



# PERCEPTRON SIMPLE Y MULTICAPA

---

Comerci, Nicolás  
Lin, Scott  
Piñeiro, Eugenia

# 01

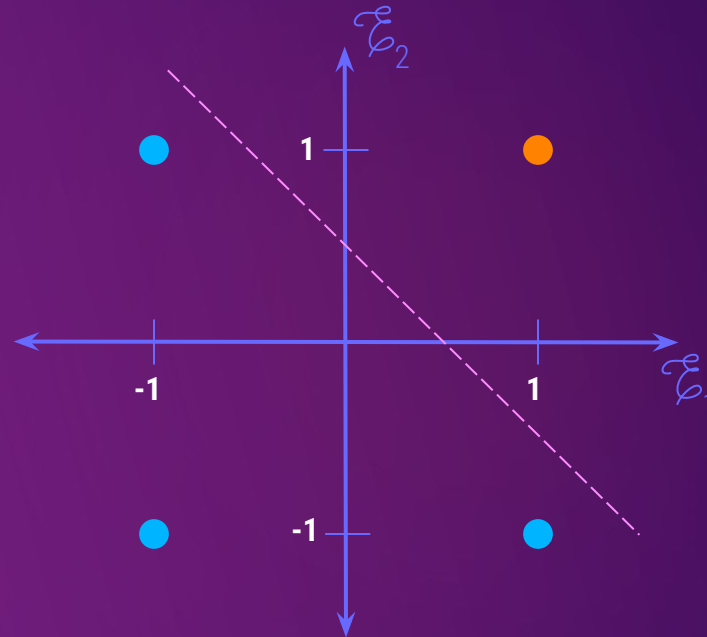
## PERCEPTRÓN SIMPLE ESCALÓN

---

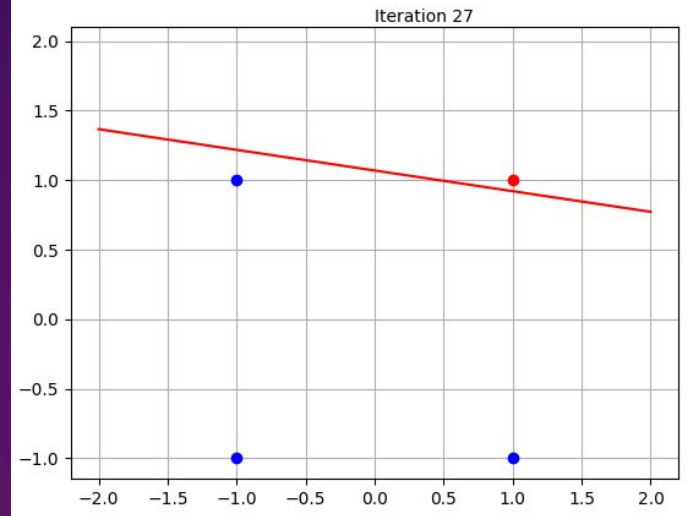
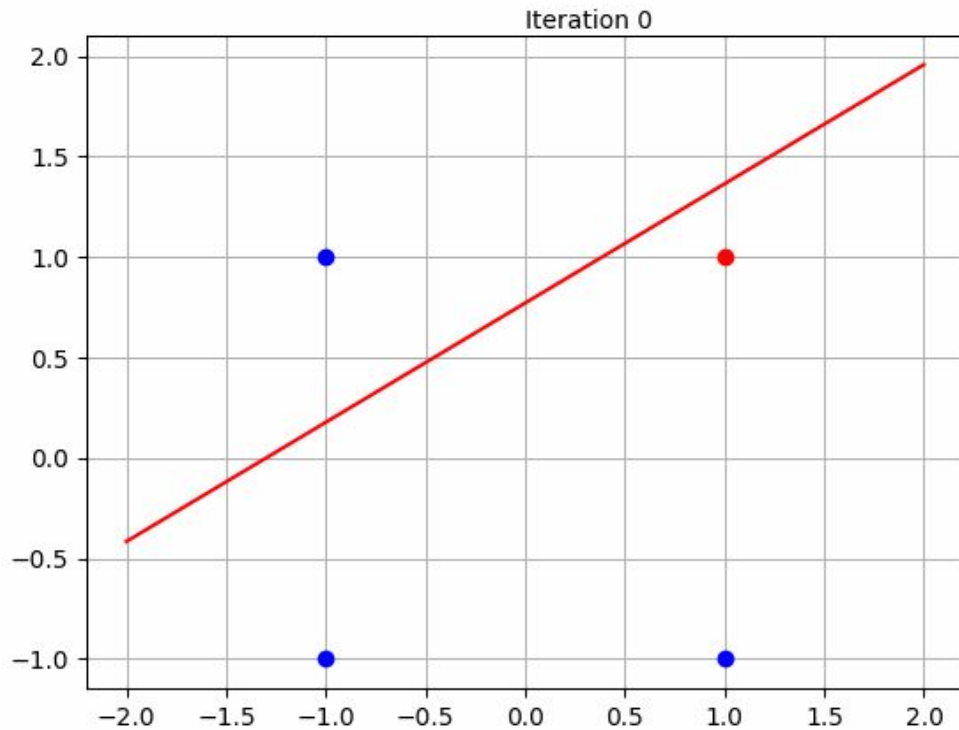


