

## 1. Resultados

Uno de los primeros experimentos realizados fue variar la cantidad de neuronas en una sola capa oculta para cada problema considerando las cantidades de épocas mencionadas anteriormente. A continuación presentamos resultados sobre dichos experimentos.

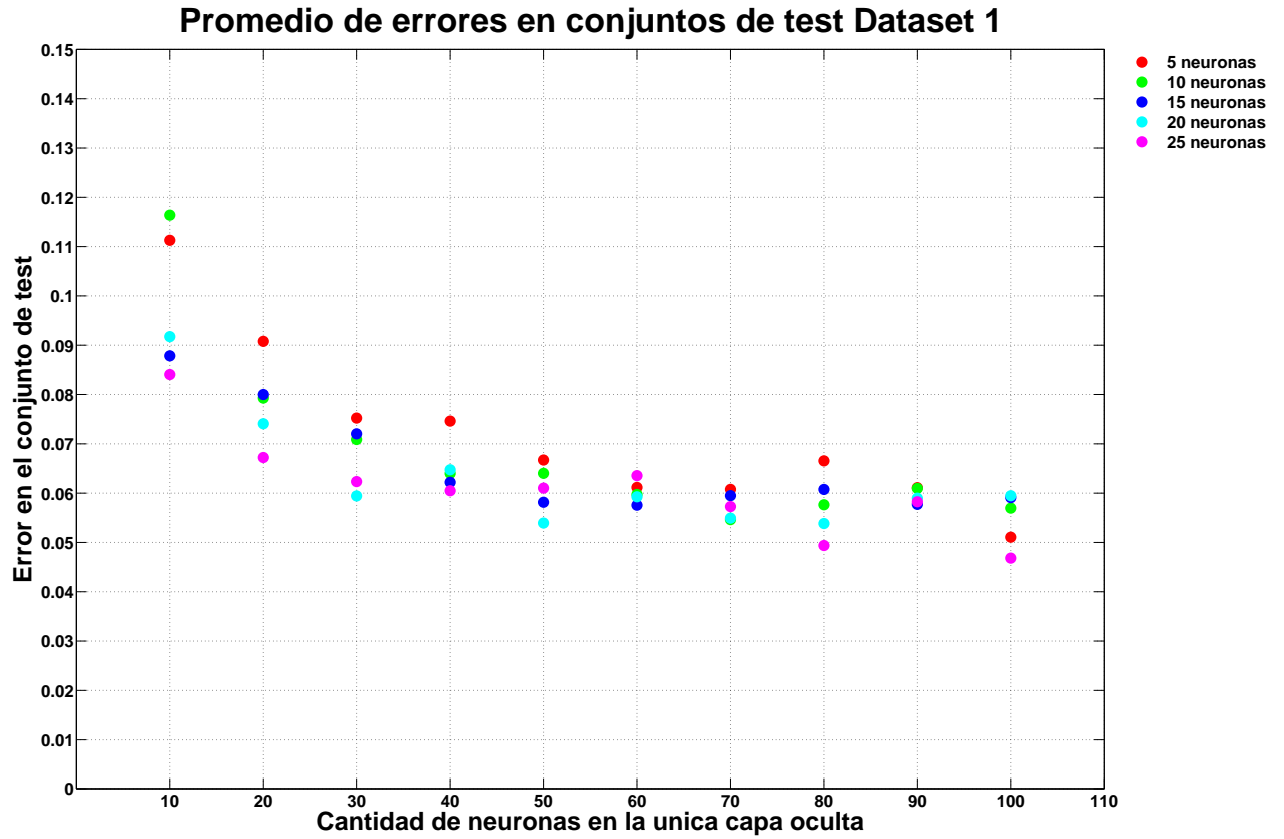


Figura 1: Promedio de error entre los 4 folds para el dataset 1 calculado sobre el conjunto de test usando  $\gamma = 0,1$

