

ATTRITION & MACHINE LEARNING

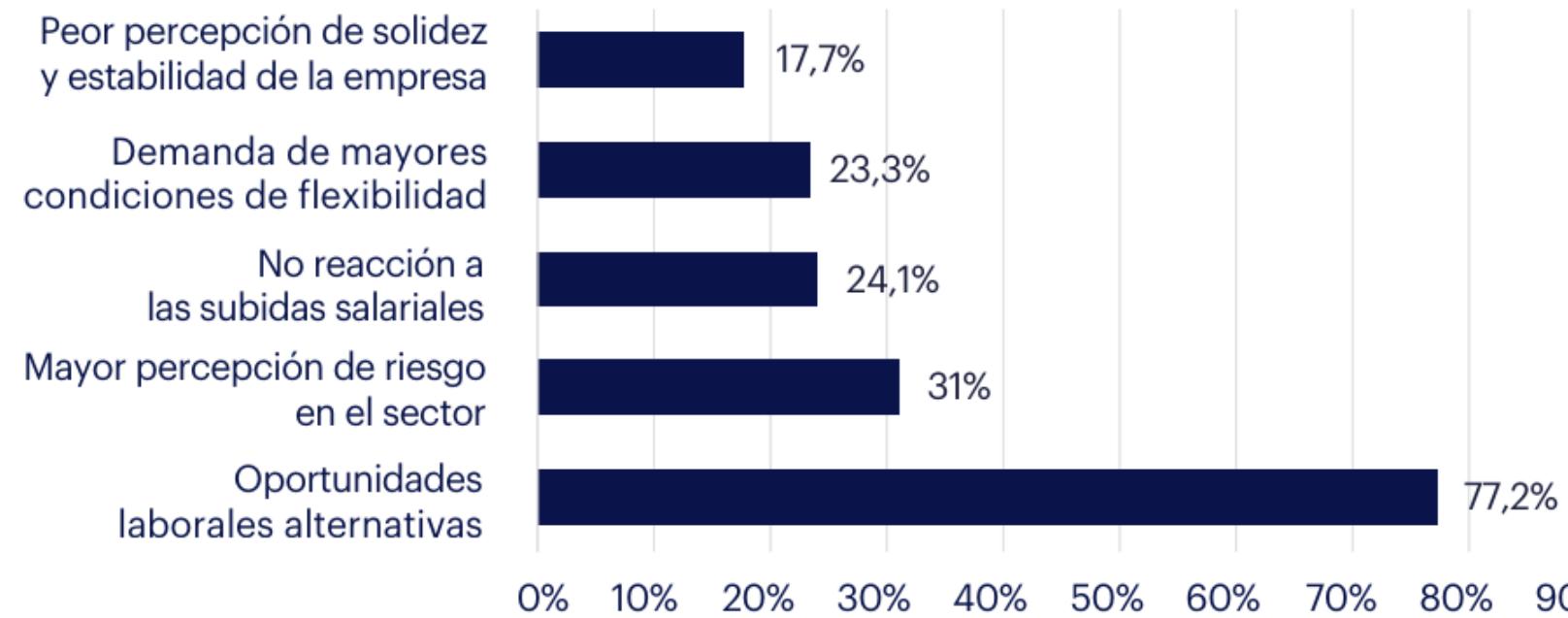
Eva Arribas



INTRODUCCIÓN

La rotación o *attrition* es una problemática creciente en el mundo de los RRHH. Las reglas del juego en el mercado laboral han cambiado desde la pandemia, y esto afecta de manera directa a la capacidad de retención de las empresas de todos los sectores. La rotación media anual española se sitúa en un **17%**, y la previsión es que se mantenga o que se incremente en los próximos años*.

Principales causas del aumento de la rotación



PREDICCIÓN DE LA ROTACIÓN

1

¿QUÉ DATOS TENEMOS?

Para poder construir nuestro modelo, tenemos los datos de **4K empleados** de una empresa con **una rotación del 15%** (en este caso, se unifican bajas voluntarias y despidos).

Utilizaremos 3 BBDD:

- **Datos generales** de empleados, incluyendo la rotación
- **Encuesta a empleados** sobre entorno laboral, trabajo, *worklife balance*...
- **Encuesta a managers** (desempeño y compromiso)

2

¿QUÉ MODELO UTILIZAREMOS?