# AND

Problema linealmente separable

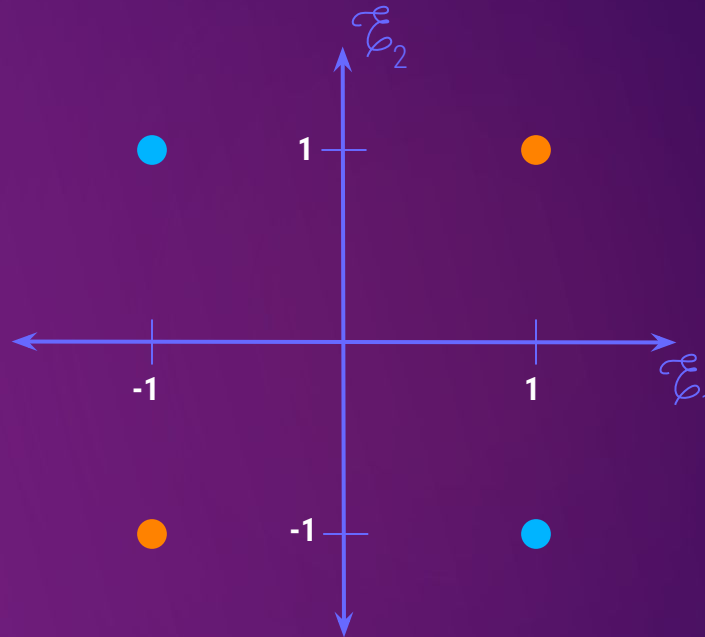


# AND



# XOR

NO es linealmente separable



# 02

## PERCEPTRÓN SIMPLE LINEAL Y NO LINEAL

---



# CÁLCULO DEL ERROR

Sea N la longitud del conjunto de datos esperados:

$$E = \frac{\sum_{i=0}^N |\zeta - 0|}{N}$$

# CAPACIDAD DE APRENDER

PERCEPTRÓN LINEAL	
TASA DE APRENDIZAJE	ERROR
0,0001	37,50995622
0,0300	10,13364206
0,0400	16,48389948
0,0700	75,50856178
0,0800	9,40E+21
0,1000	7,74E+82
0,2000	3,10E+295
<b>&gt; 0.2</b>	<b>inf</b>

