



PRESENTACIÓN

# Análisis: Probabilidad de Deserción

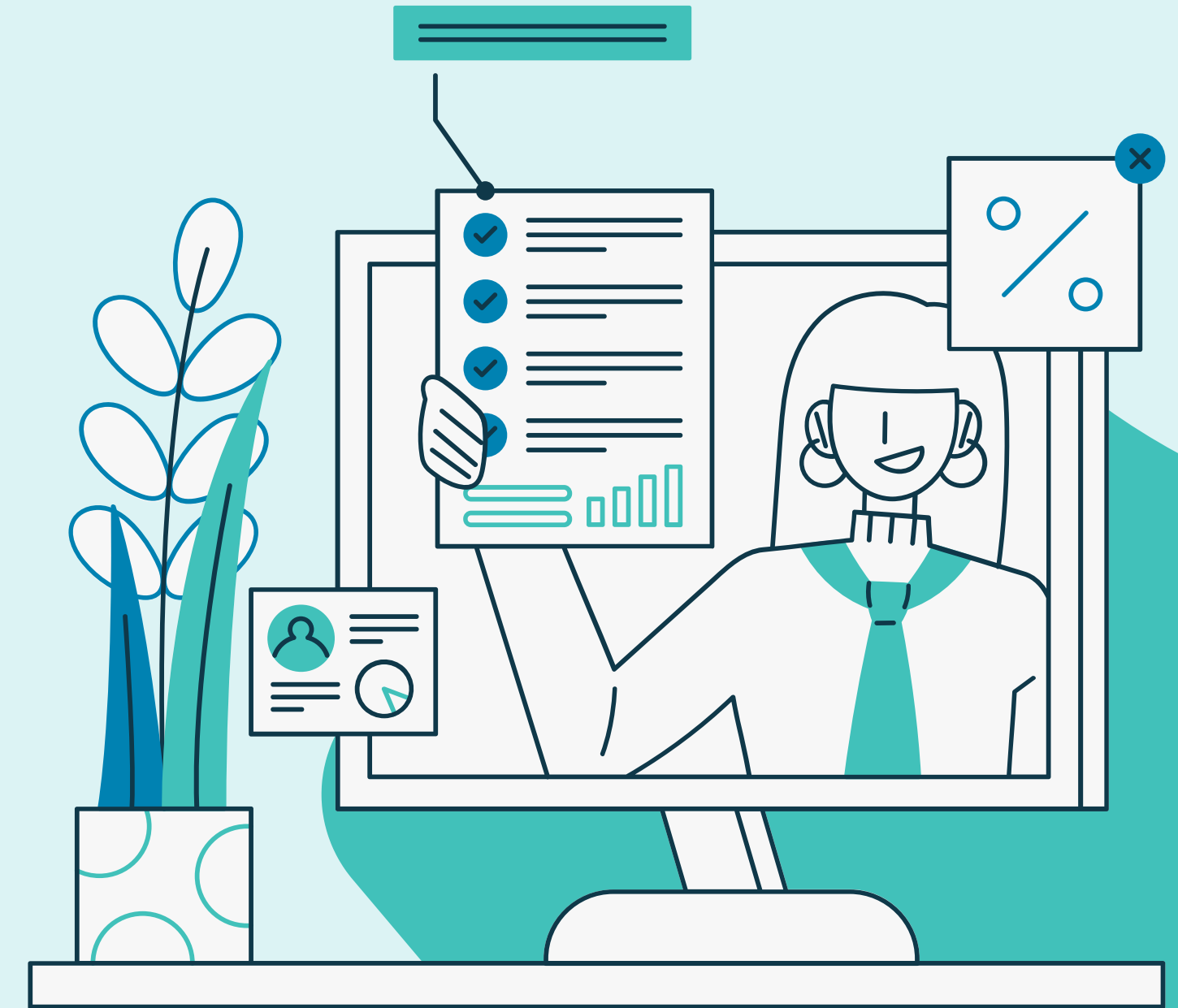
¿Cómo tomar decisiones que generen  
menos desertores universitarios?



María Gabriela Urango  
Andrea Velandia

# Contenido

- 01.** Descripción del problema
- 02.** Objetivos
- 03.** Estadísticas Descriptivas
- 04.** Selección de variables
- 05.** Modelo de Red Bayesiana
- 06.** Inferencia
- 07.** Herramientas para la toma de decisiones
- 08.** Conclusiones y recomendaciones



# Descripción del problema

3.8 billones de dólares le cuesta a la economía por año los estudiantes desertores.

- ThinkImpact Magazine

<https://www.thinkimpact.com/college-dropout-rates/>

Es de interés para las naciones garantizar educación de calidad y de acceso equitativo para todos.

- Objetivo de Desarrollo Sostenible #4

<https://sdgs.un.org/goals/goal4>

**¿Cómo ayudamos a las universidades a aumentar su tasa de retención de estudiantes con el fin de garantizar acceso a la educación y disminuir el impacto negativo que los desertores tienen en la economía de los países?**



## Objetivo general

Desarrollar una **herramienta de analítica de datos para la toma de decisiones** que permita la **detección temprana de la deserción estudiantil** para que las instituciones universitarias puedan invertir en estrategias enfocadas en el éxito académico de los estudiantes.

## Objetivos específicos

01.

Analizar la información histórica del comportamiento estudiantil recolectada por la universidad en la base de datos.

02.

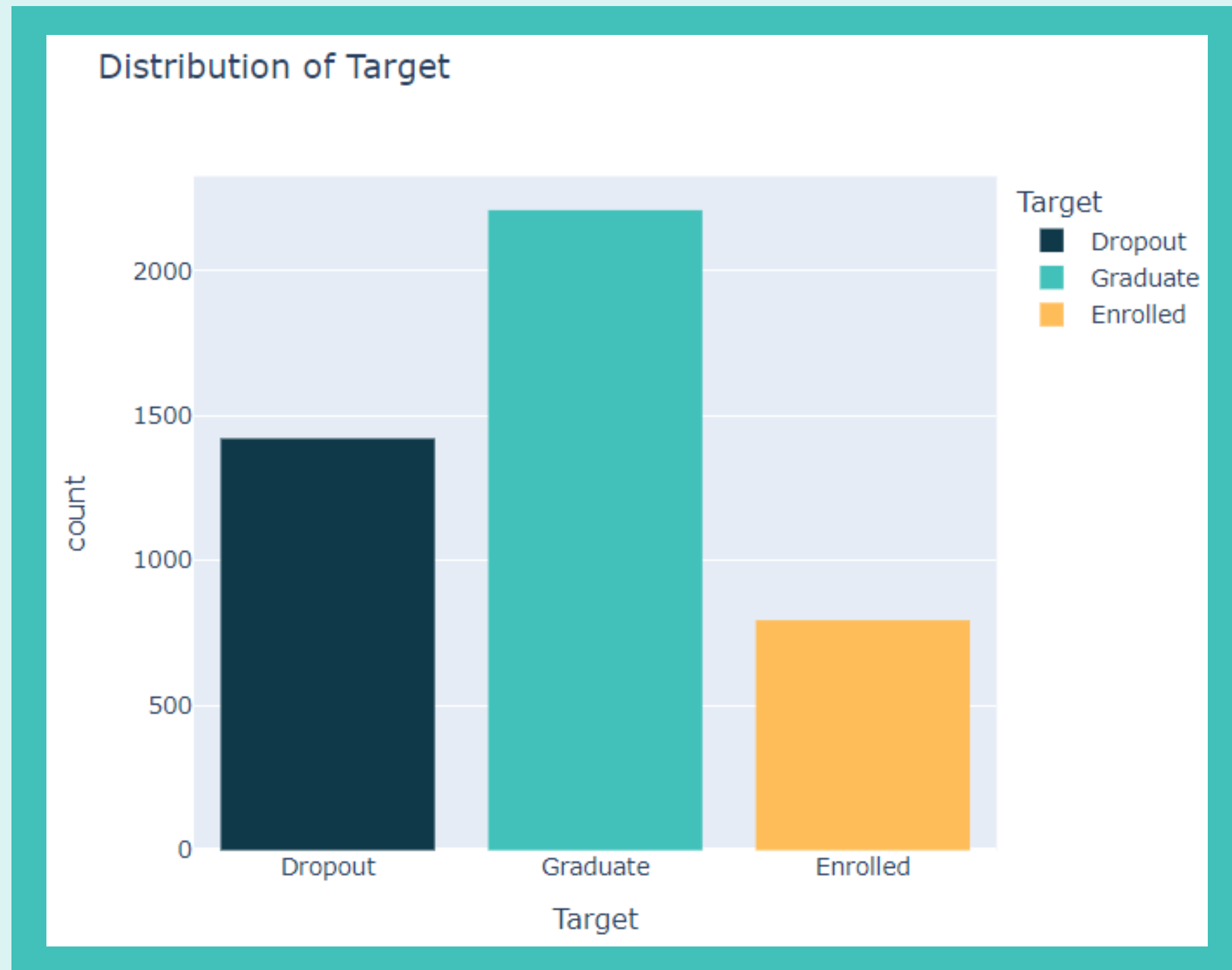
Proponer un modelo de red Bayesiana que permita inferir la probabilidad de deserción de un estudiante

03.

