

Trabajo Práctico N°2: Redes Neuronales

Introducción al Aprendizaje Automatizado



Alumno: Navall, Nicolás Uriel. N-1159/2.

a)

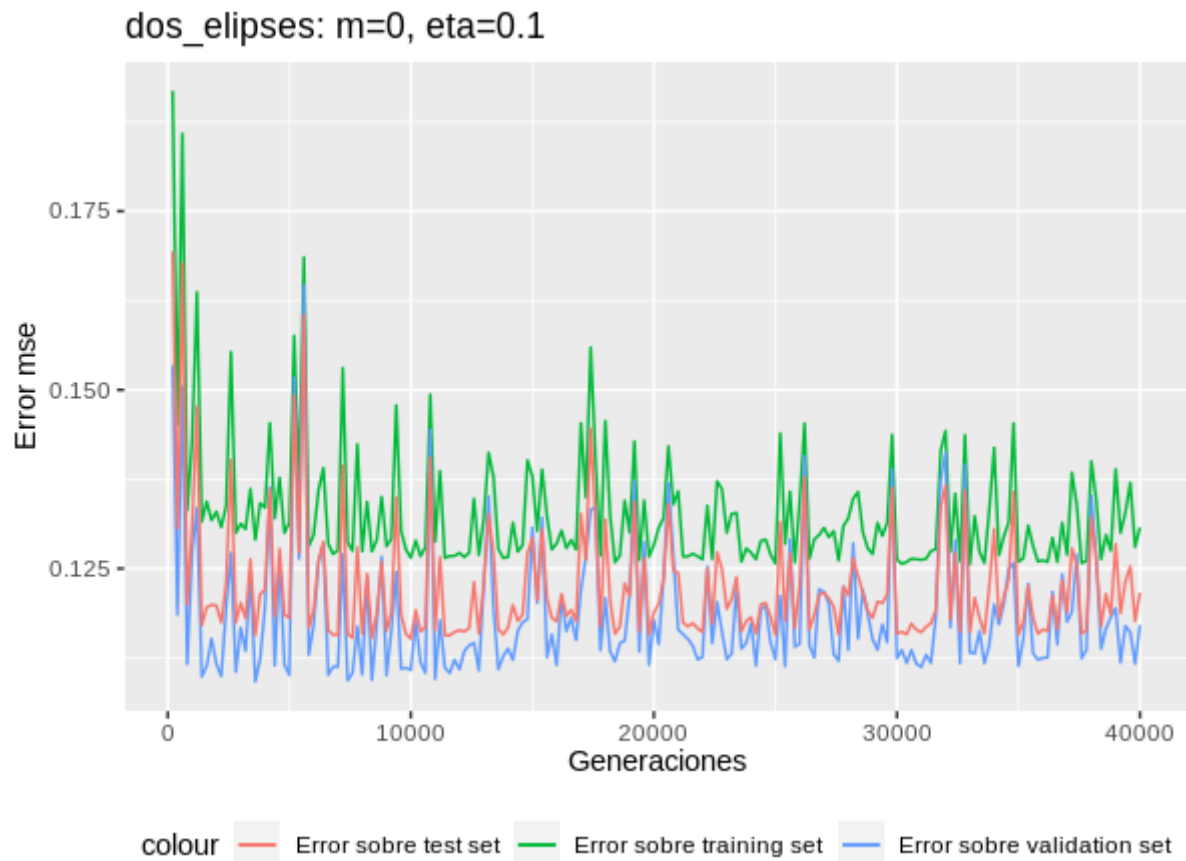


Figura 1.1

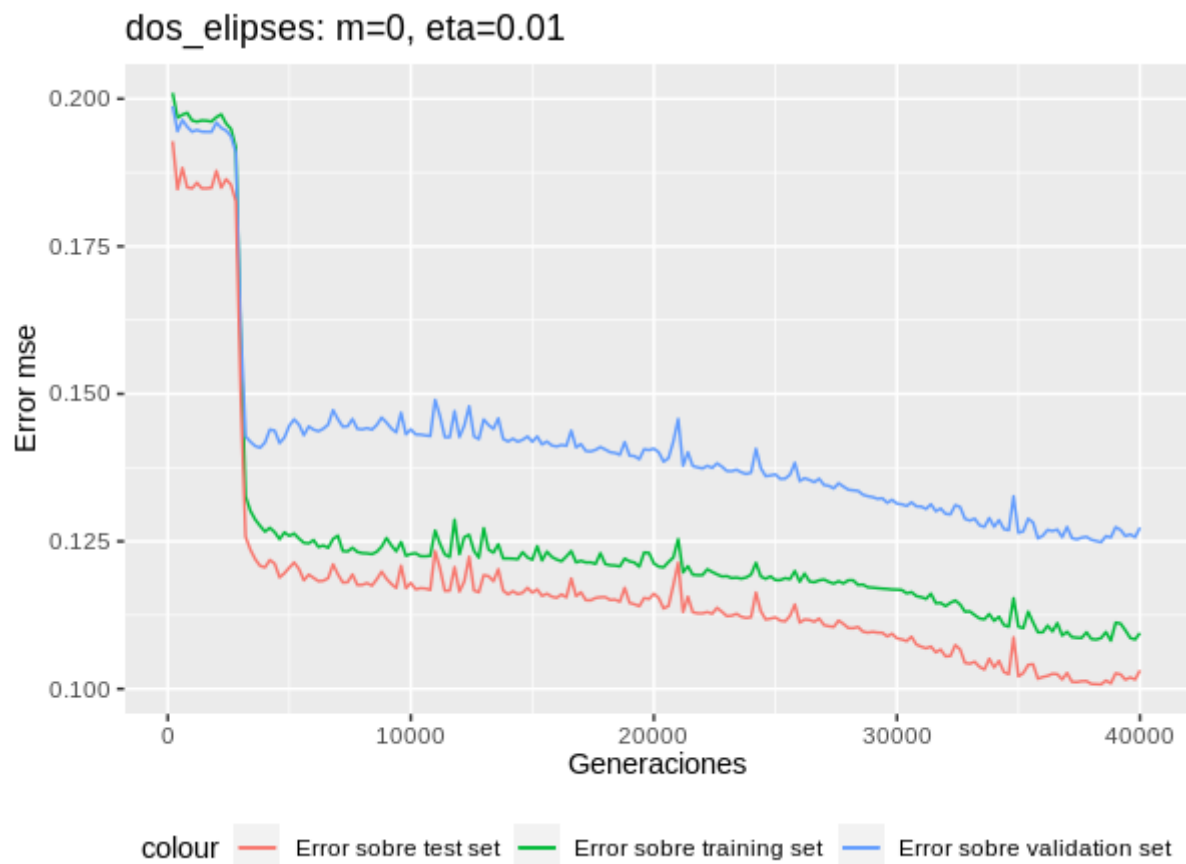