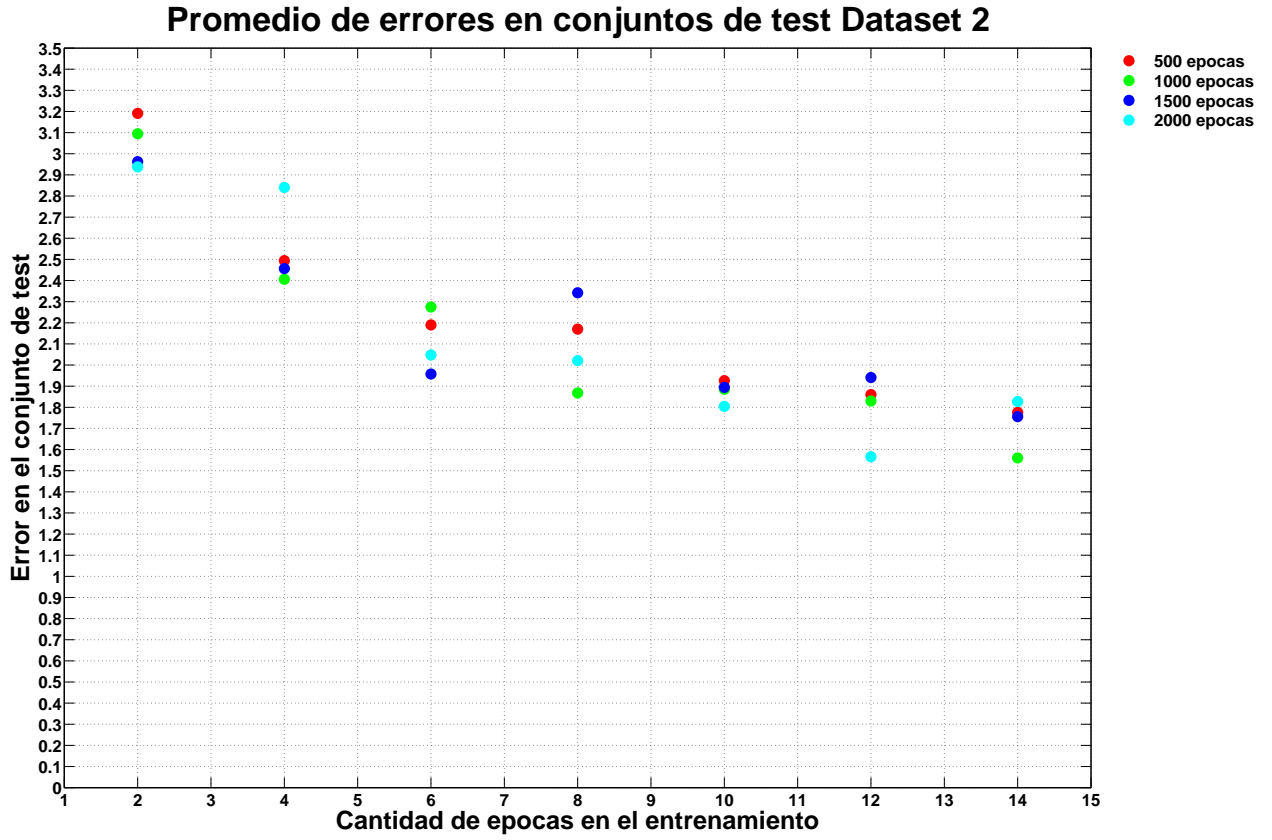
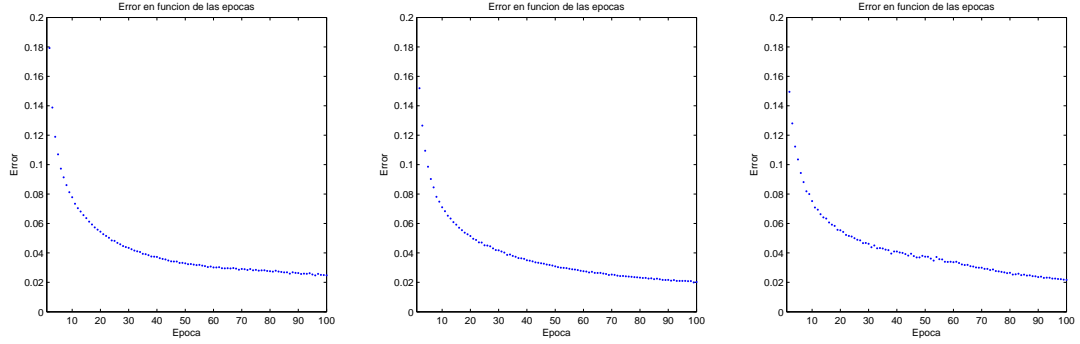