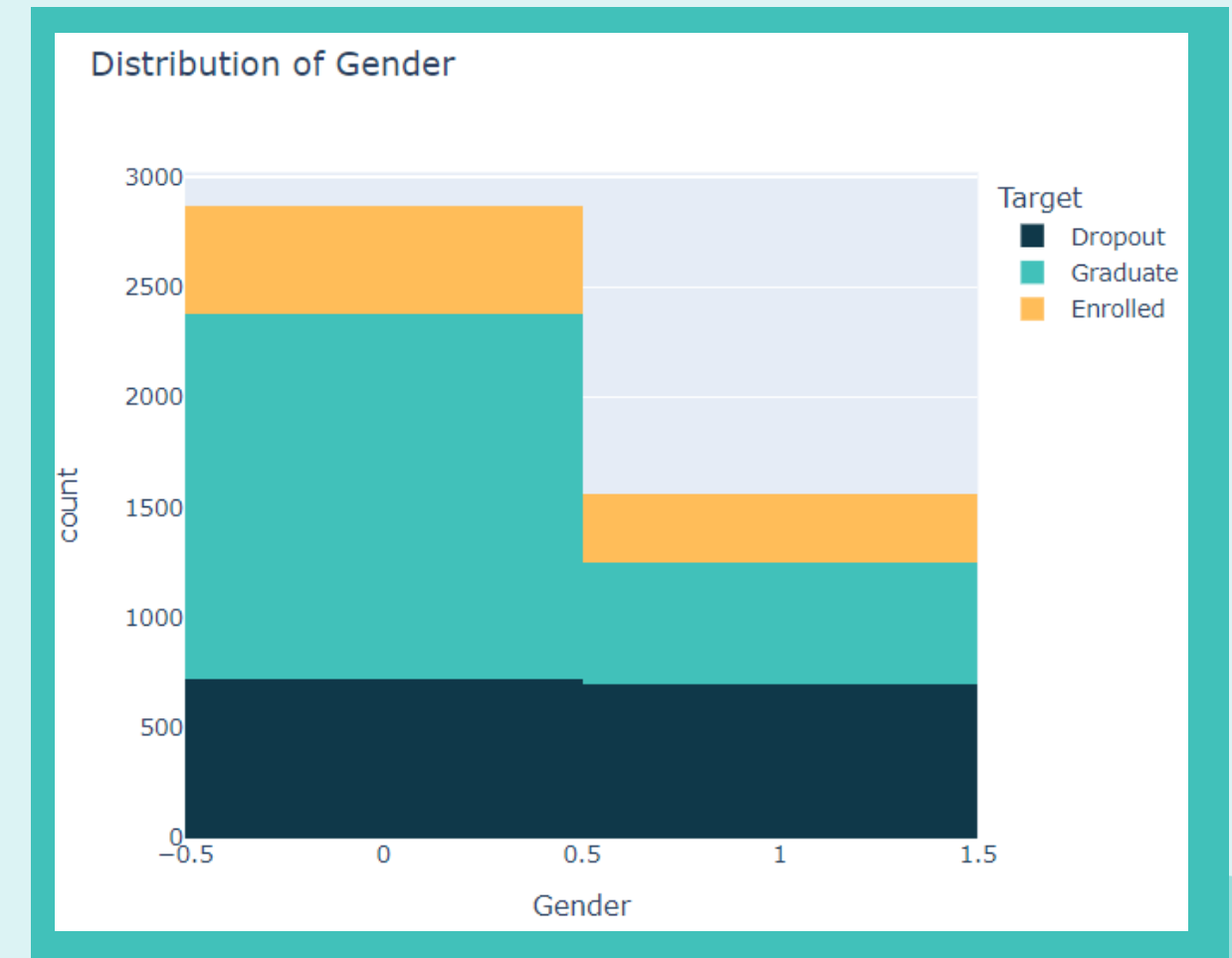
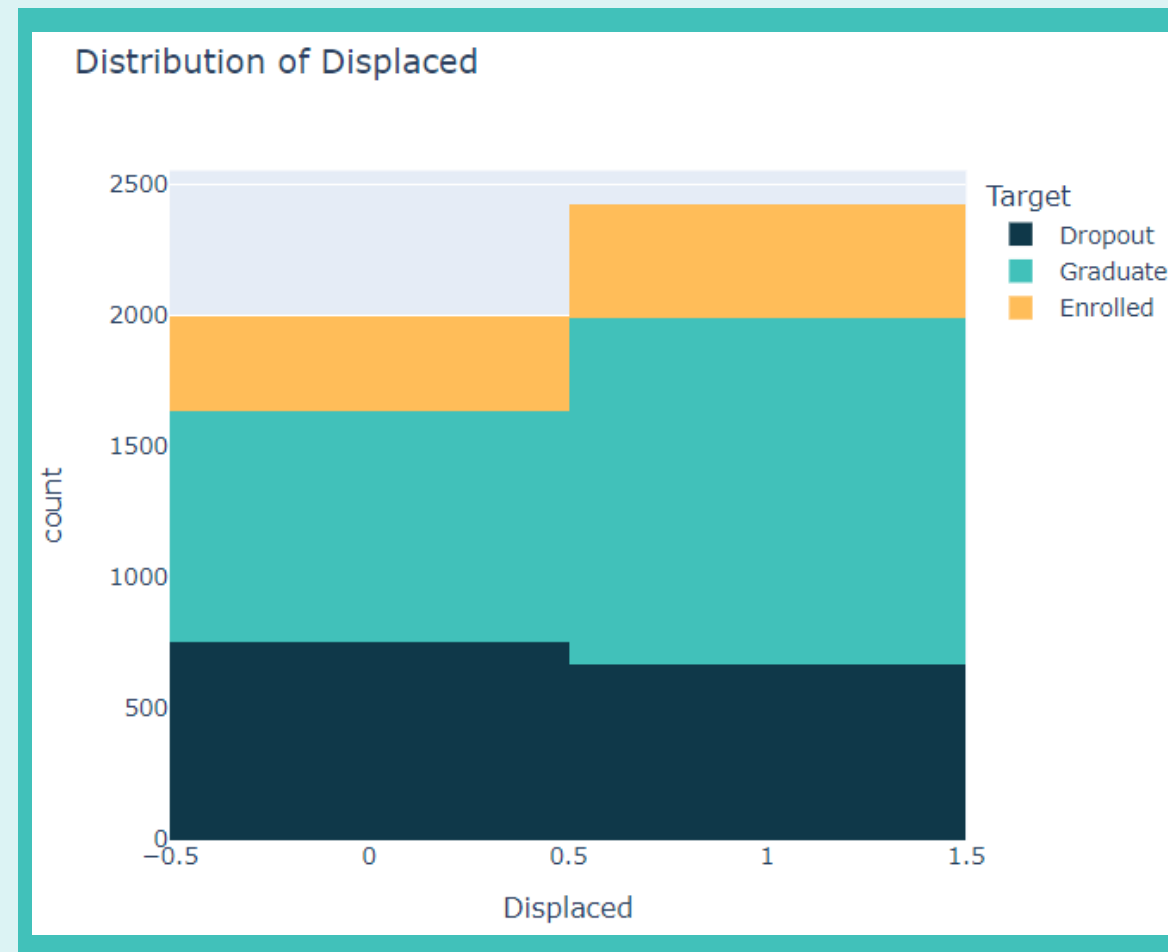
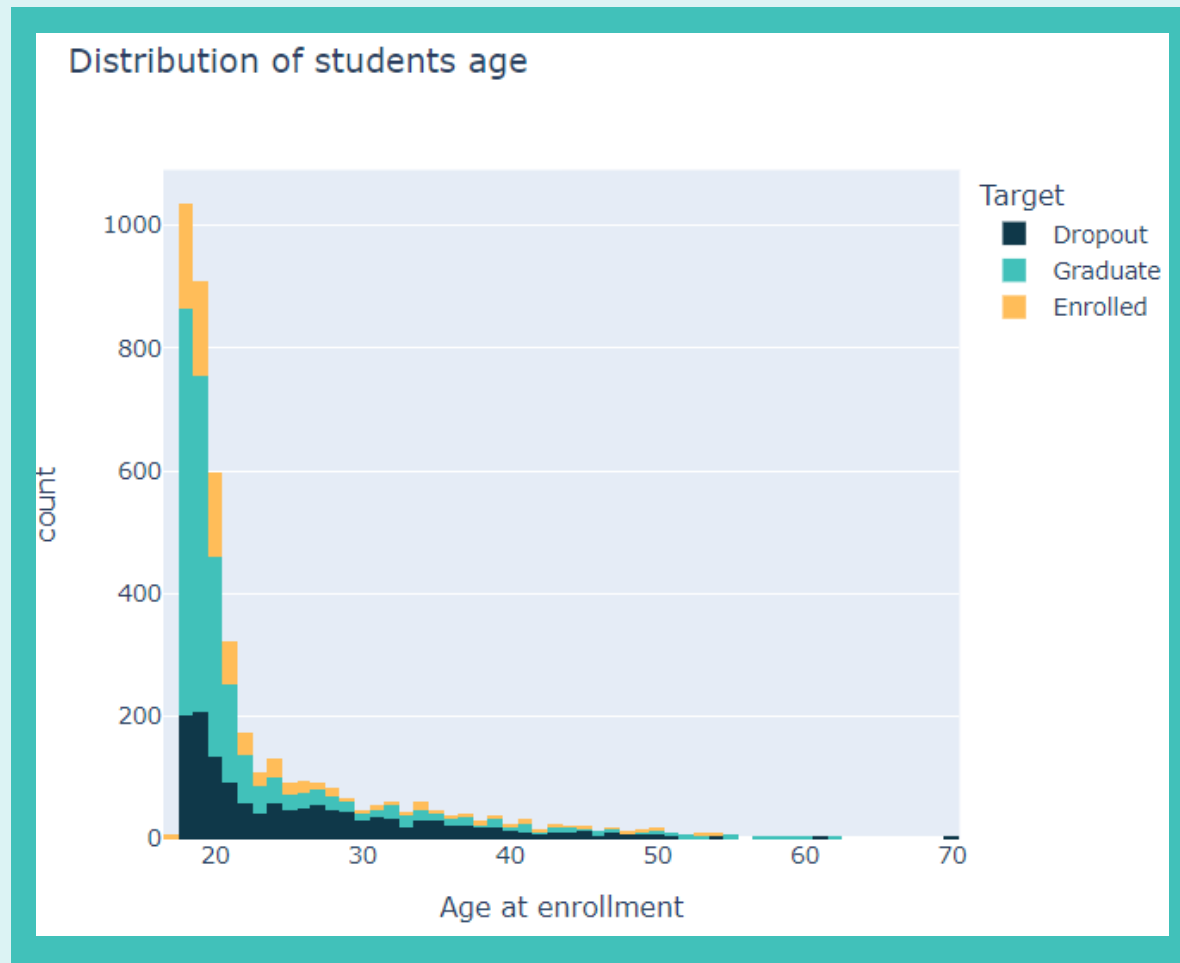
Implementar el modelo predictivo en una interfaz interactiva para la evaluación de distintos casos de estudiantes universitarios



