

**IA-ML 2025B: enero 2026. Tarea en grupo.**

**Tarea grupal: Análisis comparativo y detallado de los algoritmos de clasificación**

### **1. Objetivo general de la tarea**

Que los estudiantes:

- Comprendan **cómo funciona internamente** un algoritmo de clasificación.
  - Analicen **el efecto de los parámetros** y decisiones de diseño.
  - Comparen su comportamiento frente a otros algoritmos.
  - Desarrollen capacidad de **explicación técnica clara**, apoyada en experimentos.
- 

### **2. Organización de los grupos**

- Grupos de **3 estudiantes**
- Cada grupo **elige** un algoritmo principal:
  1. Grupo A: **k-NN**
  2. Grupo B: **Naive Bayes**
  3. Grupo C: **Árboles de decisión**
  4. Grupo D: **SVM**

Todos los grupos **trabajan con al menos dos algoritmos**, pero **uno será el foco principal**.

---

### **3. Conjunto(s) de datos**

Cada grupo debe trabajar con **al menos un dataset común** (para comparación) y puede añadir uno adicional.

**Dataset sugerido (uno obligatorio)**

- Iris
- Wine
- Breast Cancer (sklearn)

**Opcional**

- Dataset propio o de UCI
- Dataset con:
  - Clases solapadas
  - Diferente escala de variables

- Ruido

---

#### 4. Actividades obligatorias

##### Parte 1 – Comprensión teórica (común a todos)

En el cuaderno deben incluir:

- Descripción conceptual del algoritmo
- Tipo de aprendizaje (supervisado / no)
- Supuestos principales
- Cómo define la frontera de decisión
- Ventajas y limitaciones teóricas

**No copiar definiciones:** debe estar explicado con sus propias palabras.

---

##### Parte 2 – Profundización en SU algoritmo principal

Cada grupo debe realizar **experimentos sistemáticos**, por ejemplo:

###### Grupo k-NN

- Variar:
  - $k$
  - Métrica (euclídea, Manhattan)
  - Normalización vs no normalización
- Analizar:
  - Sensibilidad al escalado
  - Cambios en fronteras de decisión

###### Grupo Naive Bayes

- Comparar:
  - Gaussian NB
  - Multinomial NB (si aplica)
- Analizar:
  - Efecto de la independencia condicional
  - Qué ocurre cuando el supuesto no se cumple

###### Grupo Árboles de decisión

- Variar:

- Profundidad máxima
- Criterio (gini vs entropía)
- Analizar:
  - Overfitting
  - Interpretabilidad
  - Importancia de características

### Grupo SVM

- Comparar:
  - Kernel lineal vs RBF
- Variar:
  - $C$
  - $\gamma$
- Analizar:
  - Margen
  - Vectores de soporte
  - Sensibilidad al escalado

---

### Parte 3 – Comparación con otro algoritmo

Cada grupo debe:

- Aplicar **al menos un algoritmo adicional**
- Usar **el mismo dataset**
- Comparar:
  - Exactitud
  - Comportamiento cualitativo
  - Sensibilidad a parámetros
  - Facilidad de interpretación

No se busca “el mejor algoritmo”, sino **entender por qué se comportan distinto**.

---

### Parte 4 – Análisis crítico (muy importante)

Responder explícitamente:

- ¿En qué tipo de problema **sí usaría** este algoritmo?

- ¿En qué tipo de problema **no lo usaría**?
- ¿Qué aprendieron que **no era evidente al inicio**?
- ¿Qué les sorprendió del comportamiento del algoritmo?

---

## 5. Entregables

### Cuaderno (Notebook) – obligatorio

Debe incluir:

- Texto explicativo claro (markdown)
- Código comentado
- Gráficos bien etiquetados
- Resultados organizados
- Conclusiones reflexivas

**Debe poder leerse como un documento académico**, no solo como código.

---

### Conclusión del grupo (1–2 páginas o sección final)

Debe responder:

- Qué entendieron del algoritmo
- Qué errores conceptuales tenían al inicio
- Qué aprendieron realmente al experimentar
- Qué algoritmo usarían y por qué

---

## 6. Rúbrica de evaluación (sencilla)

Criterio	Excelente (4)	Bueno (3)	Básico (2)	Insuficiente (1)
Comprensión teórica	Explicación profunda y clara	Correcta	Superficial	Incorrecta
Experimentos	Sistemáticos y bien diseñados	Adecuados	Limitados	Incorrectos
Análisis crítico	Reflexión profunda	Buena	Débil	Ausente
Cuaderno	Claro, ordenado, didáctico	Entendible	Confuso	Desordenado
Conclusiones	Demuestran aprendizaje real	Correctas	Genéricas	Ausentes

