_test_score	rank_test_score
262	(13, 13)	0.005	2	13	0.396788	0.0	1

El error de test para el valor del hiper-parámetro seleccionado por la validación cruzada:

0.515357834076244