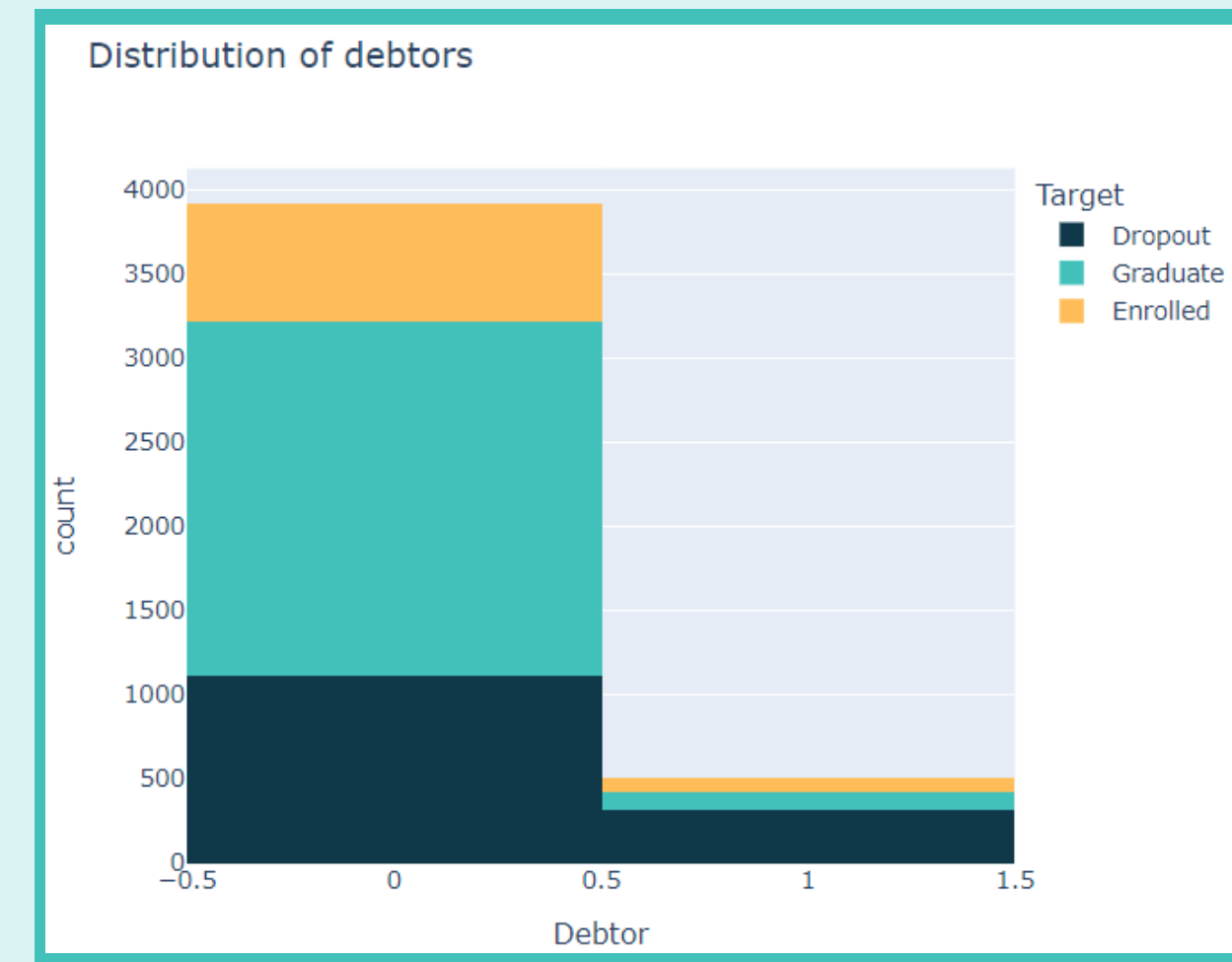
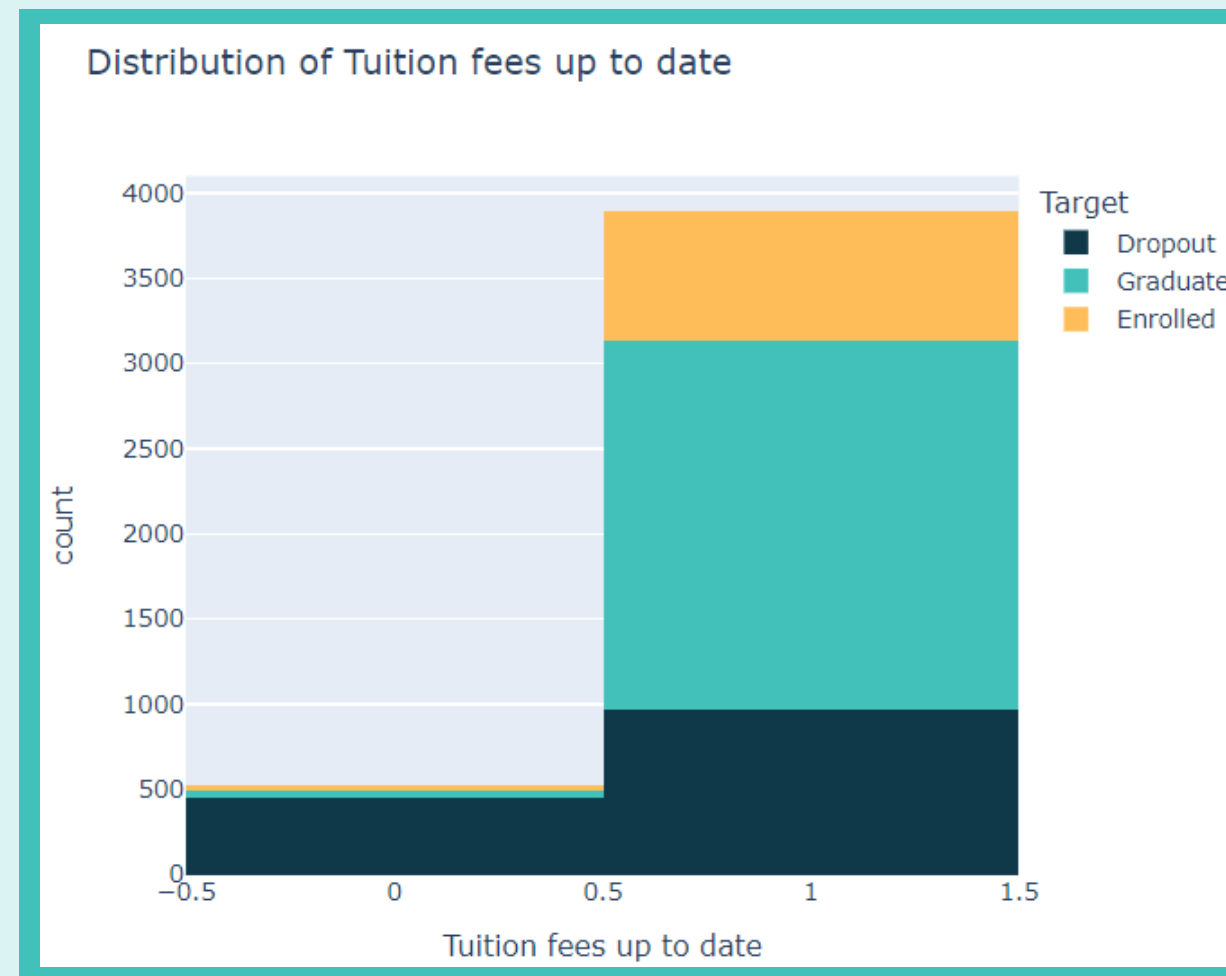
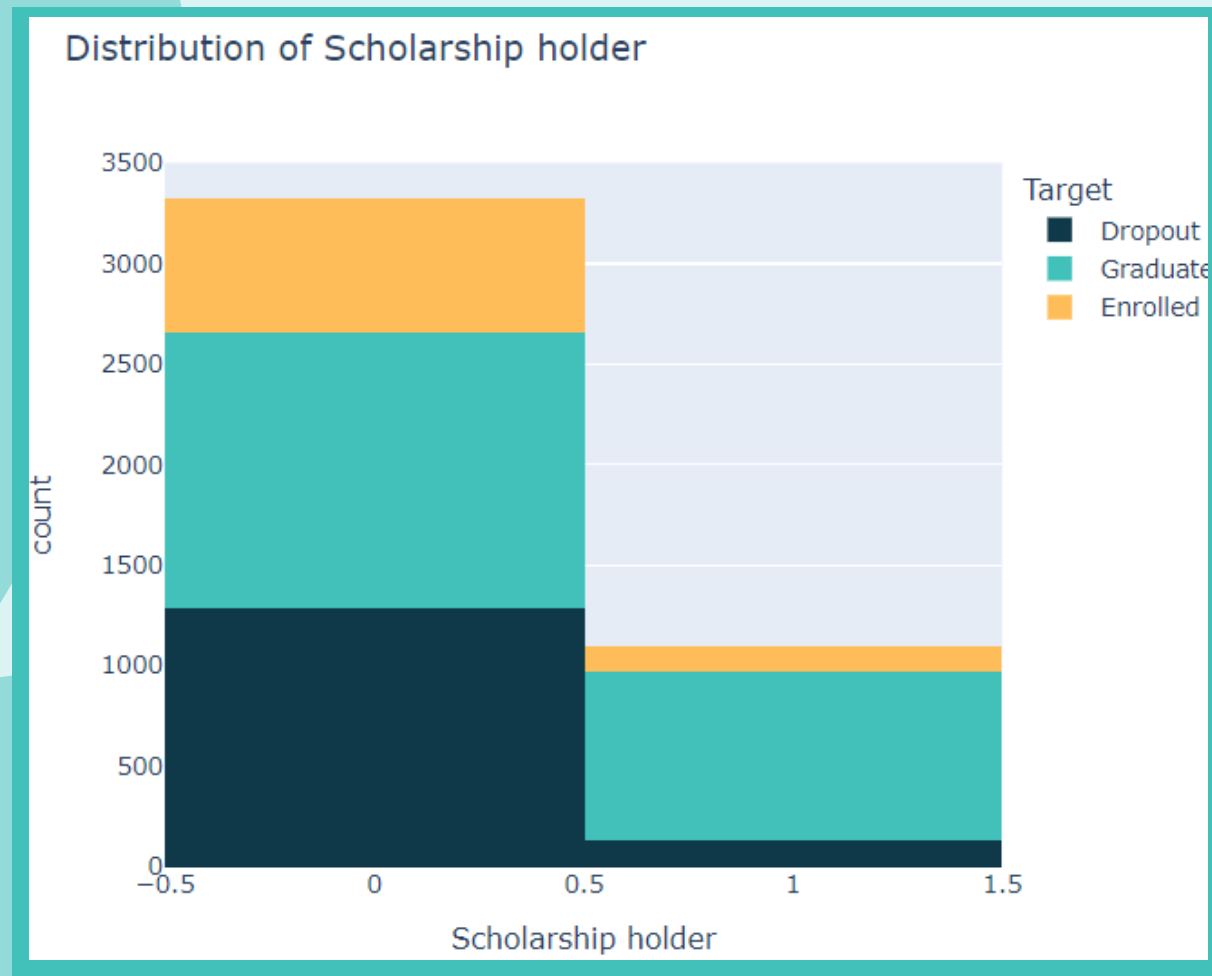
# Estadísticas Descriptivas