Figura 1.2

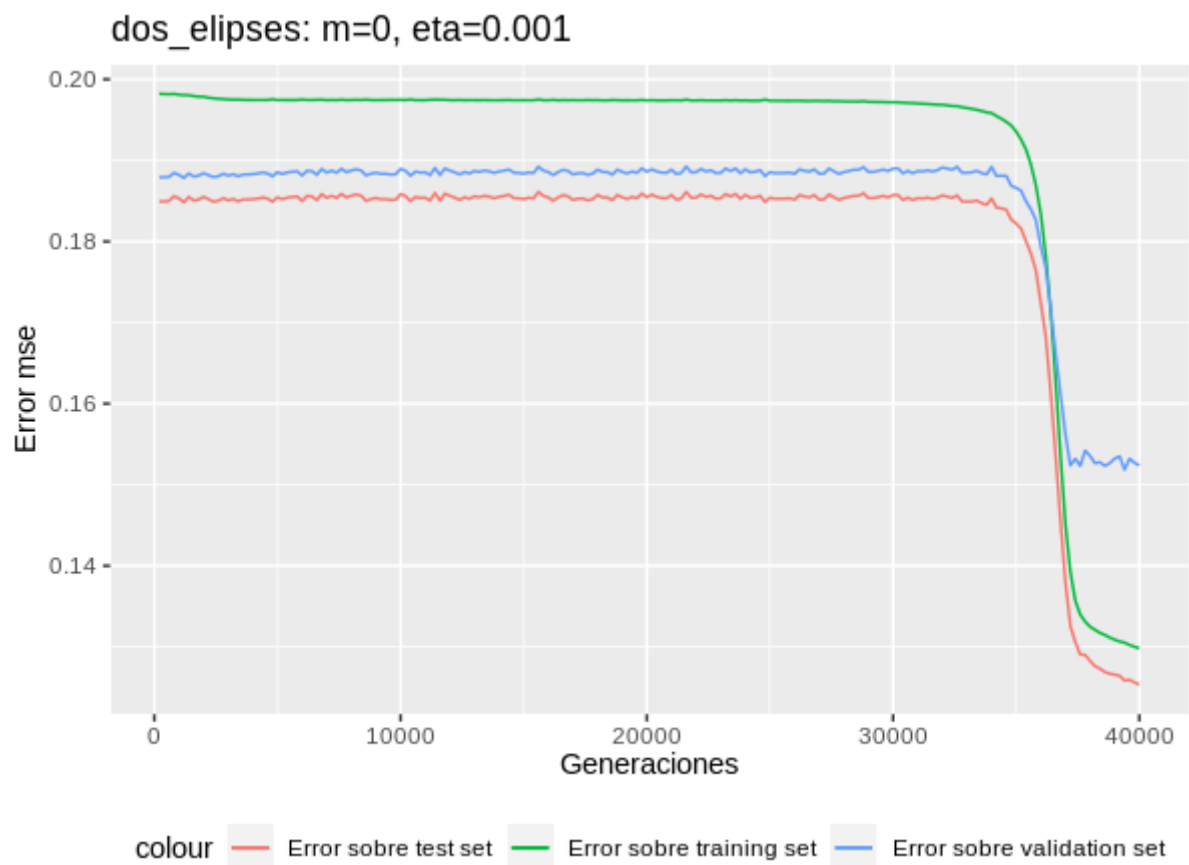


Figura 1.3

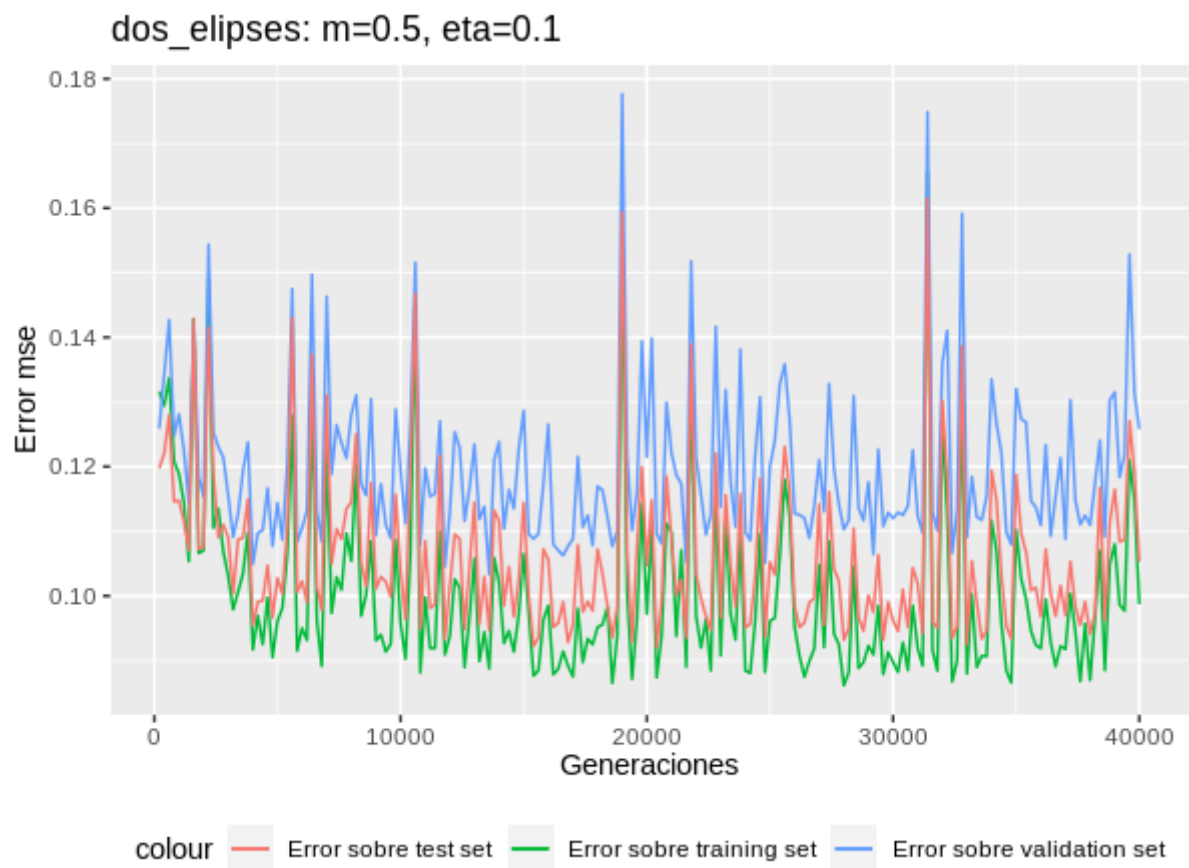


Figura 1.4

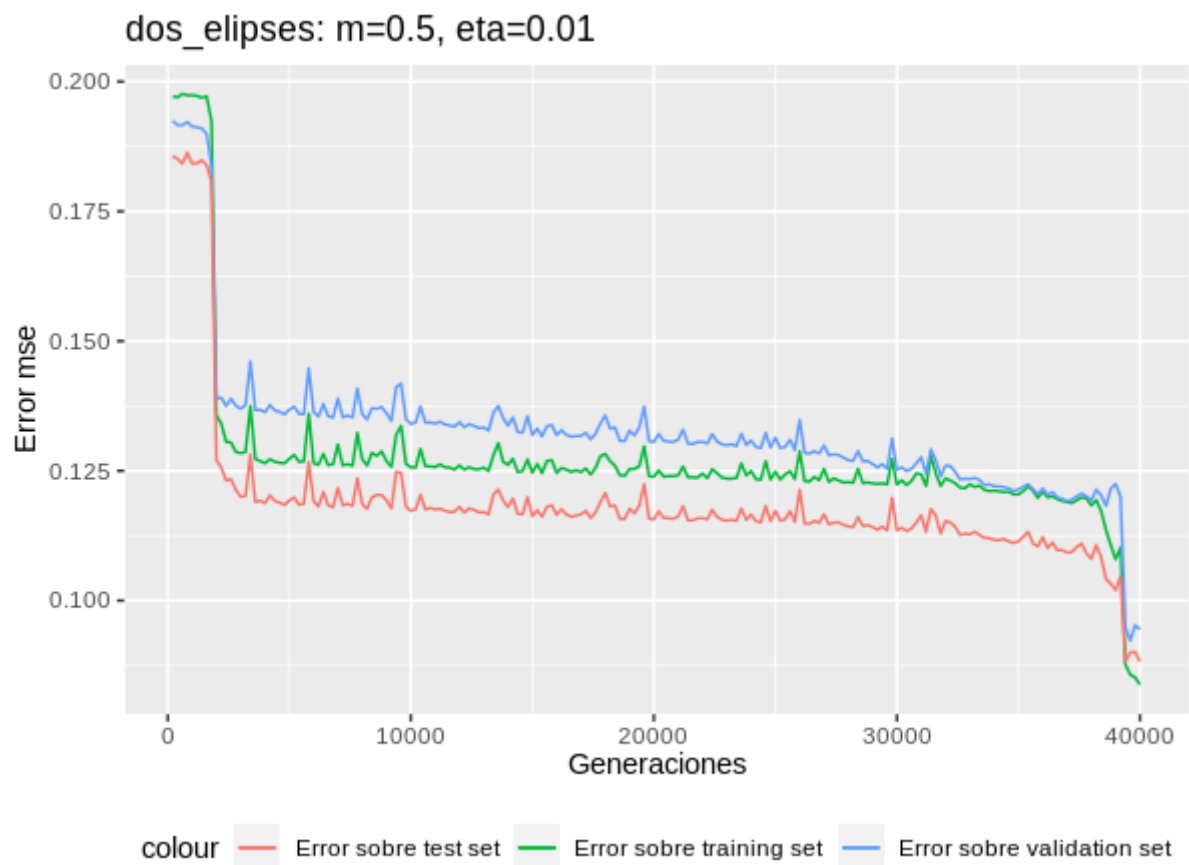


Figura 1.5

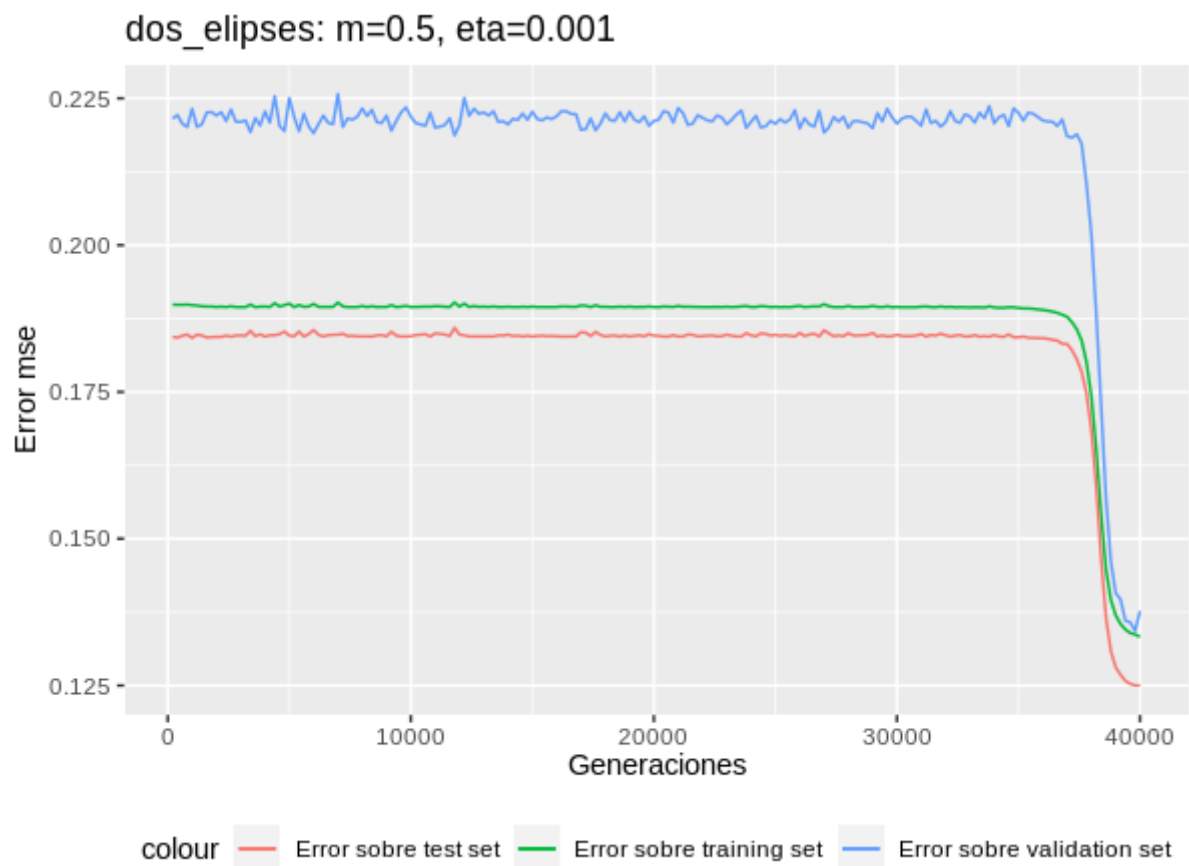