Figura 2: Promedio de error entre los 4 folds para el dataset 2 calculado sobre el conjunto de test usando  $\gamma = 0,1$

### 1.1. Problema 1

Dado que según la Figura 1 las configuraciones con 5 y 10 neuronas no presentan tan buenos resultados, mostramos a continuación (Figura 3) gráficos sobre la convergencia del error en función de las épocas usando las otras arquitecturas evaluadas. Para ello mostramos el fold 1 pero los demás presentan un comportamiento similar.



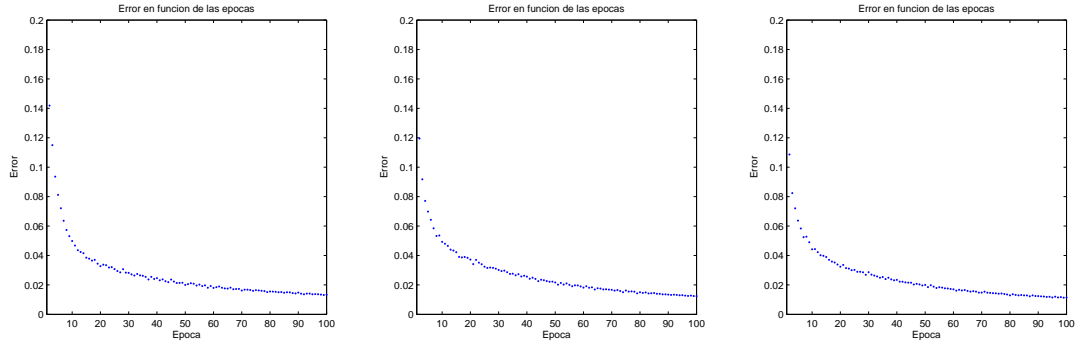
(a) Usando 15 neuronas.

(b) Usando 20 neuronas.

(c) Usando 25 neuronas.

Figura 3: Error durante el entrenamiento en el fold 1 del problema 1 usando una sola capa oculta y  $\gamma = 0,1$ .

Considerando esos resultados, decidimos analizar qué sucedía al utilizar valores de  $\gamma$  mayores ya que tal vez era posible converger más rápidamente. Los resultados se encuentran en las Figuras 4, 5, 6 y 7, nuevamente para 15, 20 y 25 neuronas.

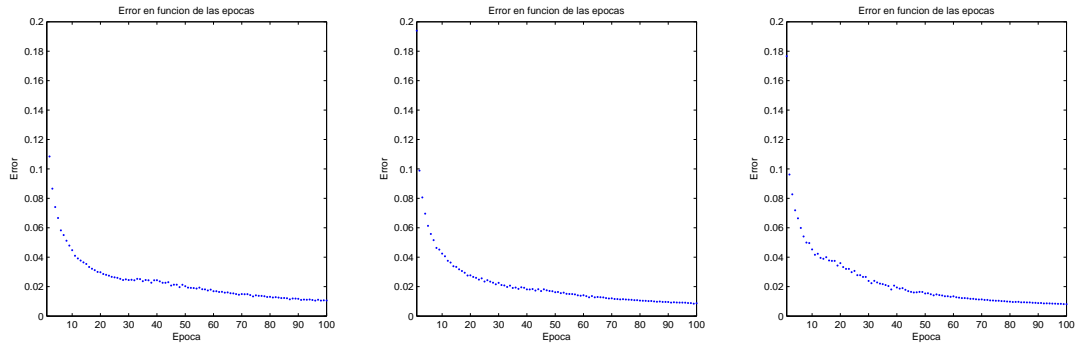


(a) Usando 15 neuronas.

