



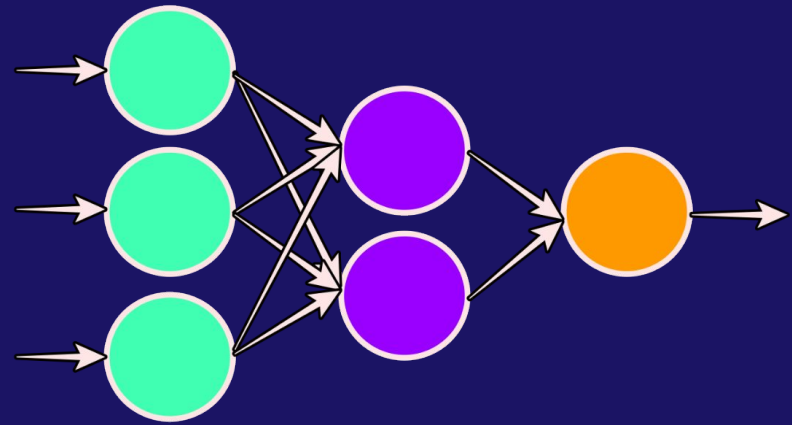
# Sistemas de Inteligencia Artificial

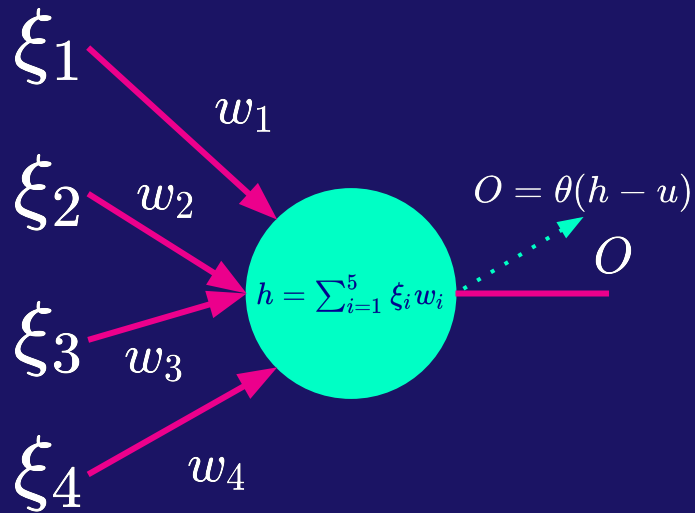
Perceptron simple y multicapa

Franco Baliarda, Joaquín Colonnello, Lucía Torrusio

# INTRODUCCIÓN

- Aprender a resolver problemas utilizando redes neuronales
- Estudiar perceptrones simples y multicapa y los problemas que resuelven
- Evaluar y comparar diferentes parámetros, métodos y métricas
- Analizar su capacidad de generalización

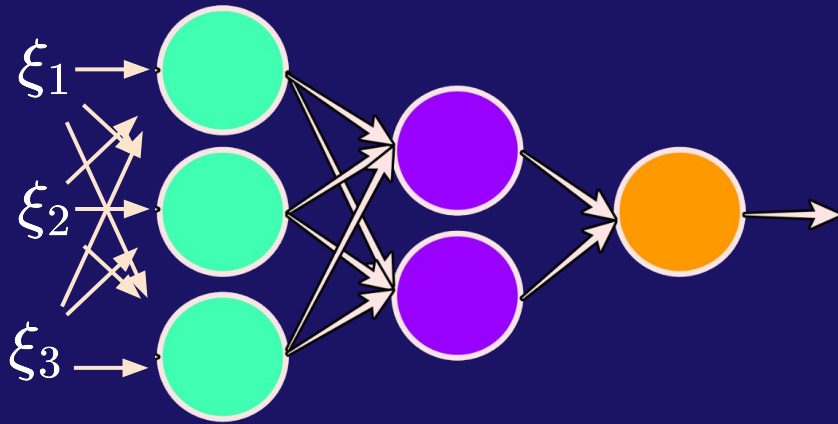




# Perceptrón Simple

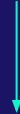
Permite resolver problemas linealmente separables por medio de un hiperplano.





# Perceptron Multicapa

El Perceptrón Multicapa es una extensión y generalización del perceptrón simple en el que se utilizan múltiples neuronas formando capas.



Permite resolver problemas no linealmente separables





# Ejercicio 1

# Aprendizaje

## Perceptrón simple escalón

“γ”

Input =  $\{-1, 1\}, \{1, -1\}, \{-1, -1\}, \{1, 1\}$

Output =  $\{-1, -1, -1, 1\}$

“0 exclusivo”