Figura 1.6

dos_elipses: $m=0.9$, $\eta=0.1$

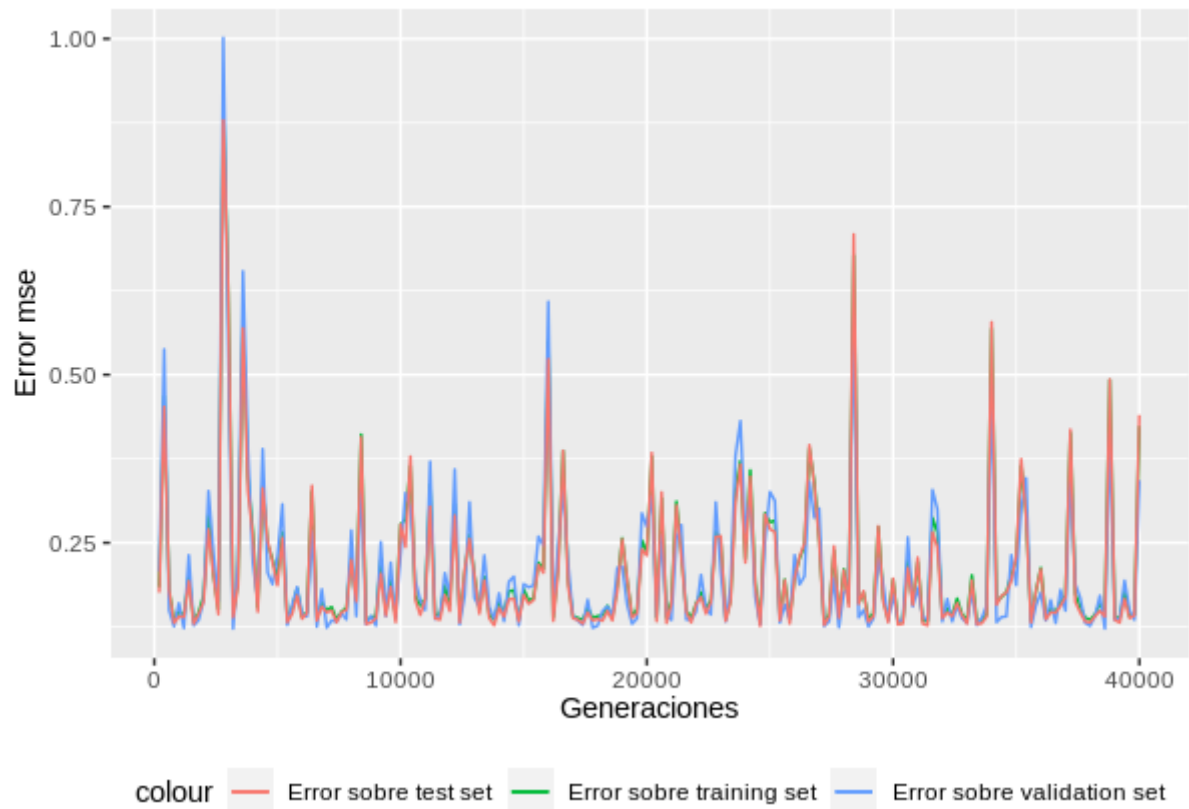


Figura 1.7

dos_elipses: $m=0.9$, $\eta=0.01$

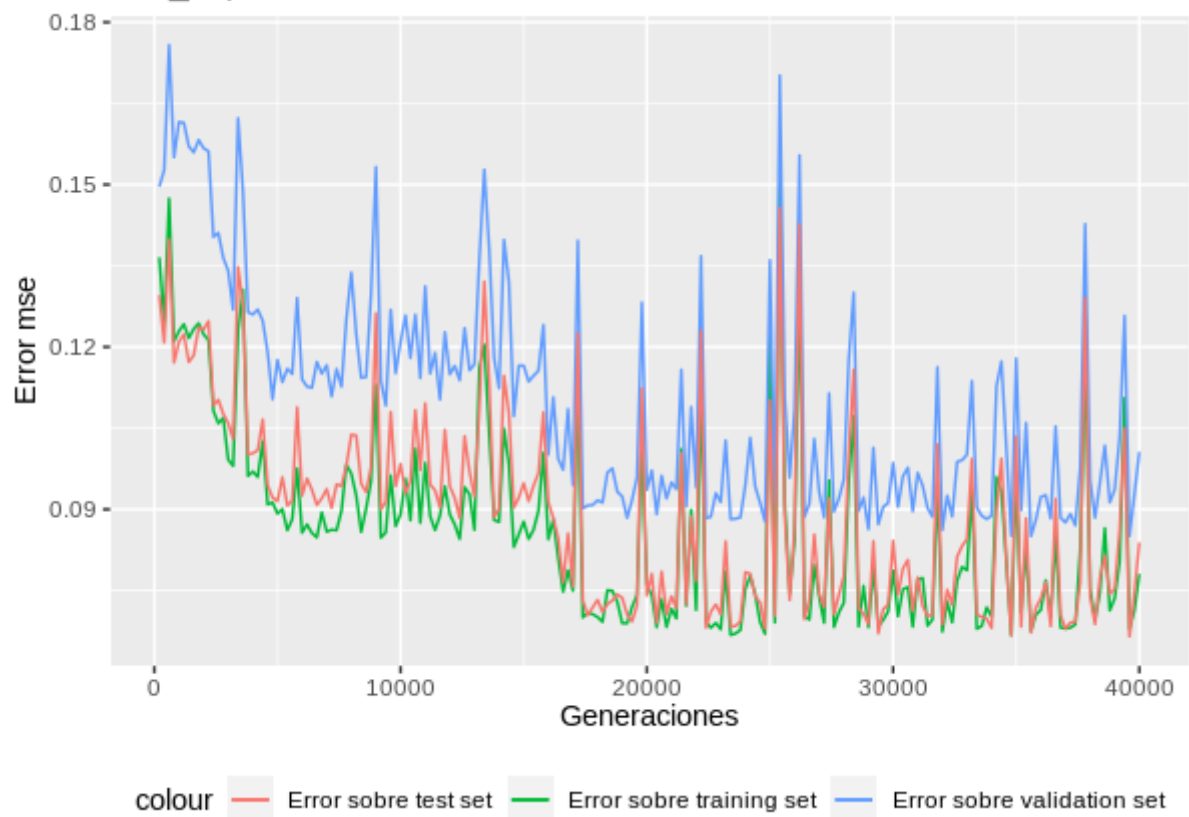


Figura 1.8