# Variables Socio-Demográficas

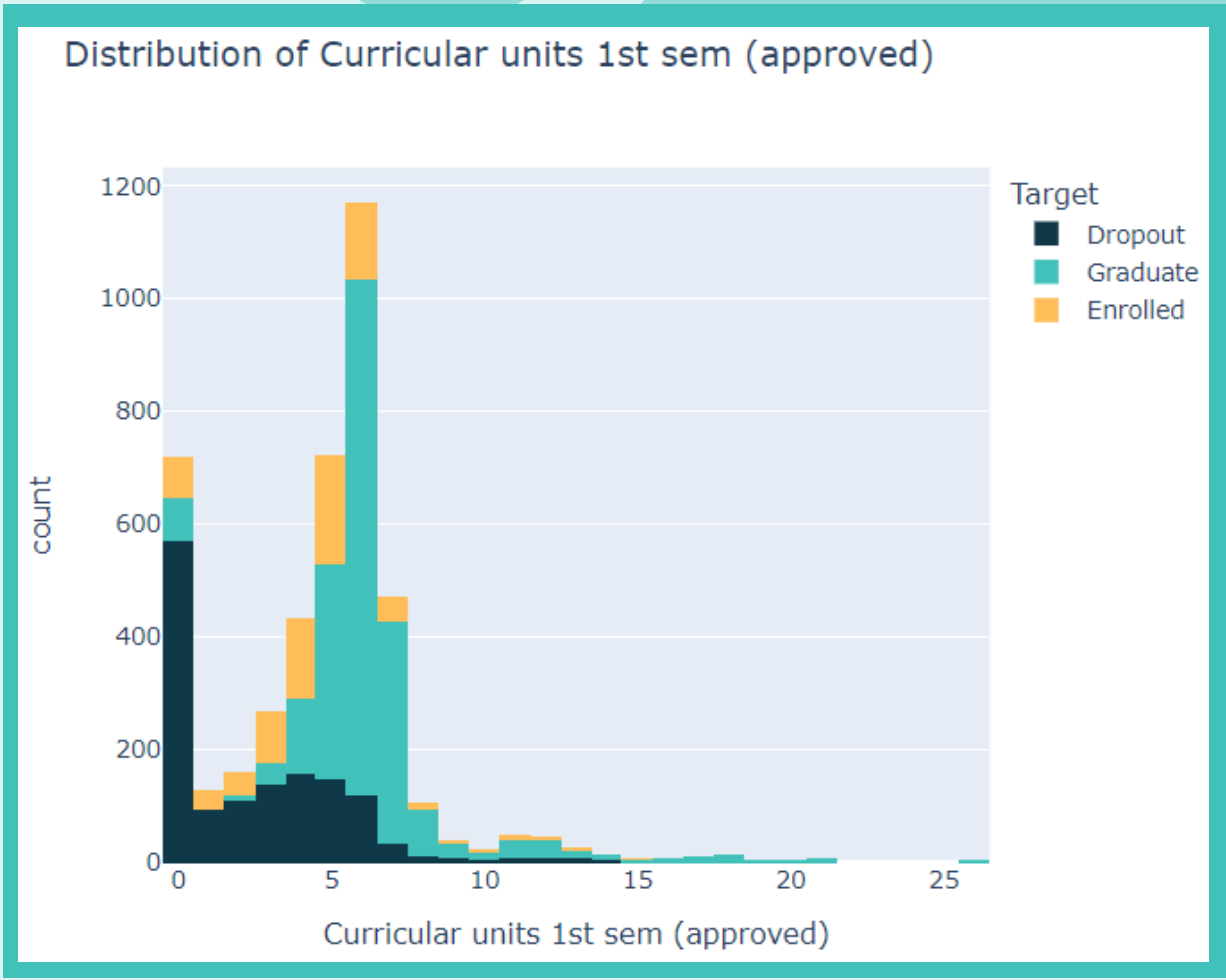
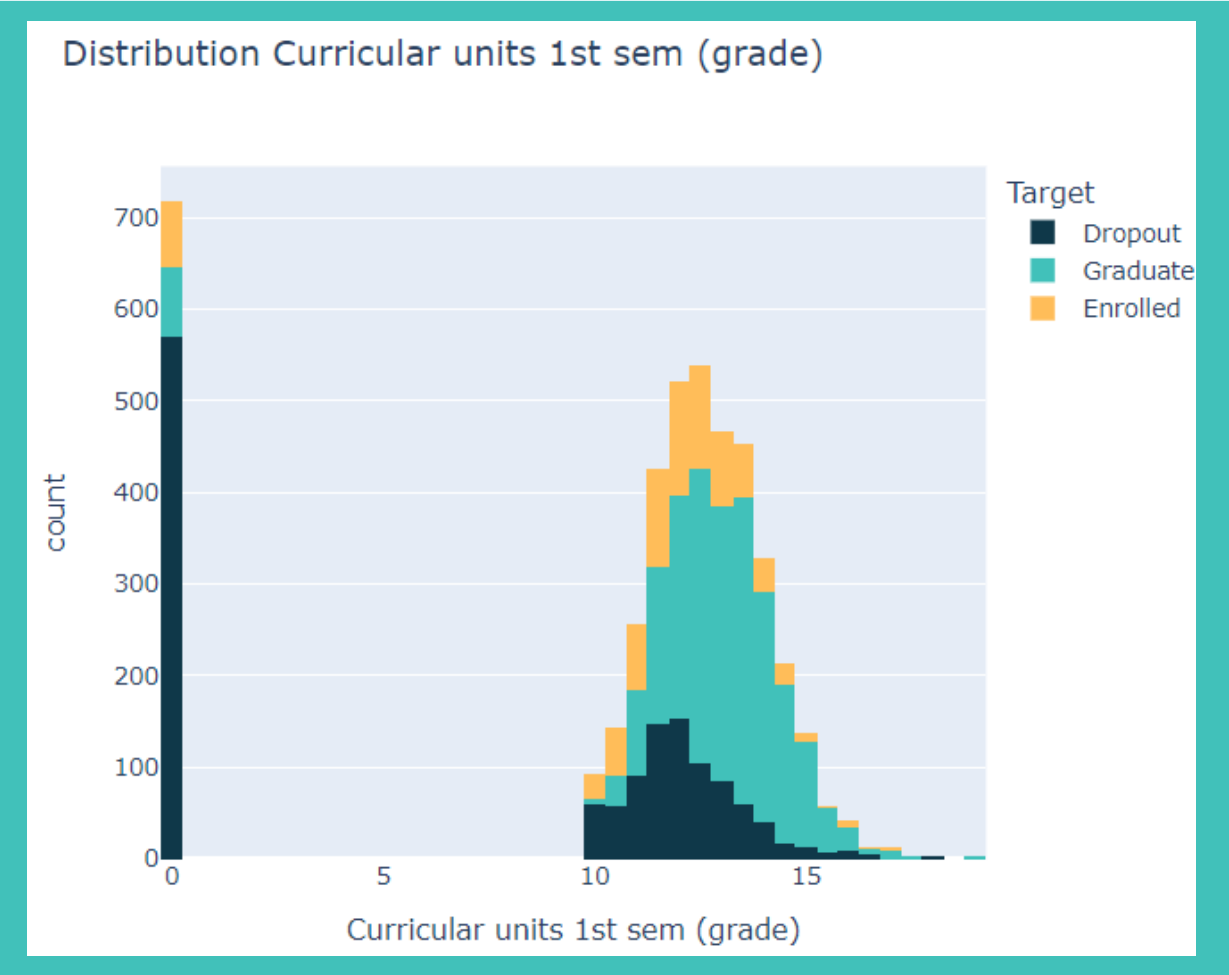
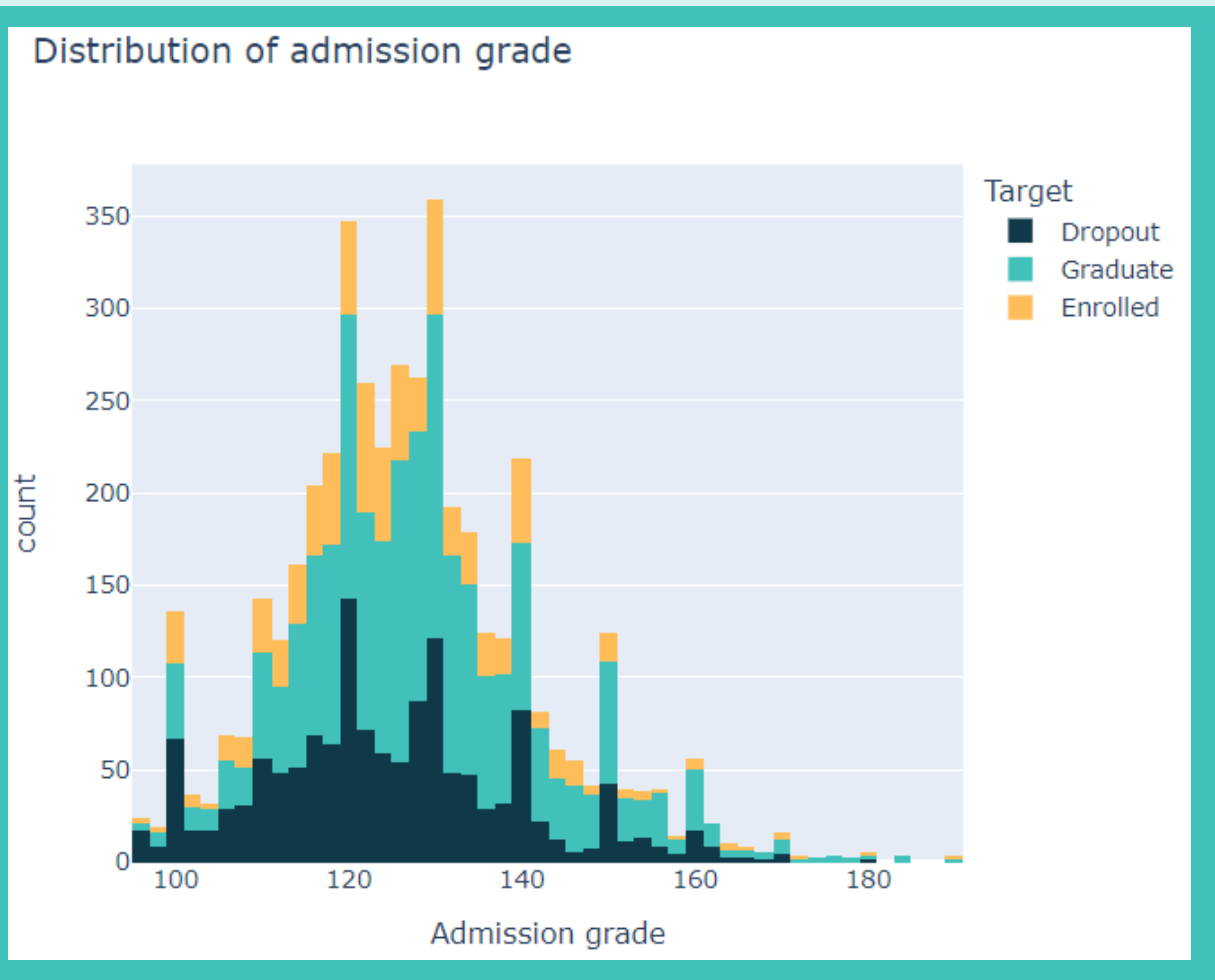