Input =  $\{-1, 1\}, \{1, -1\}, \{-1, -1\}, \{1, 1\}$

Output =  $\{1, 1, -1, -1\}$

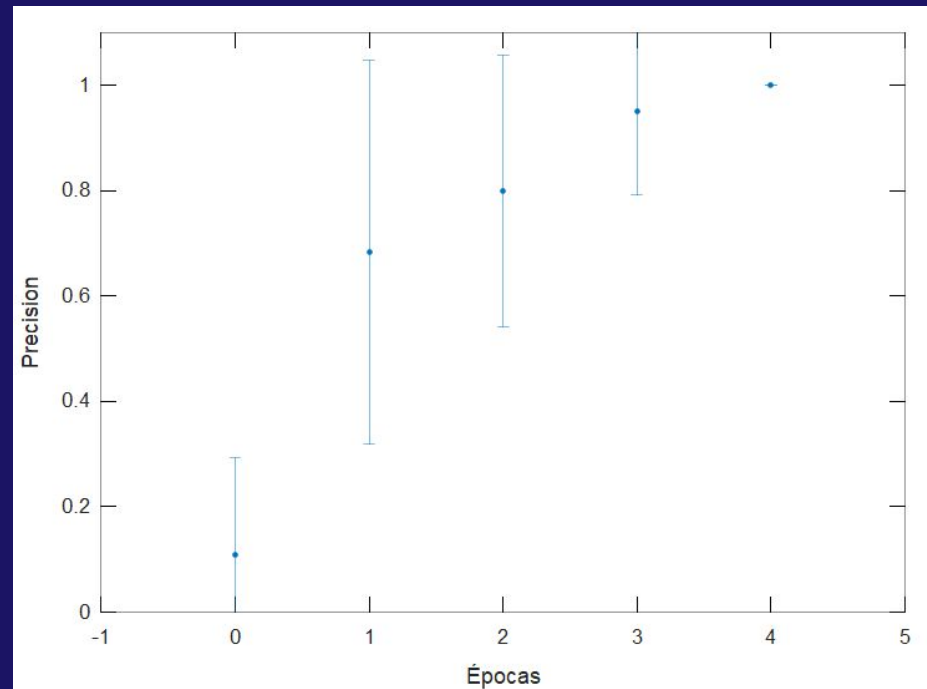
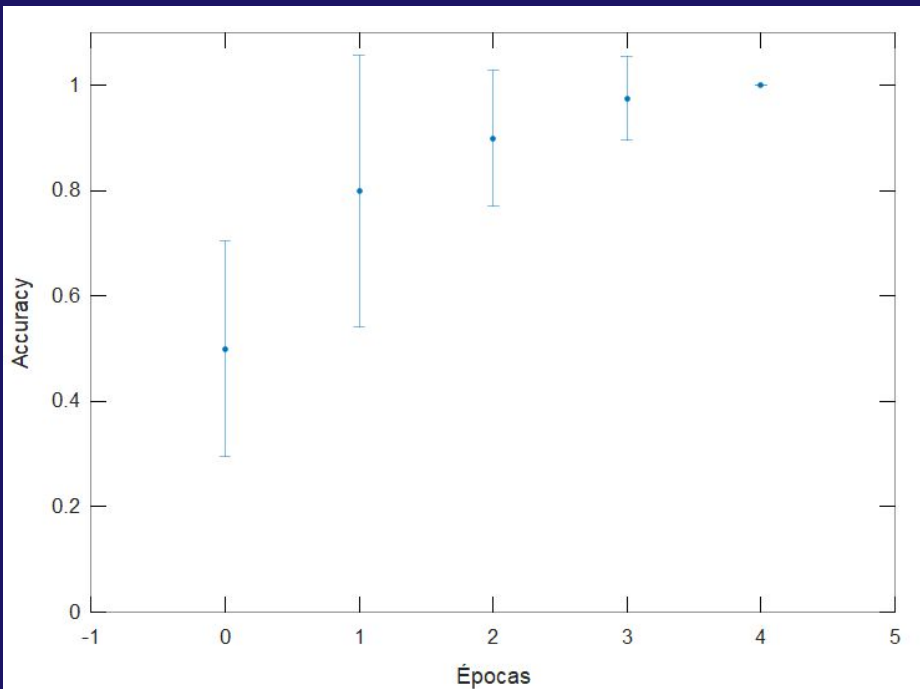
El perceptrón simple escalón conviene utilizarlo cuando hay dos outputs esperados posibles, los cuales serían 1 o -1. → clasifica los puntos en dos clases

