

Redes Neuronales

Trabajo práctico 2

Arlanti - Minestrelli - Ocamica - Hurmuz

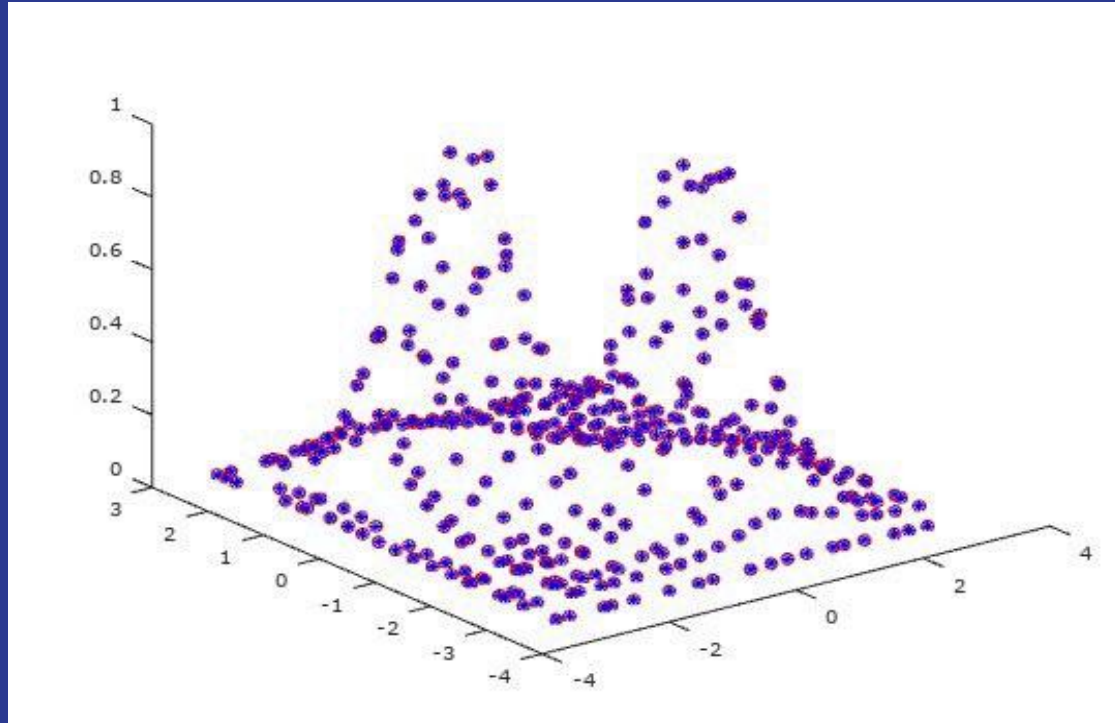
Objetivo

Objetivo

- Construir una red neuronal para que aprenda a generalizar un terreno para un videojuego.
- Analizar distintas configuraciones de red
- Evaluar variantes del algoritmo backpropagation
- Llegar a la arquitectura óptima para este problema