En el caso, nos piden utilizar el **Modelo de Regresión Logística**, por lo que nuestras medidas de éxito del modelo serán:

- **Precisión** (cuántas personas de las que digo que se van, realmente se van)
- **Recall** (cuántos de los que realmente se van he conseguido capturar con mi modelo)

3

EDA

Antes de comenzar con la fase de Machine Learning, realizamos un análisis de datos exploratorio donde:

- Creamos un **dataset** con las 3 BBDD con un merge
- Tratamos los **valores nulos** y los sustituimos por la mediana (para evitar el impacto de los outliers al usar la media)
- Analizamos la **correlación** entre las variables para intentar disminuir la colinealidad

MATRIZ DE CORRELACIÓN

Tras transformar las variables categóricas en booleanas con el *get dummies*, observamos que las principales correlaciones son:

• EDAD / ANTIGÜEDAD:

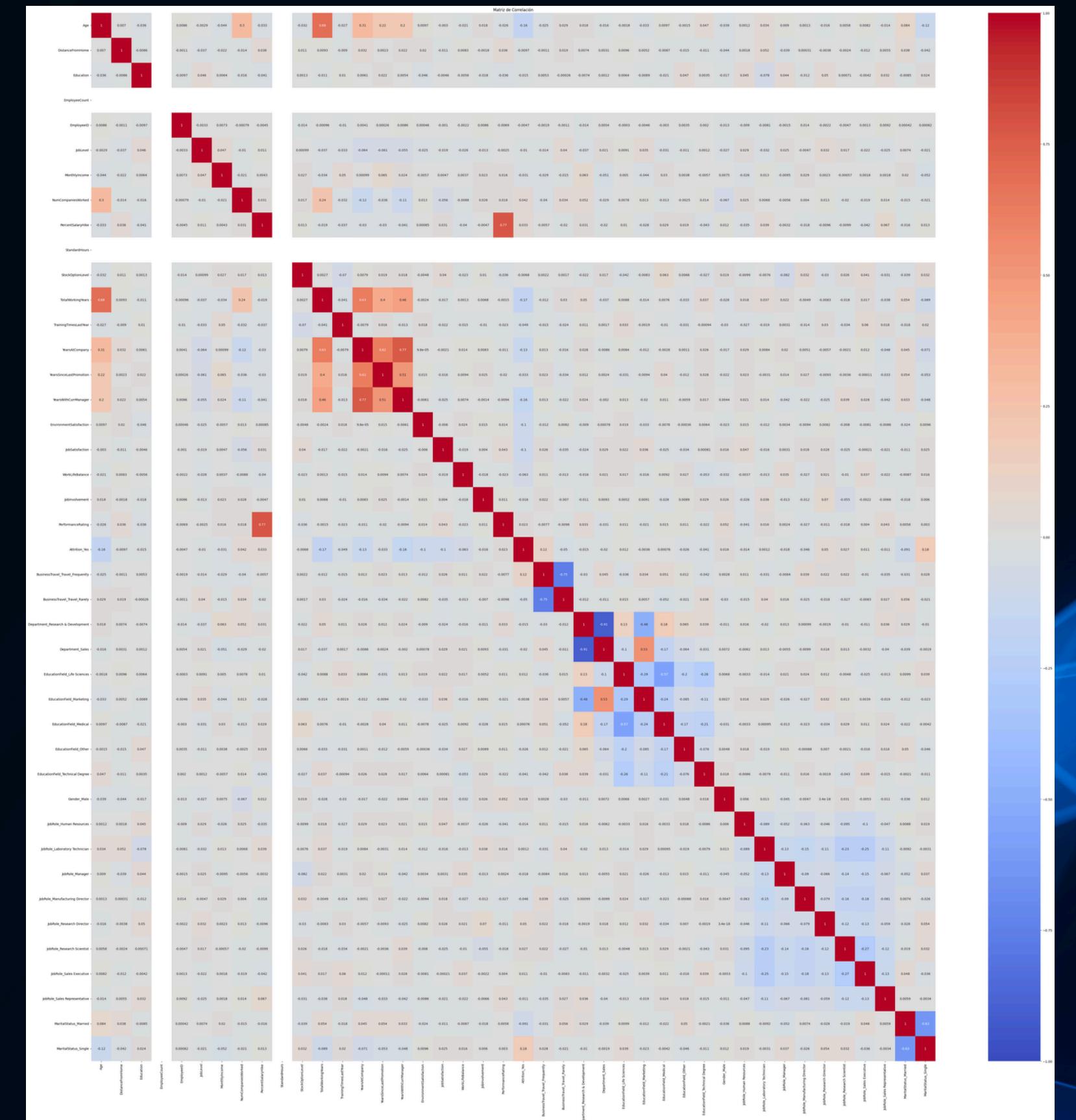
- -0.16 con la edad
 - -0.17 con años de experiencia
 - -0.13 con los años de antigüedad
 - -0.16 con años con el manager actual

• **SATISFACCIÓN:**

- -0.1 con la satisfacción con el entorno de trabajo
 - -0.1 con la satisfacción con el trabajo
 - -0.063 con el equilibrio con la vida personal

• OTRAS CORR RELEVANTES:

- -0.12 con la frecuencia con la que se viaja
 - 0.18 con estar soltero/a
 - -0.091 con estar casado



PREPARACIÓN DE LOS DATOS

Una vez analizada la correlación, preparamos los datos para la aplicación de nuestro modelo que nos ayudará a clasificar a los empleados en función del attrition (nuestra variable dependiente o variable a predecir).

