

# Ciencia de Datos

## Tarea 6

Para entregar el 5 de junio de 2020

1. Usando los datos de los dígitos escritos a mano y digitalizados (**MNIST**), complementa el ejercicio que hiciste en la tarea 4 ejercicio 2, aplicando métodos de clasificación basados en

- Regresión logística
- Redes neuronales
- Máquinas de Soporte Vectorial
- Árboles de clasificación
- AdaBoost

Utiliza  $K$ -Fold CV como criterio para elegir el mejor modelo, así como para compararlos. ¿Qué método elegirías?

Si lo consideras necesario, puedes usar otra representación de los datos, pero especifícalo en tu reporte.

Especifica los parámetros que usaste en cada método de clasificación. Incluye gráficos *informativos* sobre el desempeño de cada método. Actualiza tu aplicación interactiva, si es que la implementaste en la primera tarea.

2. **Puntos extra:** Repite el ejercicio 2 para los datos de frutas que usaste en la tarea 3. Utiliza la representación en el espacio HSV con la mediana y los cuartiles centrales.
3. **Puntos extra:** Repite el ejercicio 2 para los datos de opinión **spanish\_reviews.zip** que usaste en la tarea 4, ejercicio 3-e (análisis de sentimientos). Haz el mismo experimento (**Puntos extra++**) para la clasificación de productos (ya sean opiniones buenas o malas sobre autos, hoteles, libros, etcétera).