




# TASA DE REINGRESO Y DESERCIÓN DE ESTUDIANTES EN LA FACULTAD DE CIENCIAS



Introducción a la Ciencia de Datos

Universidad de los Andes

Alejandro Palomino



¿Es posible predecir la tasa de Deserción-Reingreso  
en la Facultad de Ciencias en la Universidad de los  
Andes?

# Datos utilizados



- La totalidad de los datos utilizados fueron extraídos del **boletín estadístico** de la Universidad de los Andes, en su versión **2021**. Donde se pueden extraer datos como el **número de estudiantes activos, los matriculados, y los graduados** de cada uno de los periodos desde el 2016.

# ¿Tasa de Deserciones-Reingresos?

- ▶ De los datos del boletín, es posible extraer el número de estudiantes activos, matriculados, y graduados para cada periodo semestral (a nivel de pregrado), sin embargo, si hacemos la operación

$$\begin{array}{ccccccc} \text{Estudiantes} & & \text{Estudiantes} & & \text{Estudiantes} & & \text{Estudiantes} \\ \text{Activos} & + & \text{Graduados} & - & \text{Matriculados} & \neq & \text{Activos} \\ & & & & & & \text{(Periodo anterior)} \end{array}$$

- ▶ ¿Dónde están los estudiantes faltantes?

# ¿Tasa de Deserciones-Reingresos?

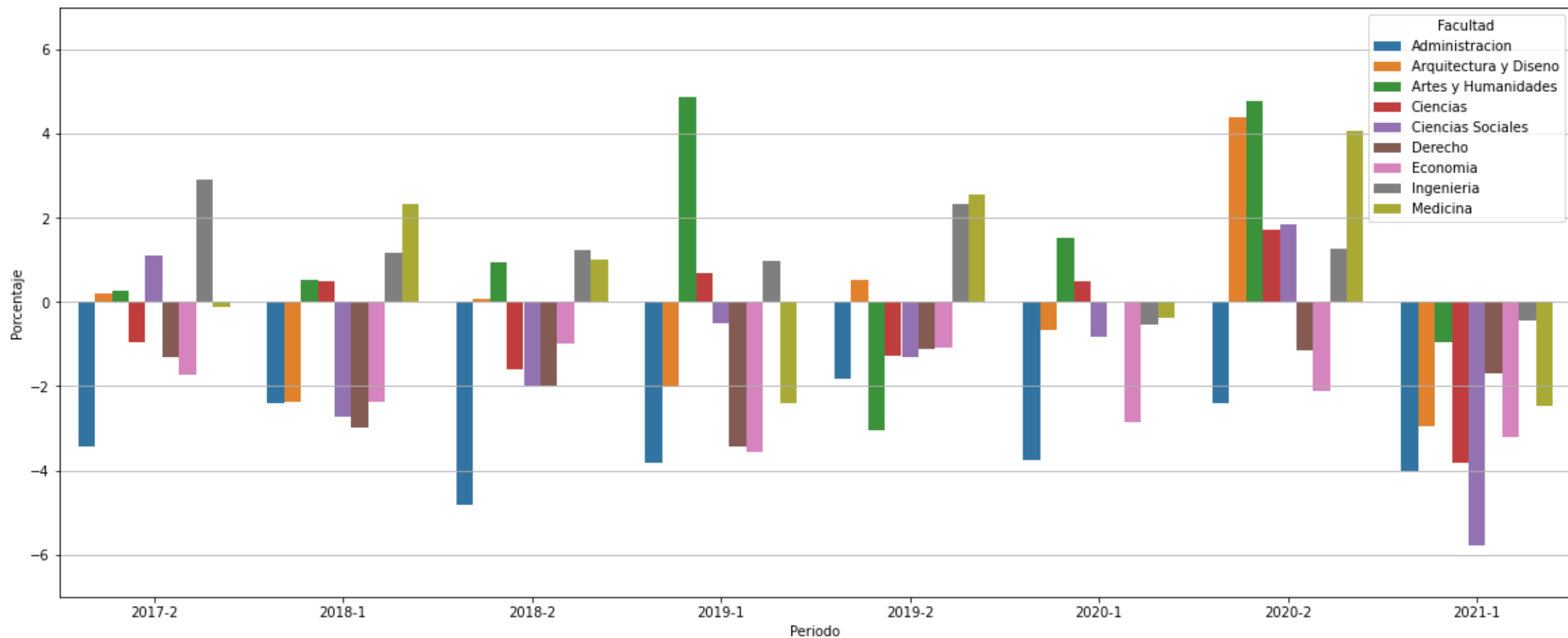
- Teorizamos que este número faltante corresponde tanto al número de estudiantes que desertaron de la carrera (o aplazaron el semestre), como al número de reingresos. Definimos entonces la cantidad:

$$\text{Estudiantes Deserción-Reingreso} = \frac{\text{Estudiantes Activos (Periodo anterior)}}{\text{Estudiantes Matriculados}} - \frac{\text{Estudiantes Graduados}}{\text{Estudiantes Activos}}$$

- Número que será
  1. **Positivo:** Si el número de deserciones es mayor que el de reingresos.
  2. **Negativo:** Si el número de reingresos es mayor que el de deserciones.
  3. **Cero:** Si ambos números son iguales.

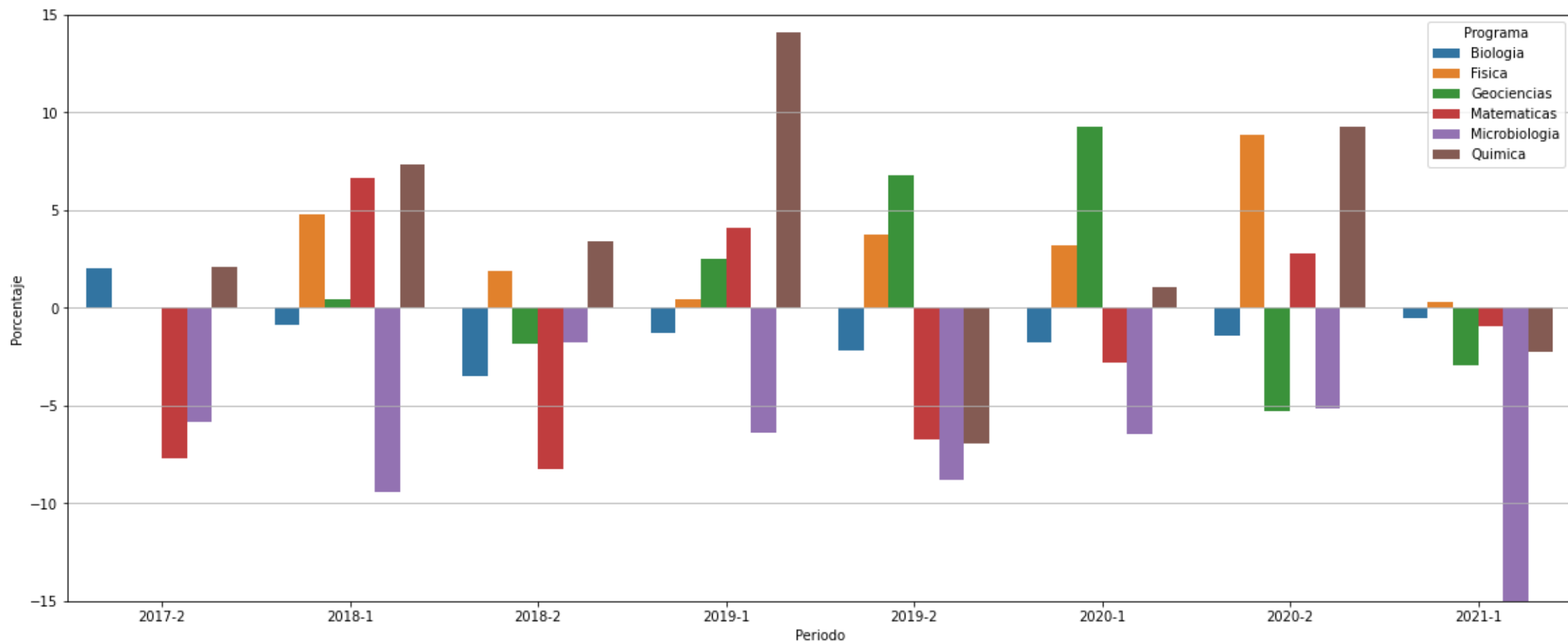
# Tasa - Facultades

- Finalmente, normalizamos esta cantidad entre el número de estudiantes activos para hallar la tasa de **Deserción-Reingreso**. Para las facultades, hallamos qué:



# Tasa – Facultad de Ciencias

- Por otro lado, para la Facultad de Ciencias específicamente observamos



# Definición del Modelo

► Escogiendo como predictores:

Estudiantes  
Activos

Estudiantes  
Matriculados

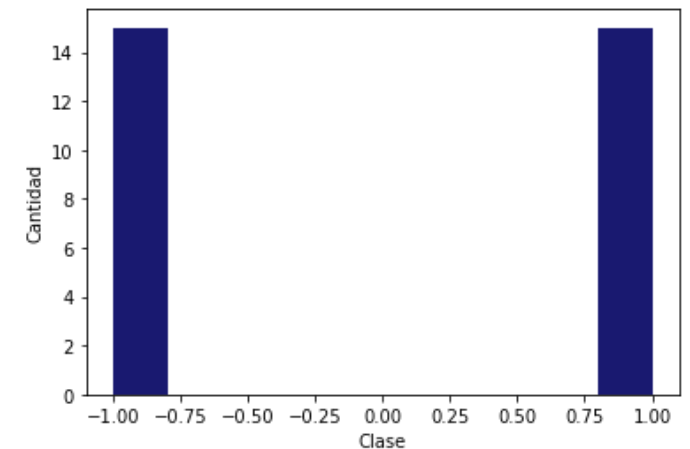
Estudiantes  
Graduados

Cursos  
Estudiante

Y como target, para ese departamento de la Facultad de Ciencias

1. **+1**: Si el número de deserciones es mayor que el de reingresos.
2. **-1**: Si el número de reingresos es mayor que el de deserciones.

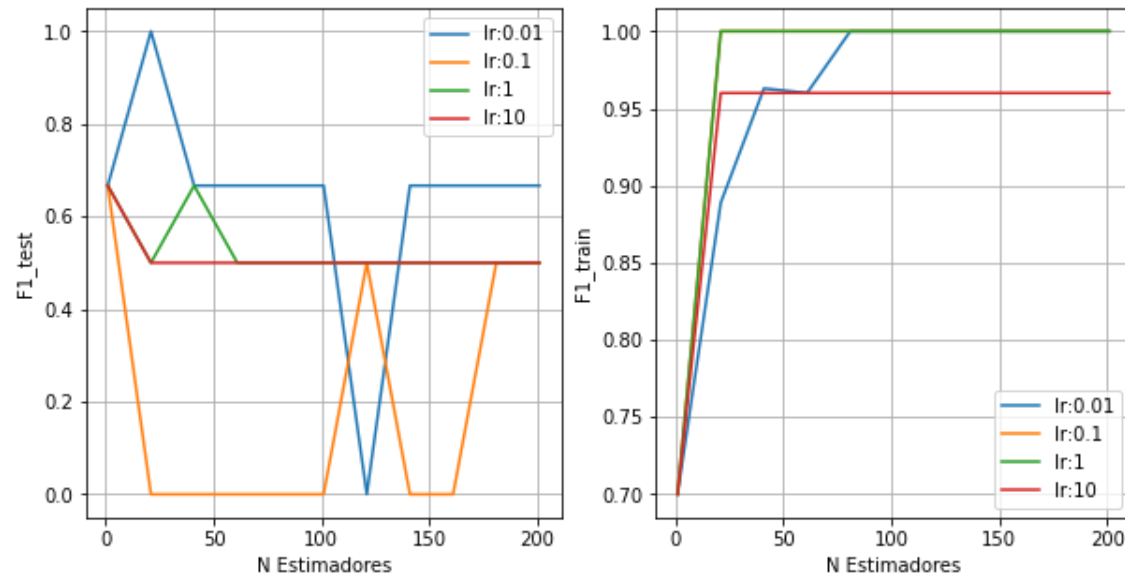
En el periodo inmediatamente siguiente.



Creamos un modelo usando **AdaBoost Classifier** para predecir si el semestre siguiente la tasa de **Deserción-Reingreso** será positiva (más deserciones) o negativa (más reingresos).

