

EGRESOS DE RECINTOS HOSPITALARIOS EN CHILE

Integrantes: Damián González
Joaquín González
Francisco Márquez
Claudio Morel
Nicolás Soto

Profesores: Felipe Bravo Márquez
Valentín Barriere

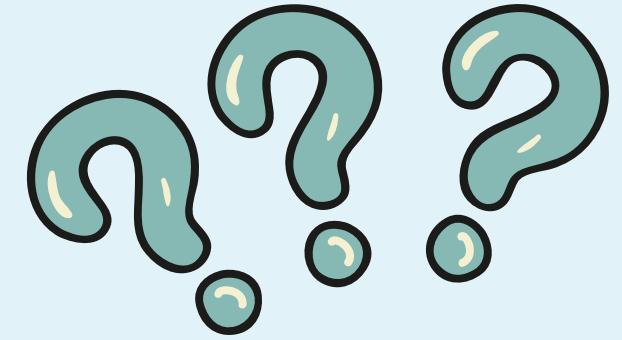
Grupo: 17

Motivación

El sistema de salud pública en Chile, está marcado por largos tiempos de espera y colapsos en la entrega de servicios básicos en situaciones críticas, por lo tanto, es crucial buscar formas de anticipar y planificar de manera más efectiva la distribución de recursos y personal.



Preguntas propuestas



- ¿Se puede predecir la cantidad de días de estadía de un paciente sabiendo su edad, sexo, región, pertenencia al sistema nacional de salud, y diagnostico/intervención quirúrgica ?
- ¿Qué variables tienen un mayor impacto en la cantidad de días de estadía de un paciente?
- ¿Se pueden encontrar agrupaciones con distinto nivel de urgencia de atención según su edad, sexo, región, pertenencia al sistema nacional de salud, y diagnostico/intervencion quirúrgica?



Pregunta 1: Clasificación de periodos de estadía

¿Se puede predecir la cantidad de días de estadía de un paciente sabiendo su edad, sexo, región, pertenencia al sistema nacional de salud, y diagnóstico/intervención quirúrgica?

Metodología a seguir

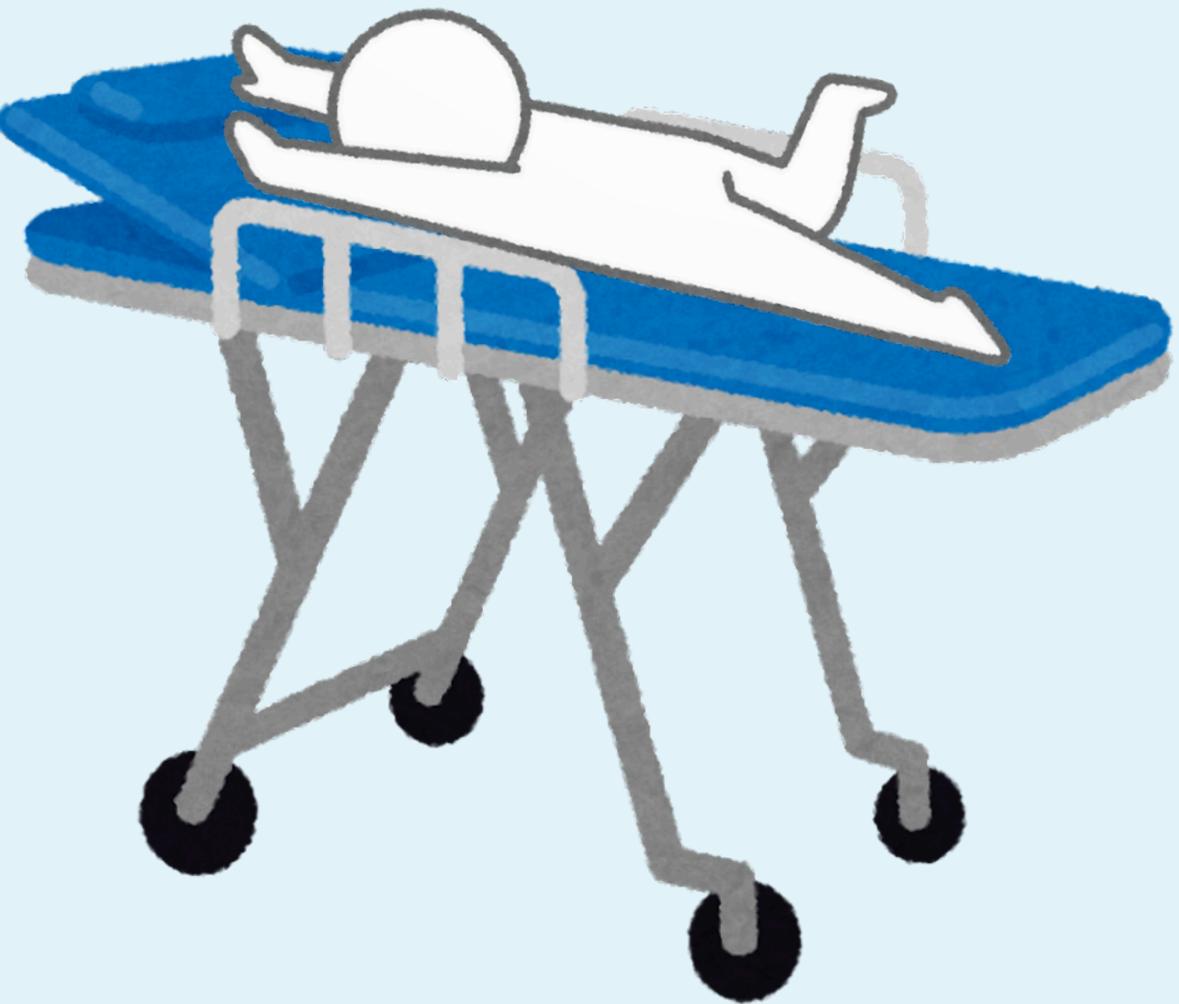
- Transformación de los datos a variables numéricas
- Balanceo de clases
- GridSearch
- Entrenamiento de modelos



Transformación de datos

Clases objetivo:

- Estancia Corta: 1 a 3 días
- Estancia Media: 4 a 8 días
- Estancia Larga: 9 o más días



