

## Tarea 2: Aplicaciones de la Matemática en Ingeniería

15 de septiembre de 2023

† **Contexto.** Considere el conjunto de datos del archivo `rice.txt` (disponible en AULA). Este es un problema de clasificación binaria donde granos de arroz pueden ser clasificados en alguna de las siguientes especies: *Cammeo* o *Osmancik*. Se tiene un total de 3810 imágenes de granos de arroz. Estas imágenes fueron procesadas y se obtuvieron 7 características morfológicas (covariables) para cada grano de arroz:

1. Área: Devuelve el número de píxeles dentro de los límites del grano de arroz.
2. Perímetro: Calcula la circunferencia a partir de la distancia entre píxeles alrededor de los límites del grano de arroz.
3. Longitud del Eje Mayor: La línea más larga que se puede trazar en el grano de arroz.
4. Longitud del Eje Menor: La línea más corta que se puede trazar en el grano de arroz.
5. Excentricidad: Mide qué tan redonda es la elipse que tiene los mismos momentos que el grano de arroz.
6. Área Convexa: Devuelve la cantidad de píxeles de la cáscara convexa más pequeña de la región formada por el grano de arroz.
7. Extensión: Devuelve la proporción de la región formada por el grano de arroz en relación con los píxeles del cuadro delimitador.

† **¿Qué se espera de esta tarea?** El propósito de esta tarea consiste en comparar el rendimiento de los métodos de vecinos cercanos y Support Vector Machine (SVM). Para lograr este objetivo, usted debe realizar estudios de validación cruzada. Tenga en cuenta la variación del número de vecinos en el caso de los vecinos cercanos y la elección del kernel en el caso de SVM.

Primero, haga un análisis descriptivo de los datos (reporte indicadores y gráficos para tener una perspectiva general del problema). Luego, proporcione una explicación detallada de los procedimientos empleados en cada uno de estos métodos y presente las conclusiones derivadas de dichos análisis. Es importante que sus estudios estén acompañados de matrices de confusión, gráficos apropiados, indicadores de precisión, etc.

Debe entregar un informe en un formato similar a la tarea 1.