

test, training

| Can leche | Cant Cafe | Cant Azucar | \$ |
|-----------|-----------|-------------|----|
| ~ | ~ | ~ | 10 |
| ~ | ~ | ~ | 20 |
| ~ | ✓ | ~ | 30 |

} → modelo

↳ Probar → 30 ✓

taller

respu.



→ Supervisado

| X | | | Y |
|-----------|-----------|-------------|----|
| Can leche | Cant Cafe | Cant Azucar | \$ |
| ~ | ~ | ~ | 10 |
| ~ | ~ | ~ | 20 |
| ~ | ✓ | ~ | 30 |

→ training - taller

→ test - examen

→ No supervisados

| | | | |
|--|--|--|--|
| | | | |
| | | | |
| | | | |

→

→

→ +

→ Regresión lineal



→ Clasificación

| N. Patas | Tipo pelo | bigotes? | 🐱 |
|----------|-----------|----------|---|
| 4 | suave | 1 | 1 |
| 2 | plumas | 0 | 0 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | ⋮ | ⋮ |

train

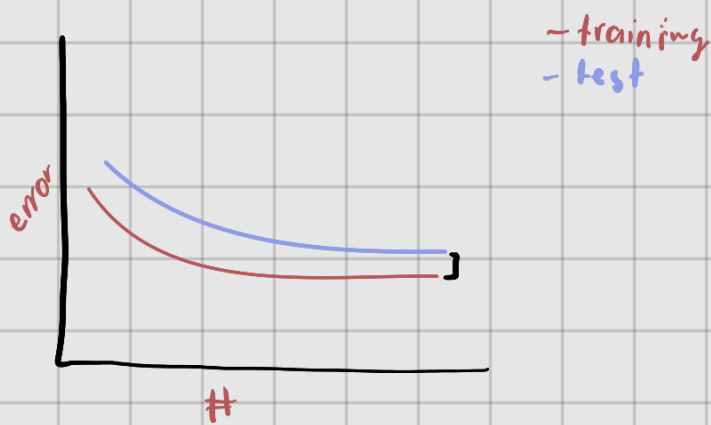
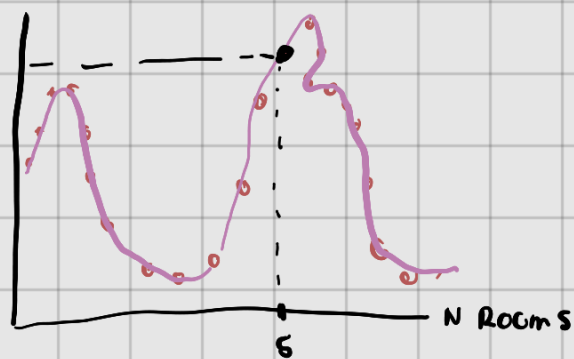
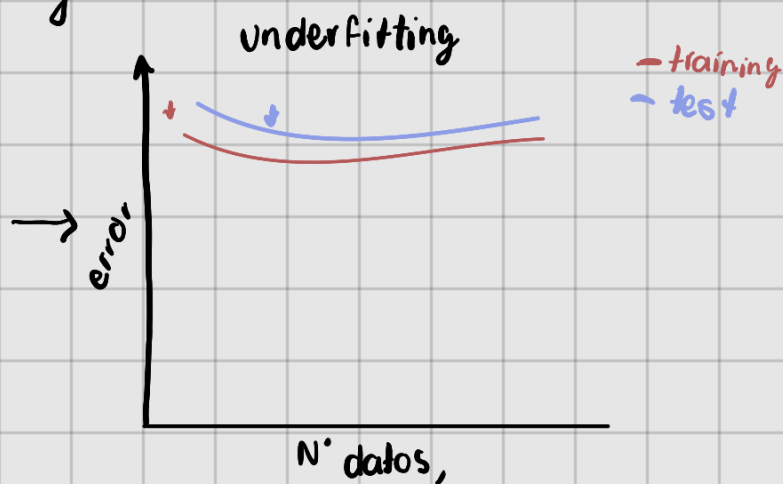
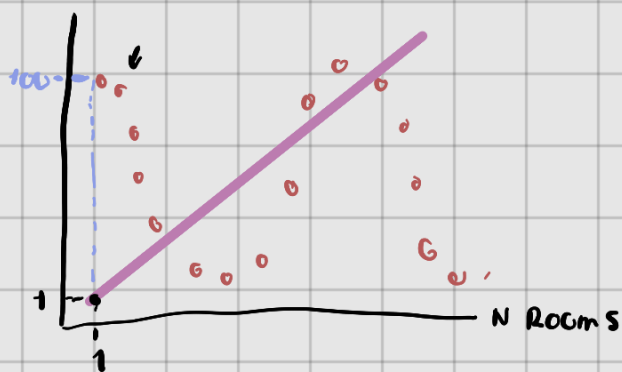
test