Terreno a Aprender

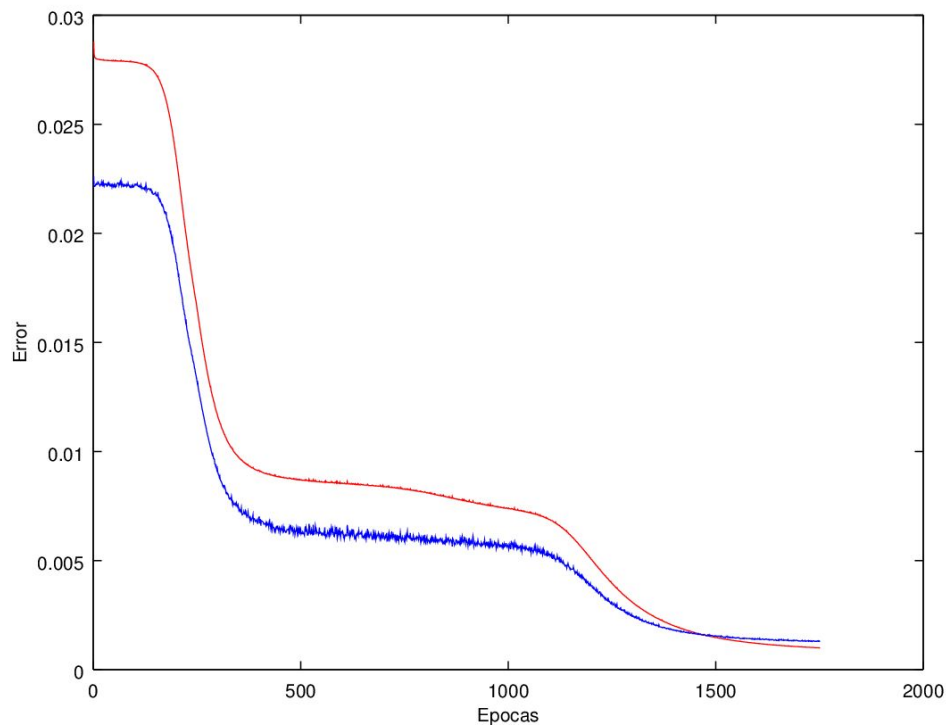


Entradas

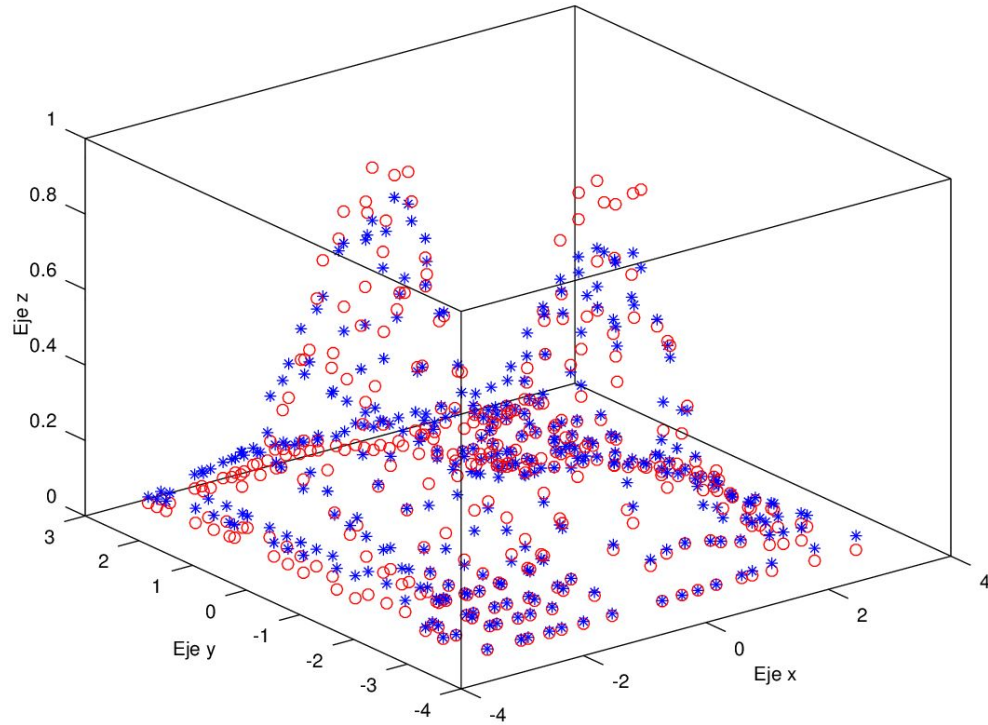
- La elección de los patrones fue al azar siguiendo una proporción de entrenamiento - testeo.
- Conjunto de entrenamiento 80%
- Conjunto de testeo 20%