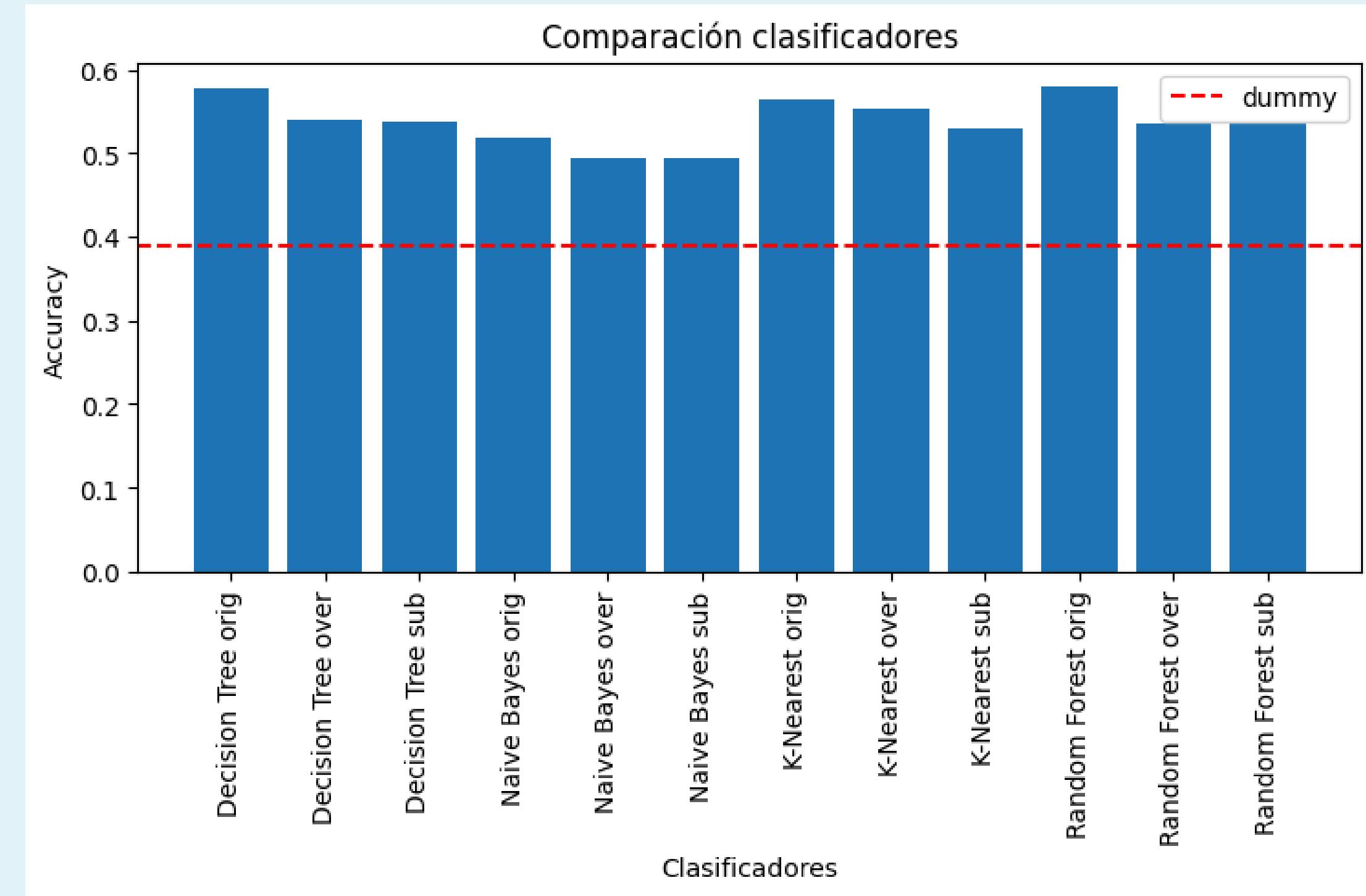
Balanceo de clases



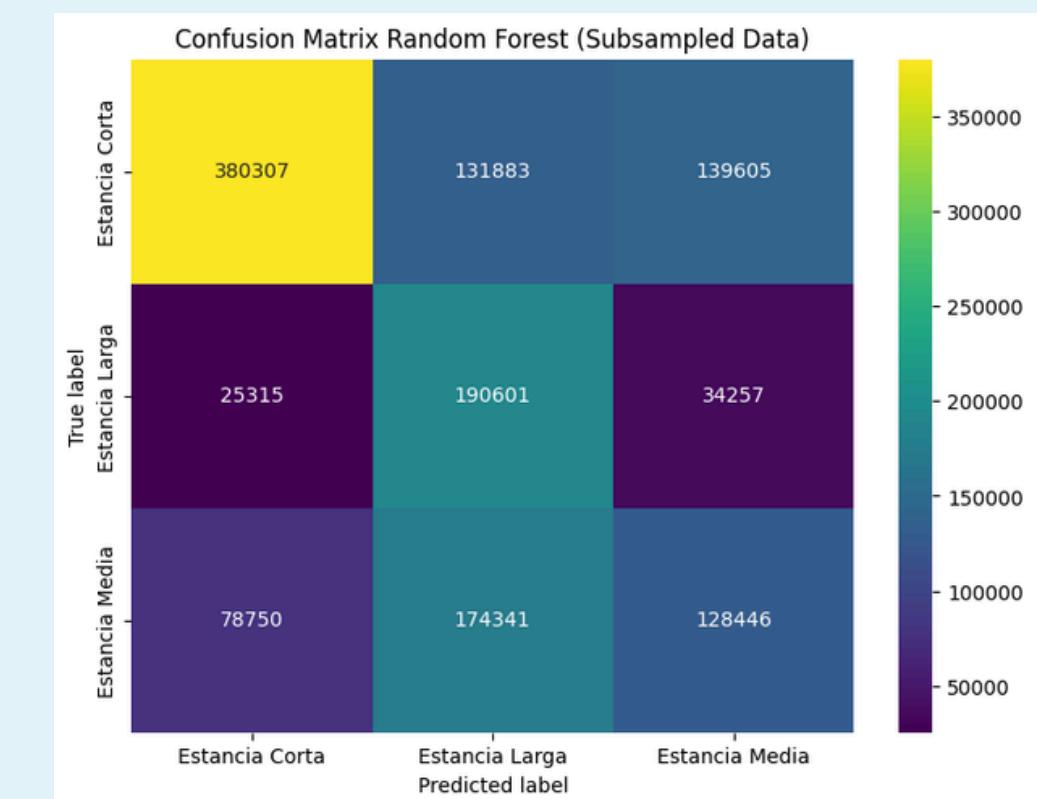
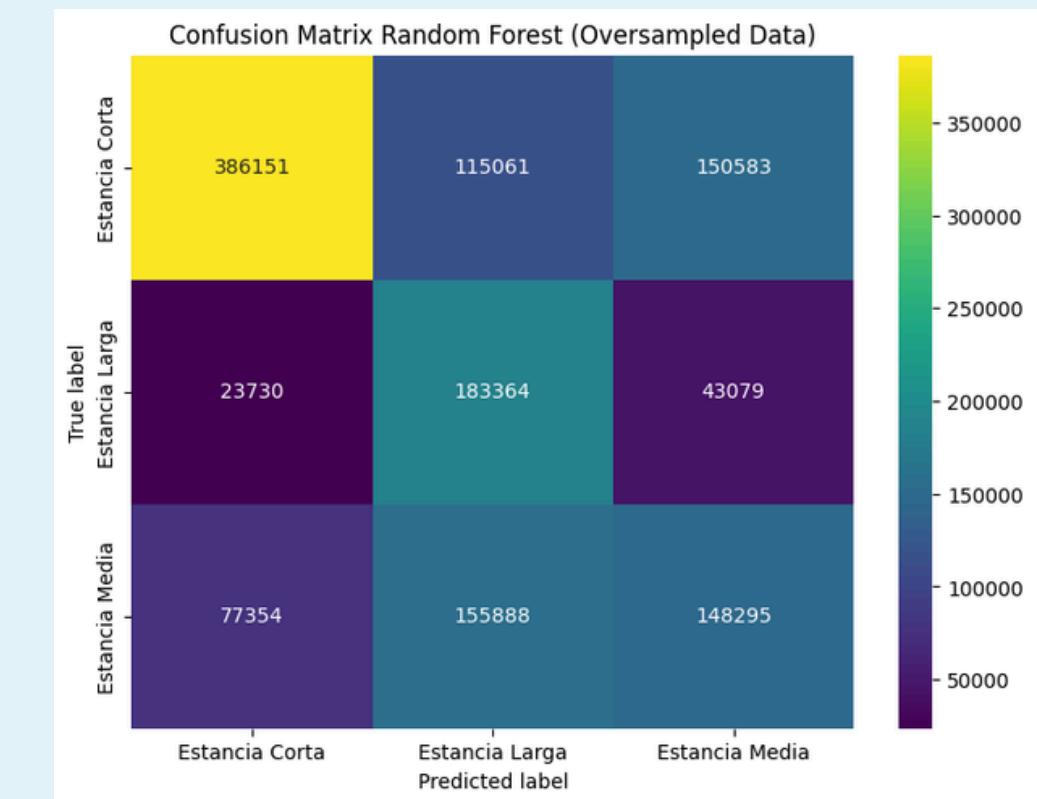
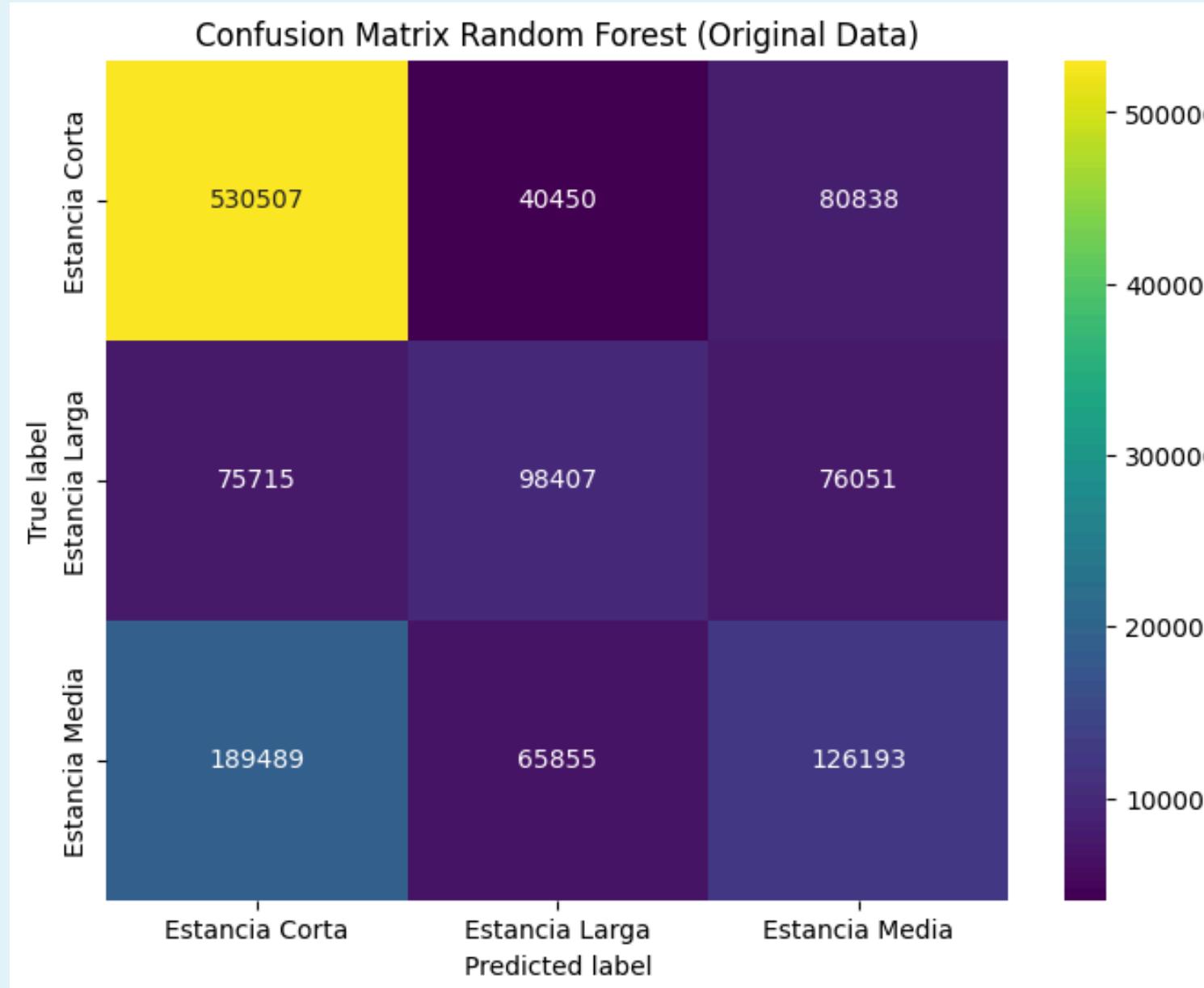
Clases objetivo:

- Estancia Corta: 3226876 datos
- Estancia Media: 904171 datos
- Estancia Larga: 150843 datos

GridSearch



Resultados



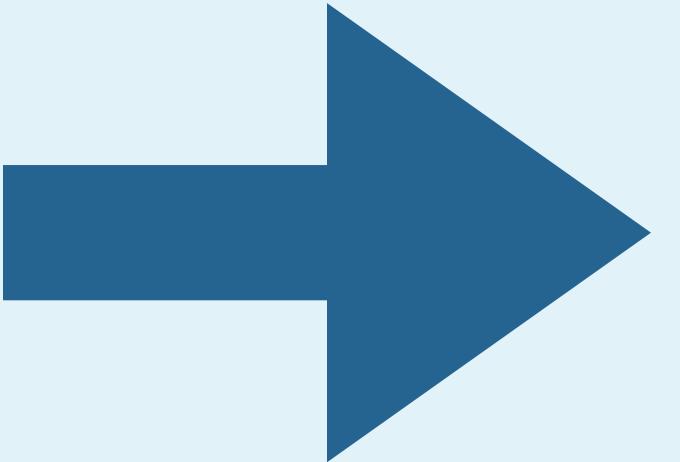
Pregunta 2: Agrupaciones por niveles de urgencia de atención

Metodología a seguir:

- Transformación de los datos a variables numéricas
- Generación de los modelos
- Validación de los clusters y selección del mejor modelo
- Caracterización de los grupos encontrados

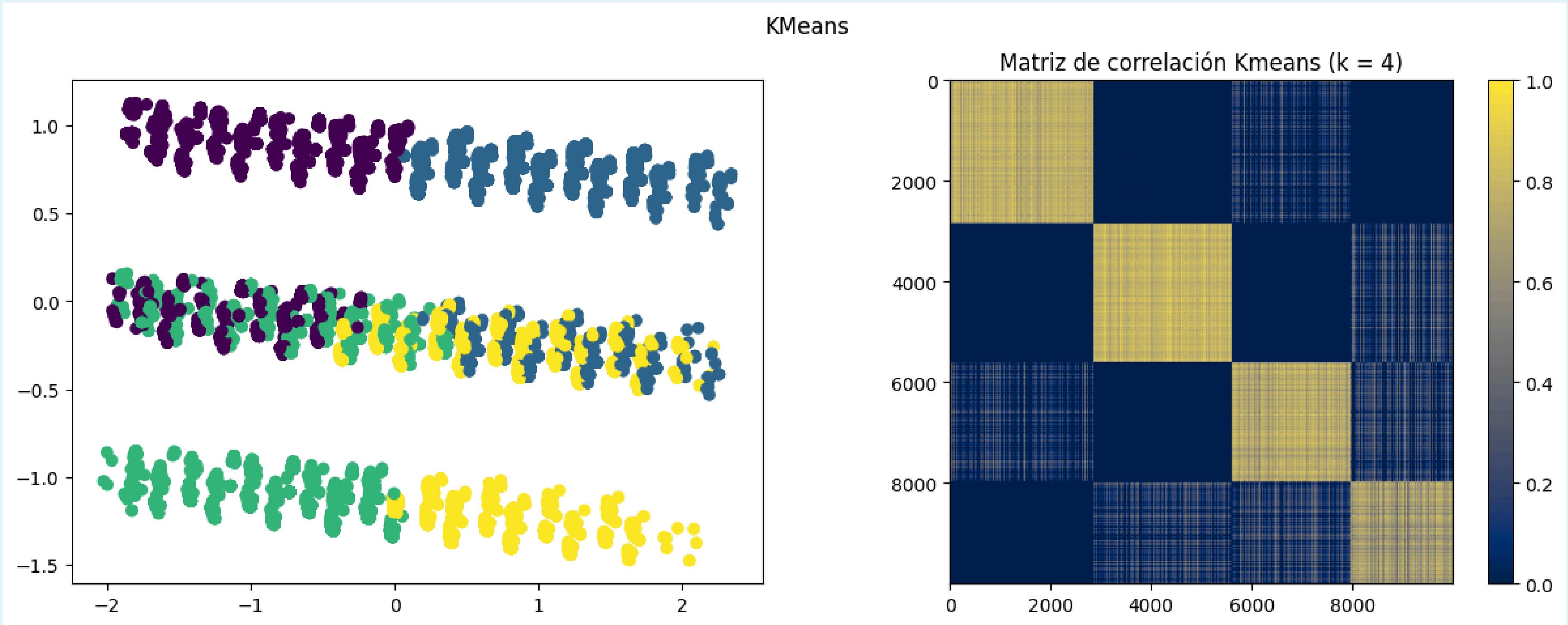
Datos utilizados:

- Label encoding:
 - - Edad
- One hot encoding:
 - - Pertenencia al SNS
 - - Intervención quirúrgica
 - - Sexo
 - - Diagnóstico
 - - Región

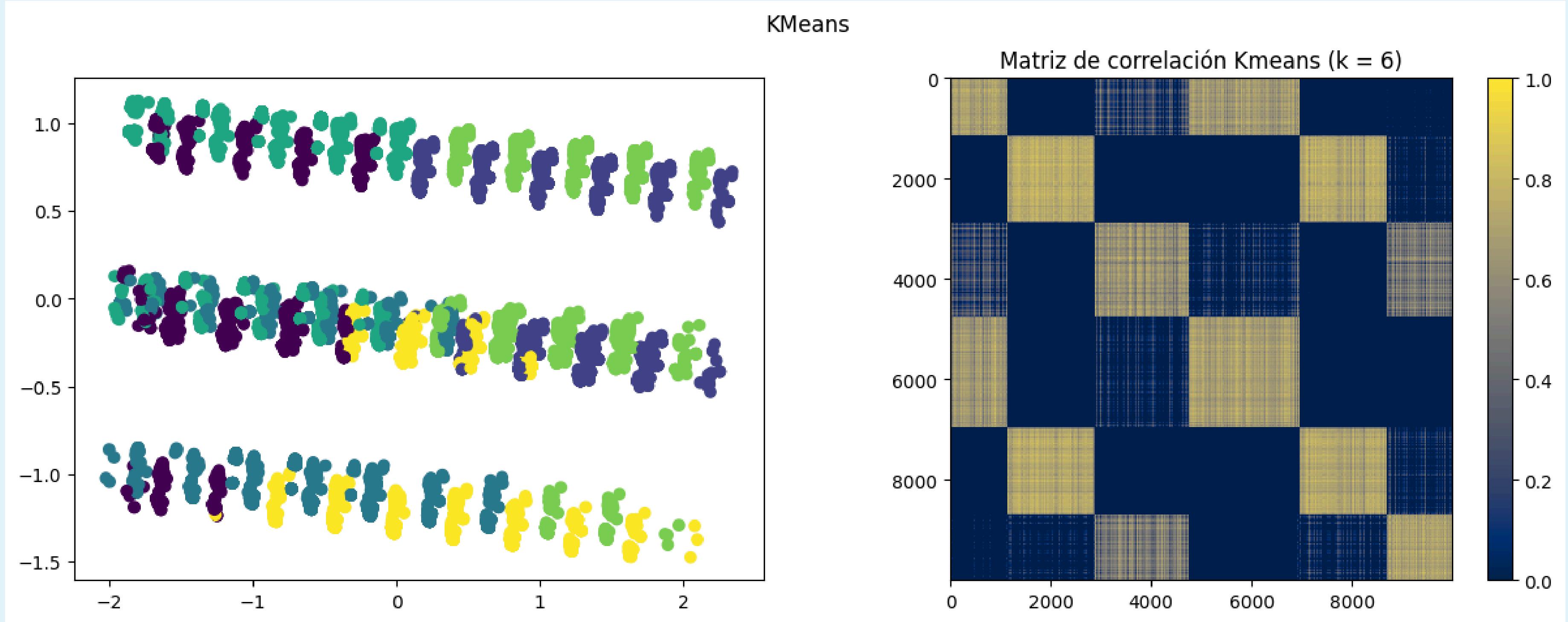


10.000 filas
44 columnas

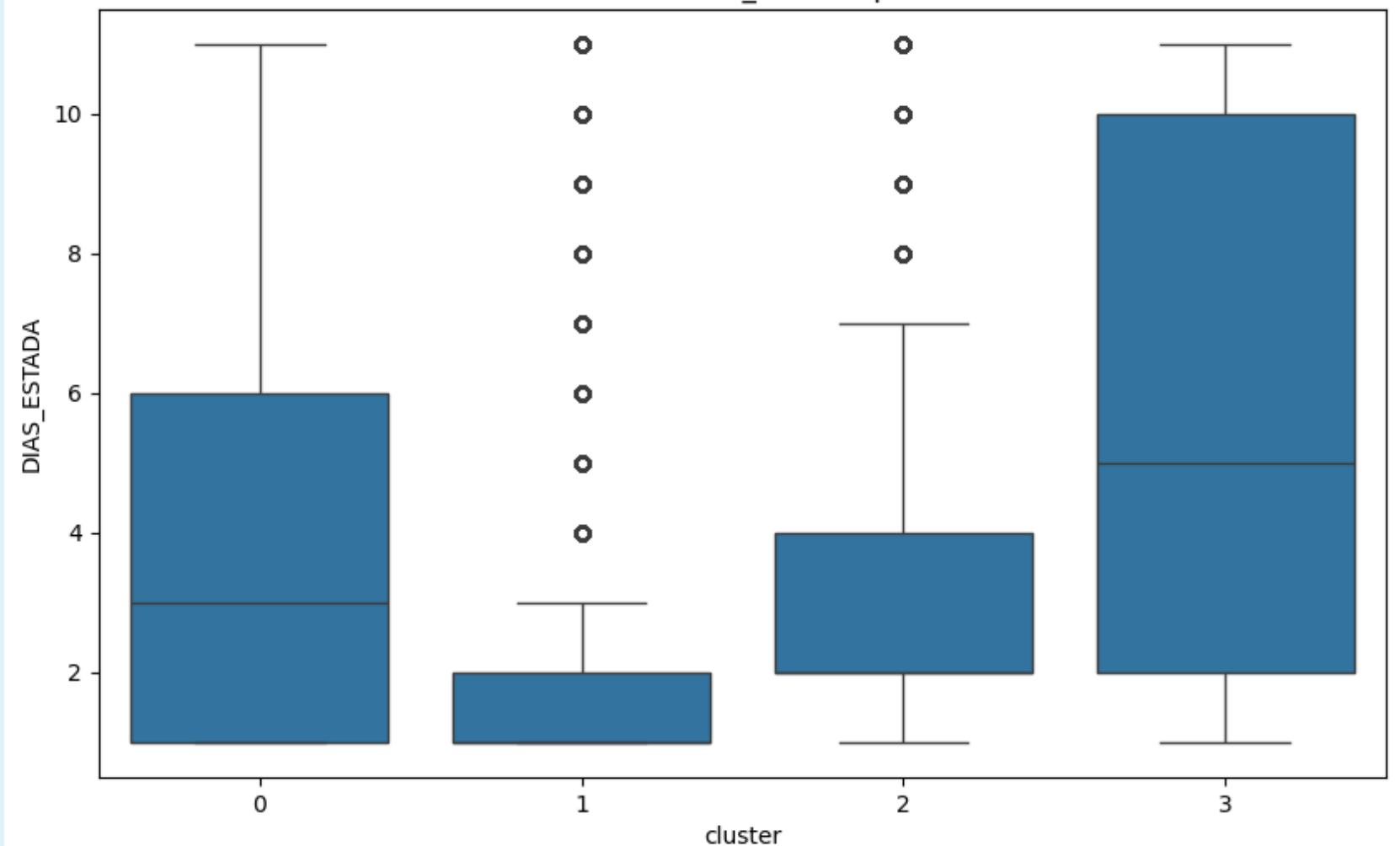
Resultados:



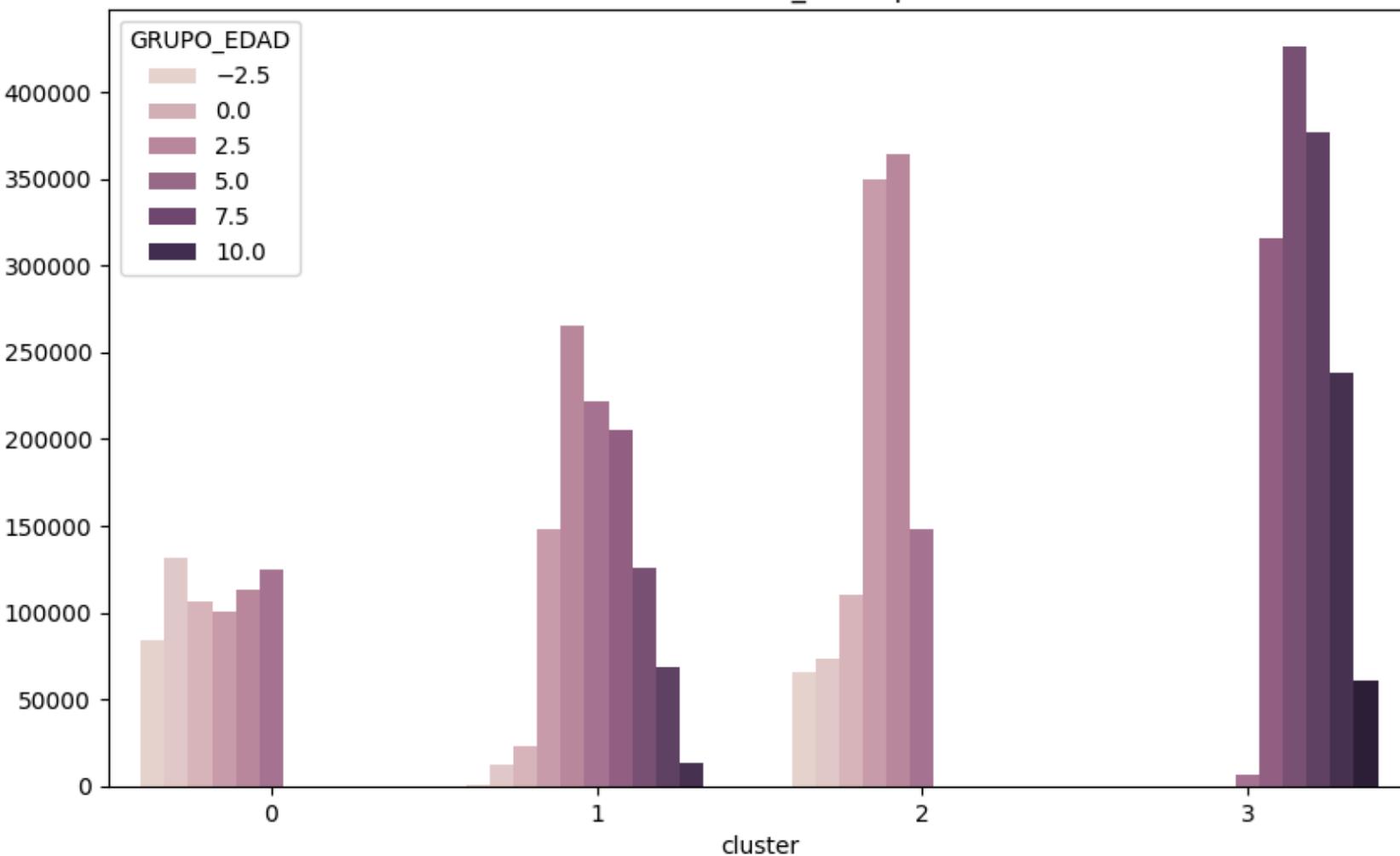
Resultados:



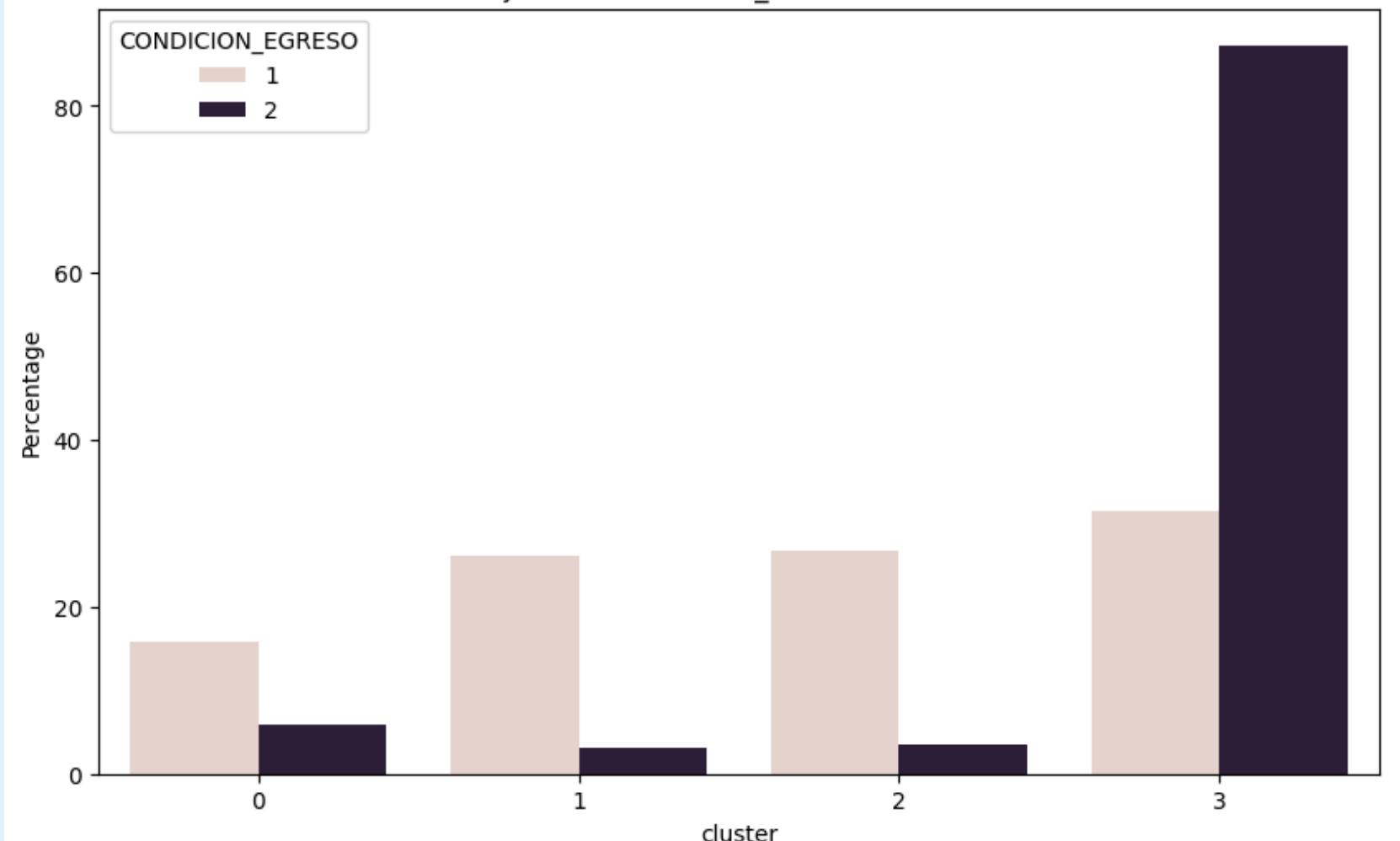
Distribución de DIAS_ESTADA por Cluster



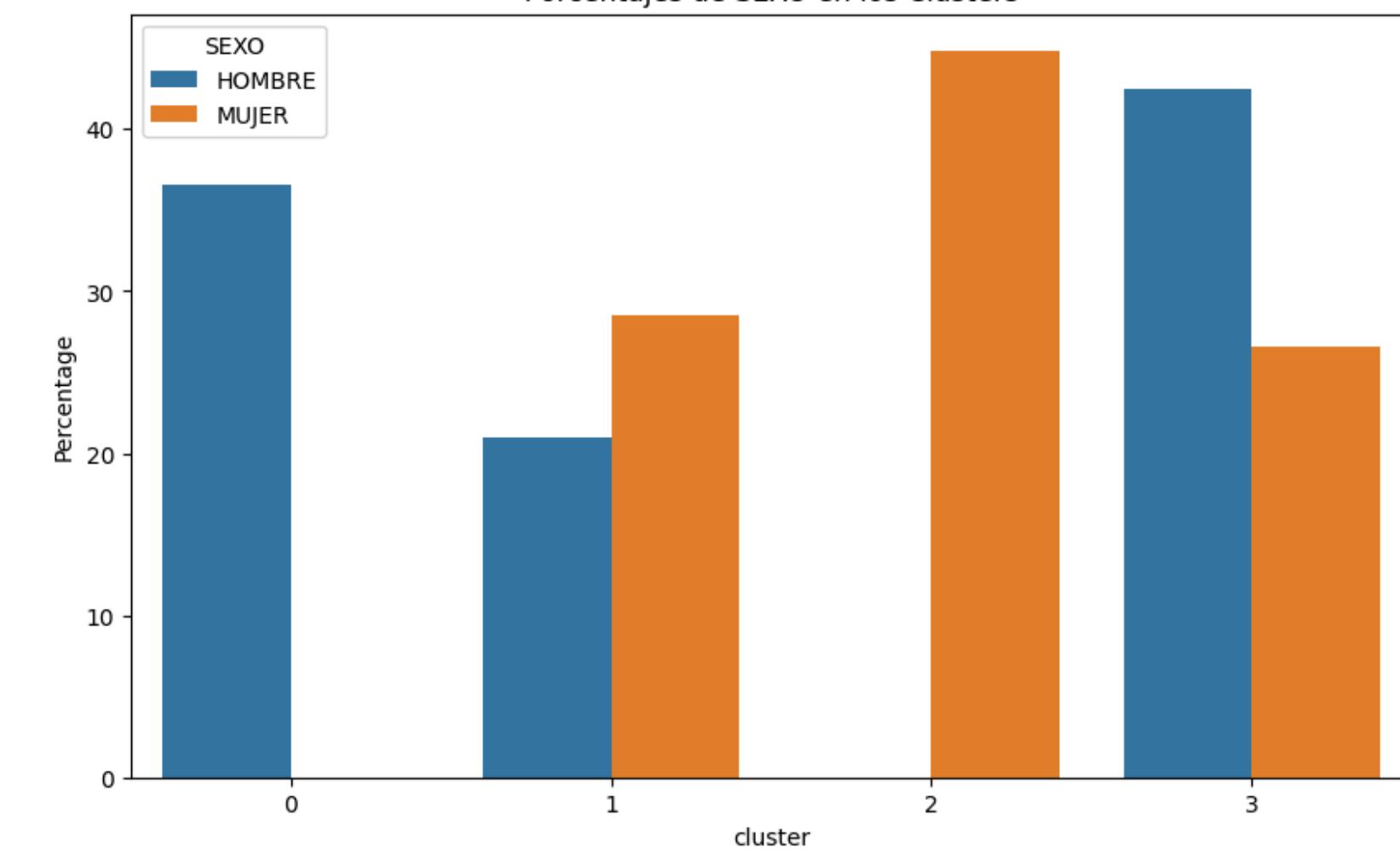
Distribución de GRUPO_EDAD por Cluster



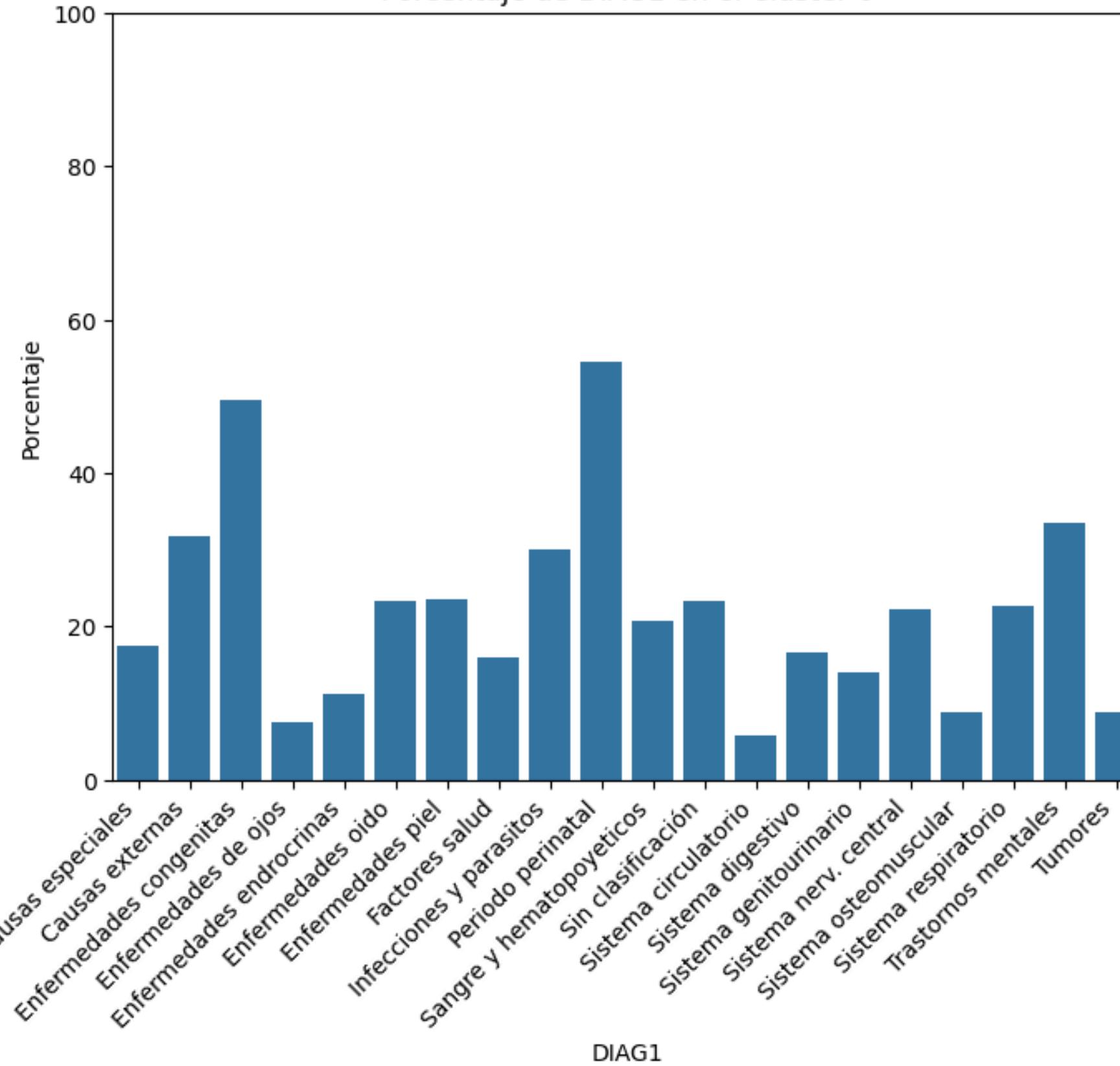
Porcentajes de CONDICION_EGRESO en los Clusters