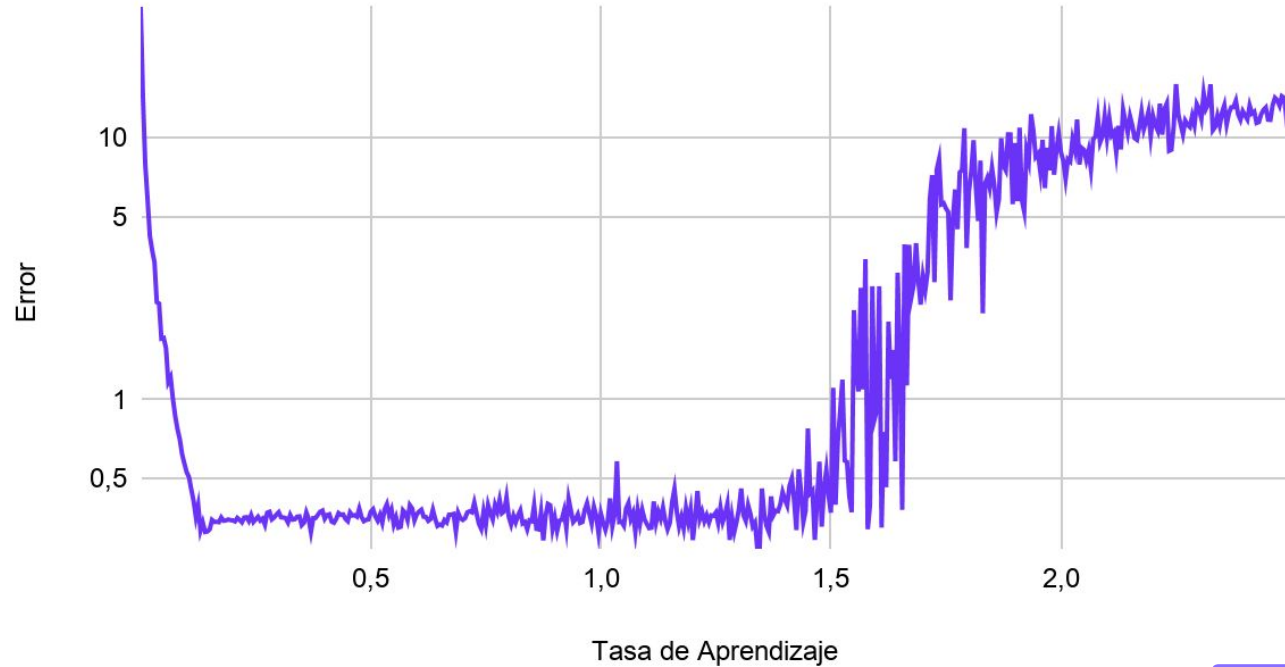
NO  
APRENDE

Promedio tras 4 ejecuciones



# CAPACIDAD DE APRENDER

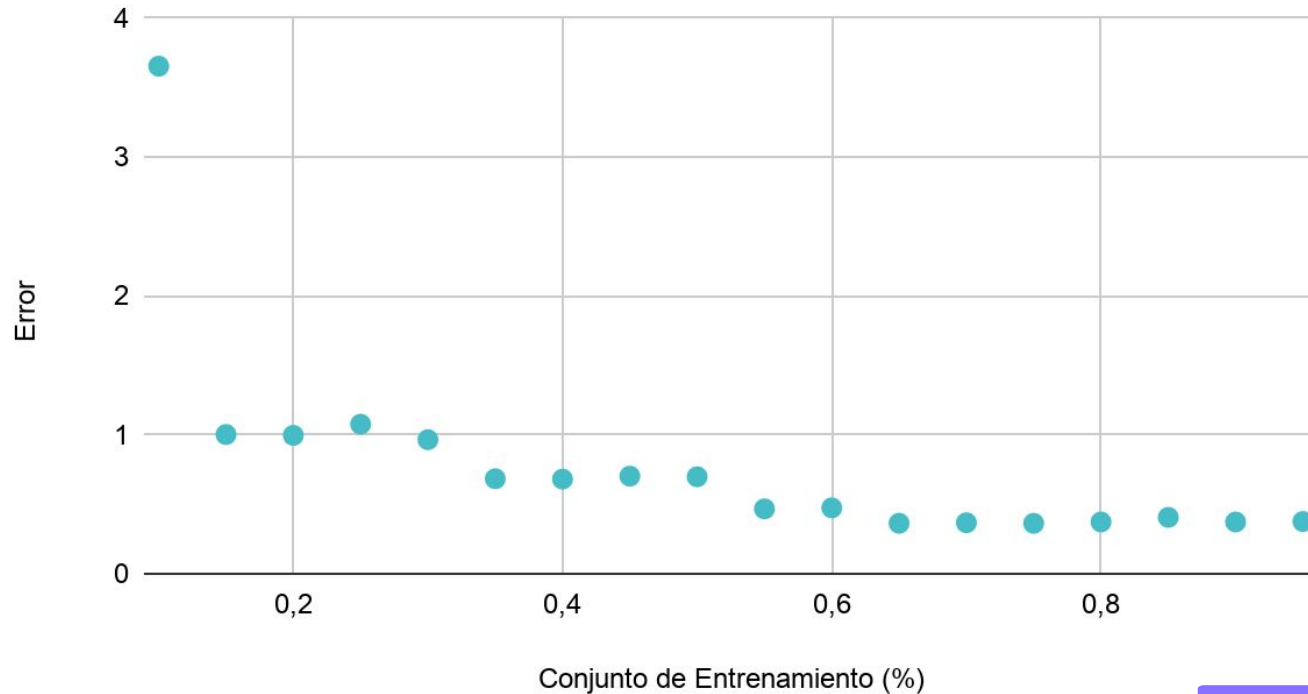
## Perceptrón No Lineal



Promedio tras 4 ejecuciones

# CAPACIDAD DE GENERALIZACIÓN

## Perceptrón No Lineal



Promedio tras 10 ejecuciones

# MÁXIMA CAPACIDAD DE GENERALIZACIÓN

**ENTRENAMIENTO**  
**65 %**

**GENERALIZACIÓN**  
**35 %**

# NO LINEAL

TAZA DE  
APRENDIZAJE

0.0001 a 0.05

ENTRENAMIENTO  
GENERALIZACIÓN

0.7  
0.3

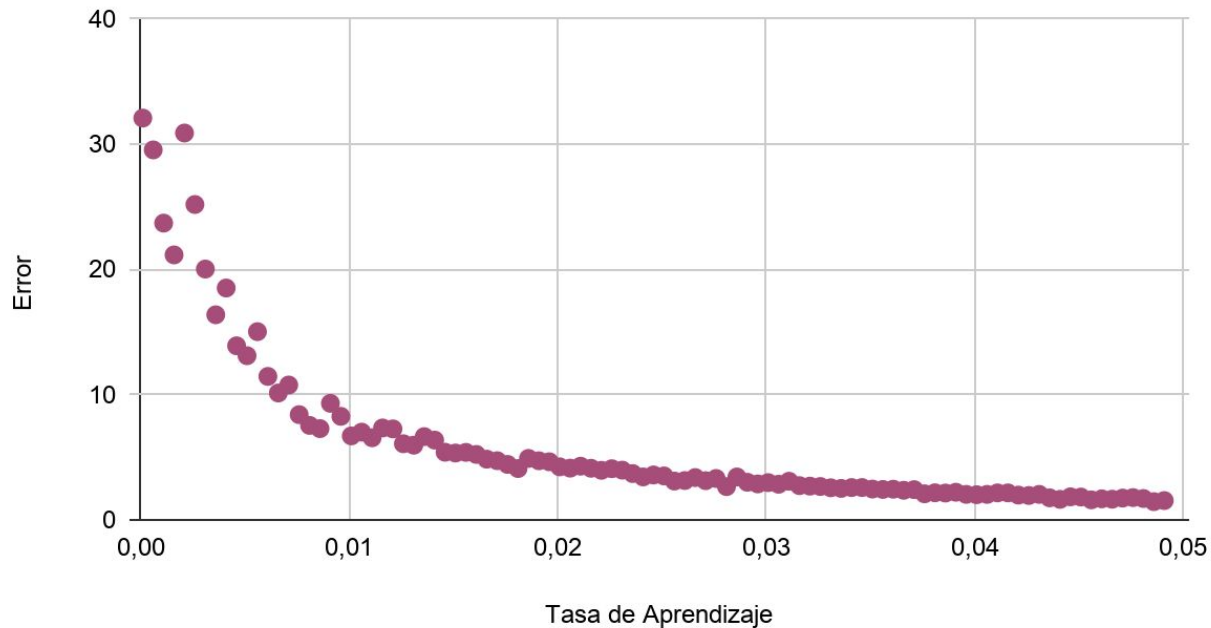
ITERACIONES

1000

**Búsqueda Lineal**

**$0.038 < \alpha < 0.062$**

No Lineal



Promedio tras 4 ejecuciones



# MEJOR CONJUNTO DE ENTRENAMIENTO

## Método de Validación Cruzada

**VARIANTE:** En vez de métricas estándar se usó el error

# 03

## PERCEPTRÓN MULTICAPA

---



TAZA DE  
APRENDIZAJE

0.03

ENTRENAMIENTO  
GENERALIZACIÓN

1  
0

ÉPOCAS

1000

CAPAS OCULTAS

[ 10, 10 ]