# AND

Tasa de Aprendizaje = 0.1

## Accuracy

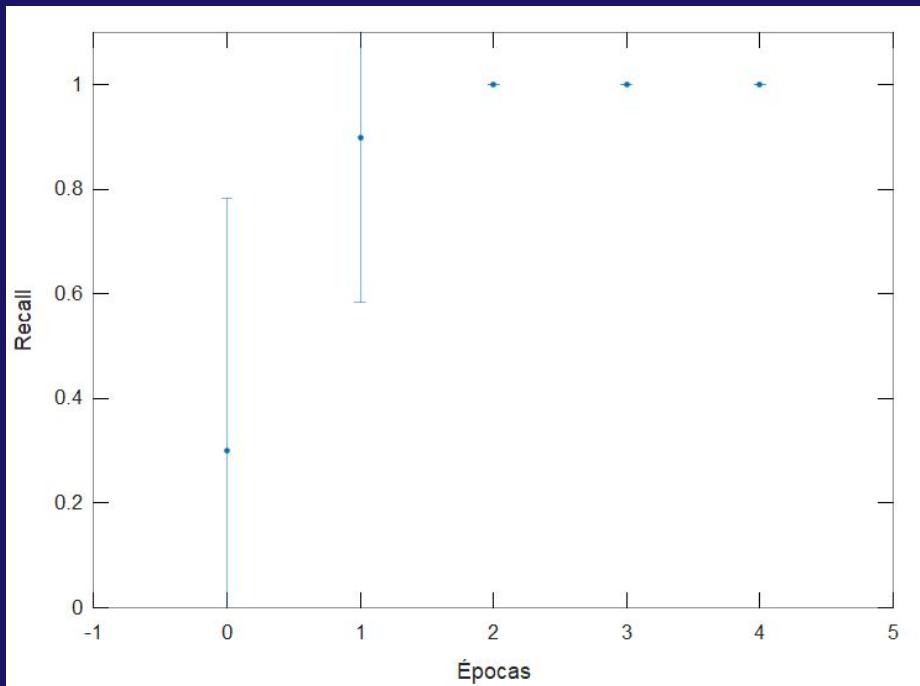
## Precision



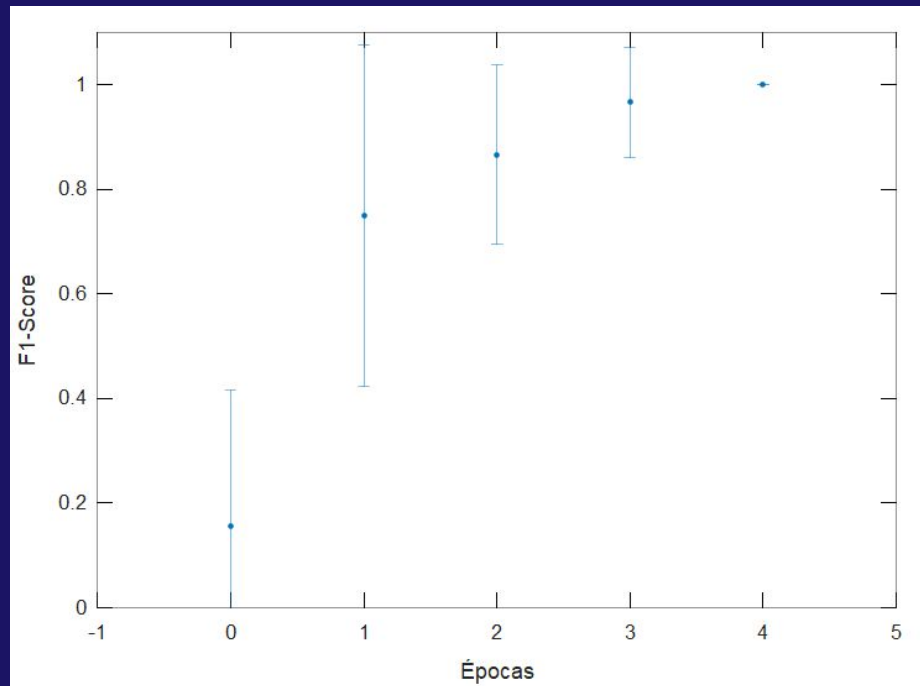
# AND

Tasa de Aprendizaje = 0.1

## Recall



## F1-Score

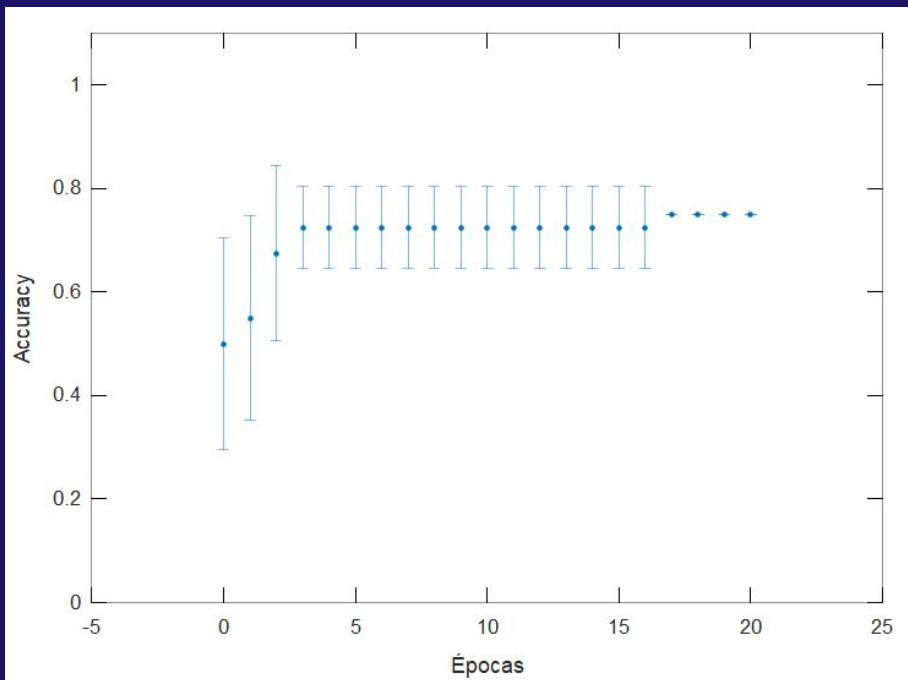




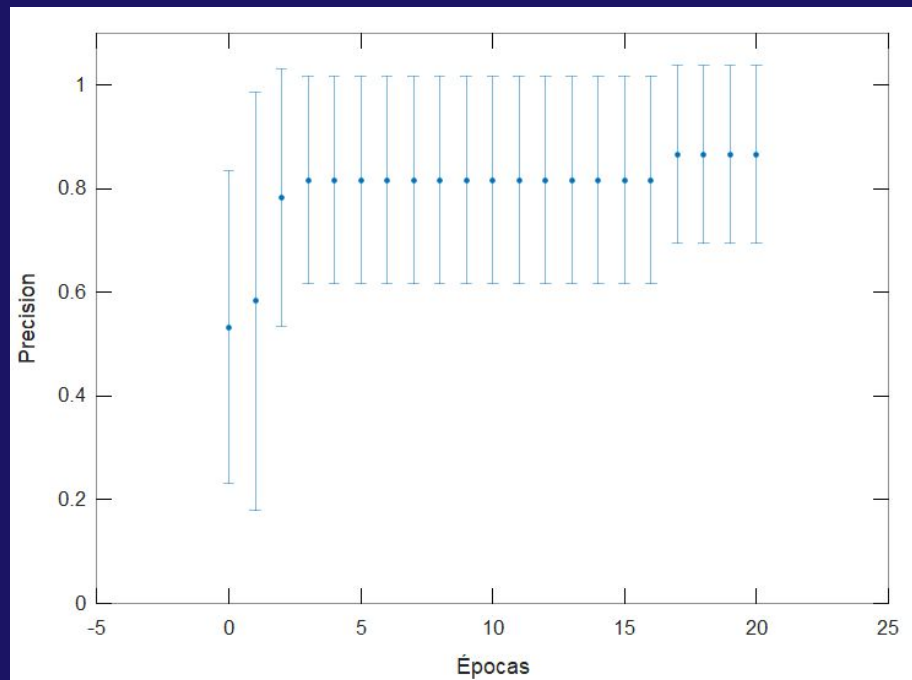
# XOR

Tasa de Aprendizaje = 0.1

## Accuracy



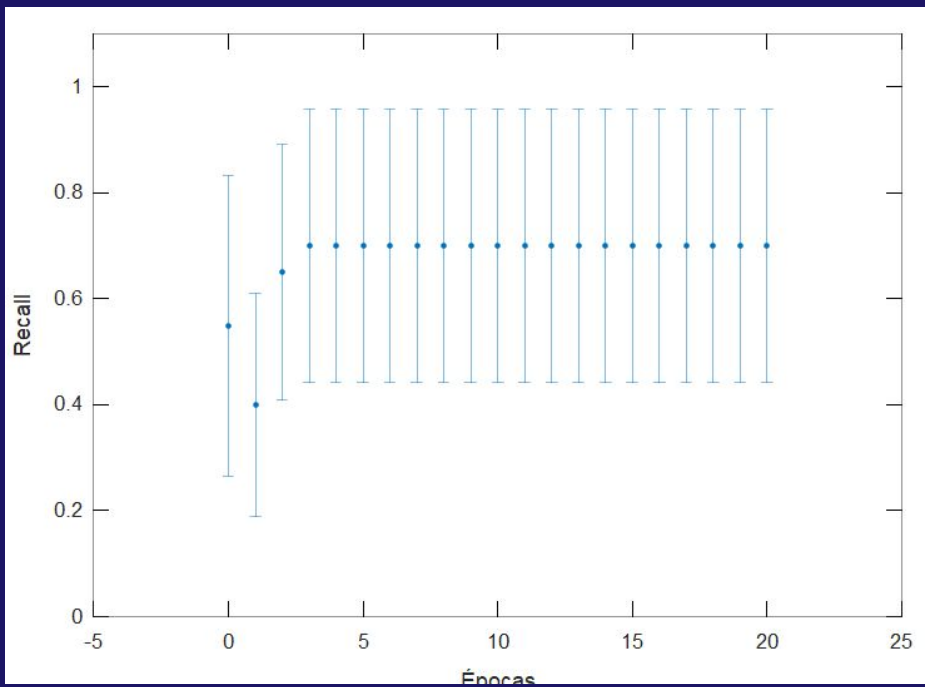
## Precision



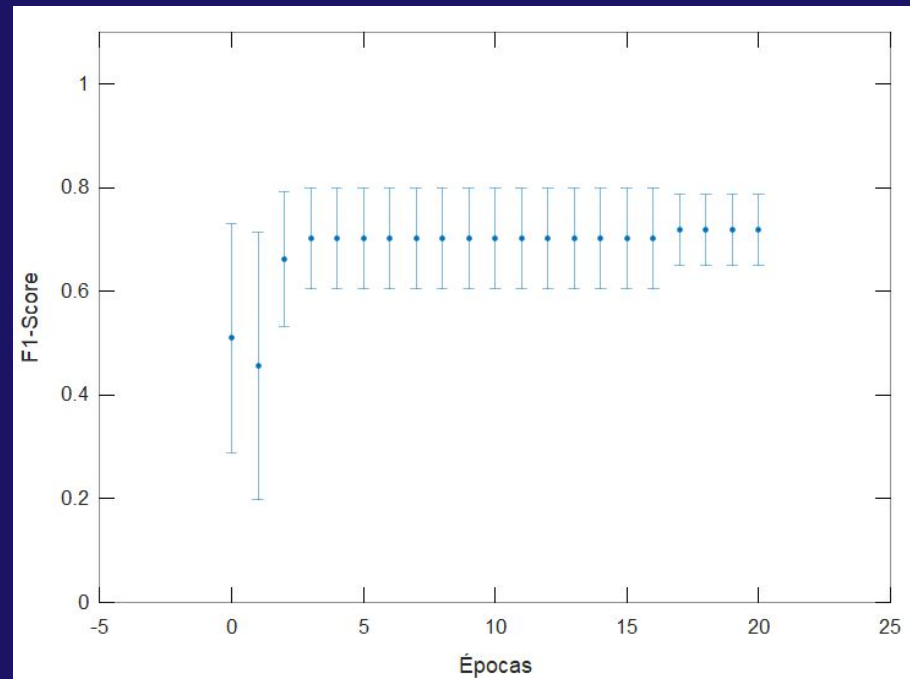
# XOR

Tasa de Aprendizaje = 0.1

**Recall**



**F1-Score**

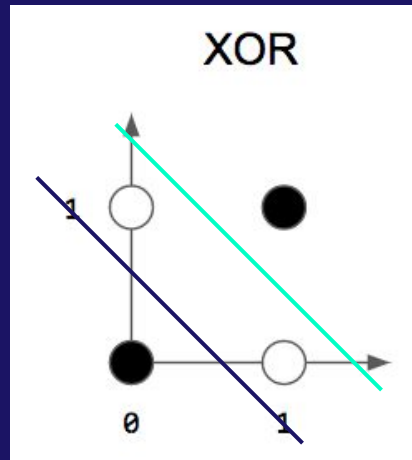
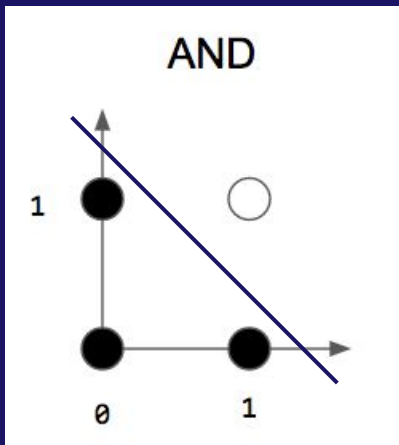


# Perceptron simple escalón

¿Qué puede decir acerca de los problemas que puede resolver el perceptrón simple escalón en relación a la resolución de los problemas que se le pidió que haga que el perceptrón aprenda?



El problema debe ser linealmente separable





# Ejercicio 2

# Perceptrón Simple

**Lineal**

**No lineal**

Aprendizaje de función

$$\mathbb{R}^3 \rightarrow \mathbb{R}$$

- Evaluar la capacidad del perceptrón para aprender la función
- Evaluar la capacidad de generalización del perceptrón