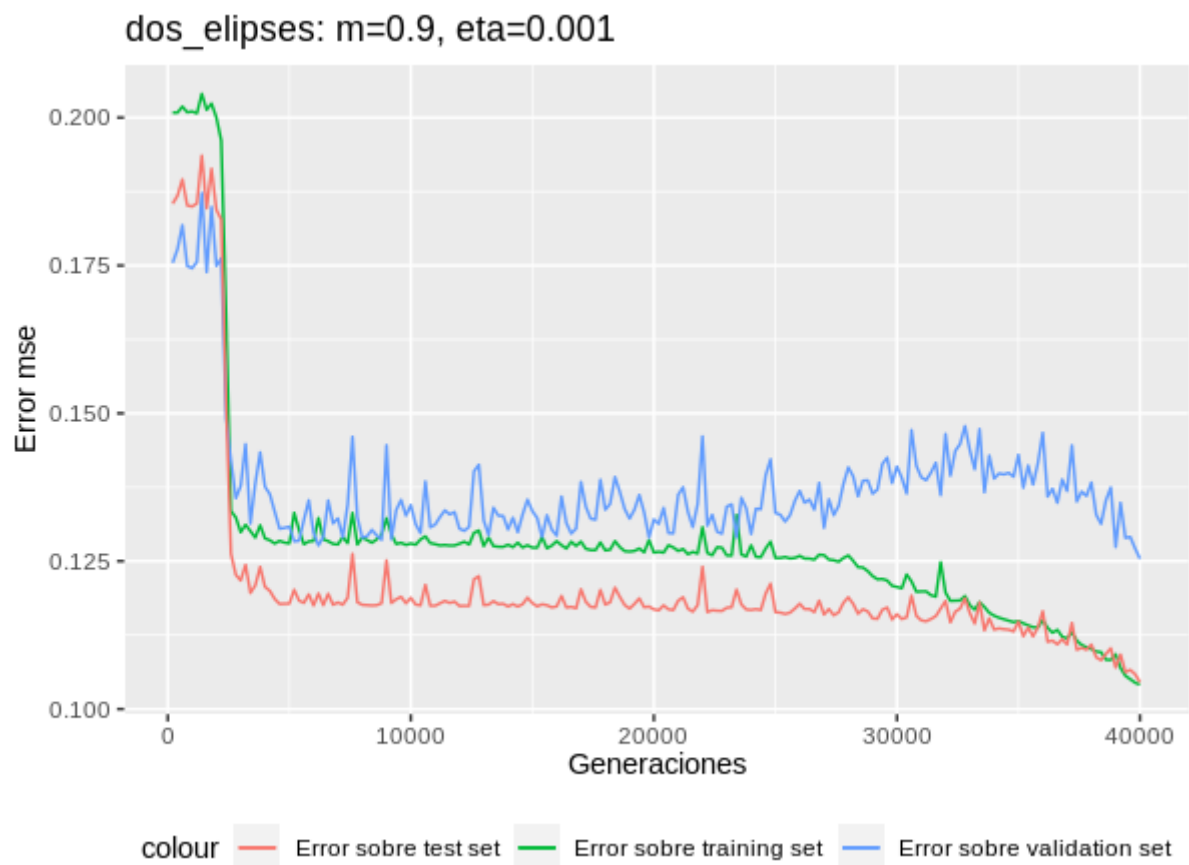


Figura 1.9

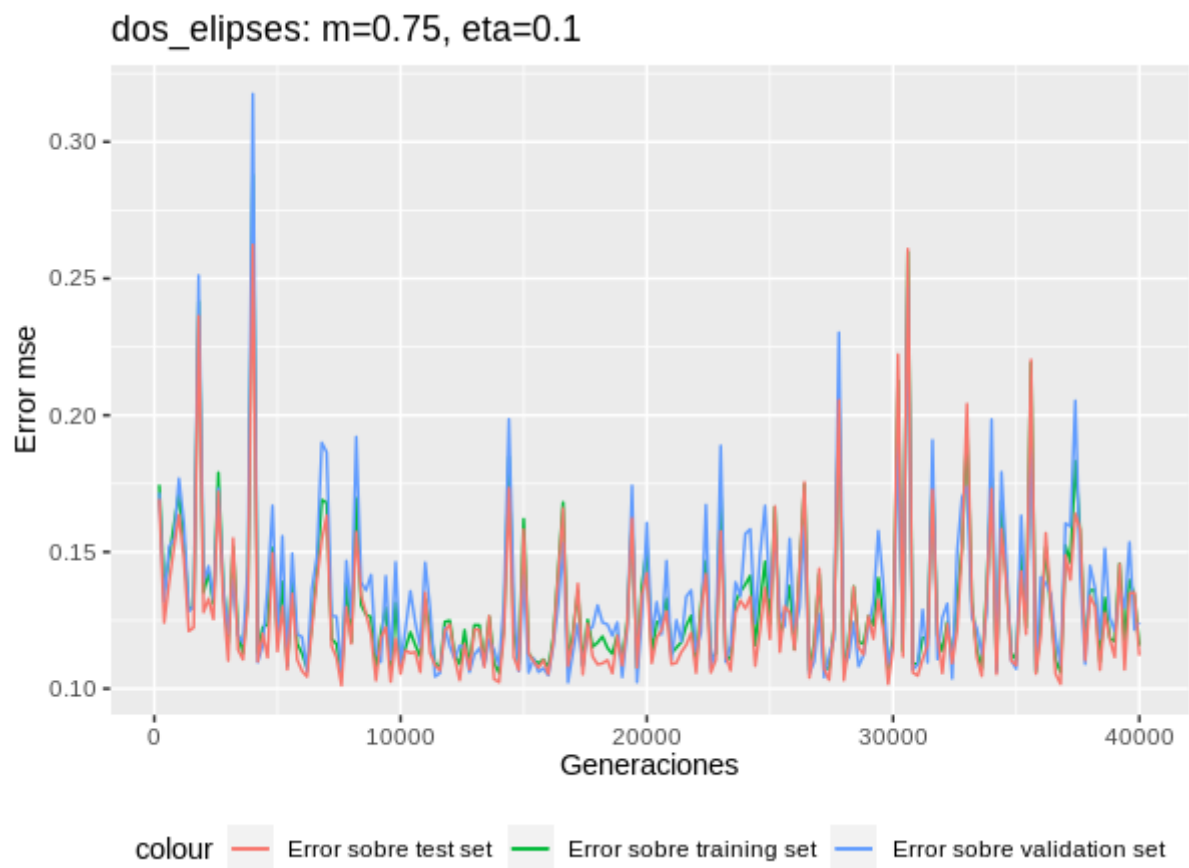


Figura 1.10

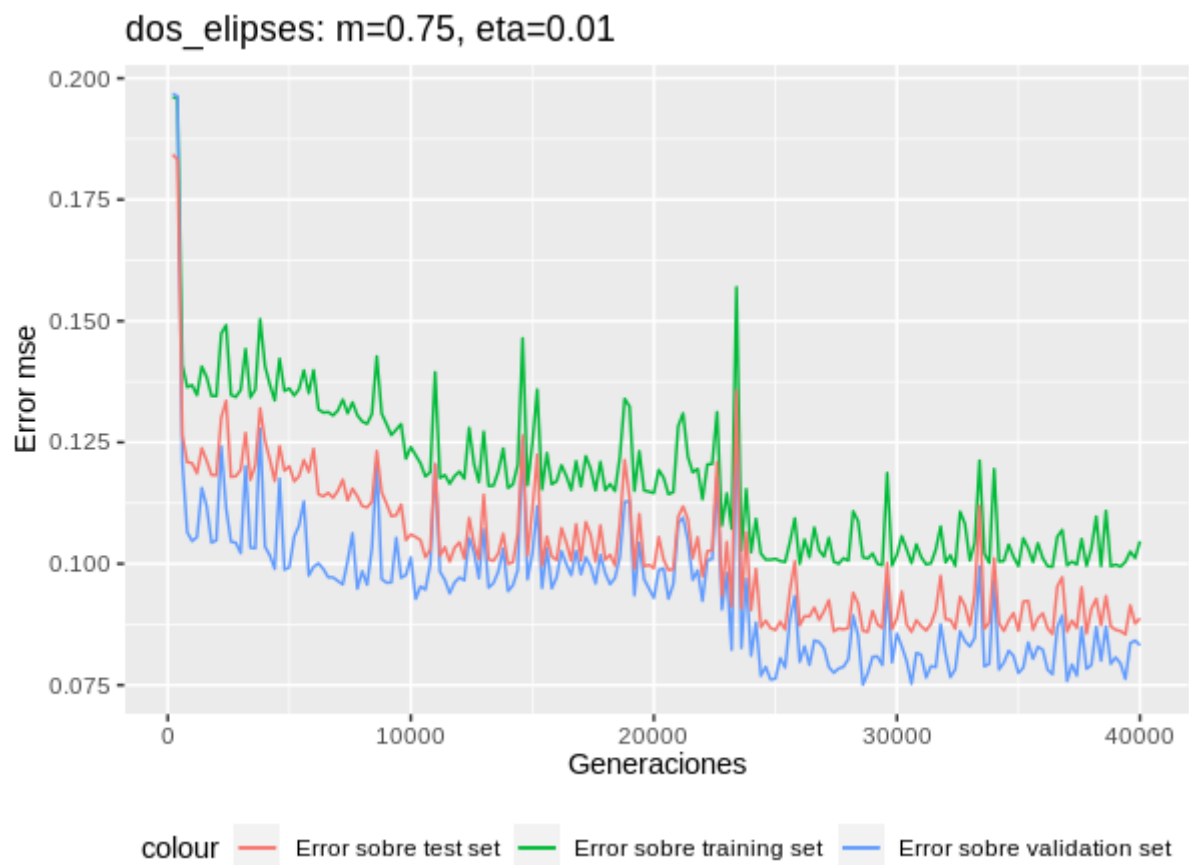


Figura 1.11

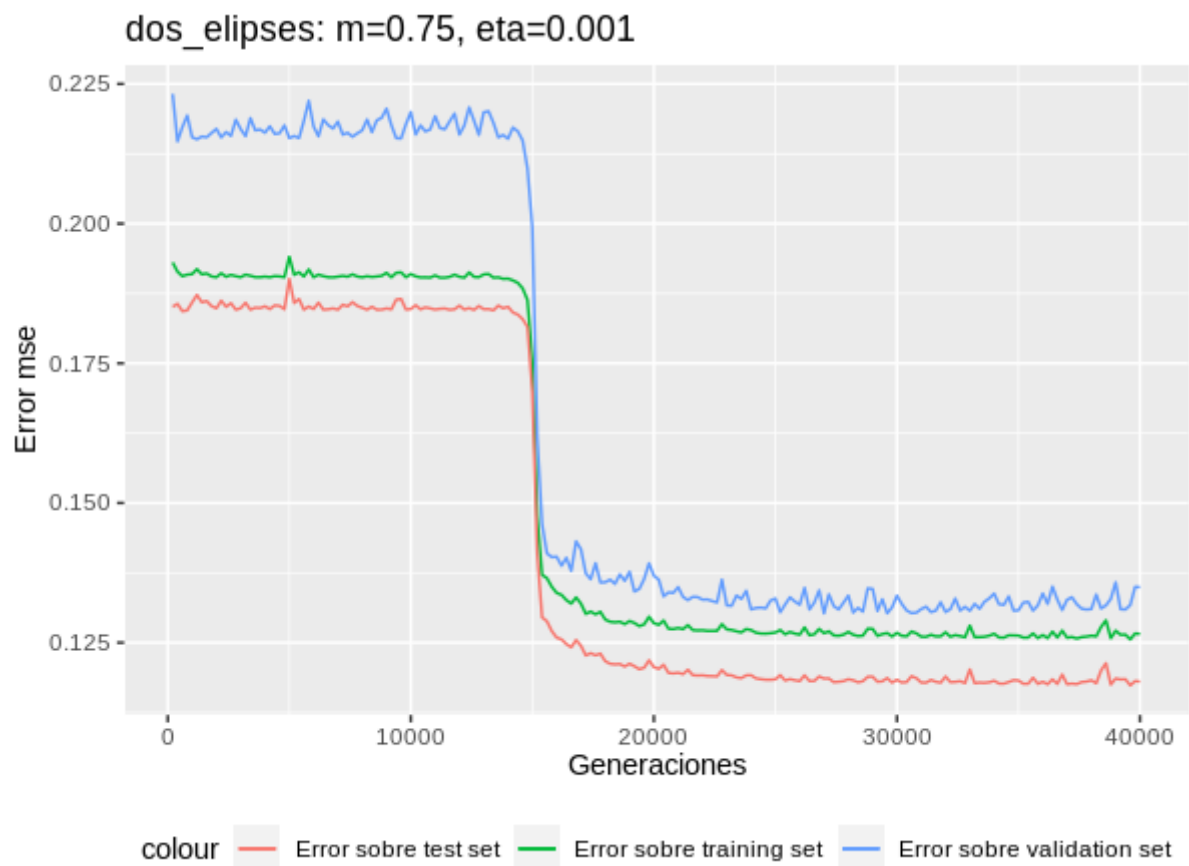


Figura 1.12