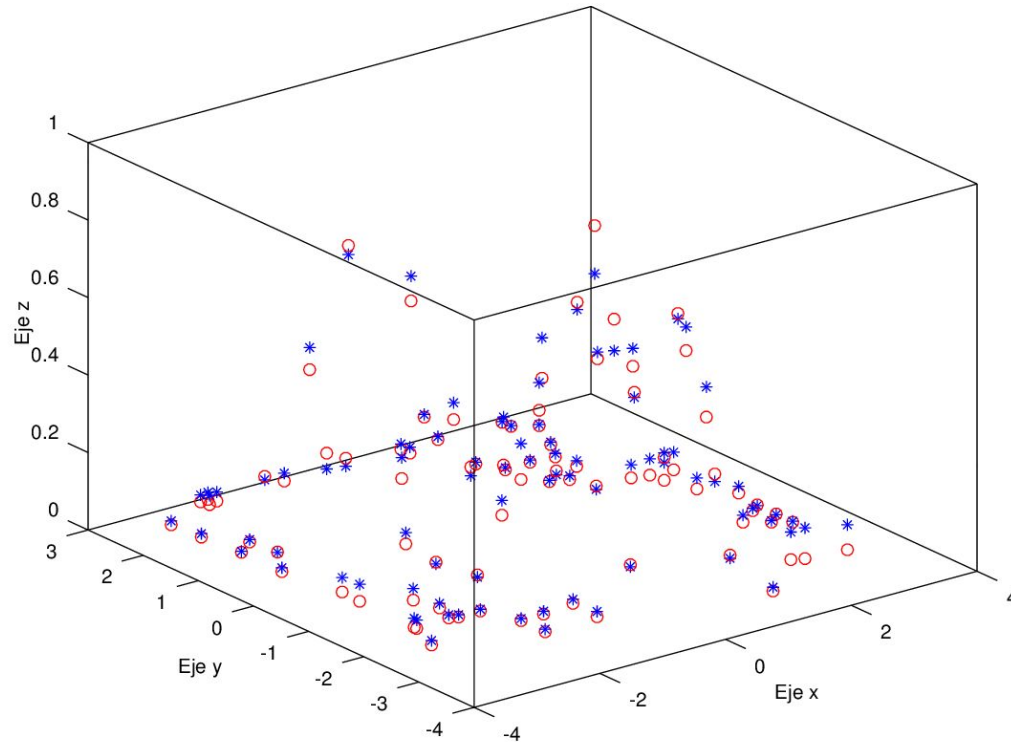
Error de Testeo vs Error de Aprendizaje



Conjunto de Entrenamiento



Conjunto de Testeo - Generalización





Arquitectura

Arquitecturas candidatas

- 1 Capa Oculta
- 2 Capas Ocultas
- 3 Capas Ocultas



Algunas Arquitecturas probadas

- 2, 50, 1
- 2, 10, 1
- 2, 5, 1
- 2,35,10,1
- 2,10,35,1
- 2, 4, 4, 1
- 2, 10, 10, 10, 1
- 2, 30, 10, 30, 1



Métricas tomadas en cuenta

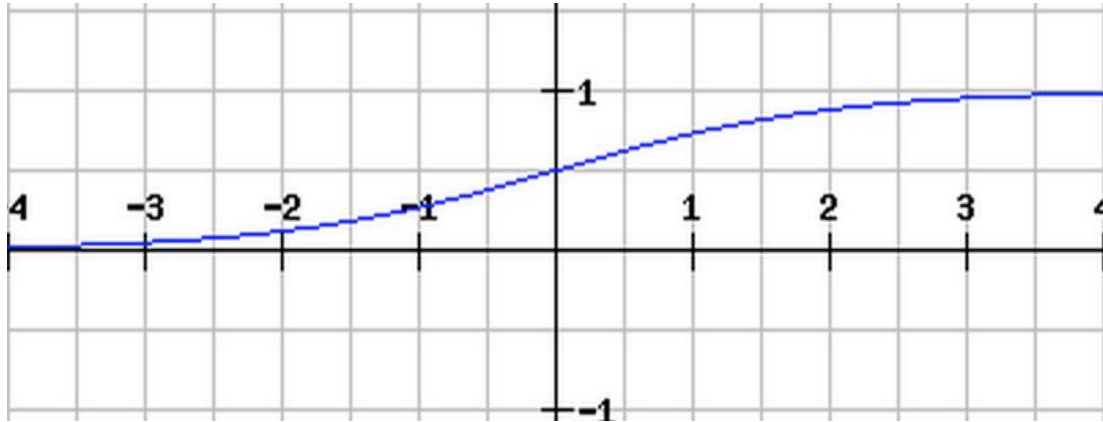
- Cota de error cuadrático medio de aprendizaje
- Error cuadrático medio de testeo
- Tiempo
- Épocas