ETA  
ADAPTATIVO

No utilizado

MOMENTUM

No utilizado

IMPLEMENTACIÓN

Incremental

# XOR

DATASET	$\zeta$	O
( -1 ; 1 )	1	0.94495444
( 1 ; -1 )	1	0.97152837
( -1 ; -1 )	-1	-0.94585528
( 1 ; 1 )	-1	-0.96588157

ERROR  $\approx$  0,0429

TAZA DE  
APRENDIZAJE

0.03 (base)

ENTRENAMIENTO

100%

ÉPOCAS

5000

CAPAS OCULTAS

[ 10 ]

ETA  
ADAPTATIVO

INCREMENTO 0.2  
DECREMENTO 0.06  
ITERACIONES 2

MOMENTUM

No utilizado

IMPLEMENTACIÓN

Incremental

DATASET

$\zeta$

O

0

1

0.9641307

1

-1

-0.968186

2

1

0.9537611

3

-1

-0.922479

4

1

0.9987612

5

-1

-0.996345

6

1

0.9915962

7

-1

-0.994246

8

1

0.9068315

9

-1

-0.999064

MAPA DE  
PIXELES

ERROR  $\approx$   
0,0304



# CAPACIDAD DE GENERALIZACIÓN

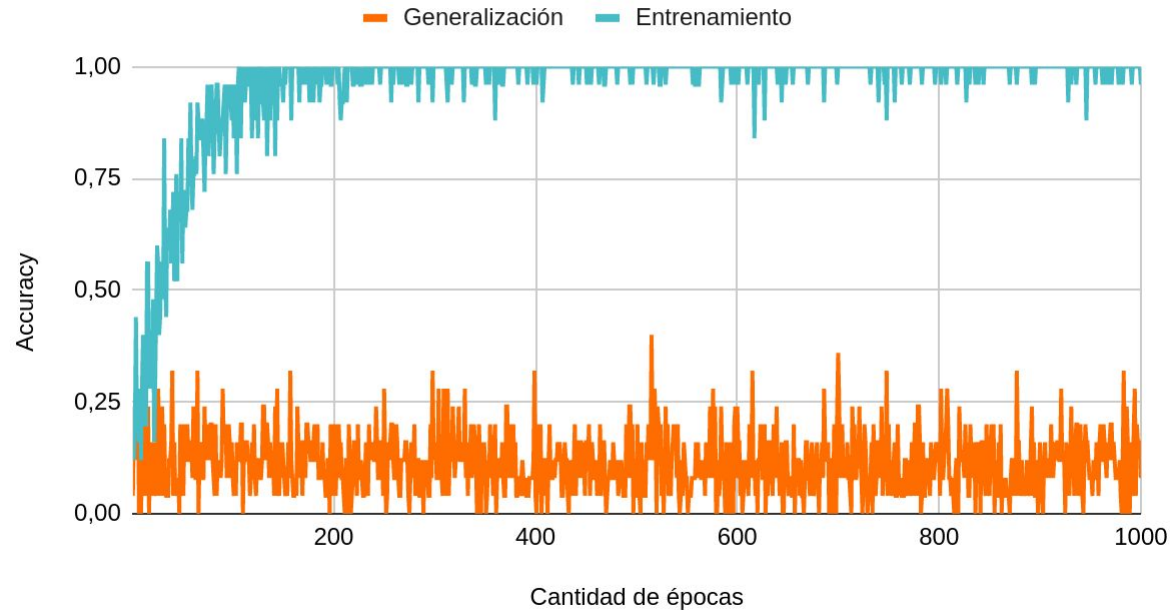
## MATRIZ DE CONFUSIÓN - MAPA DE PÍXELES

	PAR (Predecido)	IMPAR (Predecido)
PAR (Real)	VERDADERO POSITIVO	FALSO NEGATIVO
IMPAR (Real)	FALSO POSITIVO	VERDADERO NEGATIVO

Dataset: Mapa de Píxeles

# CAPACIDAD DE GENERALIZACIÓN

Mapa de píxeles (umbral = 0.08)



CAPAS  
OCULTAS

[ 10,10 ]

TAZA DE  
APRENDIZAJE

0.03

BATCH  
ETA ADAPTATIVO  
MOMENTUM

No utilizado

Promedio tras 5 ejecuciones

# CAPACIDAD DE GENERALIZACIÓN

Entreno solo con impares  
Me devuelve solo impares

Testing Set	$\zeta$	O
0	1	-0.99944
2	1	-0.99654
4	1	-0.99494
6	1	-0.99494
8	1	-0.99921