modelo

→ 🐱, 🐱 ✓

→ multi-label → 🐱, 🐱, 🐱

→ Underfitting y Overfitting



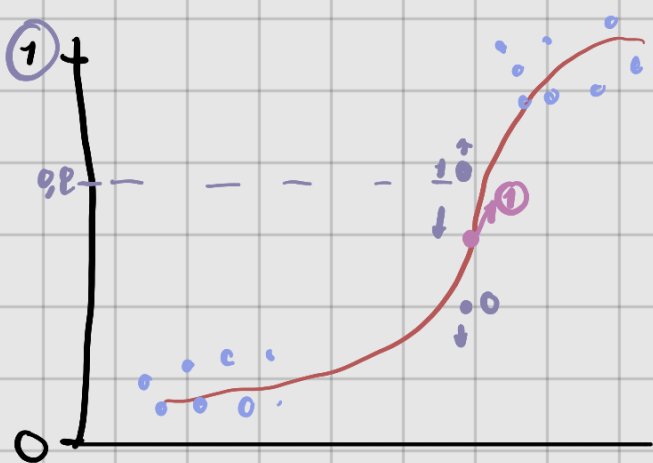
→ Validación Cruzada.

K=3



Promedio error = 0.3

→ Regresión logística



$$f(x) = \frac{1}{1 + e^{-x}}$$

$$\frac{1}{1.000...1} = 0.999...$$

$$\frac{1}{999...9} = 0.0000...$$

→ KNN → K vecinos más cercanos.

K=4



→ Matriz de confusión

Valores reales

| | | 1 | 0 |
|-------------------|---|----|----|
| Valores Predichos | 1 | TP | FP |
| | 0 | FN | TN |

① Precisión → $\frac{TP}{TP + FP}$

↓
80%

② Recall - Sensi. → $\frac{TP}{TP + FN}$

↓
60%

③ F-score →

④ Fall-out → $\frac{FP}{FP + TN}$ → 30%

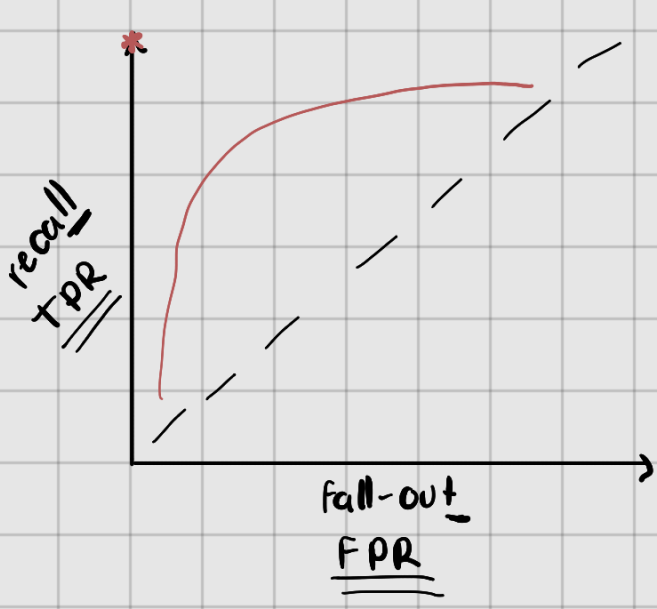
TP = Es 1 y predicho 1

TN = es 0 y " 0

FP = " 0 " 1

FN = " 1 " 0

→ Curva Roc



AUC → = 1 > 0,5 ✓
→ = 0

→ Explicación adicional

diabetes?

| # Emba | Nivel insulina | w | w | outcome |
|---------|----------------|---|---|---------|
| 0 | 0 → min | ~ | ~ | 0 |
| 6 → max | 846 → max | ~ | ~ | 0 |

→ min: 0 }
max: 1 }

min: -1
max ≈ 4

-0.6
≈ 6.5

min: -0.4
max: 5

4, 6 → 0.5

Error