# PERCEPTRON SIMPLE

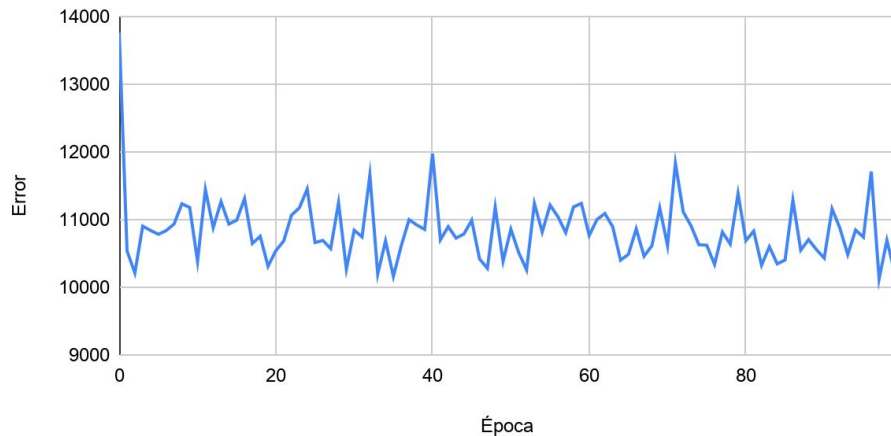
## Aprendizaje

Lineal

No lineal

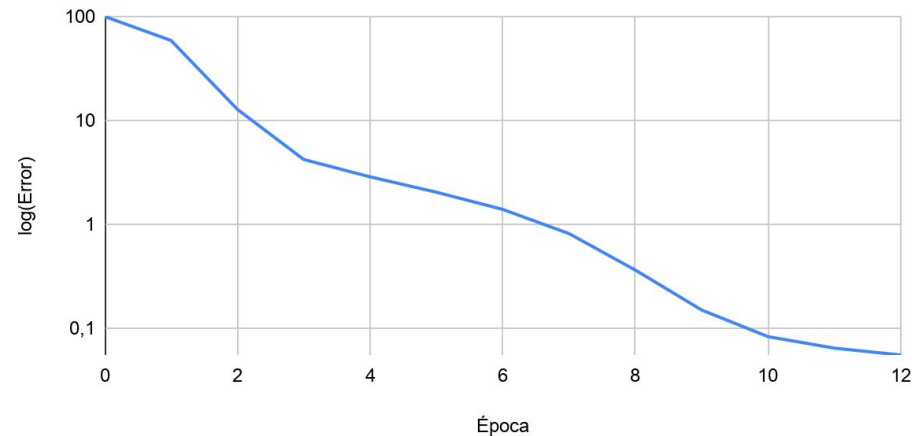
Error por época

Lineal, Learning rate = 0.01



Error por época

No lineal, Learning rate = 0.01



Error Mínimo: 0.021

# Perceptrón Simple

Tasa de aprendizaje adaptable

$$\alpha_k = \arg \min_{\alpha > 0} g(\alpha) = f(x_k + \alpha d_k)$$



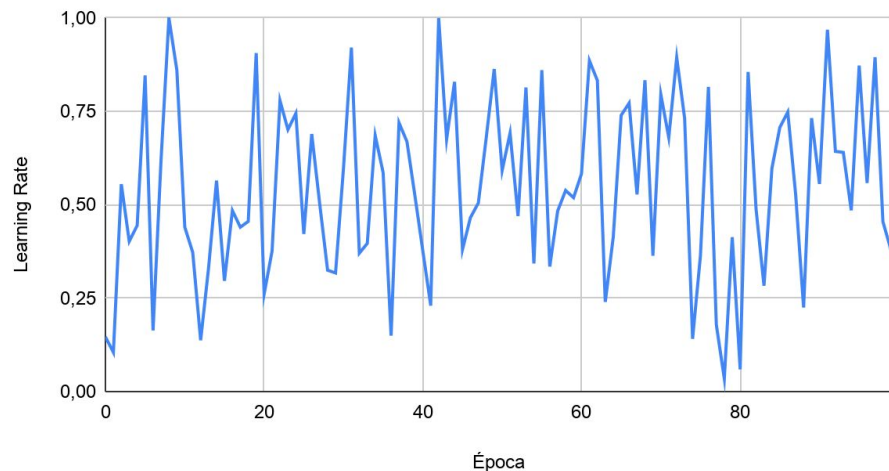
Se busca la tasa de aprendizaje óptima que obtenga el mínimo error en la dirección actual que se está moviendo

se adapta en cada época

# Perceptrón Simple Aprendizaje

No lineal

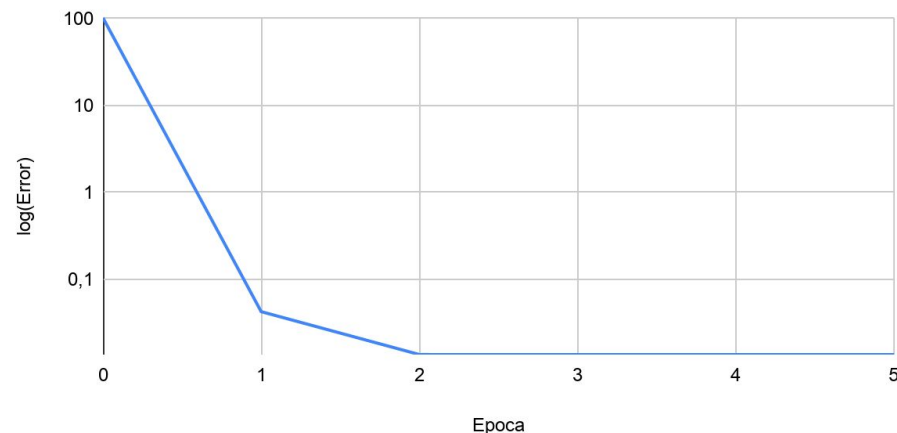
Learning rate adaptation



Error por época

Error Mínimo: 0.00118

No lineal, Learning rate adaptable





# Perceptrón Simple

## Generalización

¿Cómo podría escoger el mejor conjunto de entrenamiento?



### Validación Cruzada

Se utiliza el método de validación cruzada para obtener el conjunto de entrenamiento que más aumenta la capacidad de generalización de la red