Entreno solo con pares  
Me devuelve solo pares

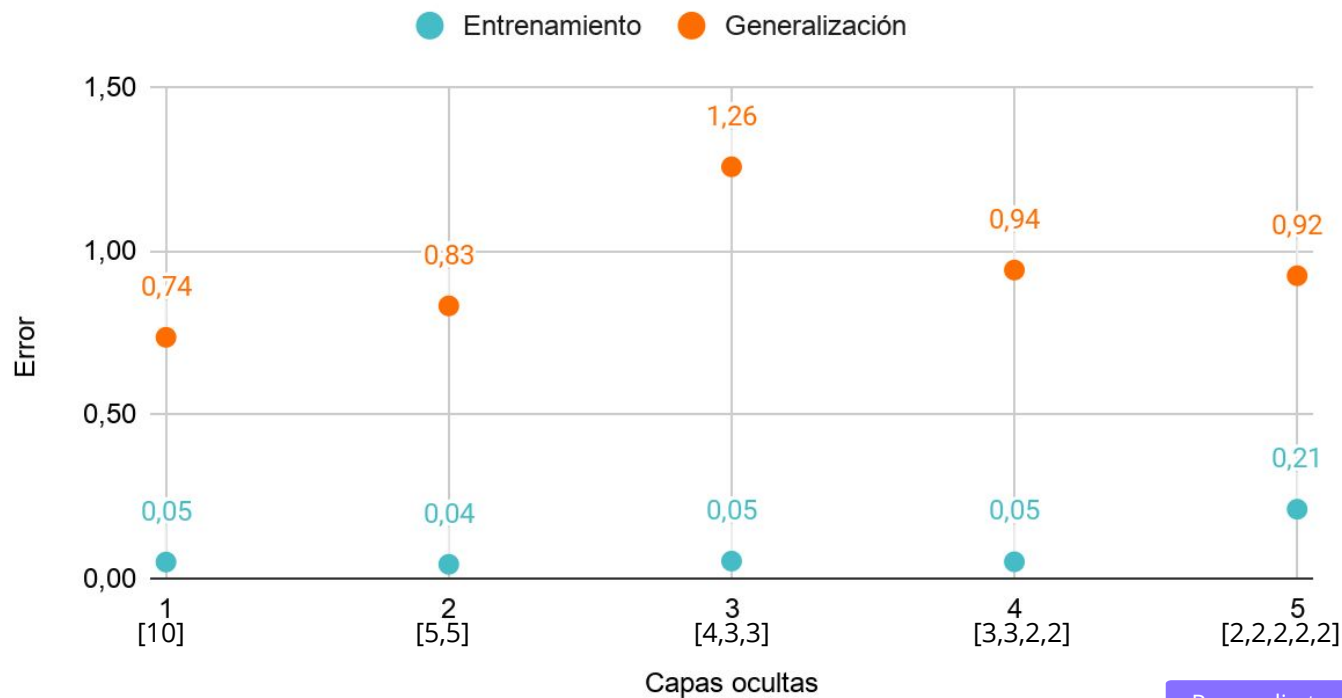
Testing Set	$\zeta$	O
1	-1	0.95995
3	-1	0.96855
5	-1	0.99311
7	-1	0.99805
9	-1	0.76661



# **EFFECTO DE CAMBIO DE PARÁMETROS**

# CAPAS OCULTAS

## Mapa de Píxeles



Promedio tras 4 ejecuciones

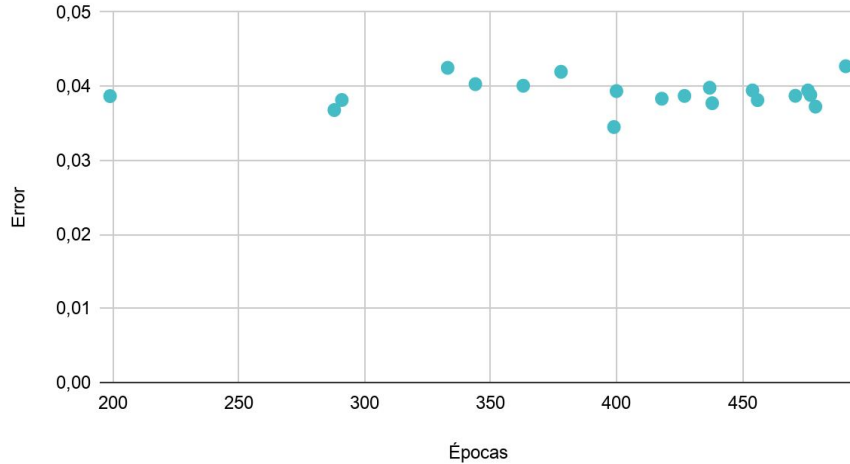
# BATCH vs INCREMENTAL

CAPAS OCULTAS

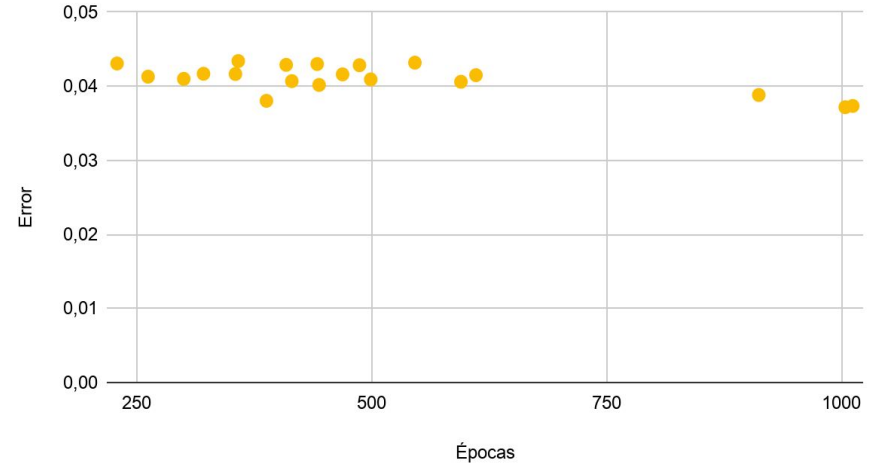
[ 10, 10 ]