(b) Usando 20 neuronas.

(c) Usando 25 neuronas.

Figura 4: Error durante el entrenamiento en el fold 1 del problema 1 usando una sola capa oculta y  $\gamma = 0,2$ .



(a) Usando 15 neuronas.

(b) Usando 20 neuronas.

(c) Usando 25 neuronas.

Figura 5: Error durante el entrenamiento en el fold 1 del problema 1 usando una sola capa oculta y  $\gamma = 0,3$ .

	15 neuronas	20 neuronas	25 neuronas
$\gamma = 0,1$	0.041432	0.043057	0.034287
$\gamma = 0,2$	0.039383	0.028796	0.045242
$\gamma = 0,3$	0.034535	0.033598	0.02312
$\gamma = 0,4$	0.034545	0.037957	0.036029
$\gamma = 0,5$	0.038391	0.03332	0.025975

Cuadro 1: Error de validación luego de 100 iteraciones en el fold 1 del problema 1.

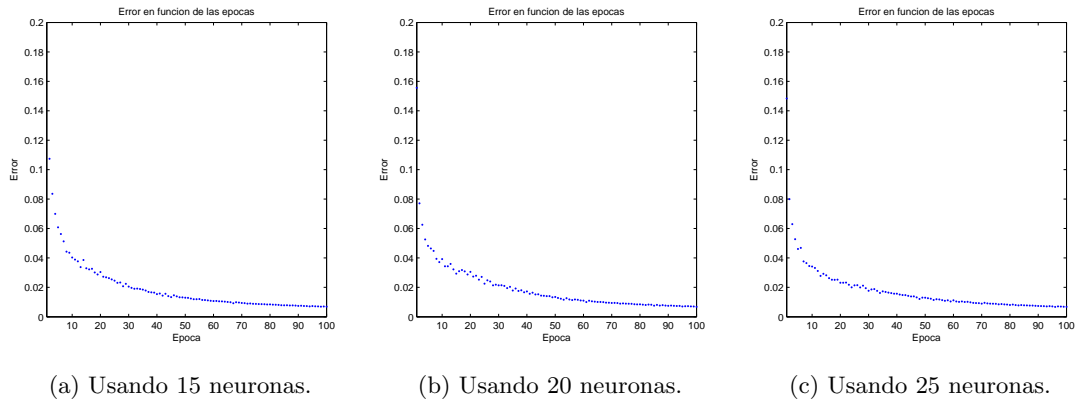


Figura 6: Error durante el entrenamiento en el fold 1 del problema 1 usando una sola capa oculta y  $\gamma = 0,4$ .

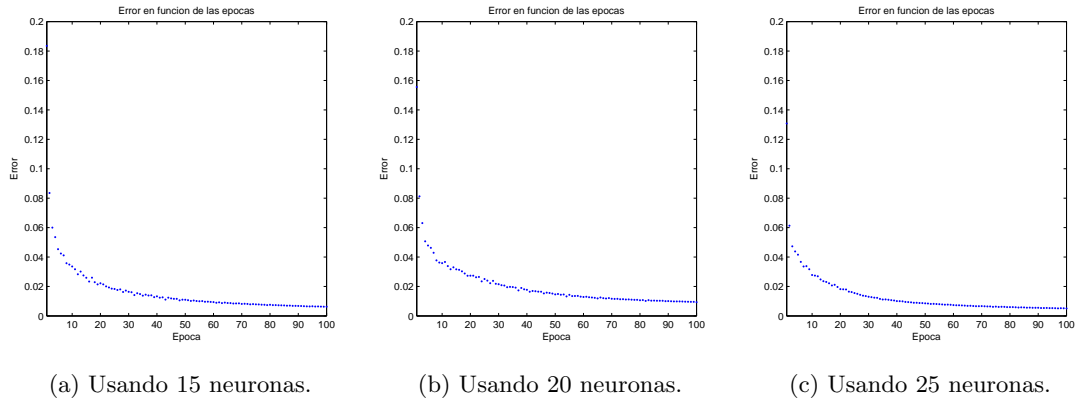
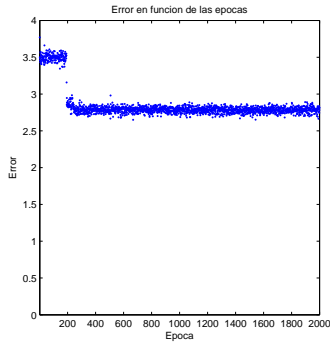