dos_elipses: $m=0.6$, $\eta=0.01$

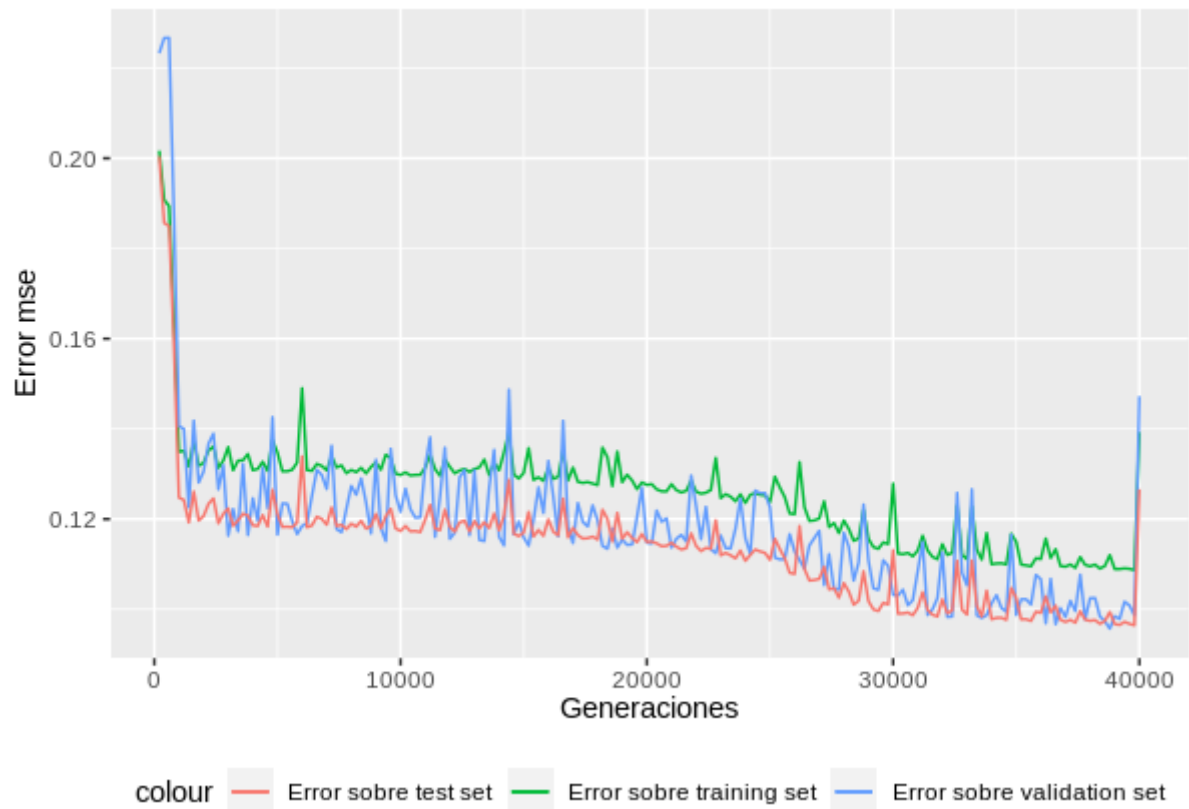


Figura 1.13

dos_elipses: $m=0.8$, $\eta=0.01$

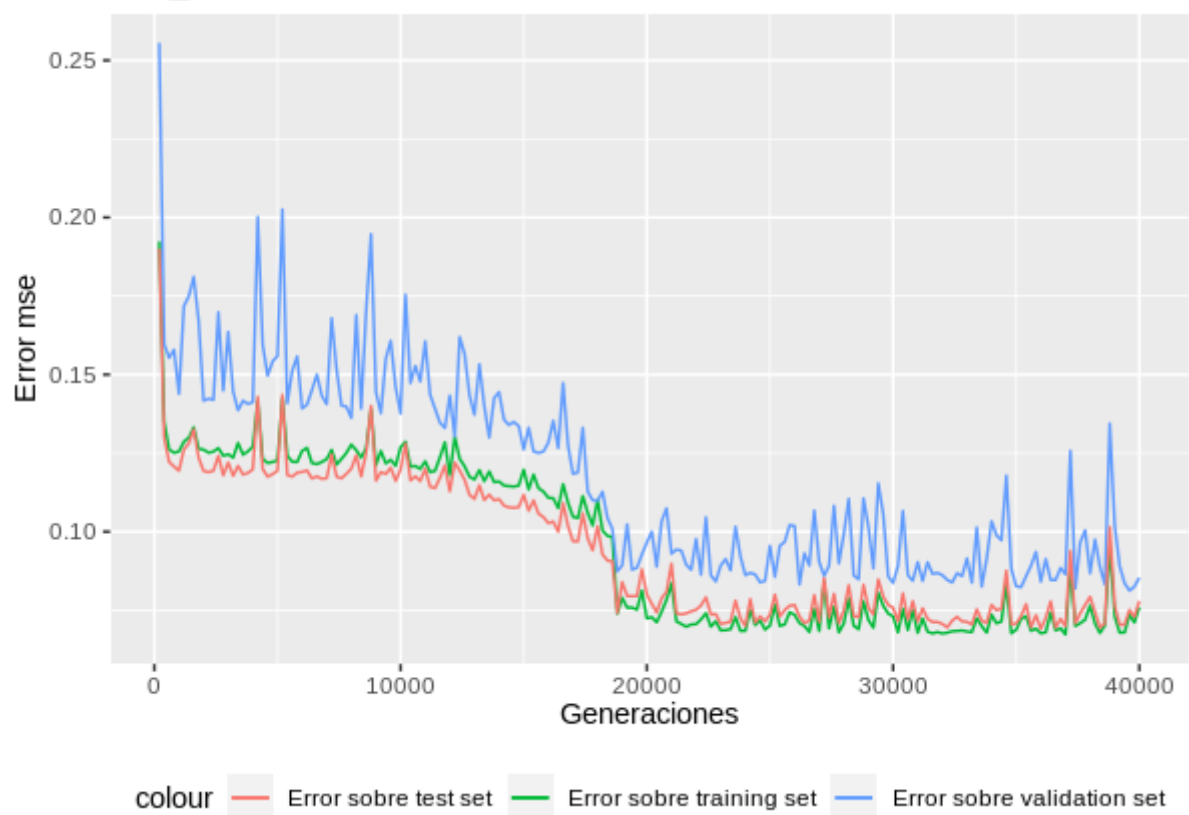
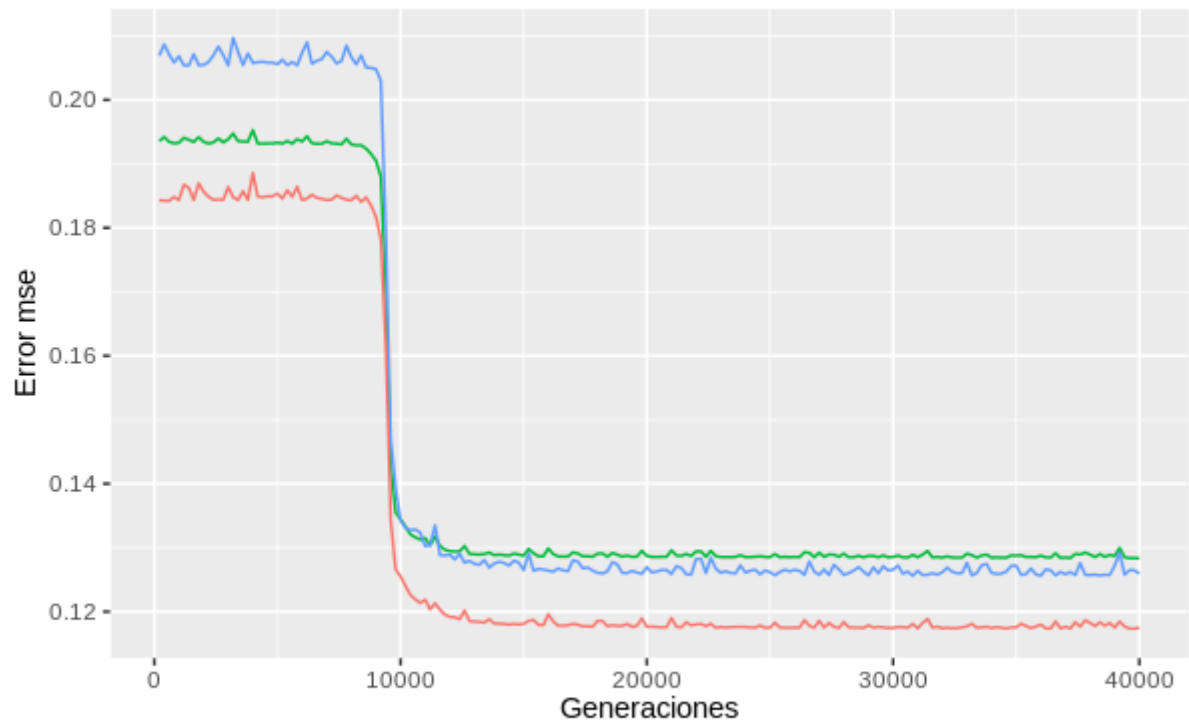