Elección de Función de Activación

Exponencial

Imagen 0 al 1



Variación de β

Variación de β

- A medida que se incrementa β , la función de activación tiende a la función escalón por lo que su salida da valores menos distinguibles y más cercanos a $|1|$.
- Por ello, a la red le cuesta más aprender, por el contrario si β fuera pequeño la red no podría distinguir los patrones de forma adecuada.
- Mediante las pruebas realizadas se llegó a que el valor que mejor performaba era $\beta = 0,5$





Mejoras Backpropagation

Mejoras backpropagation

- Momentum
- Ajuste de η




Comparación de optimizaciones

Comparación de optimizaciones

- Sin mejoras
- Ajustando el δ
- Con momentum
- Con momentum y δ





Elección de arquitectura

Criterio

- Aquella que generalizó mejor, sin importar cuánto haya tardado en aprender.
- Se tomó la mejor de 1, 2 y 3 capas ocultas con y sin mejoras.



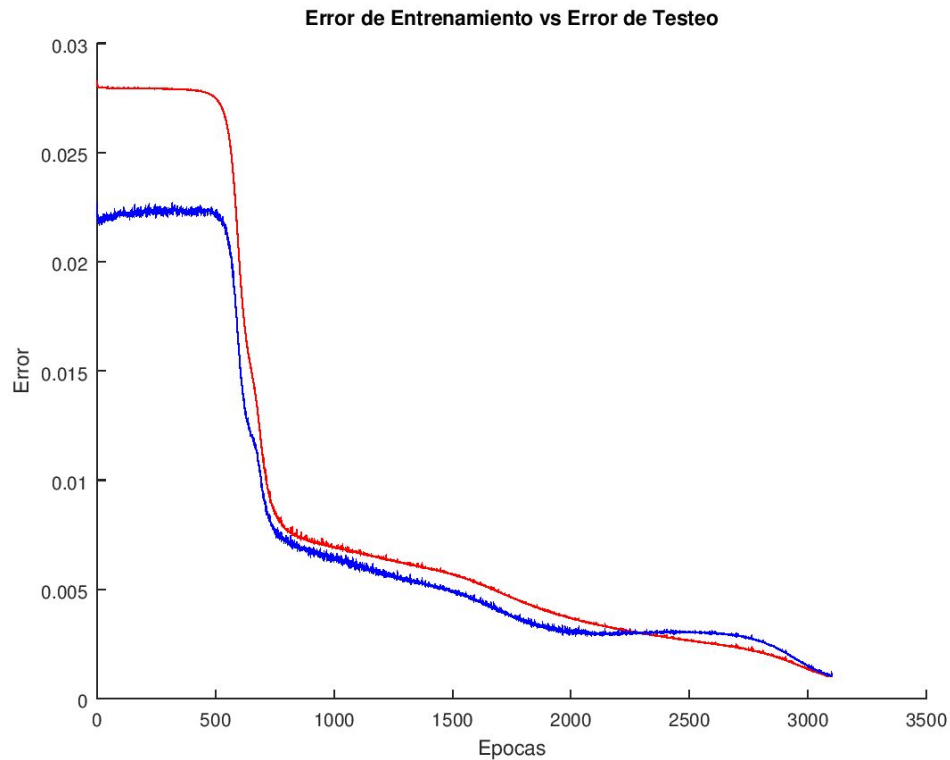
Sin mejoras

Capas	1 Oculta	2 Ocultas	3 Ocultas
Arquitectura	2,50,1	2,35,10,1	2,30,10,30,1
Épocas	3986	3098	1811
Error Cuadrático Medio Testeo	5.6847e-04	5,9098E-04	9,3602E-04