ÉPOCAS  $\approx$  395    ERROR  $\approx$  0.039

Batch



Incremental

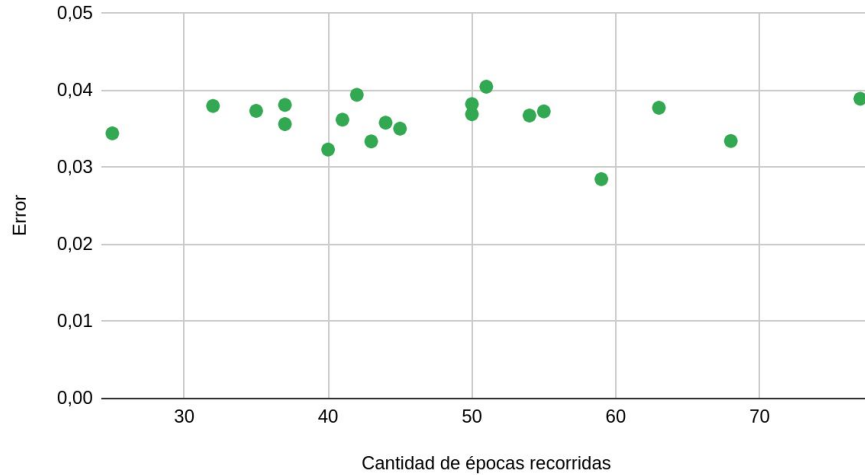


ÉPOCAS  $\approx$  502    ERROR  $\approx$  0.041

Se hicieron 20 ejecuciones  
Dataset: Mapa de Píxeles

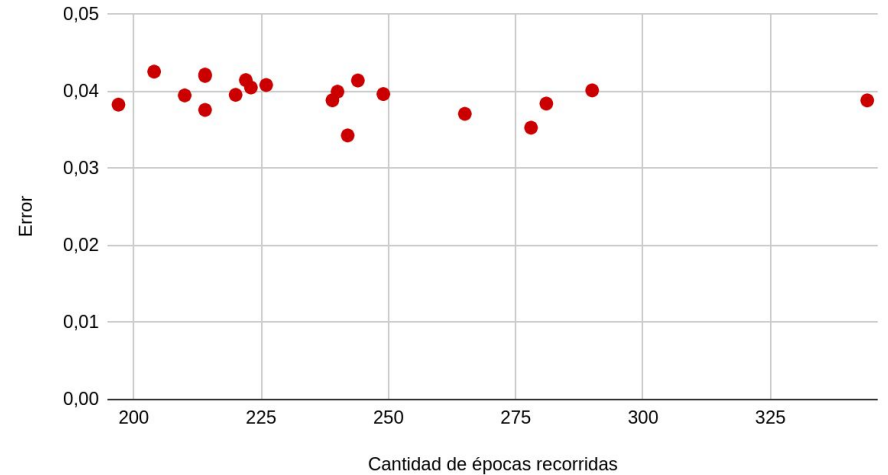
# MOMENTUM

Utilizando Momentum



ÉPOCAS ~= 47    ERROR ~= 0.036

Sin utilizar Momentum



ÉPOCAS ~= 241    ERROR ~= 0.039

CAPAS  
OCULTAS

[ 10 ]

TAZA DE  
APRENDIZAJE

0.03

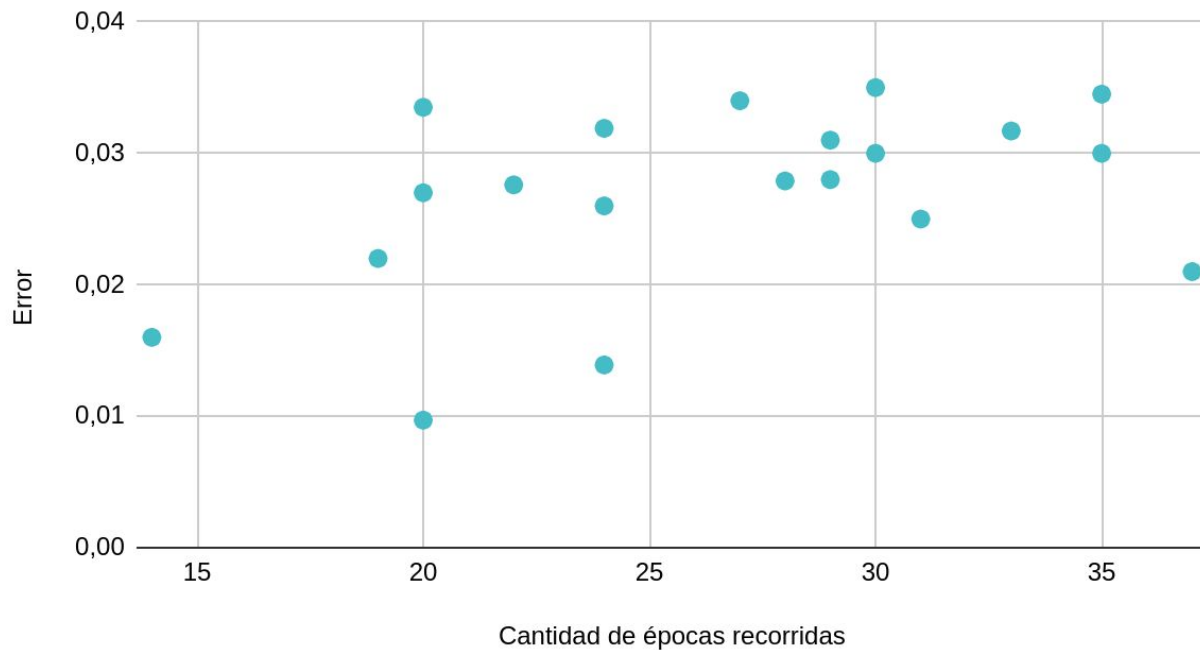
BATCH

No utilizado

Se hicieron 20 ejecuciones  
Dataset: Mapa de Píxeles

# MOMENTUM + BATCH

Batch y Momentum



CAPAS  
OCULTAS

[ 10 ]