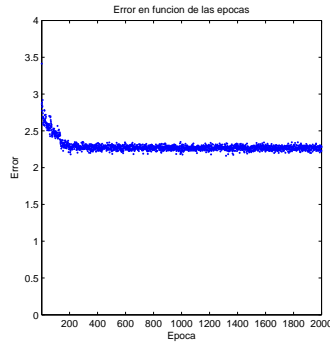
Figura 7: Error durante el entrenamiento en el fold 1 del problema 1 usando una sola capa oculta y  $\gamma = 0,5$ .

## 1.2. Problema 2

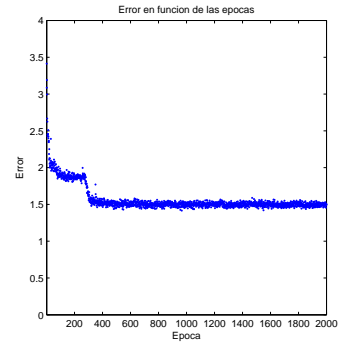
A continuación (Figura 8) mostramos los errores de entrenamiento en función de la época al entrenar las redes con diversas cantidades de neuronas en el fold 2 aunque comportamientos similares fueron observados en el resto de los folds. Se puede observar cómo la cantidad de neuronas en la capa oculta juega un papel esencial en el error obtenido. Dado que cuanto mayor es el número de neuronas, mejores resultados obtenemos, decidimos analizar los resultados usando 16, 17, 18, 19 y 20 neuronas, mostrados en la Figura 9.



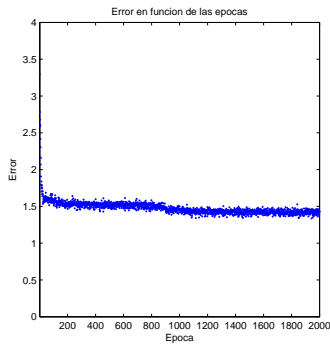
(a) Usando 2 neuronas.



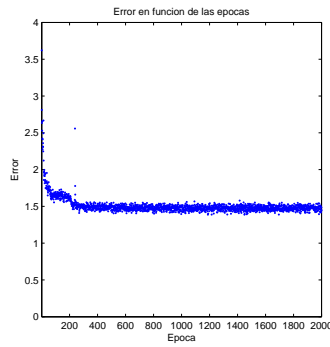
(b) Usando 4 neuronas.



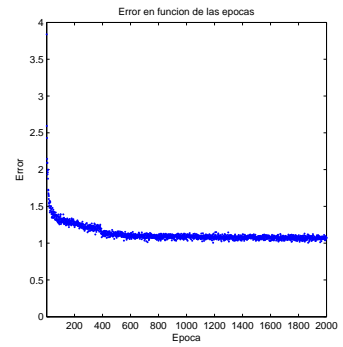
(c) Usando 6 neuronas.



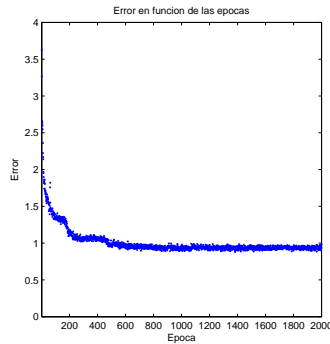
(d) Usando 8 neuronas.



(e) Usando 10 neuronas.