# VALIDACIÓN CRUZADA

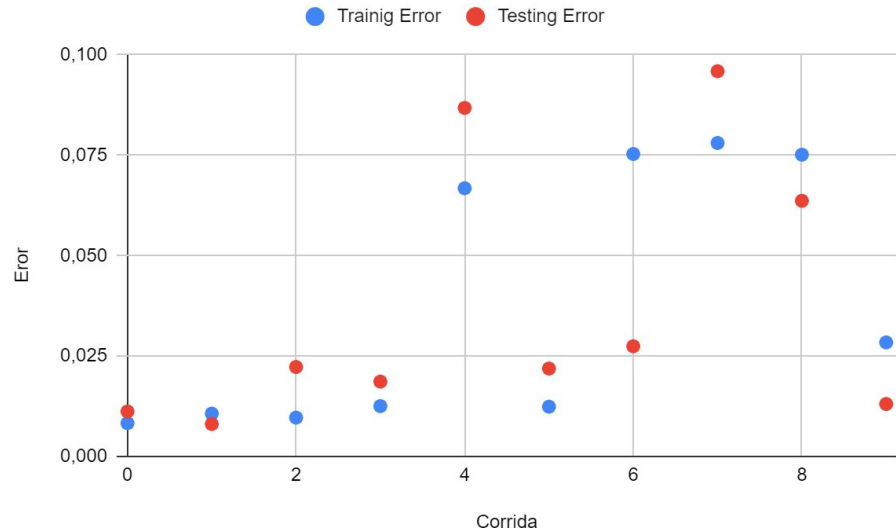
## APLICADA



Diferencia de error máximo: **0,02424**

Diferencia de error promedio: **0,00777**

## NO APLICADA



Diferencia de error máximo: **0,04788**

Diferencia de error promedio: **0,01463**



# Ejercicio 3

# Perceptrón multicapa

## “0 exclusivo”

Input =  $\{\{-1, 1\}, \{1, -1\}, \{-1, -1\}, \{1, 1\}\}$

Output =  $\{1, 1, -1, -1\}$

## “Número par”

Input = Imágenes de 5 x 7 píxeles

Output = 1 si es par, -1 sino

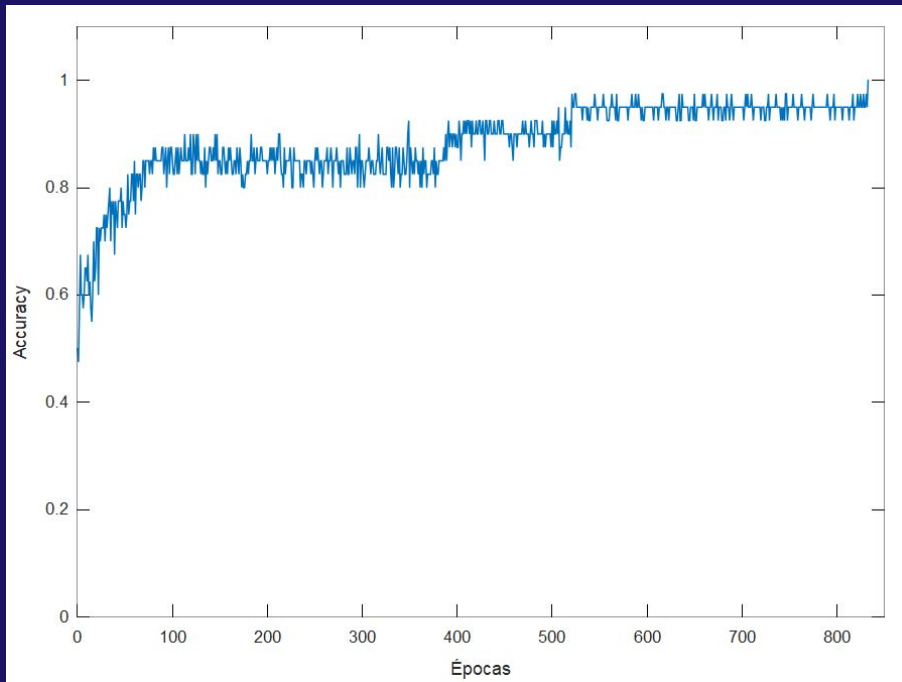


Estudiar la capacidad de generalización

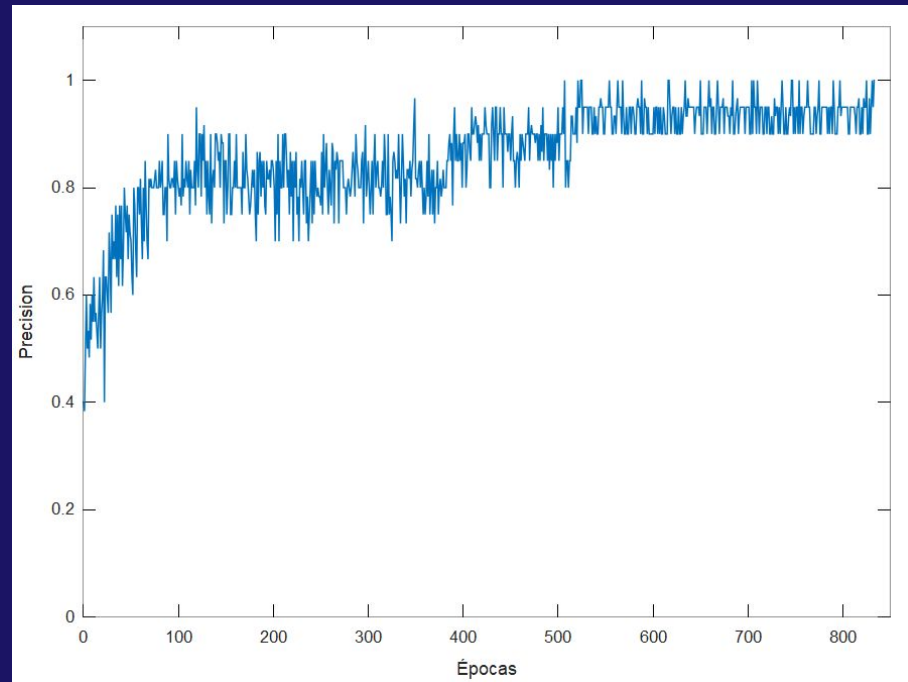
# XOR

Tasa de Aprendizaje = 0.1

**Accuracy**



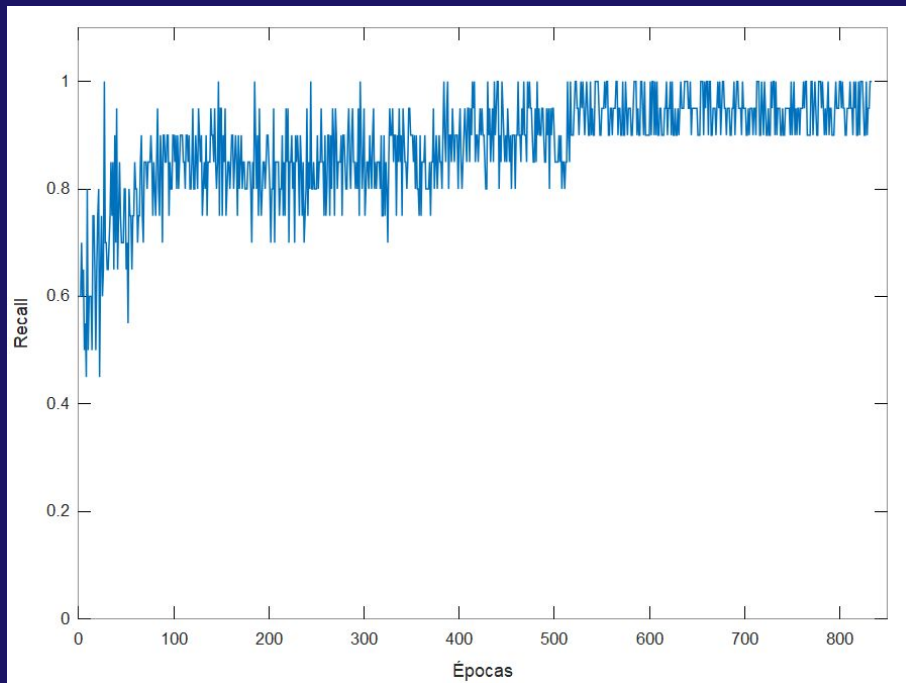
