

Mineria de Datos

Proyecto 3.2: “Clasificación”

Integrantes:

- Perez Uribe

- Hernandez Ortiz

Caso de negocio

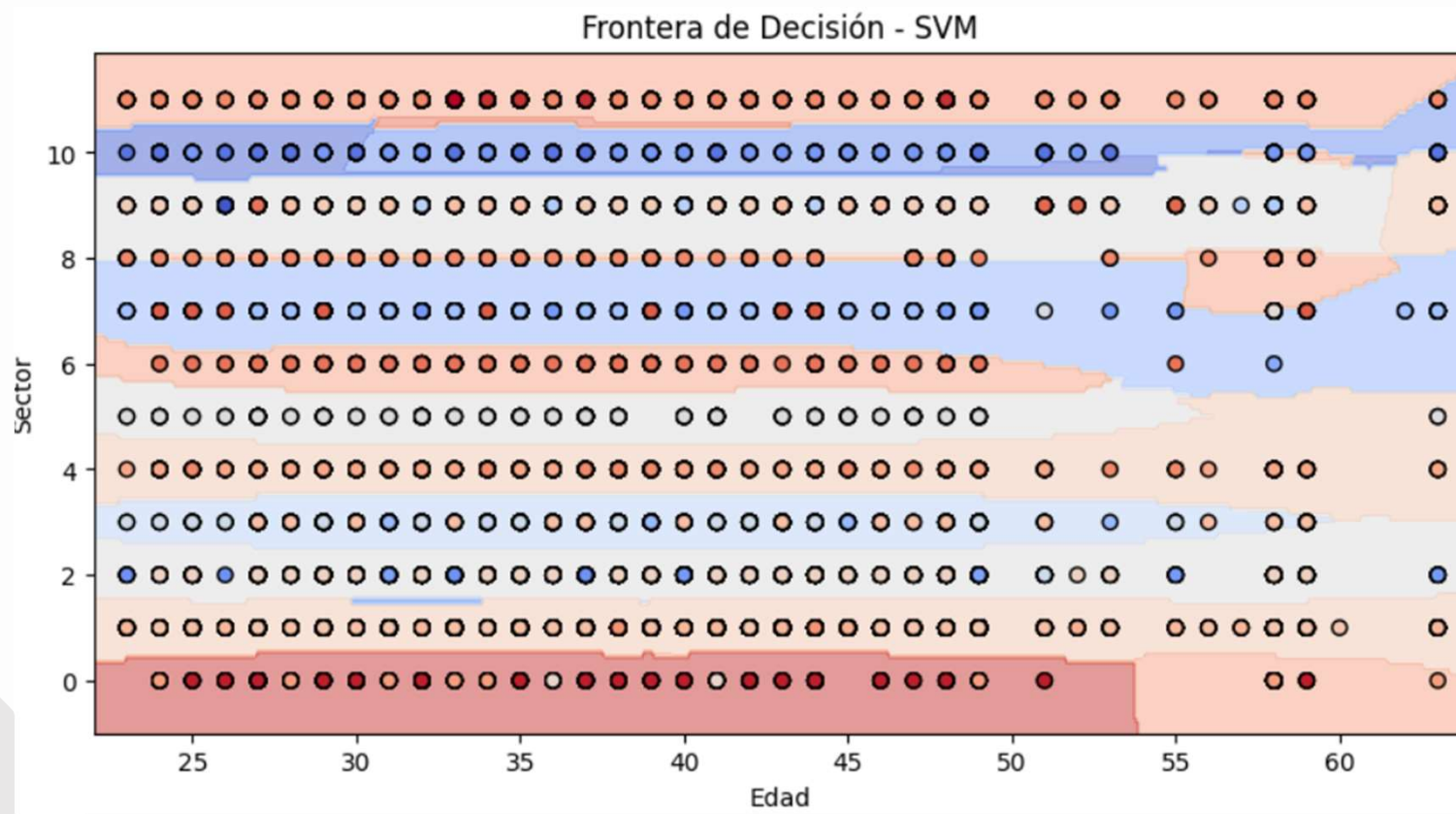
Una compañía quiere optimizar sus campañas de marketing clasificando el tipo de cliente con base en la edad y el sector con el fin de identificar oportunidades del sector en la región.