# Variables Financieras





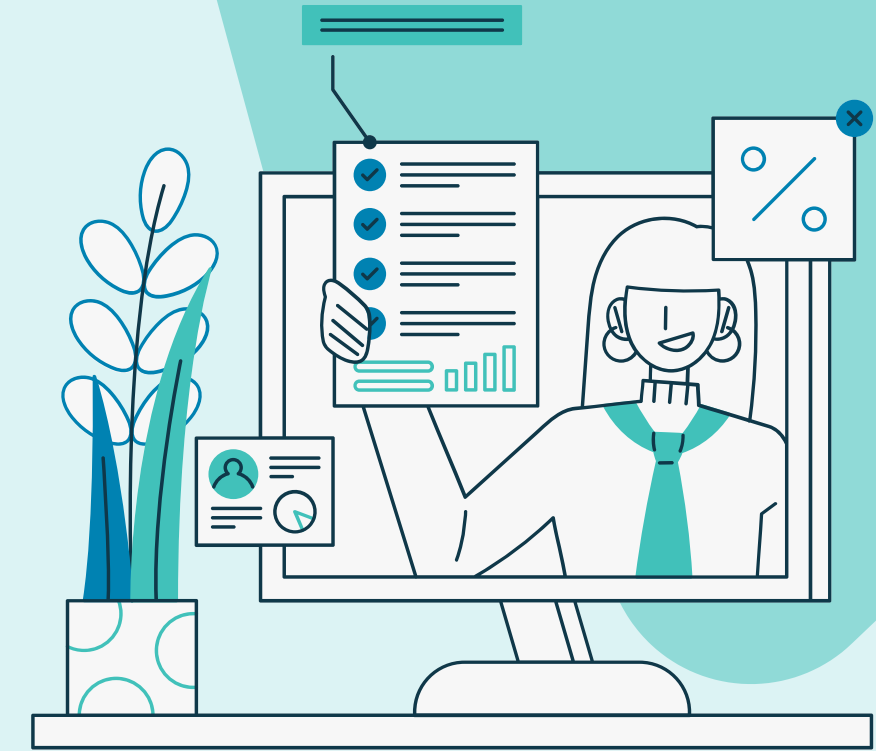
# Variables Académicas





# Modelización

# Métodología



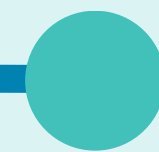
## FASE 1



### Selección de Variables

Selección de factores influyentes en la deserción universitaria.

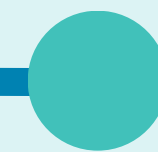
## FASE 2



### Transformación de Datos

Discretización de variables seleccionadas.

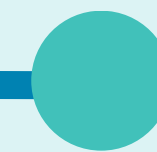
## FASE 3



### Modelo Red Bayesiana

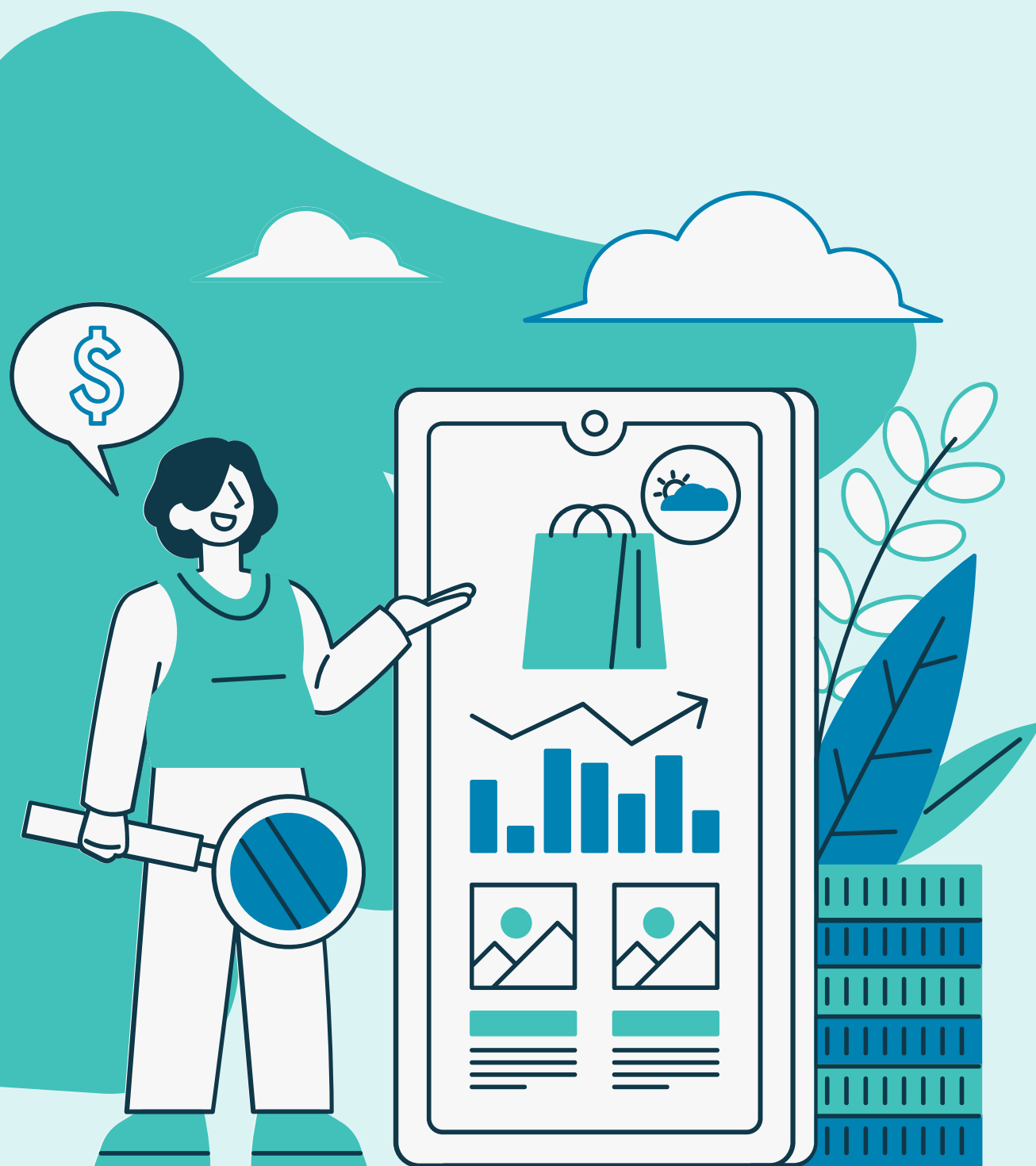
Diseño de estructura de modelo causal a través de Red Bayesiana.