**Precision**



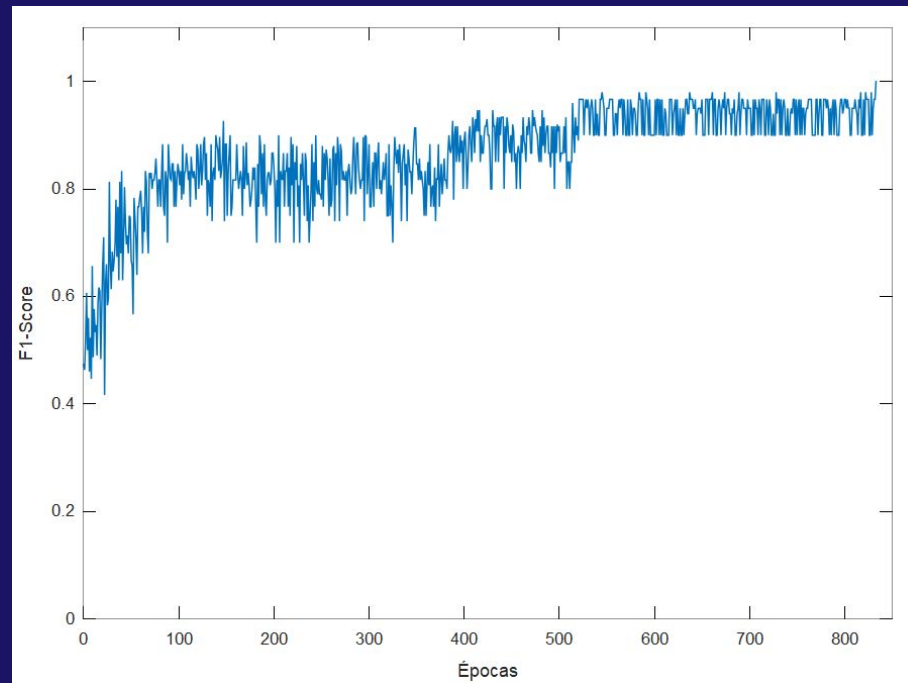
# XOR

Tasa de Aprendizaje = 0.1

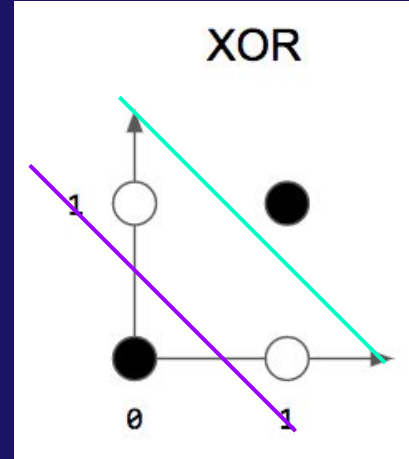
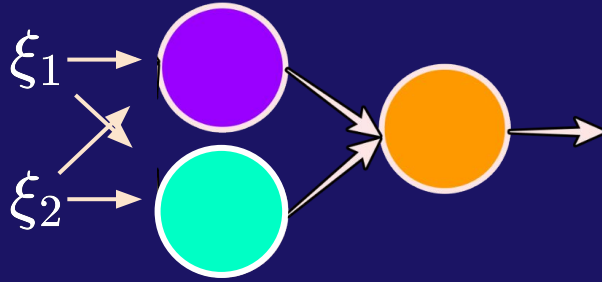
**Recall**



**F1-Score**



# XOR



Logra separar las clases por medio de dos hiperplanos

# Número Par

¿Qué podría decir acerca de la capacidad para generalizar de la red?

## Validación Cruzada

Se utiliza el método de validación cruzada para obtener el conjunto de entrenamiento que más aumenta la capacidad de generalización de la red



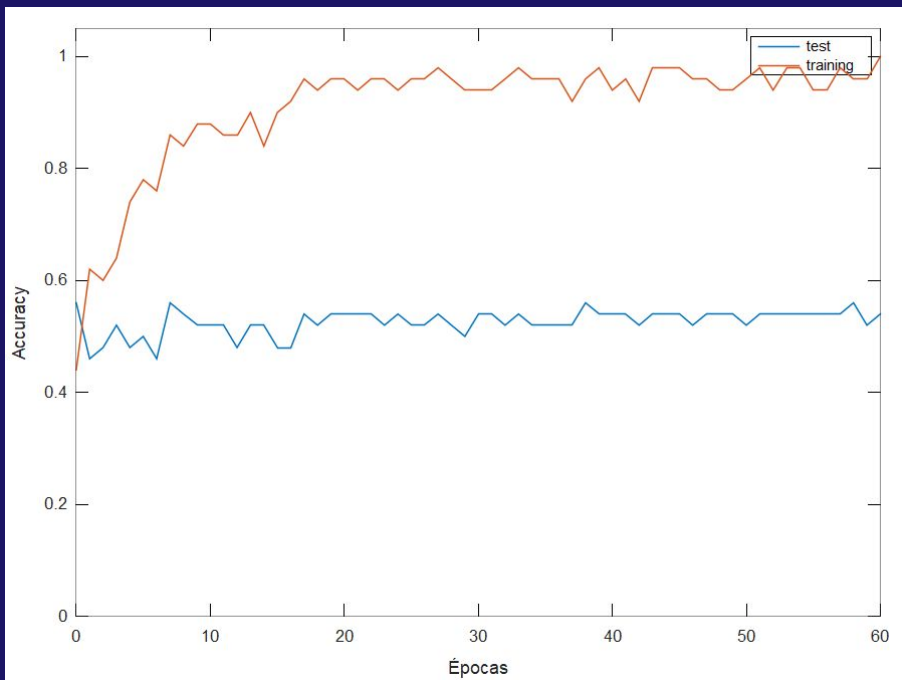
Luego se calculan métricas de la red para evaluar la mejor capacidad de generalización que la red logró obtener