Figura 1.14

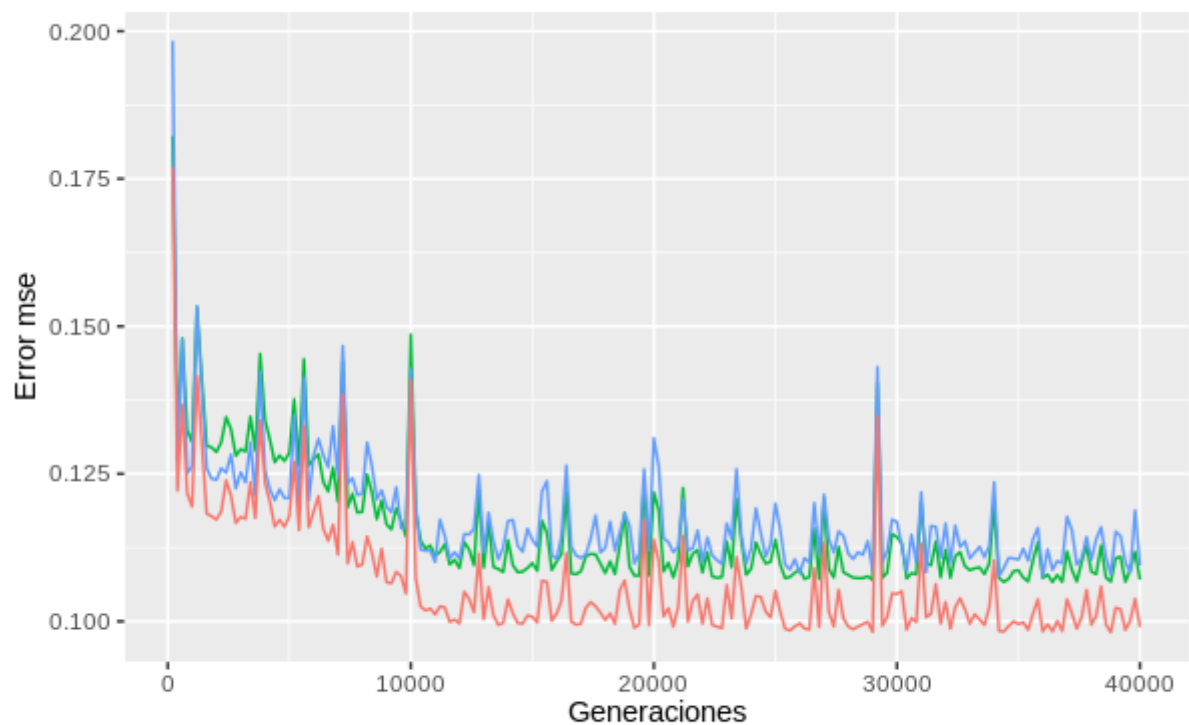
dos_elipses: $m=0.8$, $\eta=0.001$



colour — Error sobre test set — Error sobre training set — Error sobre validation set

Figura 1.15

dos_elipses: $m=0.85$, $\eta=0.01$



colour — Error sobre test set — Error sobre training set — Error sobre validation set

Figura 1.15

Error de test mínimo de cada red entrenada:

eta\m	0	0.25	0.5	0.6	0.75	0.8	0.85	0.9
0.1	0.115179	-	0.092022	-	0.101186	-	-	0.126191
0.01	0.100752	0.105953	0.088163	0.096443	0.085448	0.069087	0.098129	0.066529
0.001	0.125349	-	0.125021	-	0.117349	0.117344	-	0.104514

En las redes entrenadas con un $\eta=0.1$ podemos observar como estas realizan saltos de error más bruscos. Esto se debe a que los pesos de la red se ven modificados en mayor medida que el resto de las redes, por lo cual los cambios entre cada generación son más grandes en comparación.

Podemos observar como en las figuras 1.2, 1.3, 1.5, 1.9, 1.12, 1.15 la red sale del mínimo local en que se encuentra para llegar a otro mínimo. Podemos apreciar cómo en estos casos el η es un número bajo en comparación con el de los casos observados anteriormente, y tiene sentido que pasos chicos en los pesos nos dejen varados en mínimos locales, al menos durante unas épocas.

En 1.9 se ve lo diferente que puede ser la gráfica de error de validación con la de test

En la figura 1.4 se puede observar como la combinación de un $\text{momentum}=0.5$ con una $\eta=0.1$ puede resultar en que el error generado por la red no converja a ningún valor, ya que los cambios son tan bruscos y el momentum tan fuerte que sigue recorriendo el espacio de solución sin establecerse en ningún mínimo. Puede observarse algo similar en las figura 1.10.

b)