TAZA DE  
APRENDIZAJE

0.03

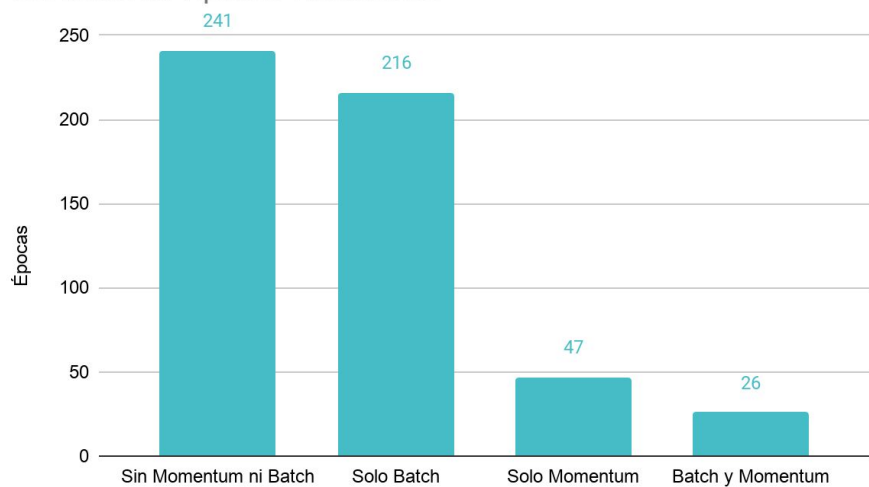
ÉPOCAS  $\approx$  26  
ERROR  $\approx$  0.030

Se hicieron 20 ejecuciones  
Dataset: Mapa de Píxeles



# MOMENTUM + BATCH

Cantidad de Épocas Recorridas



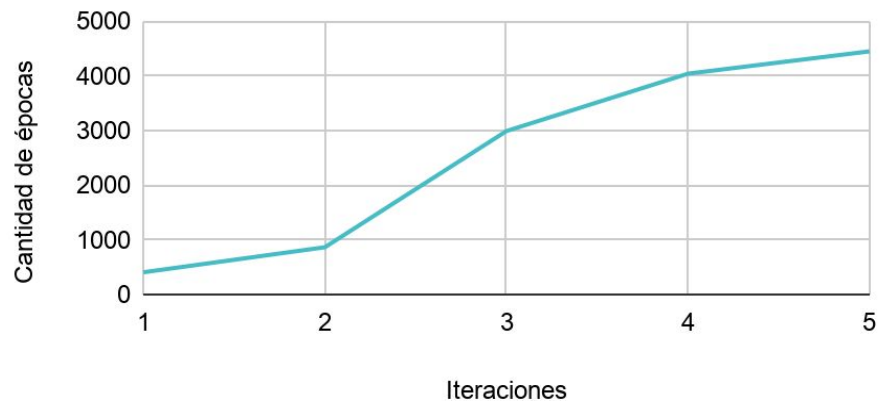
Error



# ETA ADAPTATIVO

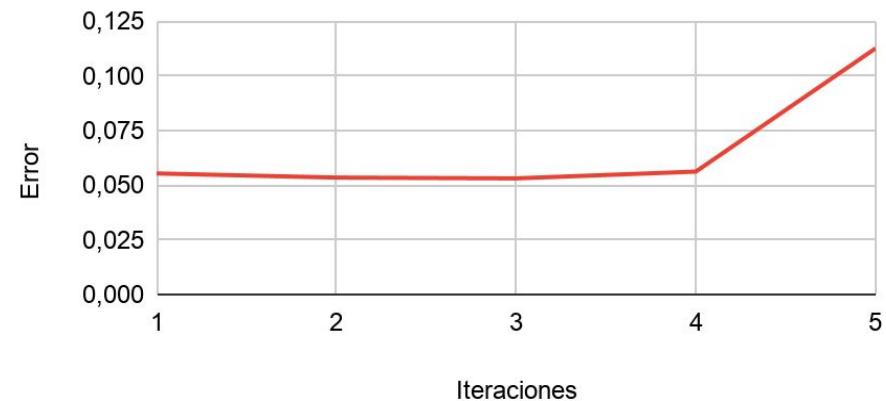
Variación de iteraciones máximas

Épocas



Variación de iteraciones máximas

Error



CAPAS  
OCULTAS

[ 10 ]

TAZA DE  
APRENDIZAJE

0.001 (base)