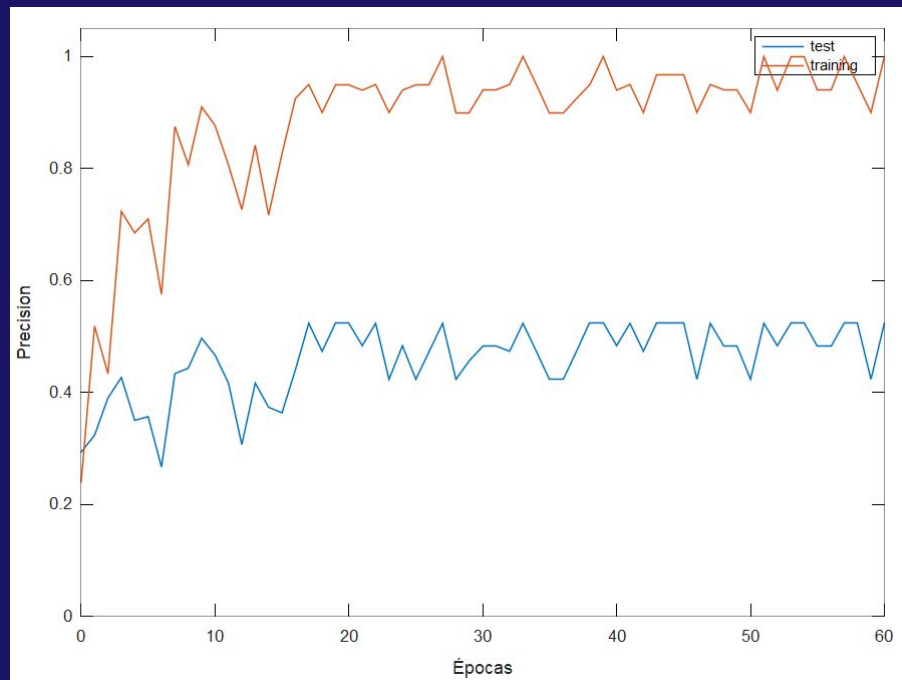
# Número Par

Conjunto de prueba = 50%

## Accuracy



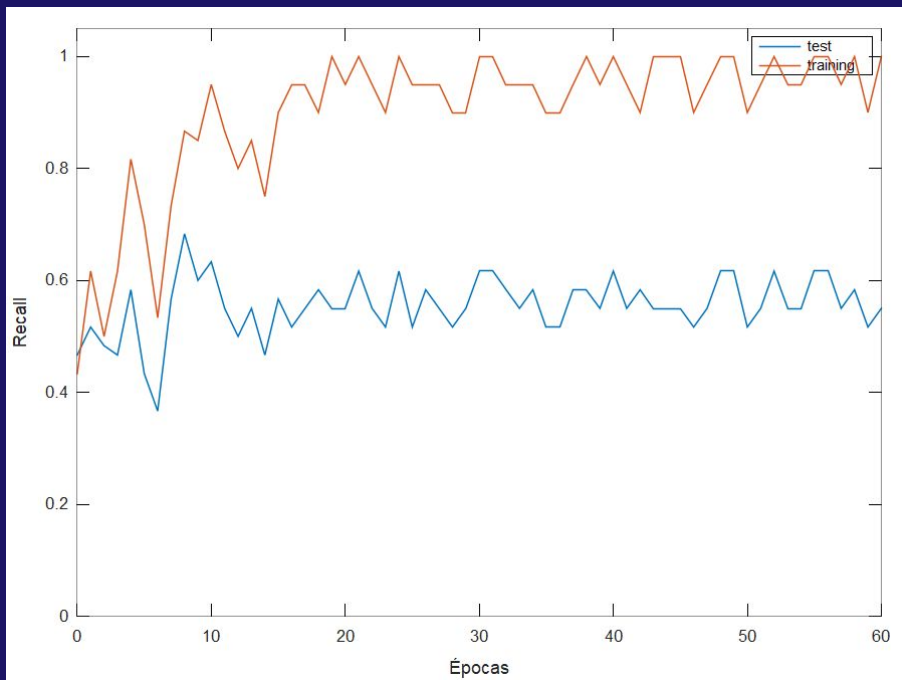
## Precision



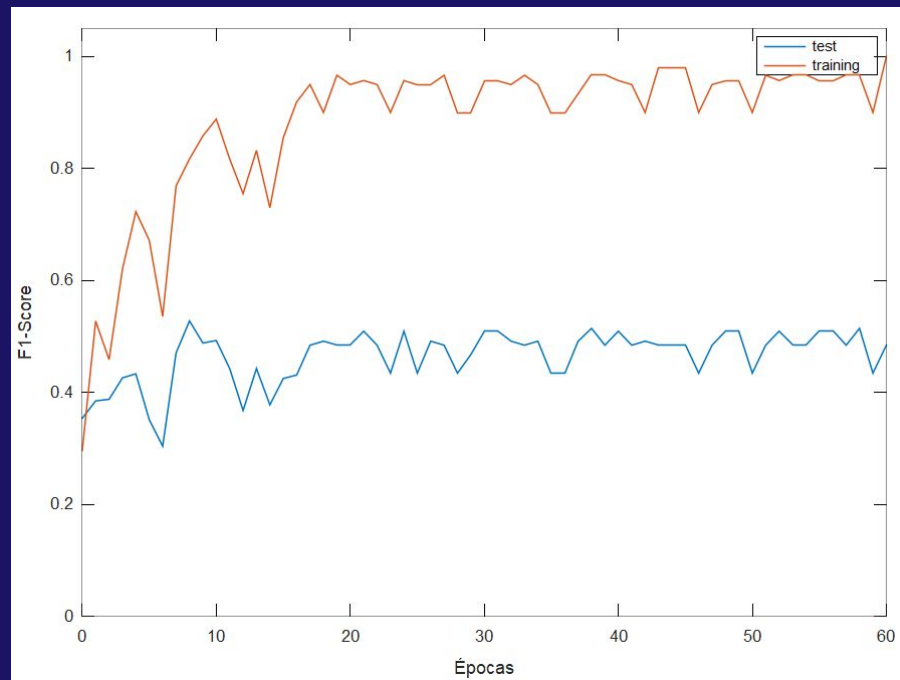
# Número Par

Conjunto de prueba = 50%

## Recall



## F1-Score



# Número Par

¿Qué podría decir acerca de la capacidad para generalizar de la red?

Aún con el método de validación cruzada no se puede encontrar un conjunto de entrenamiento que permita a la red generalizar de manera precisa

- ¿Múltiples registros por número?



Reconocer formas parecidas y determinar cuál número es (y por lo tanto si es par)



# Conclusiones

# Conclusiones

- No todos los problemas pueden ser resueltos por cualquier tipo de perceptrón
- La capacidad de generalización de la red depende de los datos del conjunto de entrenamiento



La validación cruzada permitió encontrar conjuntos de entrenamiento que maximicen la capacidad de generalización

- El método de la tasa de aprendizaje adaptable permite obtener un menor error y converger más rápidamente



Fin