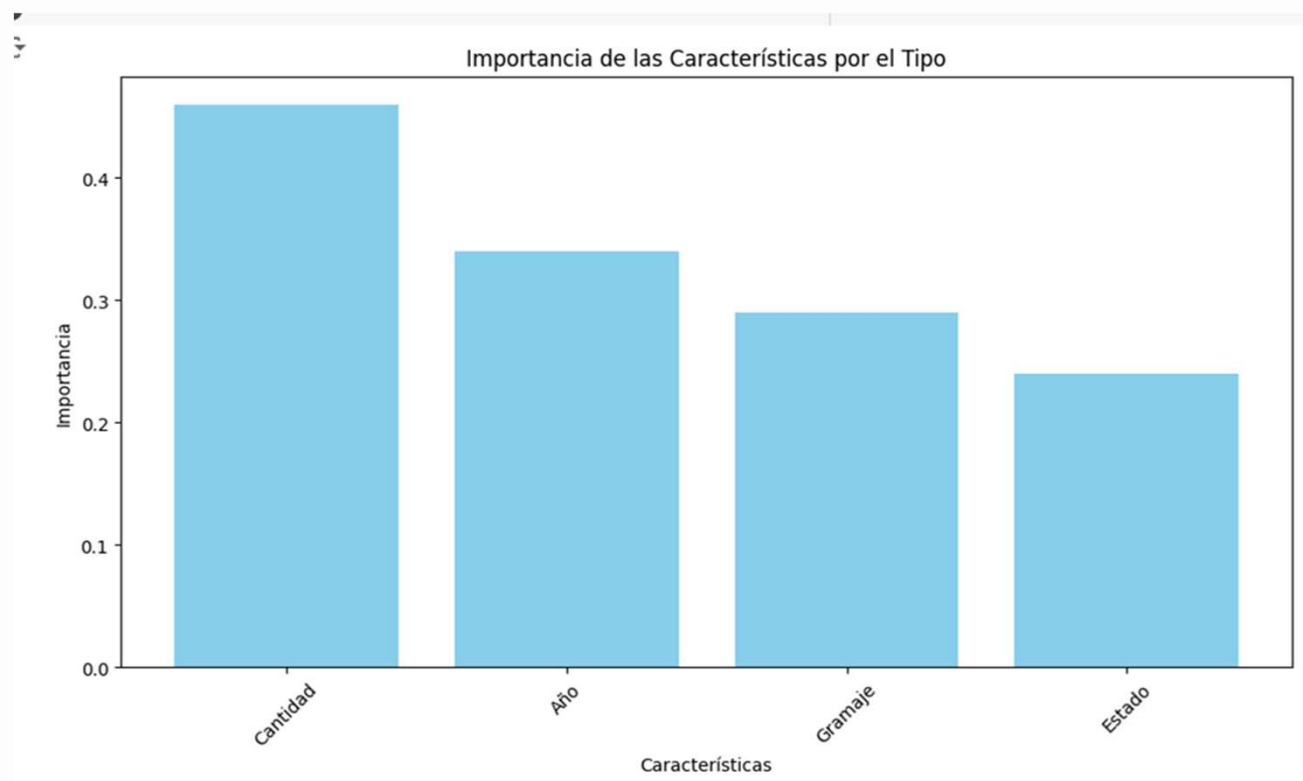
La variable objetivo es el "Tipo", mientras que las variables independientes son "Edad" y "Sector".