## FASE 4



### Estimación de Parámetros

Parámetroización mediante estimación de máxima verosimilitud.



# Selección de variables

## Edad



15% de los estudiantes de 19 años o menores desertan  
54% de los estudiantes mayores a 30 años desertan

## Género



Las tasas de deserción son 20% más altas para estudiantes hombres que estudiantes mujeres

## Desplazado



El 4% de los estudiantes que desertan lo hacen por la distancia.  
El 13% de los estudiantes consideran difícil encontrar vida social en la Universidad.

## Becado



# Selección de variables



**Deudor**



**Matrícula al día**



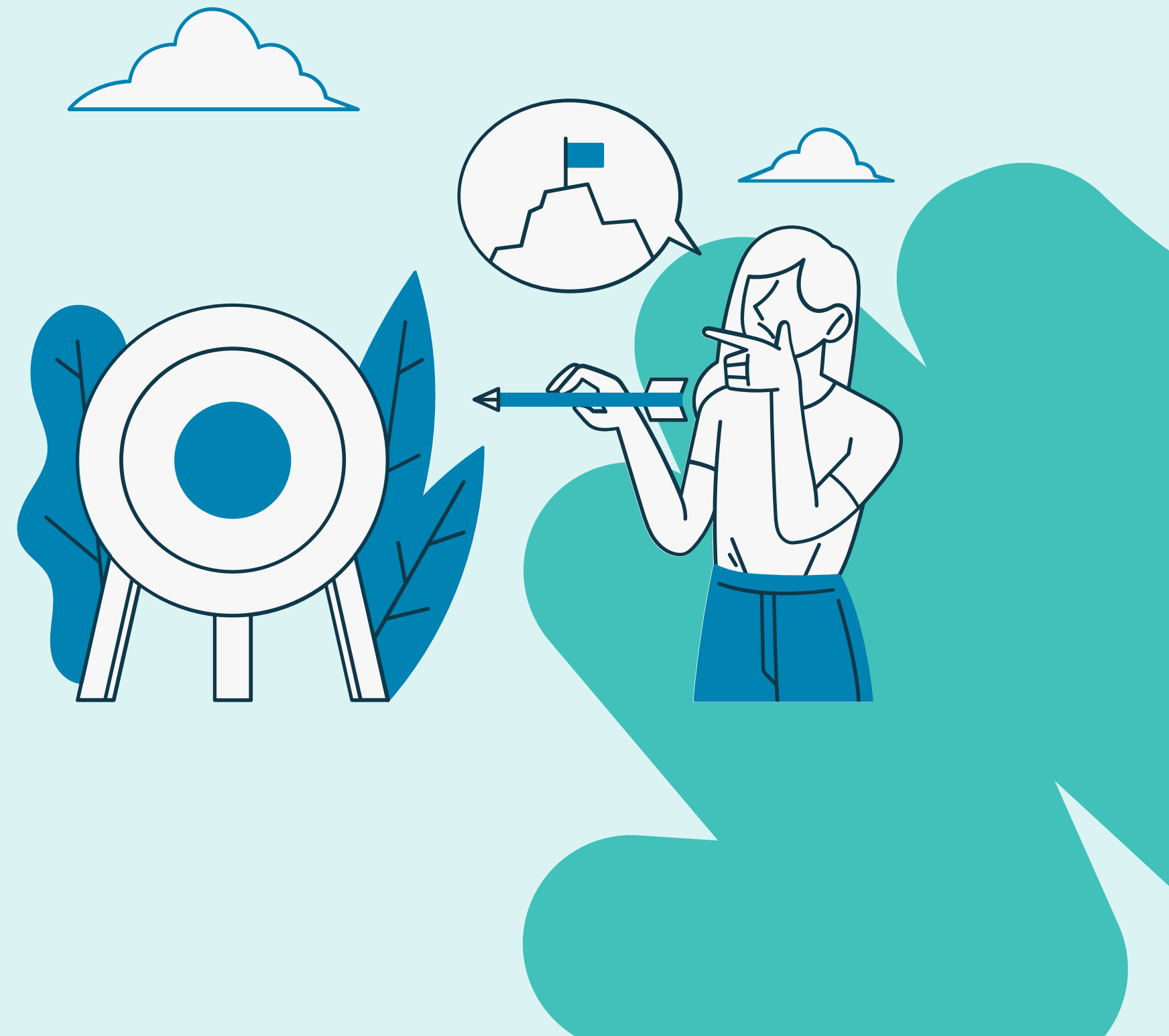
**Créditos aprobados en primer semestre**

El 28% de los estudiantes que desertan lo hacen por sentirse descalificados o incapaces de alcanzar los requerimientos académicos



**Nota de admisión**

Las Universidades con menor tasa de aceptación reportaron las mayores tasas de retención de estudiantes



# Modelo de Red Bayesiana

## 01.

Modelo Naive Bayes con dimensiones sociodemográficas, económicas y académicas.

Variables sociodemográficas

*Age at enrollment*

*Gender*

*Displaced*

Variables económicas

*Scholarship holder*

*Debtor*

*Tuition fees up to date*

Variables académicas

*Curricular units 1st sem (grade)*

*Curricular units 1st sem (approved)*

*Admission grade*

Target

```
graph LR; A[Age at enrollment] --> T[Target]; B[Gender] --> T; C[Displaced] --> T; D[Scholarship holder] --> T; E[Debtor] --> T; F[Tuition fees up to date] --> T; G[Curricular units 1st sem (grade)] --> T; H[Curricular units 1st sem (approved)] --> T; I[Admission grade] --> T;
```

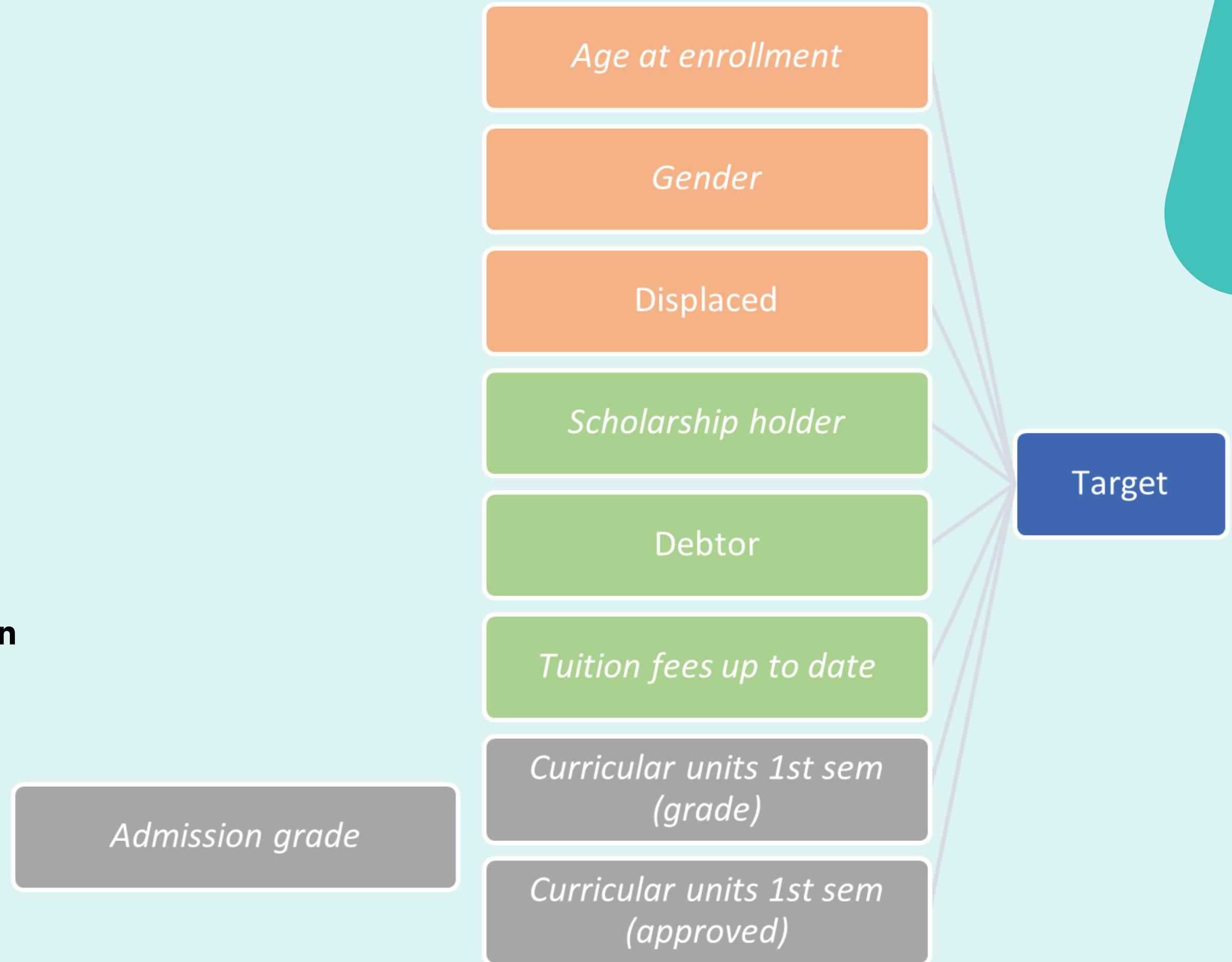
The diagram illustrates a Bayesian Network for a Naive Bayes model. It features nine input variables on the left, grouped into three categories: sociodemographic (orange), economic (green), and academic (grey). Each variable is represented by a rounded rectangle. Lines connect each of these nine variables to a single 'Target' node on the right, which is also a rounded rectangle. The variables are: 'Age at enrollment', 'Gender', 'Displaced', 'Scholarship holder', 'Debtor', 'Tuition fees up to date', 'Curricular units 1st sem (grade)', 'Curricular units 1st sem (approved)', and 'Admission grade'. The 'Target' node is the final outcome predicted by the model.