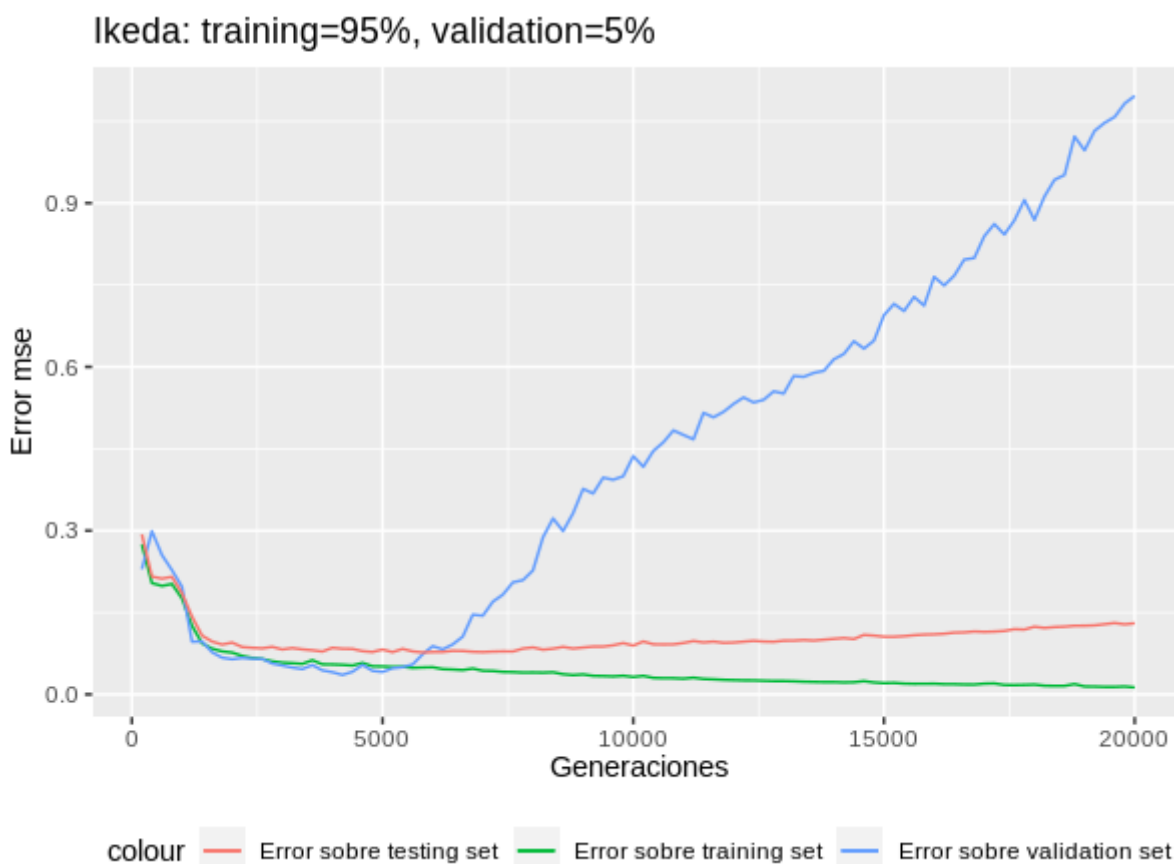


Figura 2.1

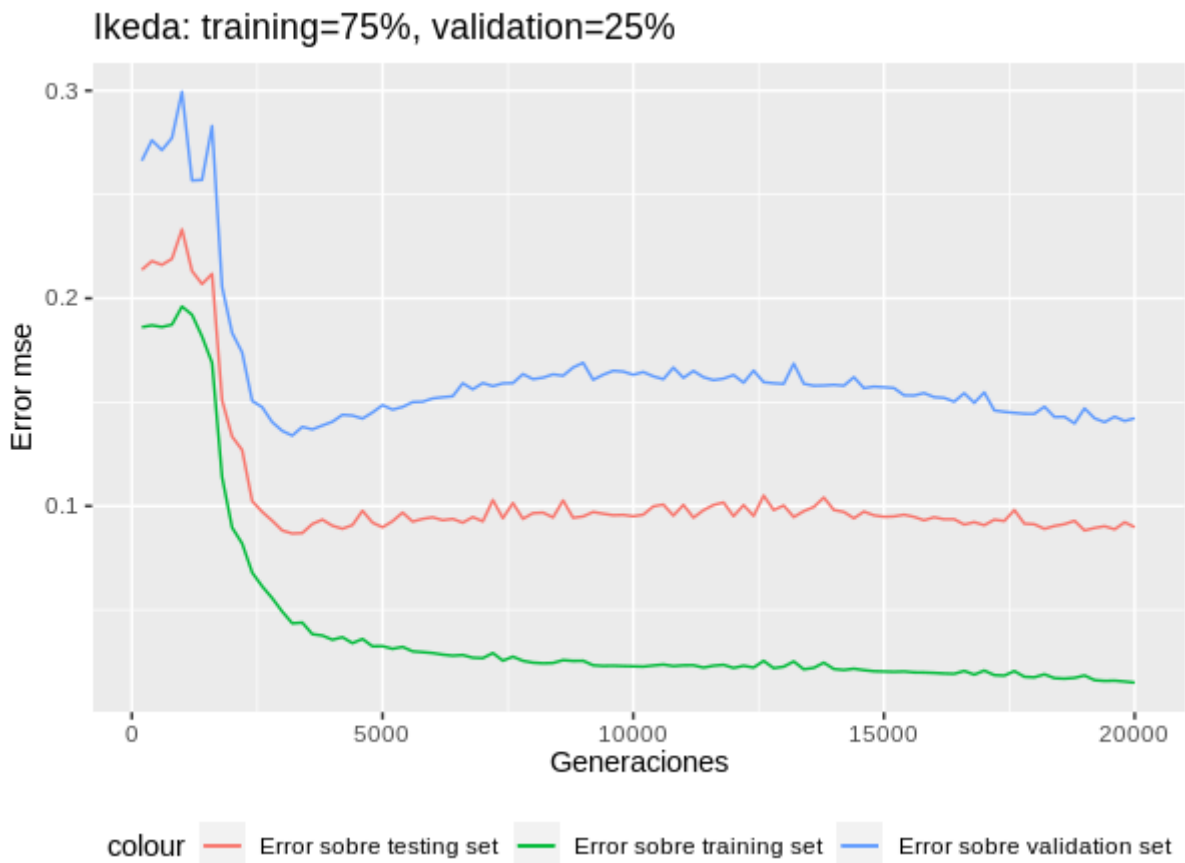


Figura 2.2

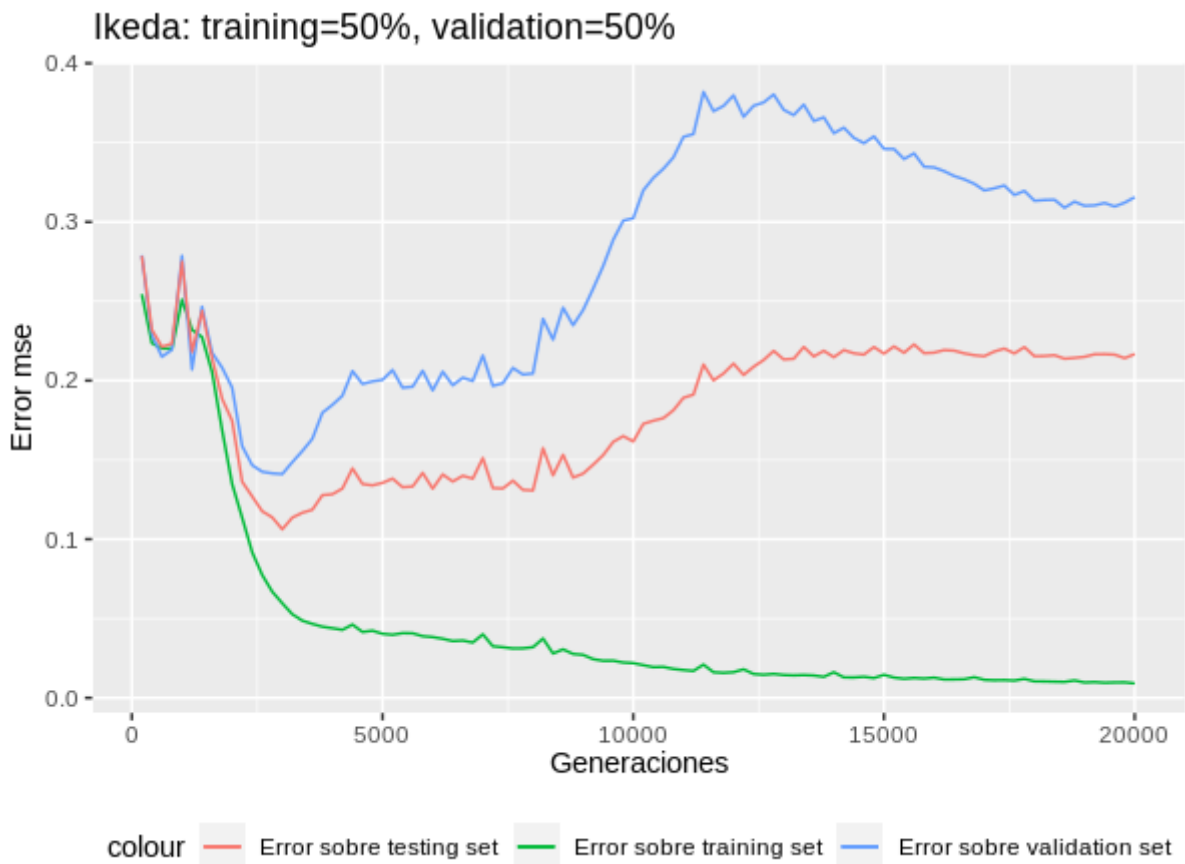


Figura 2.3

En la figura 2.1 se puede observar la diferencia abismal entre el error sobre el conjunto de testing en comparación con el error sobre el conjunto de validación, lo cual me da a entender que 5% de nuestro total de patrones que tenemos no es suficiente para representar el error sobre el universo.

En la figura 2.3 se puede ver como los errores de validación y testing comienzan a aumentar a partir de cierta generación. Esto se debe a que la complejidad de la red aumentó más de lo adecuado. Un sistema de early stopping tomaría la red en el momento donde su error de validación es menor, en este caso antes de la época nº 3700.

c)