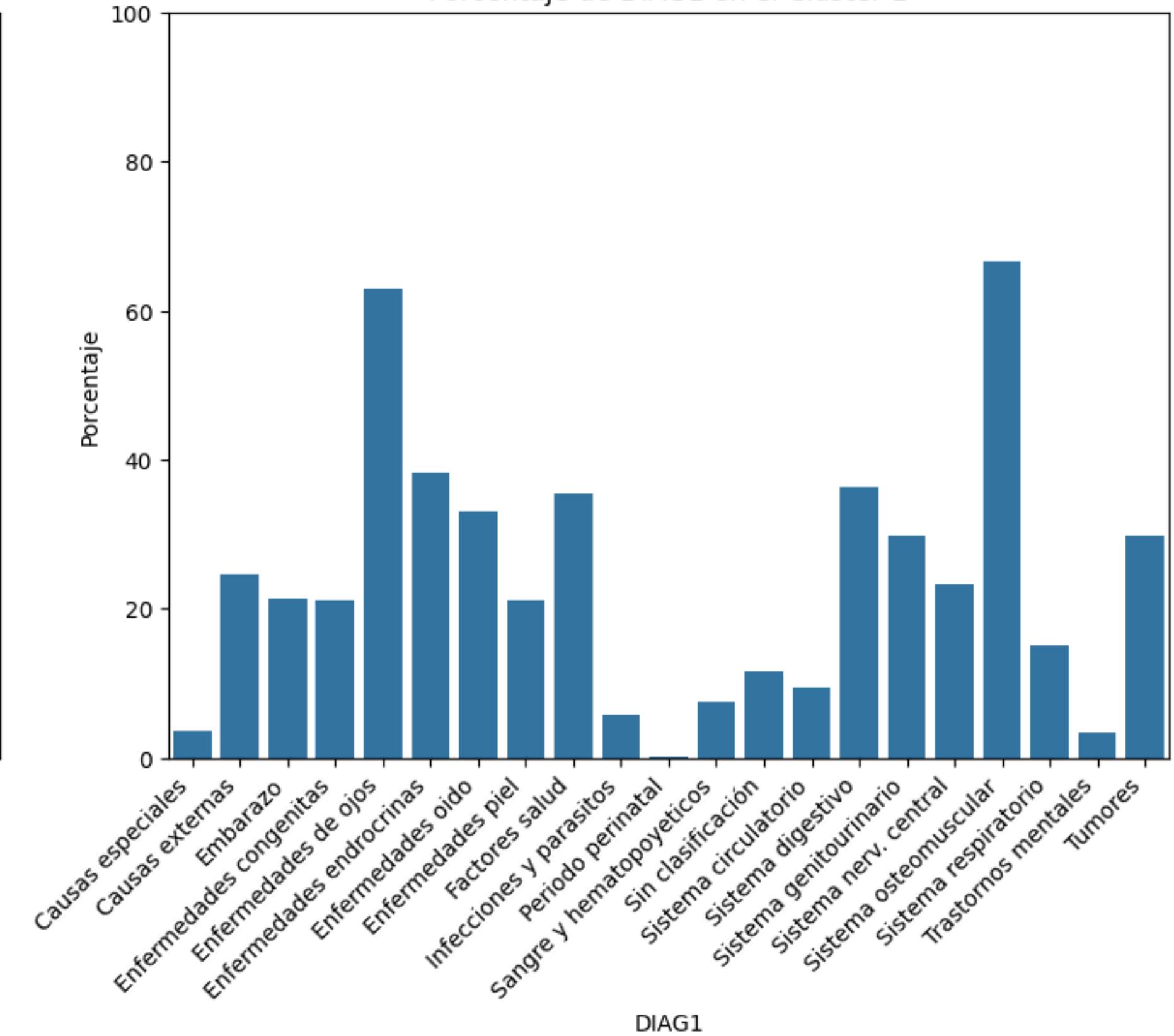
Porcentajes de SEXO en los Clusters



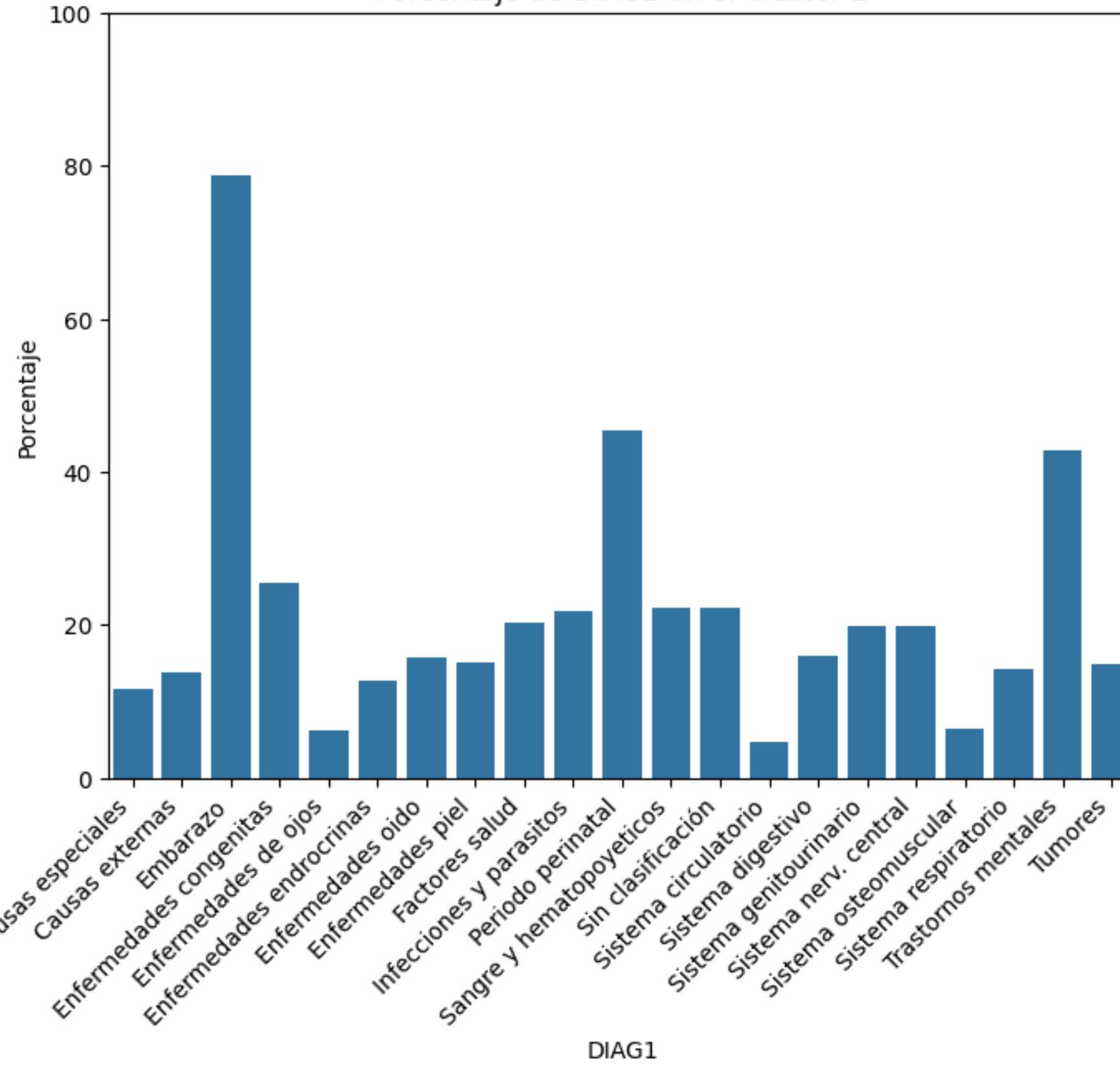
Porcentaje de DIAG1 en el Cluster 0



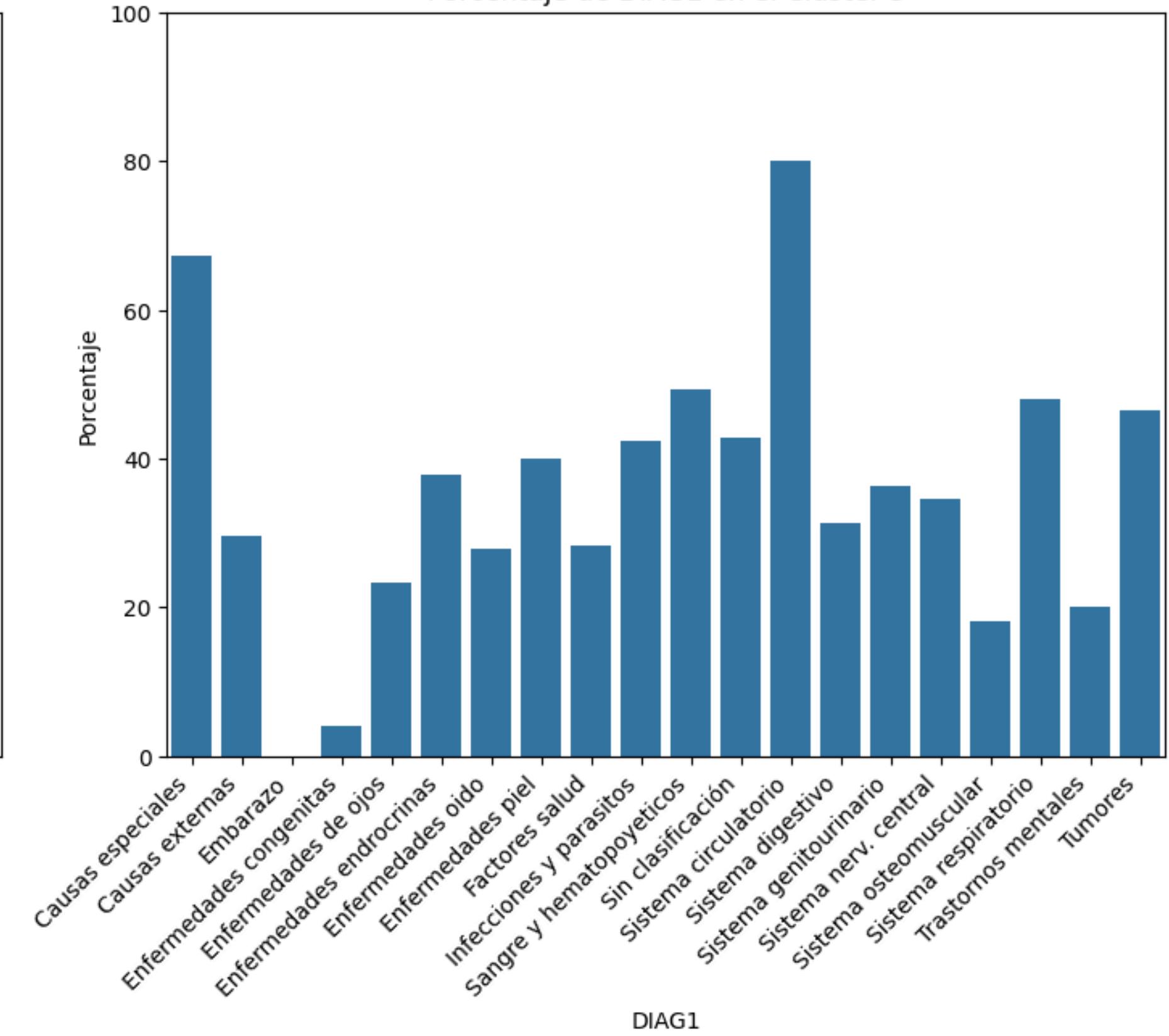
Porcentaje de DIAG1 en el Cluster 1

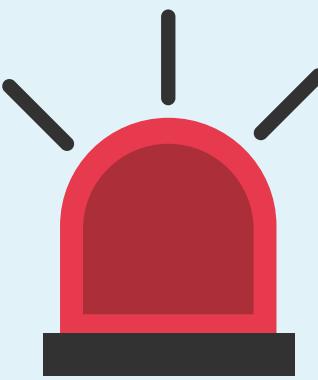


Porcentaje de DIAG1 en el Cluster 2



Porcentaje de DIAG1 en el Cluster 3



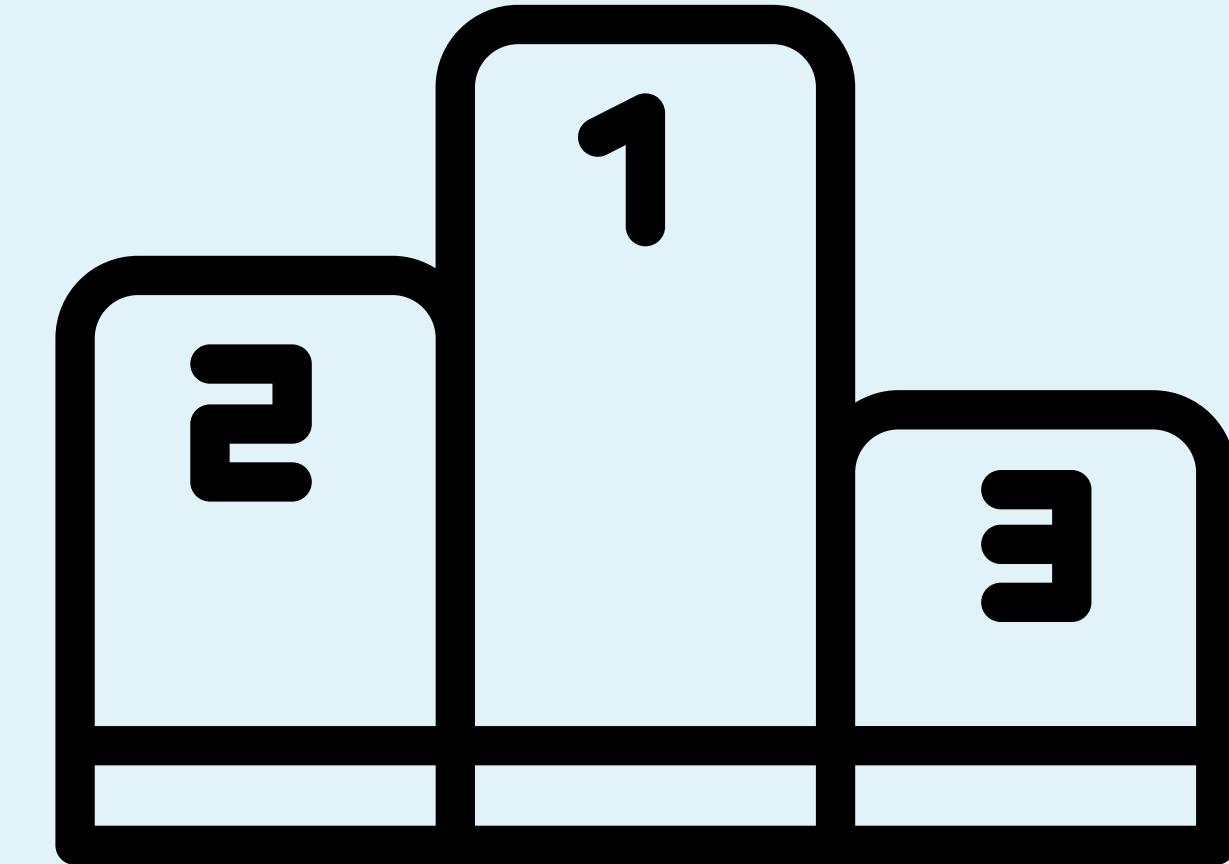


Asignación de niveles de urgencia de atención:

- Cluster 0: nivel de urgencia alto