# Modelo de Red Bayesiana

## 02.

**Modelo Red Bayesiana con cambios en las variables académicas.**

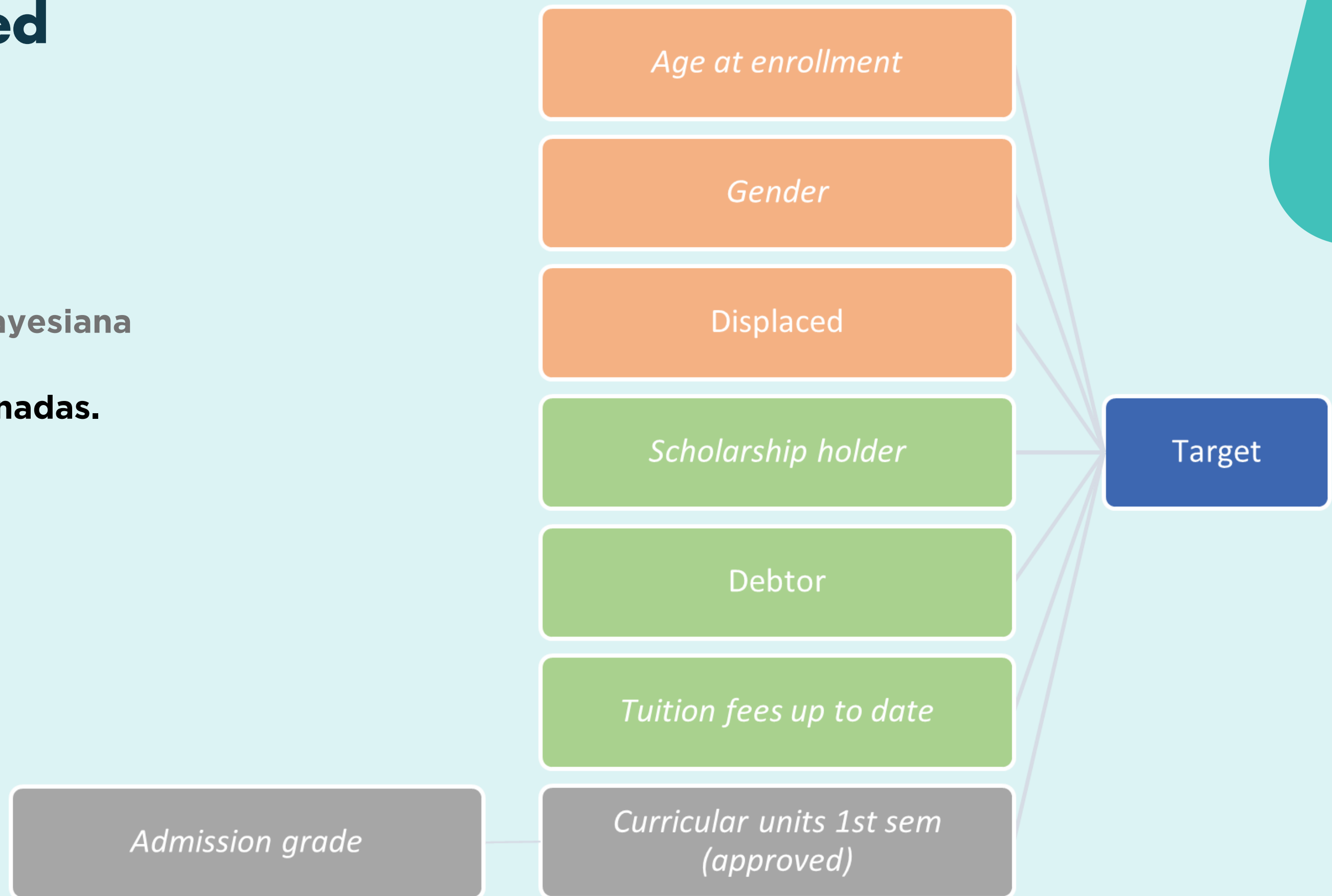
**Predecir tanto los resultados del primer semestre como la probabilidad de deserción con solo contar con la nota de admisión a la universidad.**



# Modelo de Red Bayesiana

## 03.

Modelo final de Red Bayesiana eliminando variables académicas correlacionadas.





# Estimación de Parámetros y Evaluación del modelo

**80%**

**Datos de  
Entrenamiento**

Estimación de por Máxima  
Verosimilitud

Estimación de Probabilidades  
Condicionales

**20%**

**Datos de  
Prueba**

Evaluación del Modelo

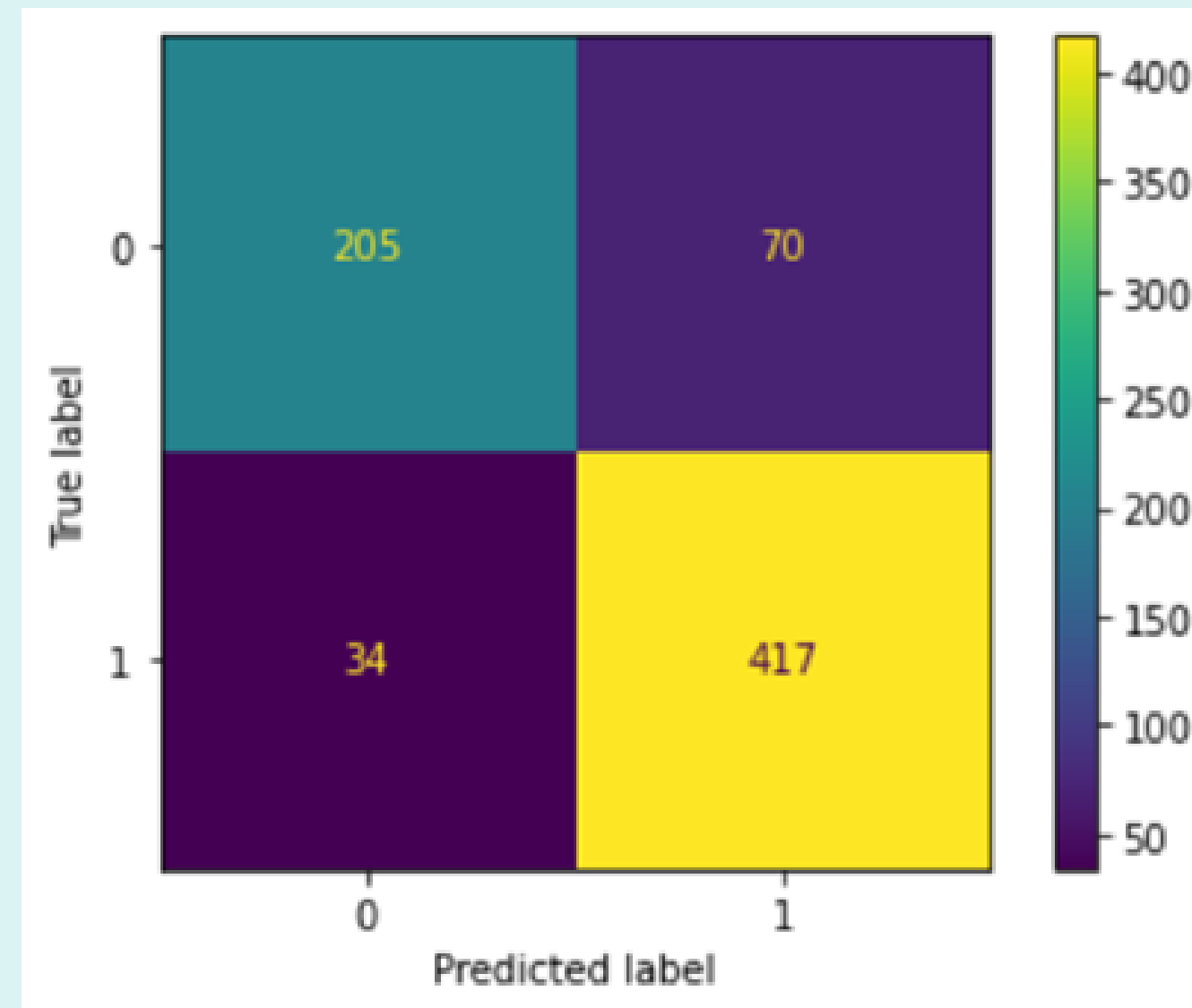
Predicción empleando la Red  
Bayesiana



# Resultados de la Predicción

01. Predicción del modelo utilizando **TODAS** las variables predictoras:  
Estudiante con su primer semestre completado.