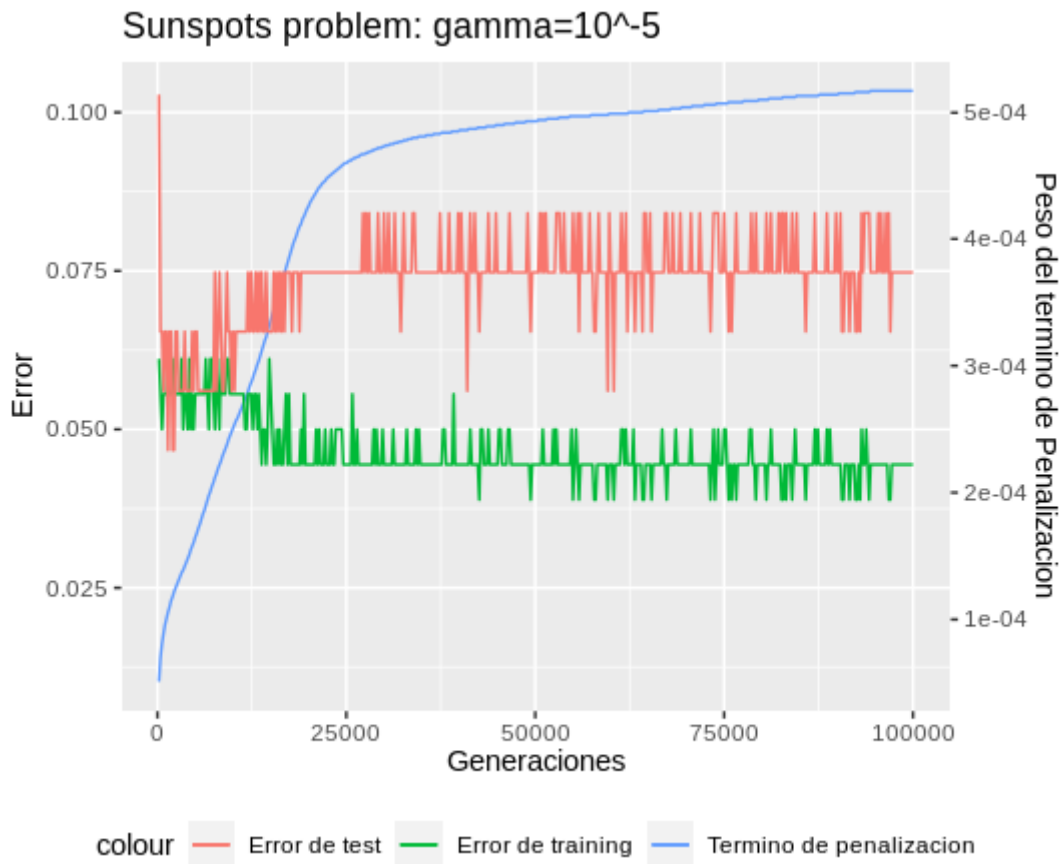


Figura 3.1

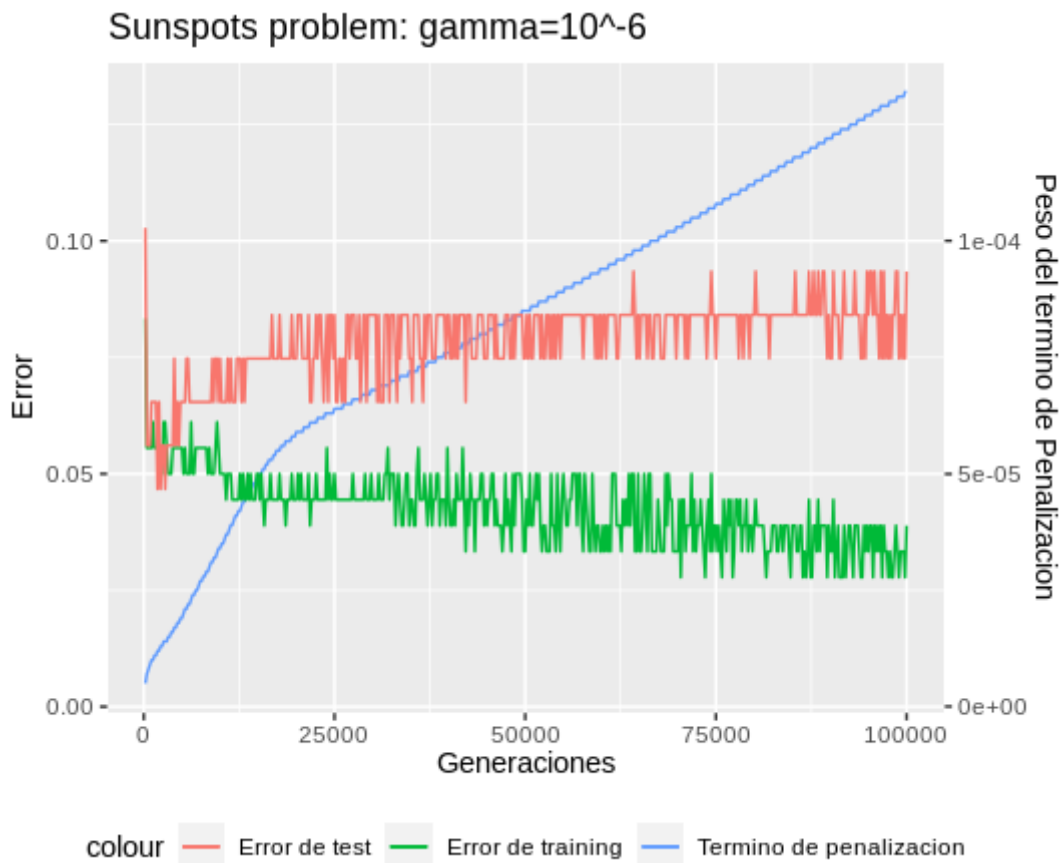


Figura 3.2

Podemos ver como el término de penalización comienza a aumentar en la figura 3.1, mientras el error de test aumenta, lo cual tiene sentido ya que el término de penalización busca evitar el overfitting, y un aumento del error de test es un signo de que se estaba produciendo un sobreajuste (el otro es que el error de training baje, lo cual también se puede observar en la gráfica). Algo similar puede observarse en la figura 3.2.

d)

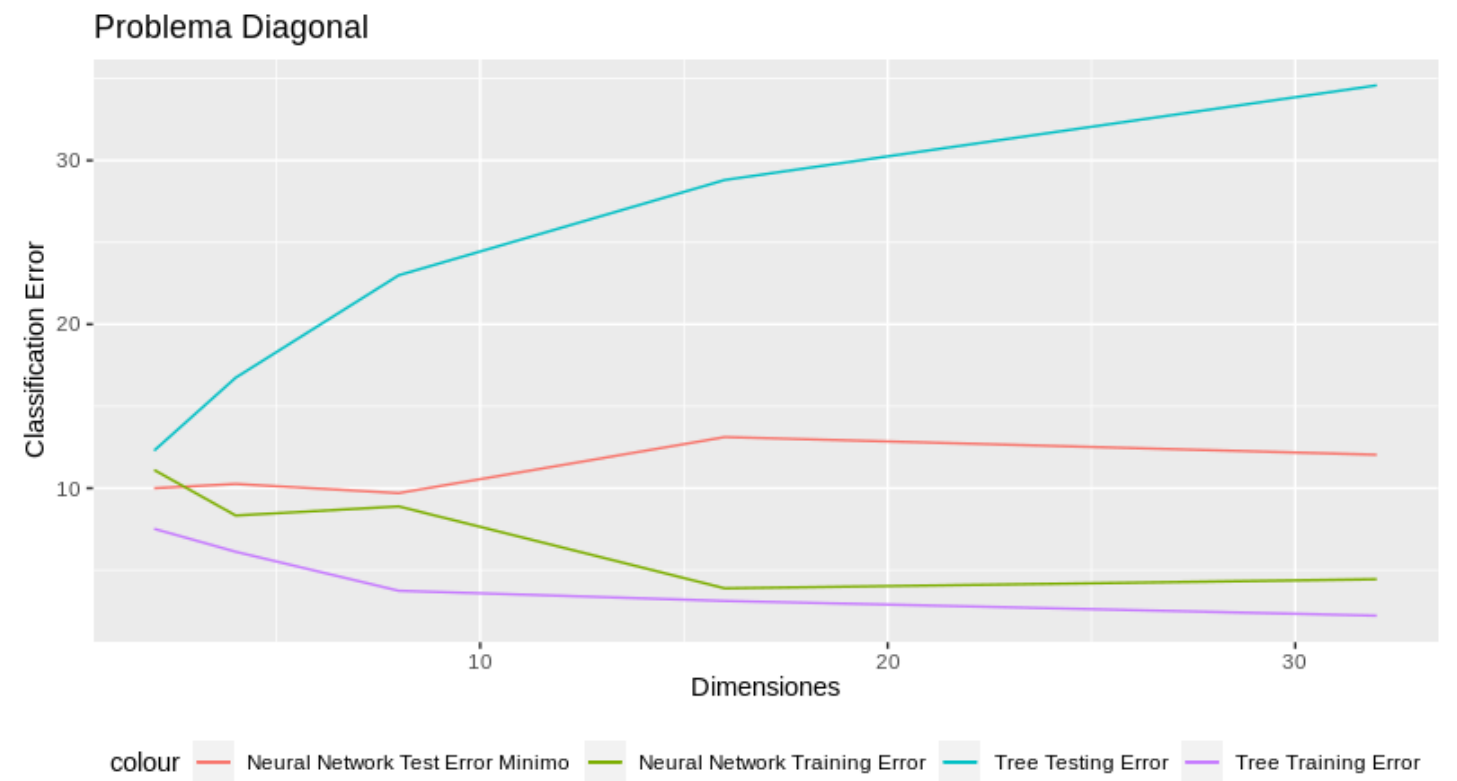


Figura 4.1

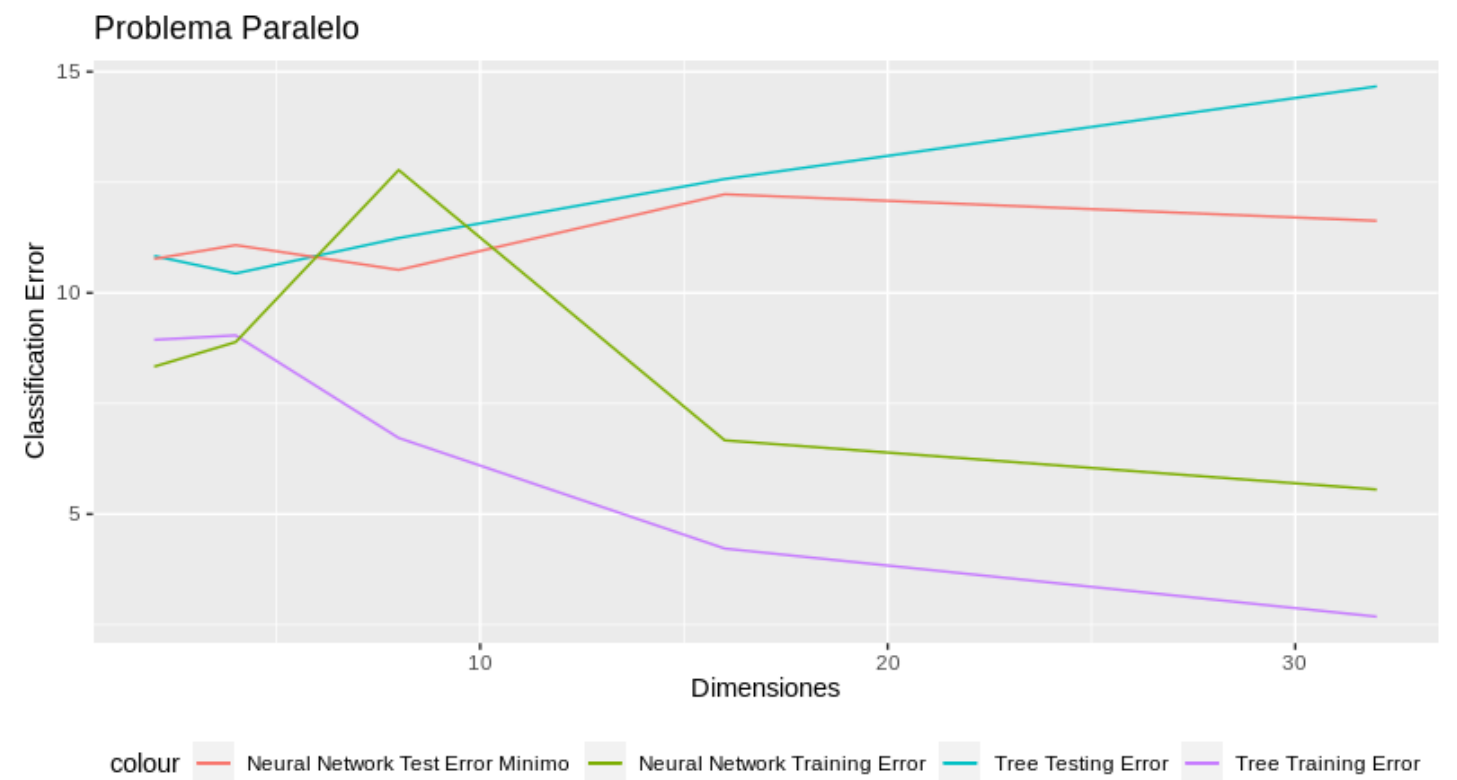


Figura 4.2

Obtenemos mejores resultados con las redes neuronales que con los árboles de decisión debido a que las redes neuronales tienen mayor resistencia al ruido que estos últimos, característica la cual es muy importante en estos dos problemas por la cantidad de error que sus datos contienen.

Además, podemos apreciar como cuando aumenta la dimensión del problema, el error de test en las redes neuronales se mantiene relativamente estable, caso que no sucede con el error de testing utilizando árboles de decisión.