Para el árbol de decisión, utilizamos variable objetivo "Estado_Civil" y variables independientes "Antigüedad", "Edad", "No._Hijos", "Sexo",

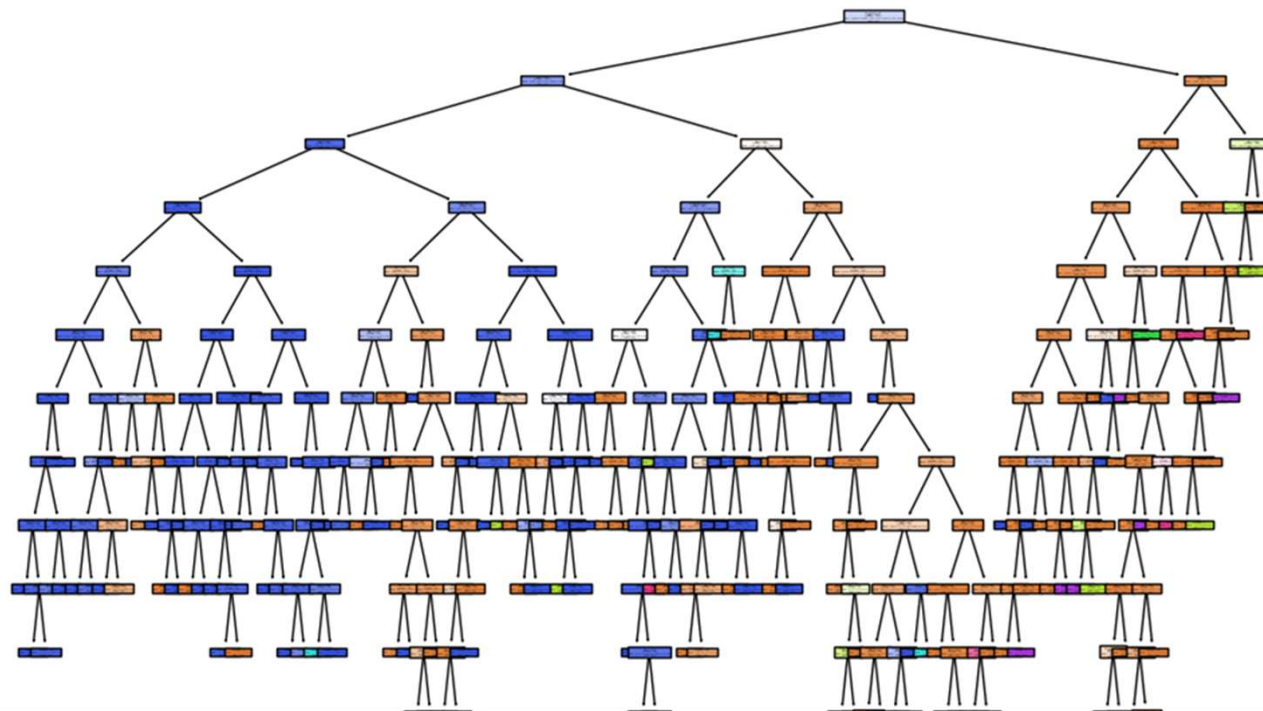
Algoritmos de clasificación

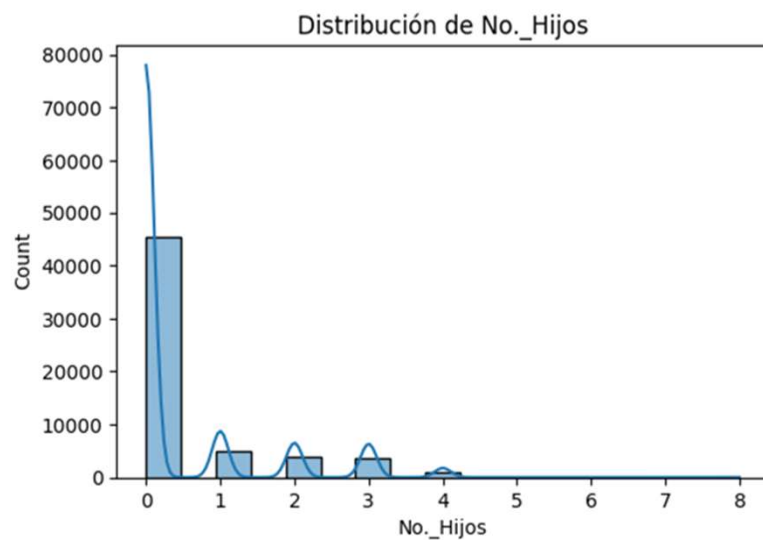
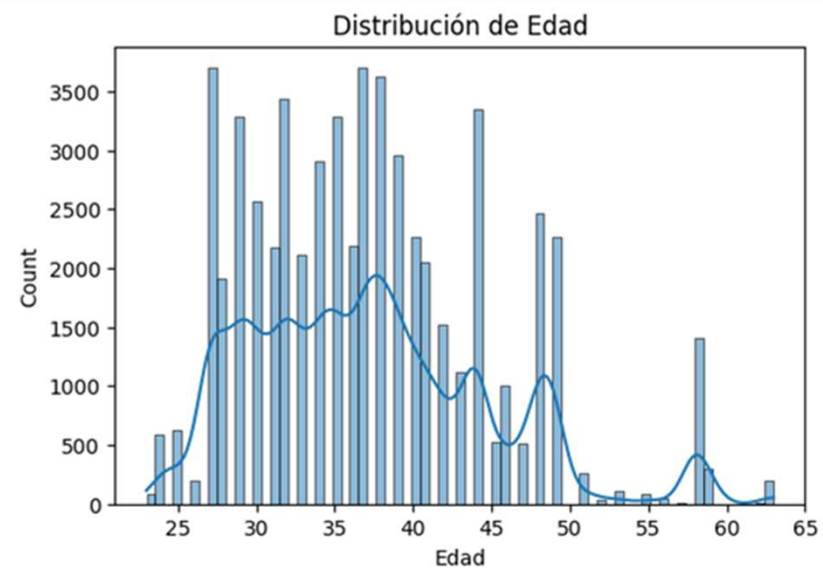
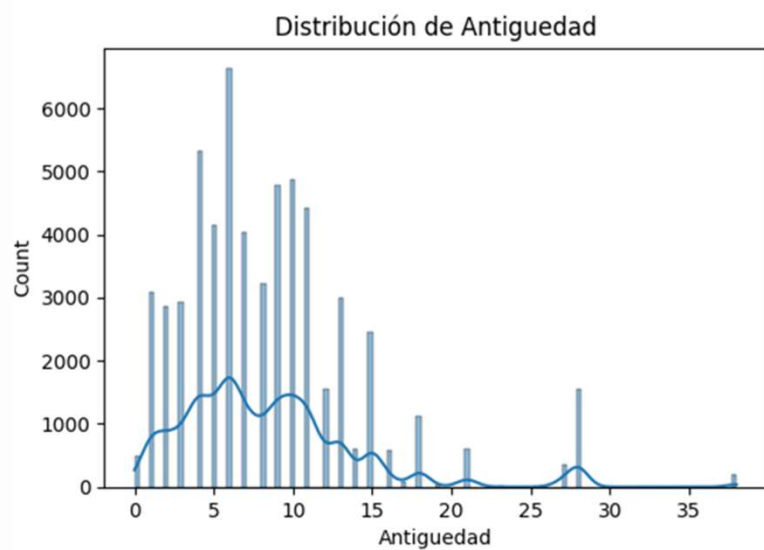
- Modelo SVM
- Árbol de decisión





Árbol de decisión





Conclusiones

La precisión del modelo VSM fue aproximadamente del 0.50, lo que indica que, de todas las predicciones realizadas en el conjunto de prueba, esta proporción fue correcta.

SVM es ideal para problemas no lineales, ya que usa funciones de kernel para transformar los datos. Es eficiente en espacios de alta dimensión y tiene buena generalización al maximizar el margen entre clases.

Árboles de decisión destacan por su interpretabilidad y facilidad para visualizar reglas de decisión, siendo útiles para entender cómo se llega a una predicción. Además, manejan bien tanto datos categóricos como numéricos.