85.6%  
Accuracy

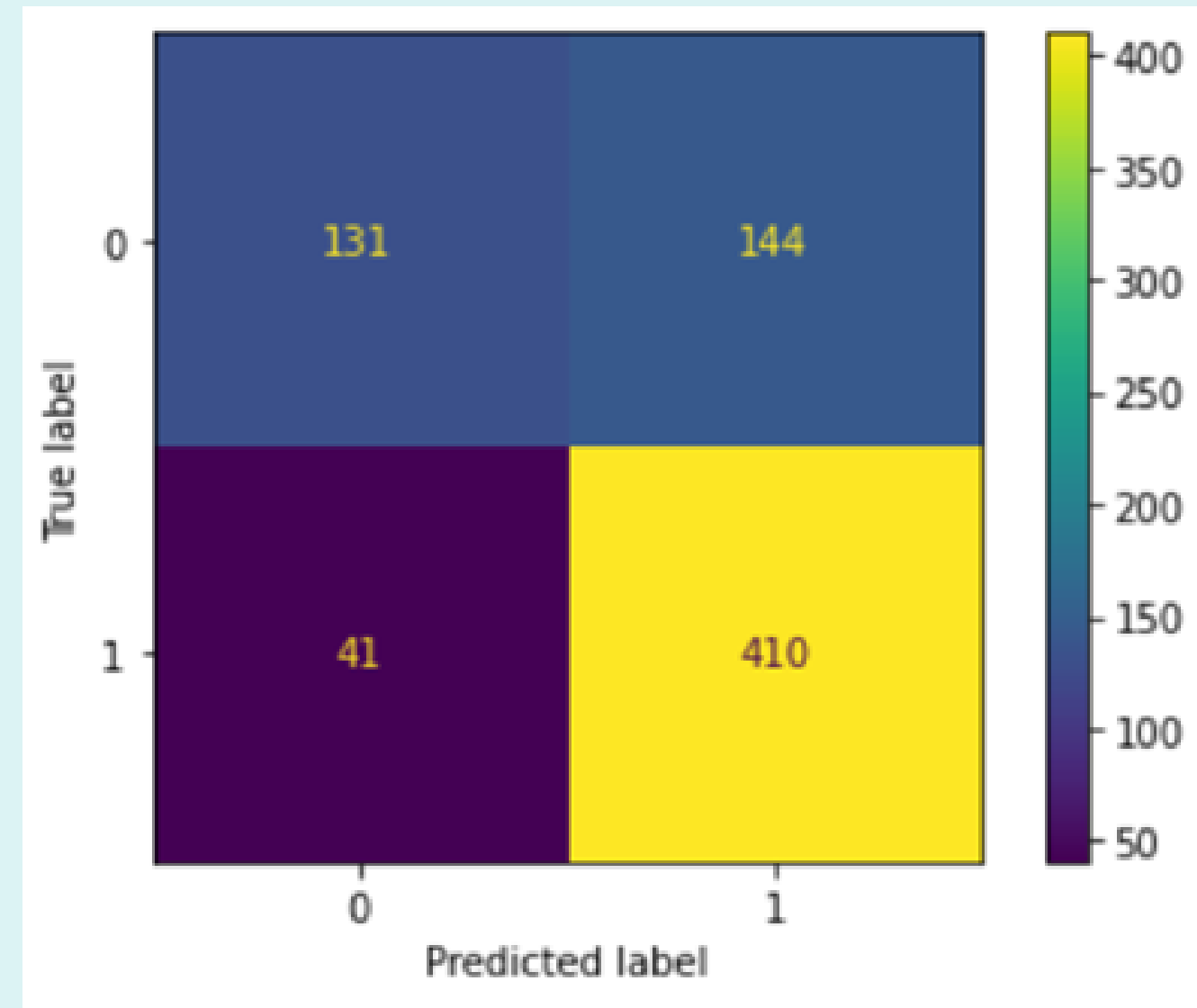


# Resultados de la Predicción

**02.** Predicción para un aspirante o un estudiante sin completar su primer semestre.

**74.5%**  
Accuracy

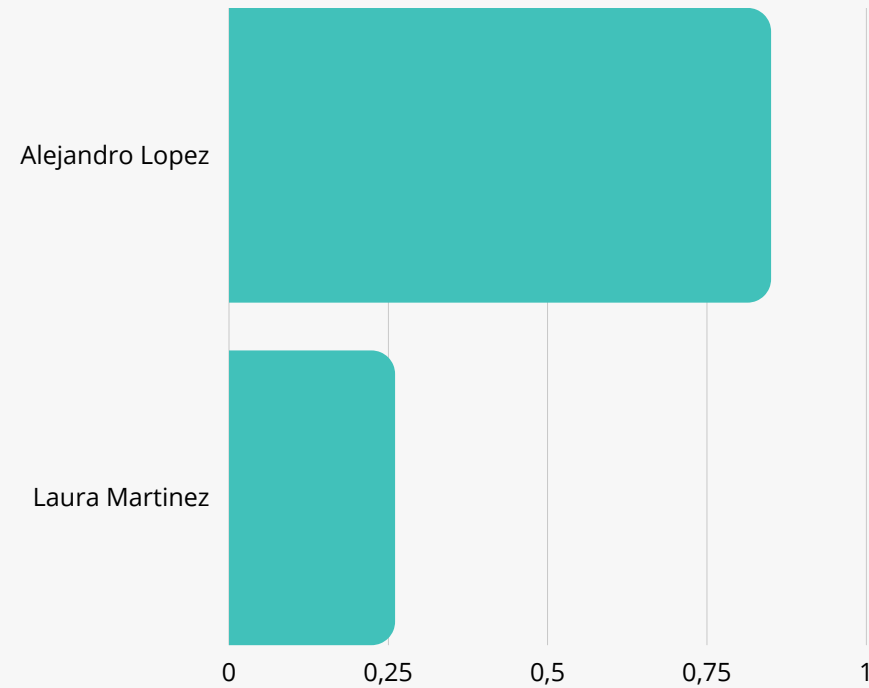
**-11.1% exactitud**





# Inferencia

# Veamos algunos casos ...



**Alejandro Lopez**

Aspirante nuevo:

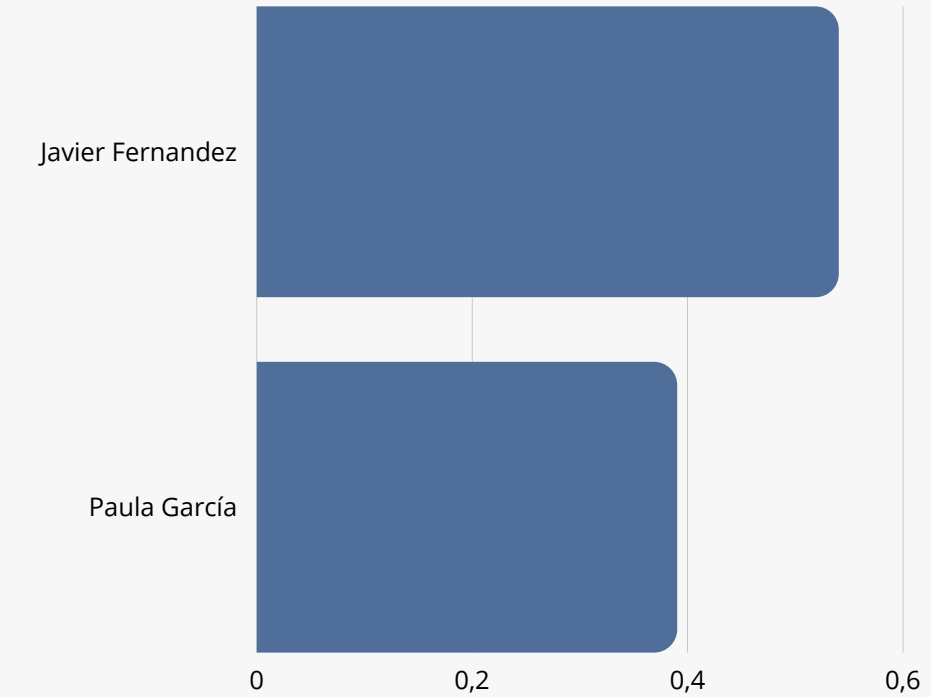
- Foraneo
- Deudor
- **Con deudas en la matrícula**
- Entre 16 a 18 años



**Laura Martínez**

Aspirante nuevo:

- Foraneo
- **Becado**
- **Sin deudas en la matrícula**
- Entre 16 a 18 años



**Javier Fernandez**

Estudiante antiguo:

- Deudor
- Edad entre 19-20
- Sin deudas en la matrícula
- **Nota de admisión baja**



**Paula García**

Estudiante antiguo:

- Deudor
- Edad entre 19-20
- Sin deudas en la matrícula
- **Nota de admisión alta**



# Herramienta para la toma de decisiones



# Conclusiones y Recomendaciones

# Conclusiones y Recomendaciones

## VARIABLES

Tanto los factores de las dimensiones demográficos, sociales, financieros y académicos influyen en la deserción de los estudiantes universitarios.

## MODELO

El modelo de Red Bayesiana permite predecir la deserción universitaria de los estudiantes desde el inicio de su trayectoria universitaria con un exáctitud de predicción en el 85.6% de los casos.

## RECOMENDACIÓN

La Universidad puede enfocar sus esfuerzos en:

- Apoyo financiero
- Regular el proceso de admisión
- Acompañar a los estudiantes más jóvenes y aquellos más adultos



# ¡ Muchas Gracias!