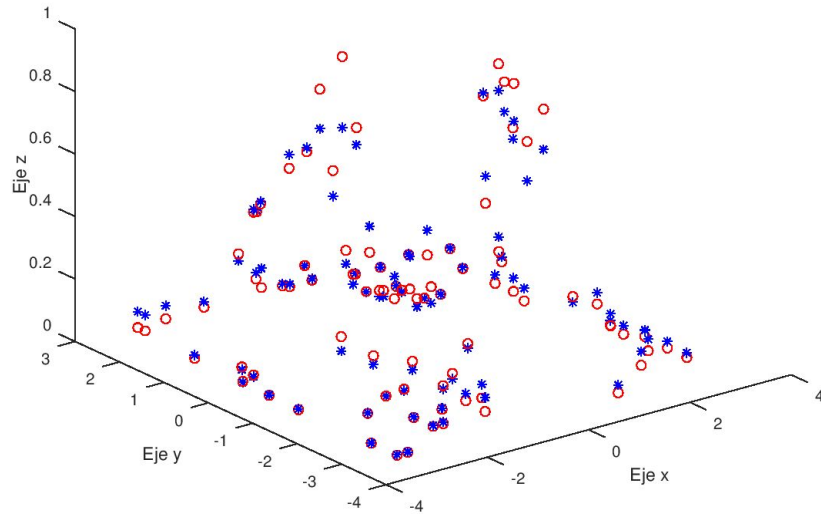
Con Delta y Momentum

Capas	1 Oculta	2 Ocultas	3 Ocultas
Arquitectura	2,50,1	2,35,10,1	2,30,10,30,1
Épocas	3466	2303	2382
Error Cuadrático Medio Testeo	4,456E-04	0,001221	0,0010997

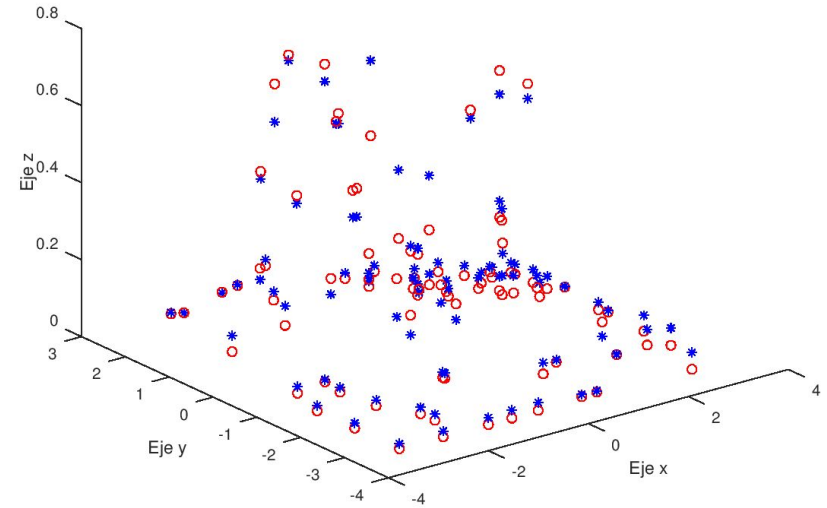
Gráfico de error configuración óptima



[2, 10, 1] con
optimizaciones



[2, 50, 1] con
optimizaciones



Conclusiones

- Conjunto de entrenamiento mayor a 80% → No generaliza bien
- Conjunto de entrenamiento menor a 80% → No aprende bien
- Momentum con $\alpha = 0.9$ → Generaliza Mejor
- Arquitectura elegida 1 capa oculta con 50 neuronas (2,50,1) → Generalizó mejor