- Cluster 1: nivel de urgencia bajo
- Cluster 2: nivel de urgencia moderado
- Cluster 3: nivel de urgencia crítico



Pregunta 3: Importancia de atributos en condición de egreso



**¿Cuáles son los factores que
más influyen en la condición
de egreso del paciente (si
sobrevive o no)?**

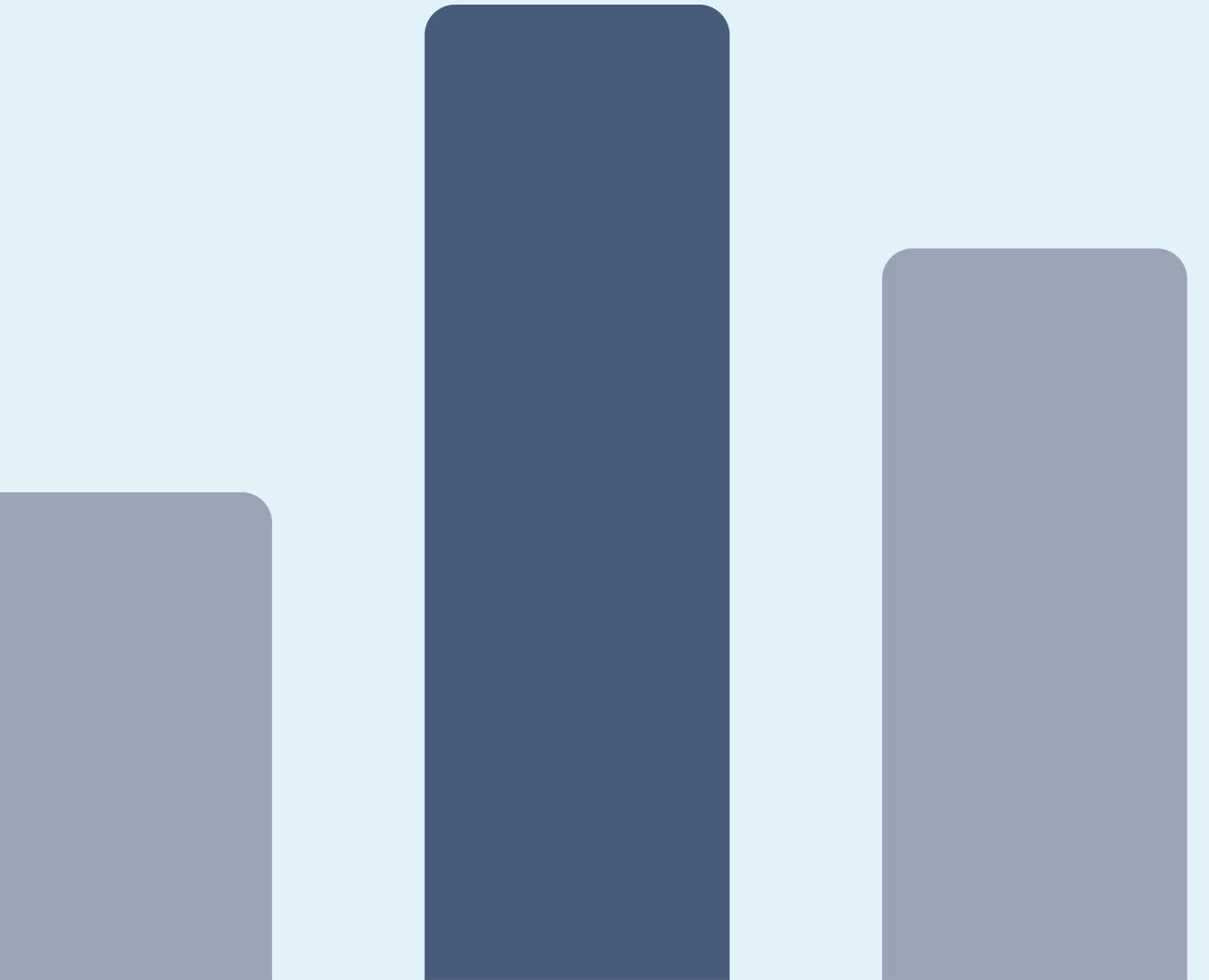
Metodología a seguir:

- Preprocesamiento y balanceo de clases
- Selección y entrenamiento de modelos
- Evaluación de Modelos
- Cálculo de importancia de atributos