Dividimos entre train y test

`test_size = 0.2, random_state = 42, shuffle = True`

Estandarizamos las variables

Hacemos el fit

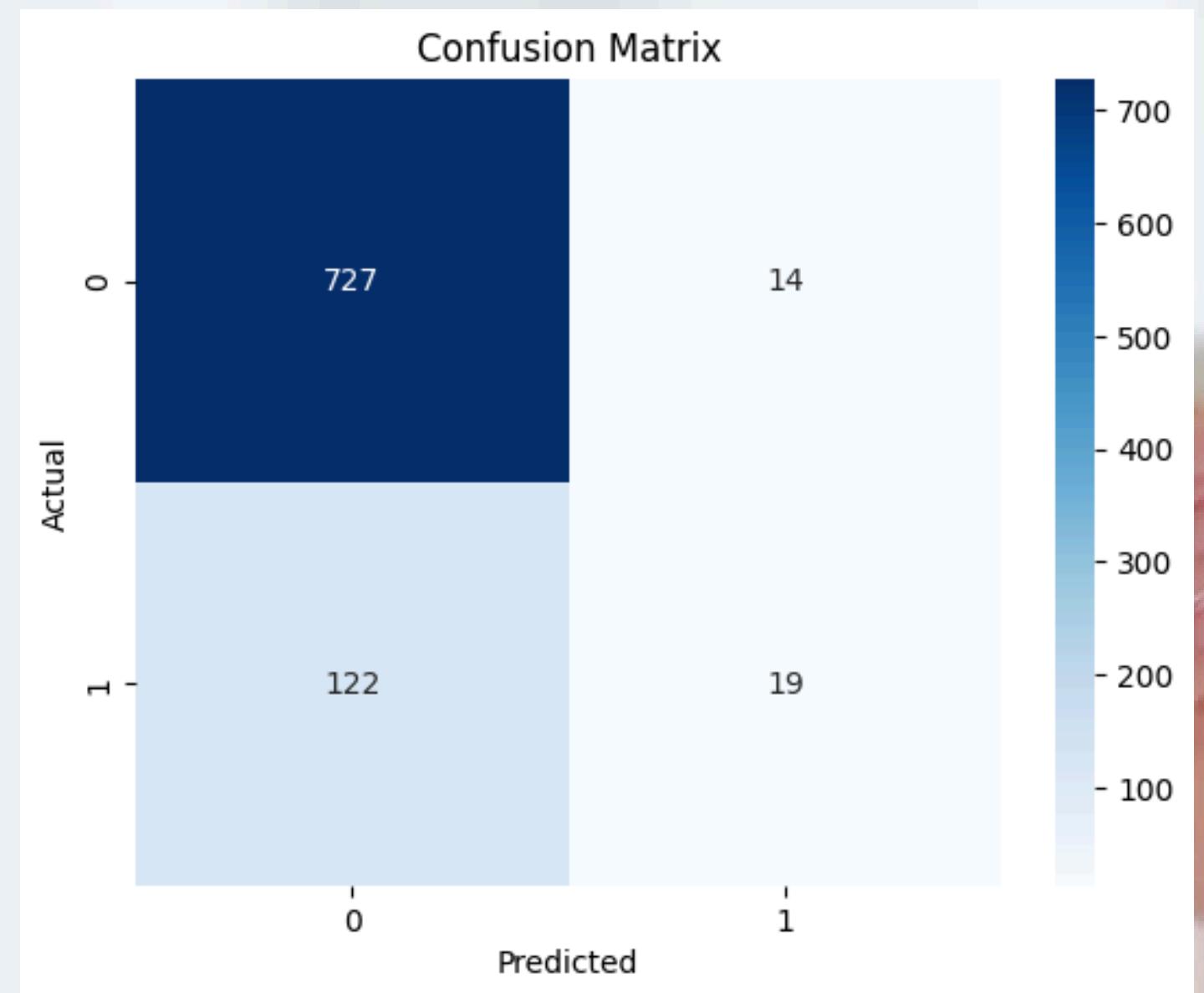
A man with short dark hair, wearing a yellow and black martial arts gi, stands in a dynamic pose against a dark, smoky background. He holds two wooden nunchaku in his hands, one in each hand, pointing them upwards. The lighting highlights his torso and the wooden texture of the nunchaku.

LET'S
TRAIN

RESULTADOS

Precisión: 0.58

Recall: 0.13



CONCLUSIONES

Cambio de modelo

Categorización attrition

Organigrama

Sector

Salarios y beneficios sociales



GRACIAS