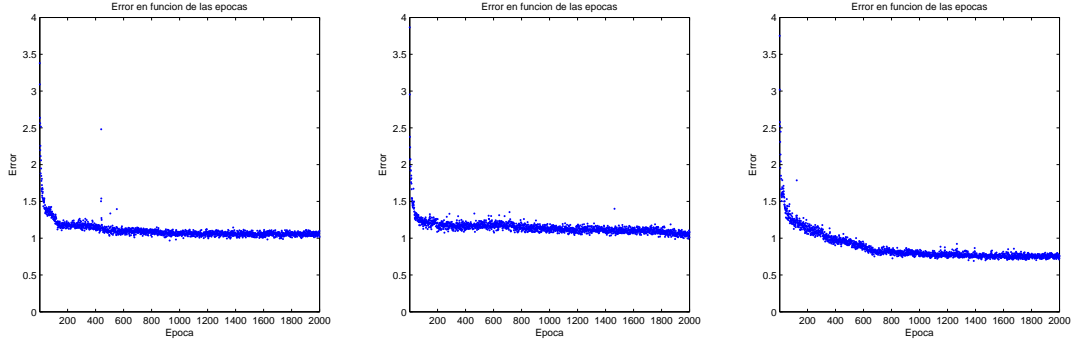


(f) Usando 12 neuronas.



(g) Usando 14 neuronas.

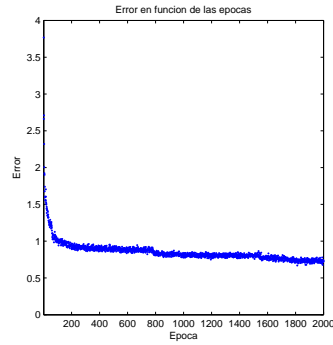
Figura 8: Error durante el entrenamiento en el fold 2 del problema 2 usando una sola capa oculta y  $\gamma = 0,1$ .



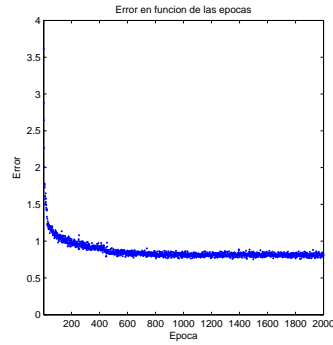
(a) Usando 16 neuronas.

(b) Usando 17 neuronas.

(c) Usando 18 neuronas.



(d) Usando 19 neuronas.

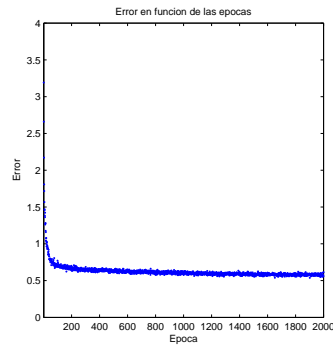


(e) Usando 20 neuronas.

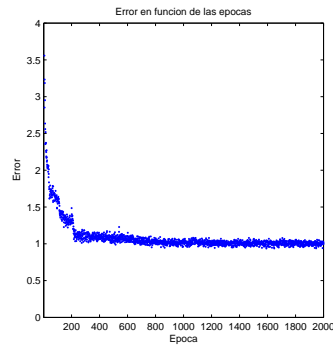
Figura 9: Error durante el entrenamiento en el fold 2 del problema 2 usando una sola capa oculta y  $\gamma = 0,1$  segunda parte. Error en el conjunto de test (promedio entre los cuatro folds): 1,758475, 1,926875, 1,82305, 1,4711275 y 1,38239 respectivamente.

Tomando los resultados de las Figuras 8 y 9 se puede ver que los errores (tanto en entrenamiento como en validación) más bajos se obtienen usando 20 neuronas, además de un comportamiento estable por lo que decidimos no buscar configuraciones con más neuronas.

Considerando esa arquitectura, decidimos analizar otros valores de learning rate a fin de observar si es posible converger más rápido. Los resultados se encuentran en la Figura 10.



(a) Con  $\gamma = 0,05$ .



(b) Con  $\gamma = 0,15$ .

Figura 10: Error durante el entrenamiento en el fold 2 del problema 2 usando una sola capa oculta de 20 neuronas variando el valor de  $\gamma$ .