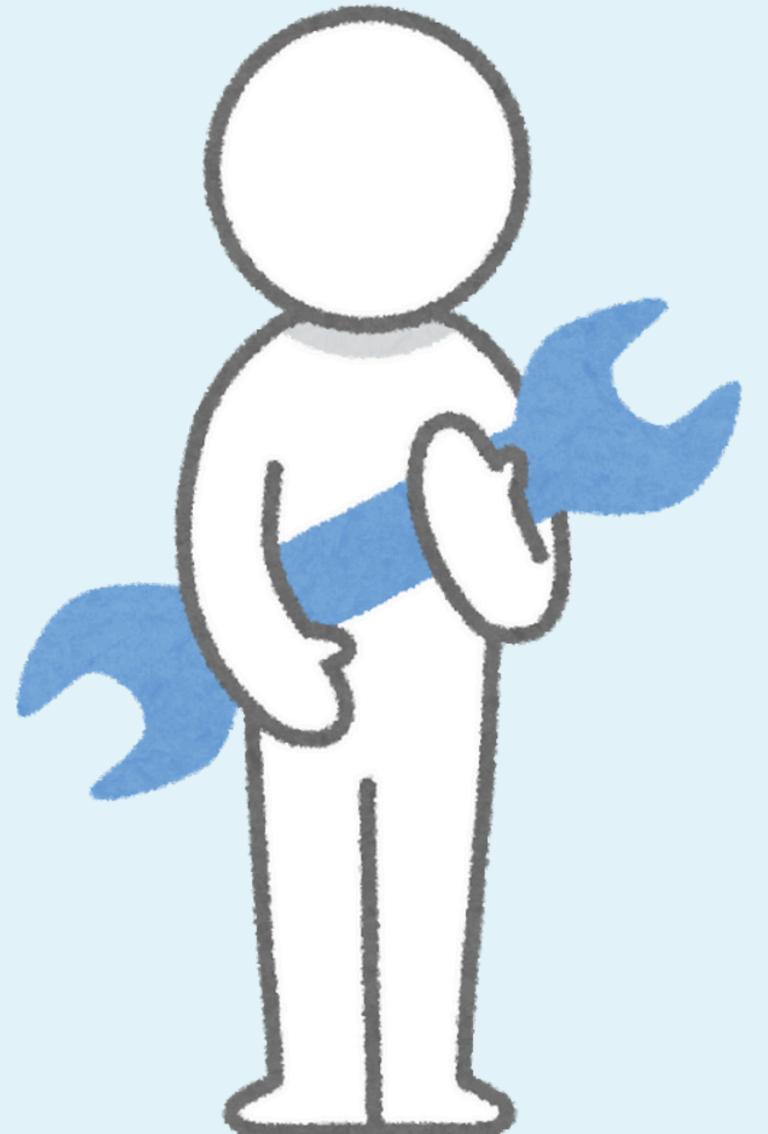
¿Qué es la *feature importance*?

- Variables Relevantes: Identifica las características más significativas.
- Mejora del Modelo: Optimiza el rendimiento eliminando variables irrelevantes.
- Interpretabilidad: Facilita la comprensión de las decisiones del modelo.



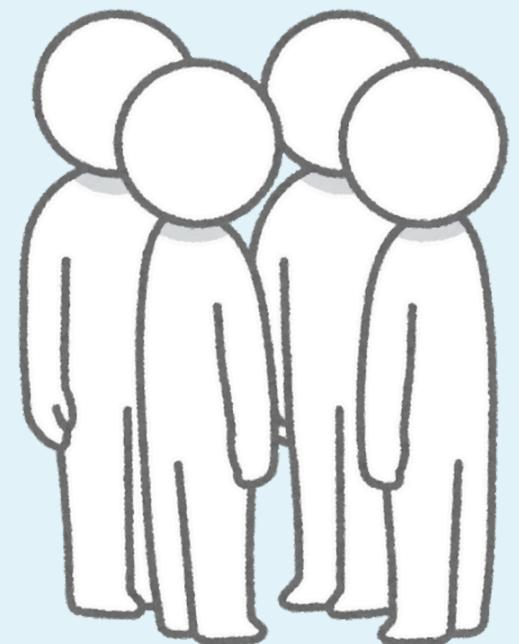
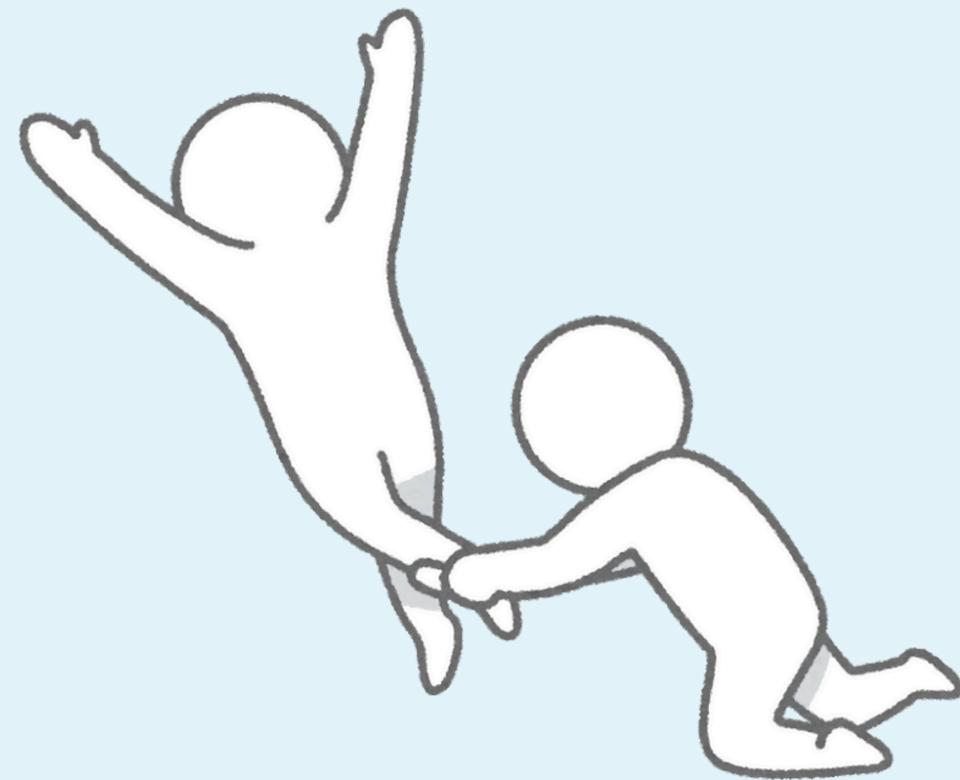
¿Cómo se mide?

- Reducción de impurezas (Entropía, Gini): Árboles de decisión
- Importancia de permutación: KNN
- Magnitud de coeficientes: Modelos lineales