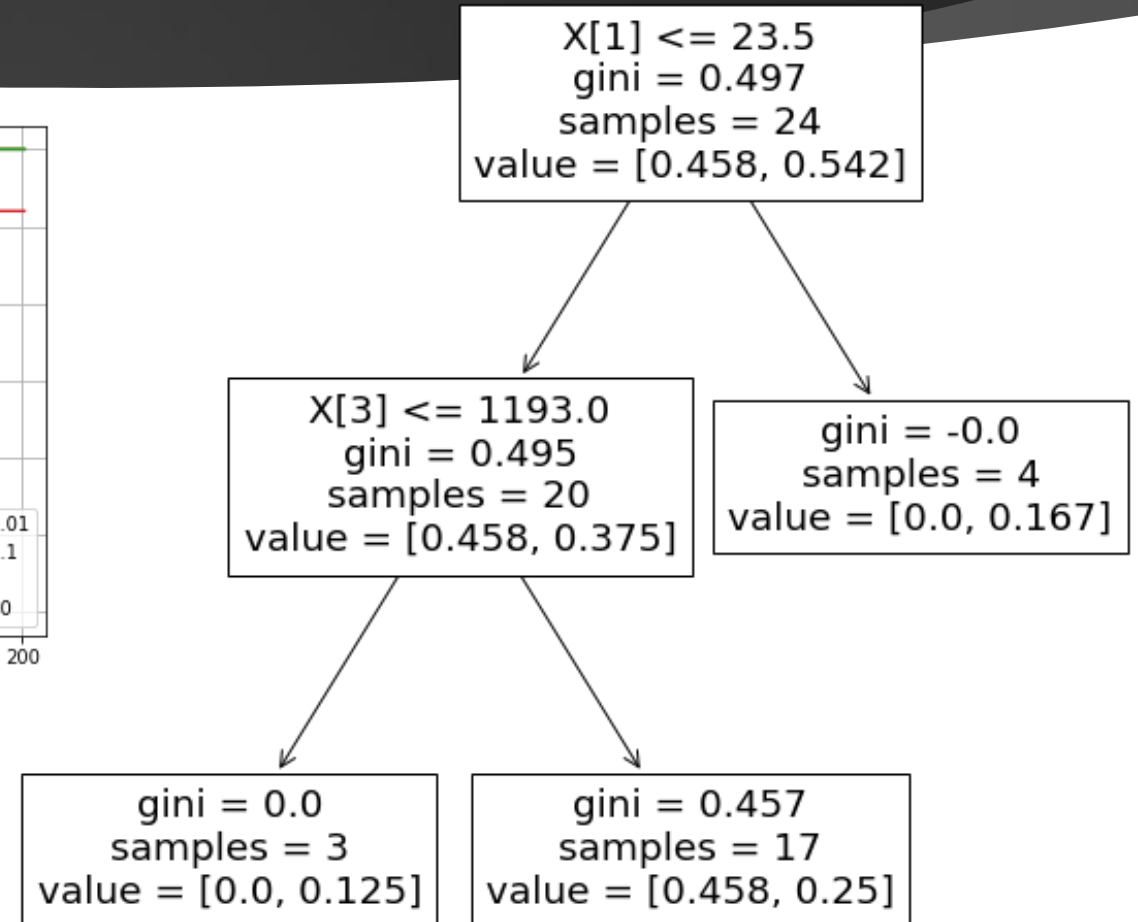


# Grid de metaparámetros

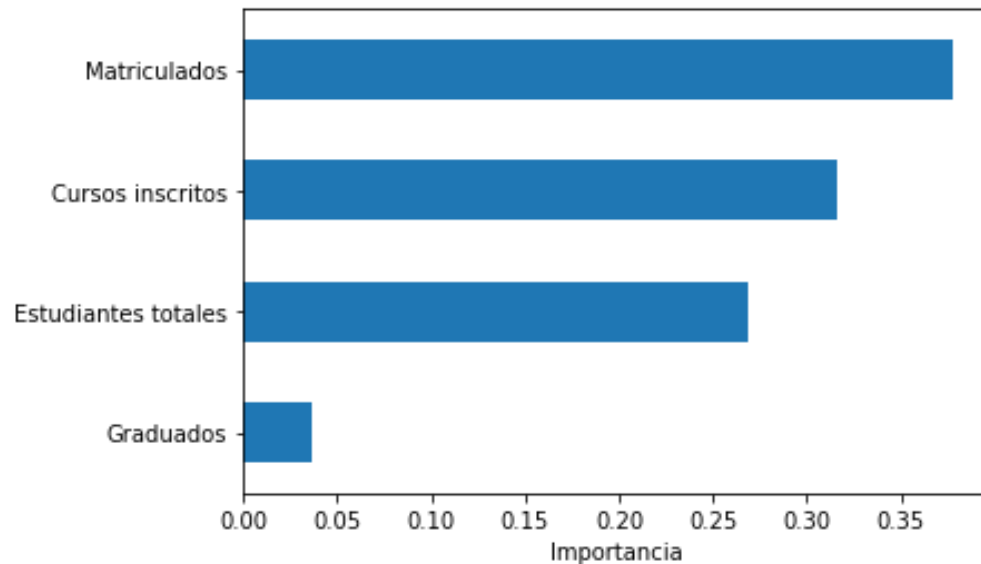


► Se terminaron escogiendo:

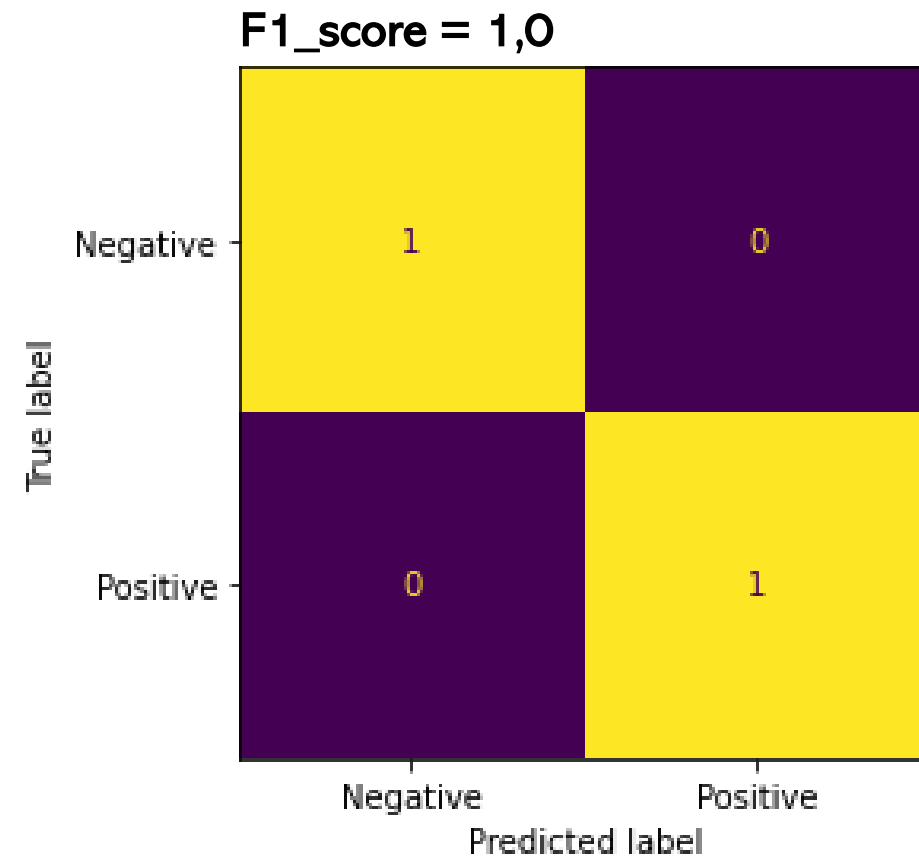
1. N\_estimators = 61
2. Learning\_rate = 0,01



# Resultados



- Al observar los pesos se encontró que el número de matriculados y los cursos inscritos son los **predictores** más importantes.



# Conclusiones

- ▶ Se pudo entrenar un modelo que predecía si en un periodo semestral la tasa de Deserciones-Reingresos sería positiva o negativa con los datos de estudiantes **matriculados, admitidos, graduados y peso académico** del periodo anterior, con un F1\_Score de 1.00.
- ▶ Obtener datos tanto de **años anteriores** como más actuales para realizar predicciones a futuro con las **tasas** de por pandemia.
- ▶ Generalizar el modelo para ser capaz de predecir no solamente el **signo** de la tasa de Deserción-Reingreso, sino el **valor** de esta.

Gracias