MOMENTUM  
BATCH

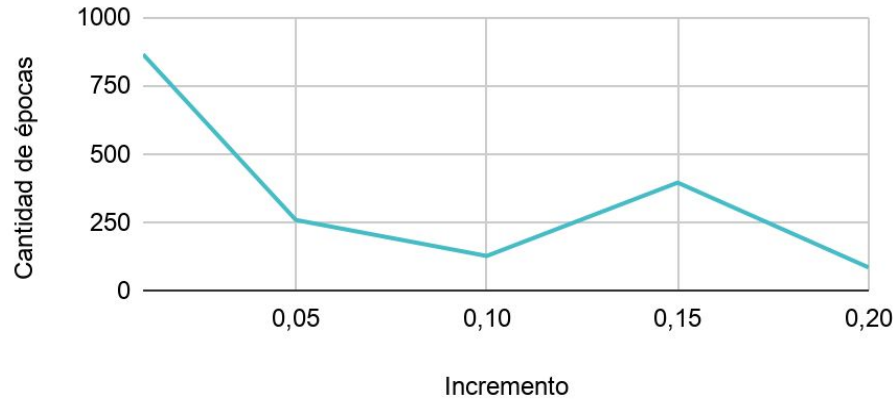
No  
utilizados

Promedio tras 5 ejecuciones  
Dataset: Mapa de Píxeles

# ETA ADAPTATIVO

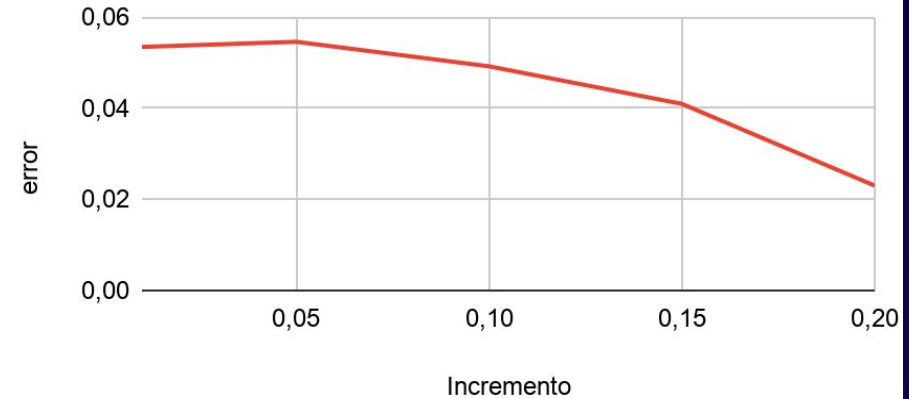
## Variación del incremento

Épocas



## Variación del incremento

Error



CAPAS  
OCULTAS

[ 10 ]

TAZA DE  
APRENDIZAJE

0.001 (base)

MOMENTUM  
BATCH

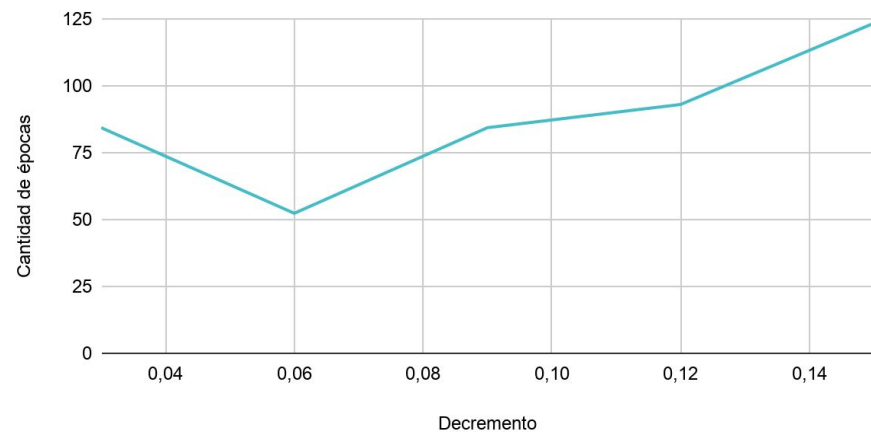
No  
utilizados

Promedio tras 5 ejecuciones  
Dataset: Mapa de Píxeles

# ETA ADAPTATIVO

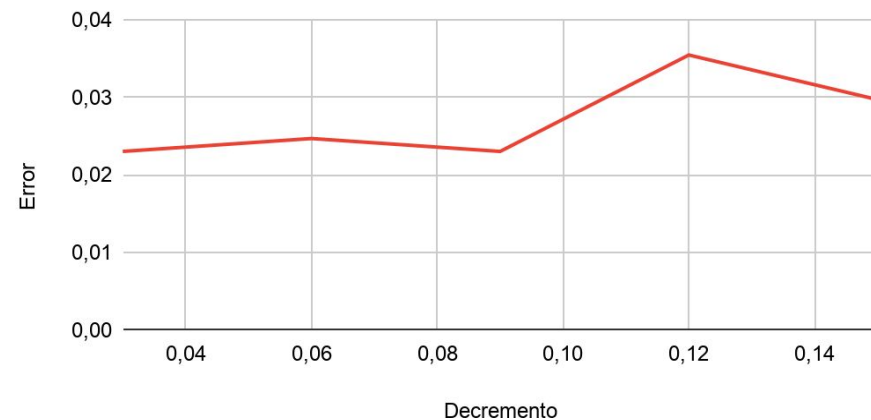
Variación del decremento

Épocas



Variación del decremento

Error



CAPAS  
OCULTAS

[ 10 ]

TAZA DE  
APRENDIZAJE

0.001 (base)

MOMENTUM  
BATCH

No  
utilizados

Promedio tras 5 ejecuciones  
Dataset: Mapa de Píxeles

# ETA ADAPTATIVO

Sea  $\eta$  la tasa de aprendizaje y  $E()$  la función de error:

$\Delta\eta$

+0.2

-0.06 $\eta$

si  $E()$  decrece **2** iteraciones seguidas

si  $E()$  aumenta **2** iteraciones seguidas

# CONCLUSIÓN

Depende del problema el tipo de red a utilizar

Métodos de Optimización: Muy útiles pero sólo para problemas de clasificación

Modificación de parámetros (Batch, Eta adaptativo, Momentum): rápida convergencia

Aumentar cantidad de neuronas y capas ayuda a que aprenda mejor