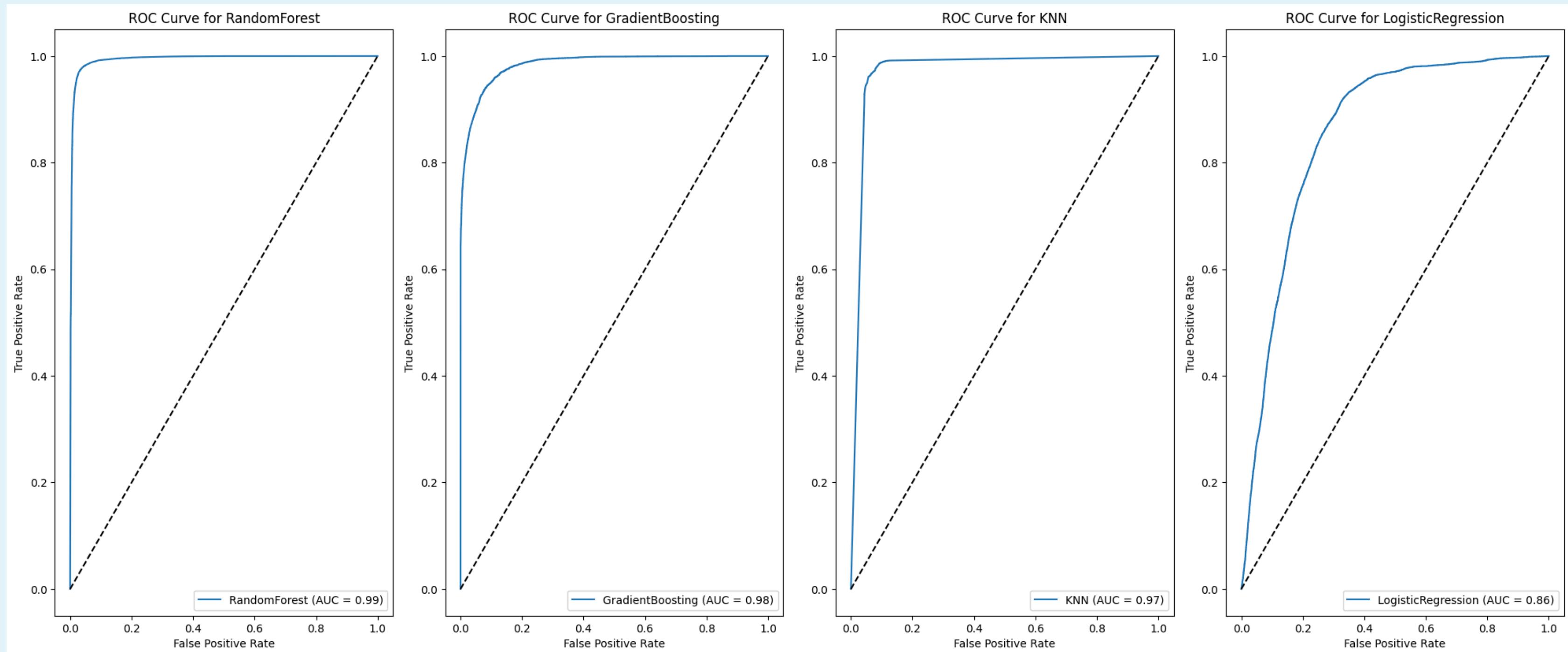




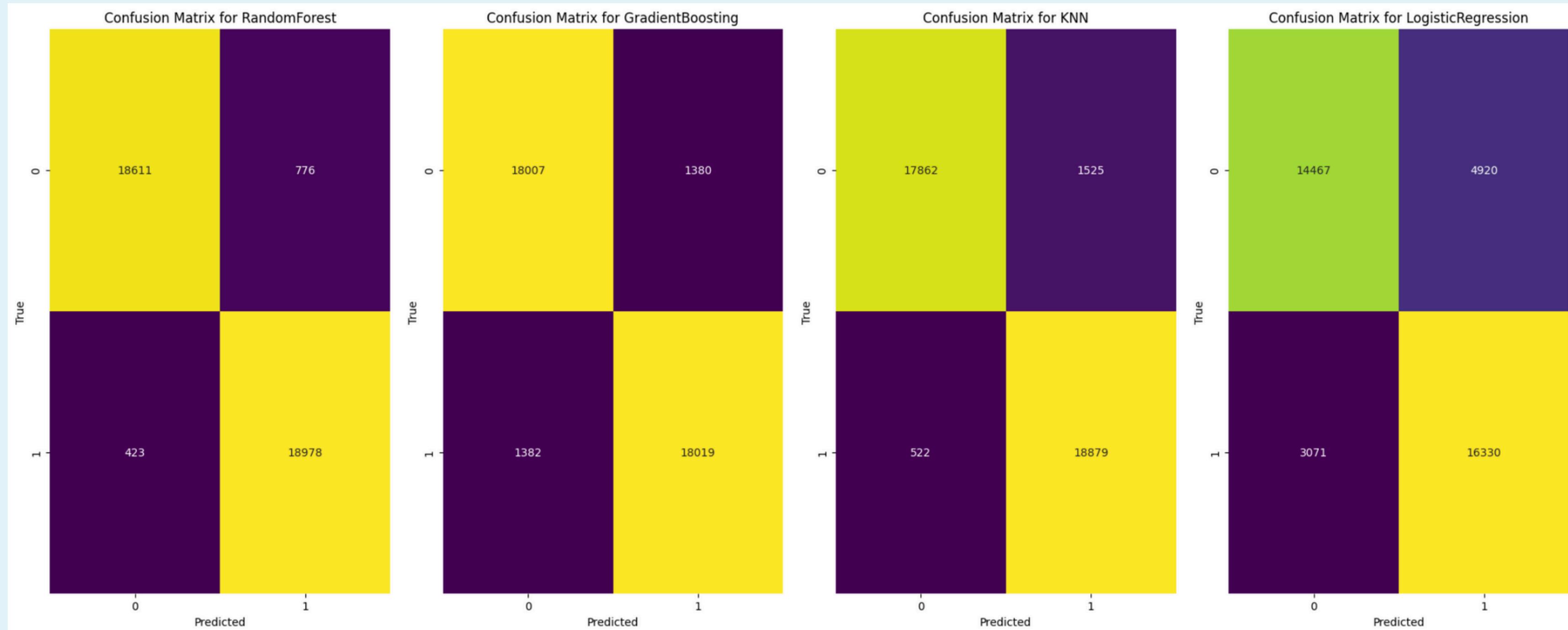
Evaluando los modelos



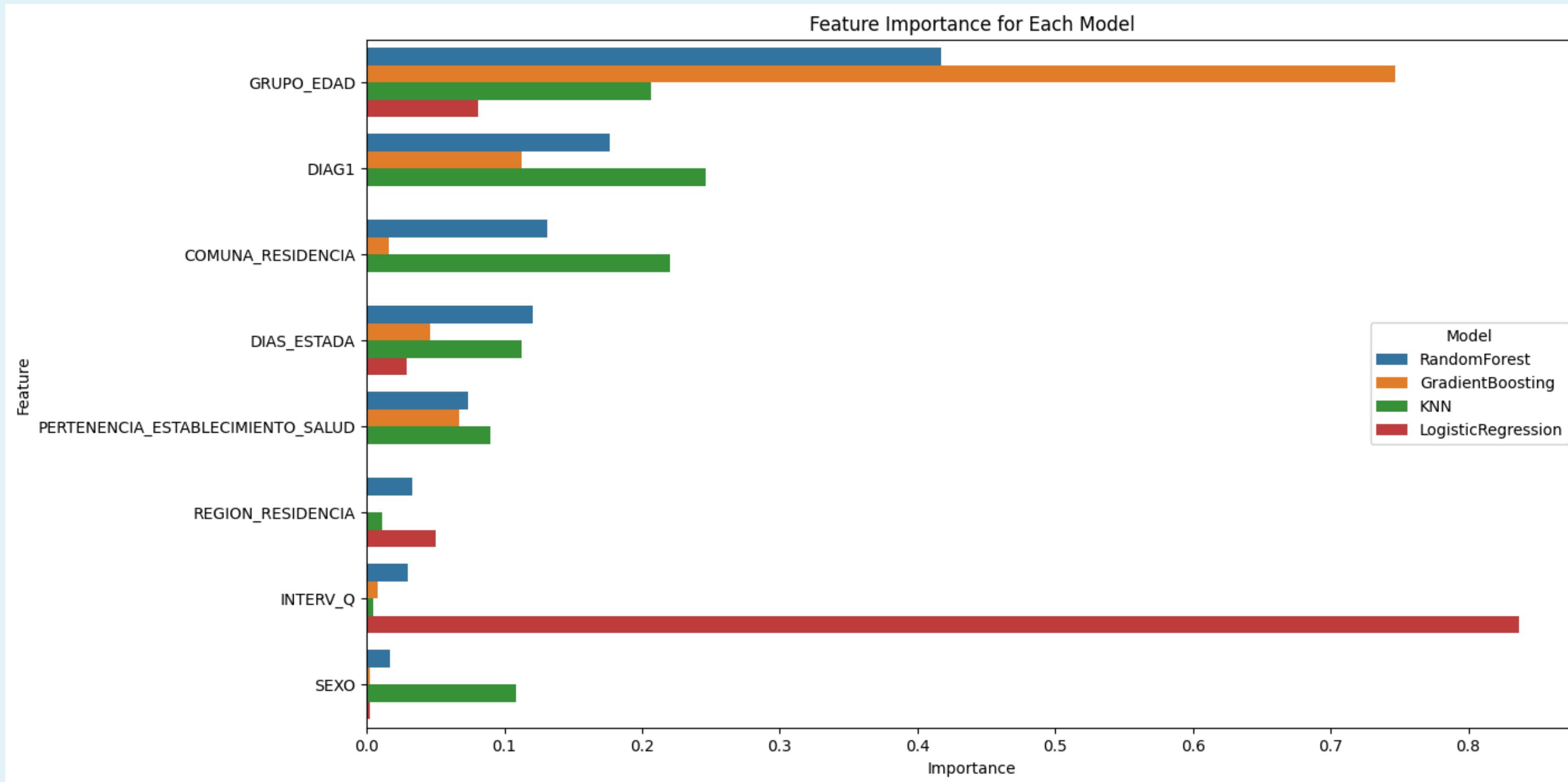
Desempeño de los modelos



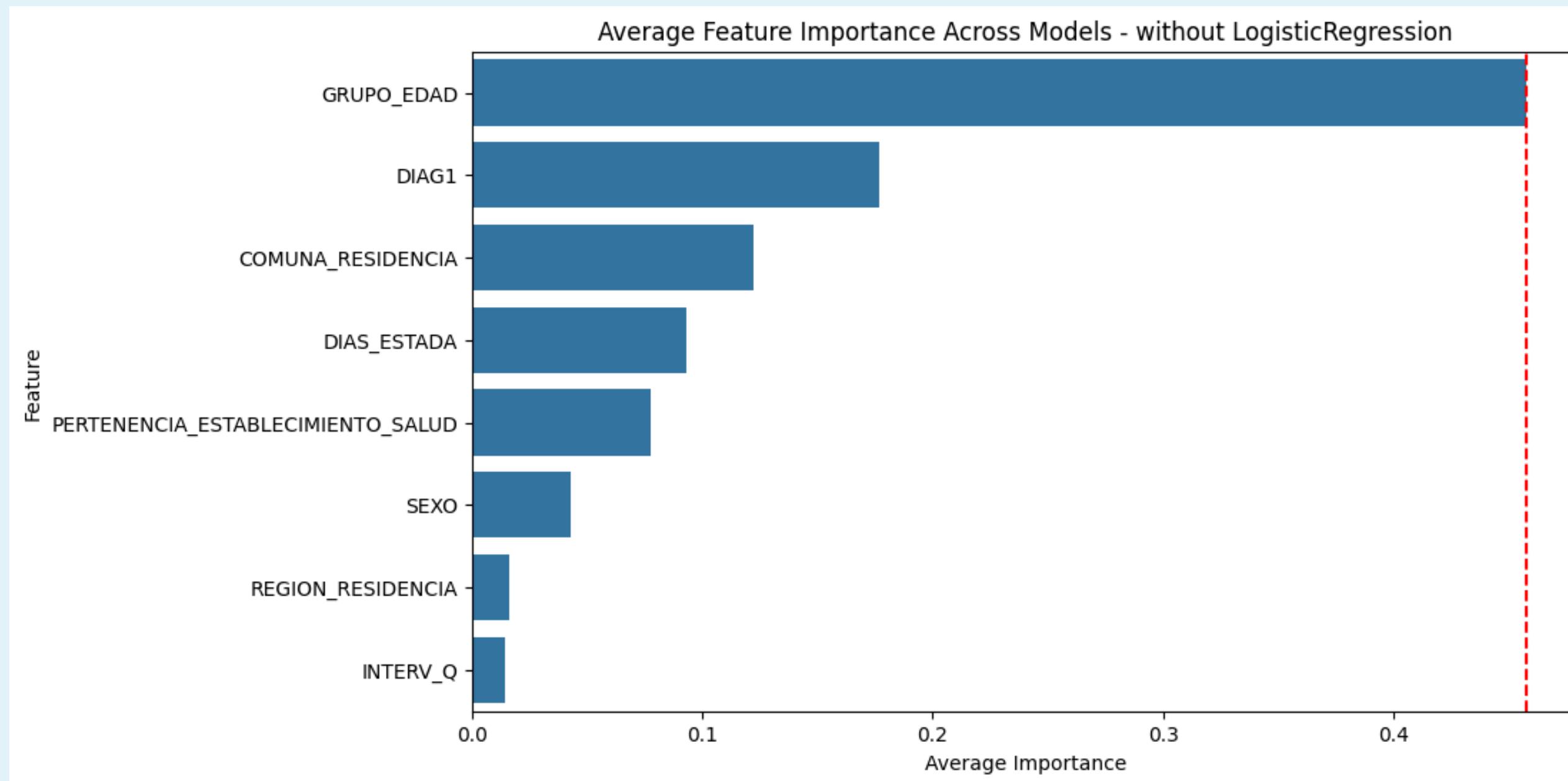
Desempeño de los modelos



Importancia de atributos



Importancia de atributos

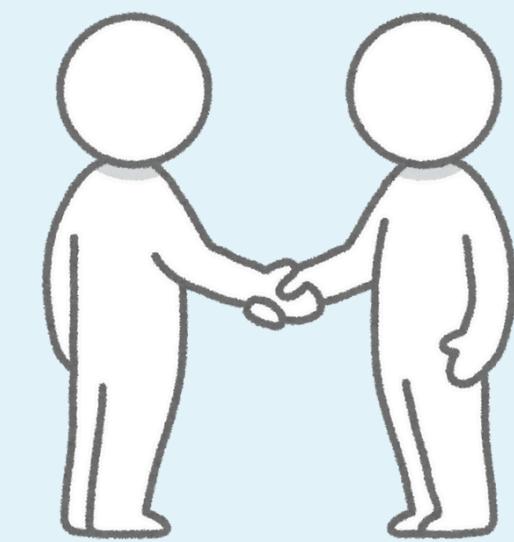
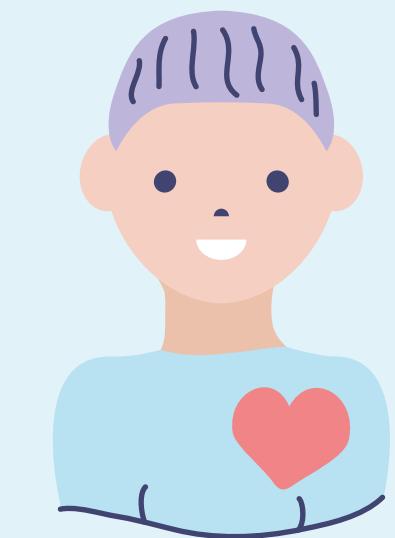


Por lo tanto...

- La **edad** es el factor más importante
- La **comuna y diagnóstico** tienen resultados variados pero relevantes
- Podemos descartar **sexo y sistema de salud**
- Problemas regresión logística



Conclusiones y proyecciones



**¡MUCHAS
GRACIAS!**

