

# Examen Final

[¿A quienes hemos preguntado?](#)

[¿Cómo están las finanzas de los encuestados?](#)

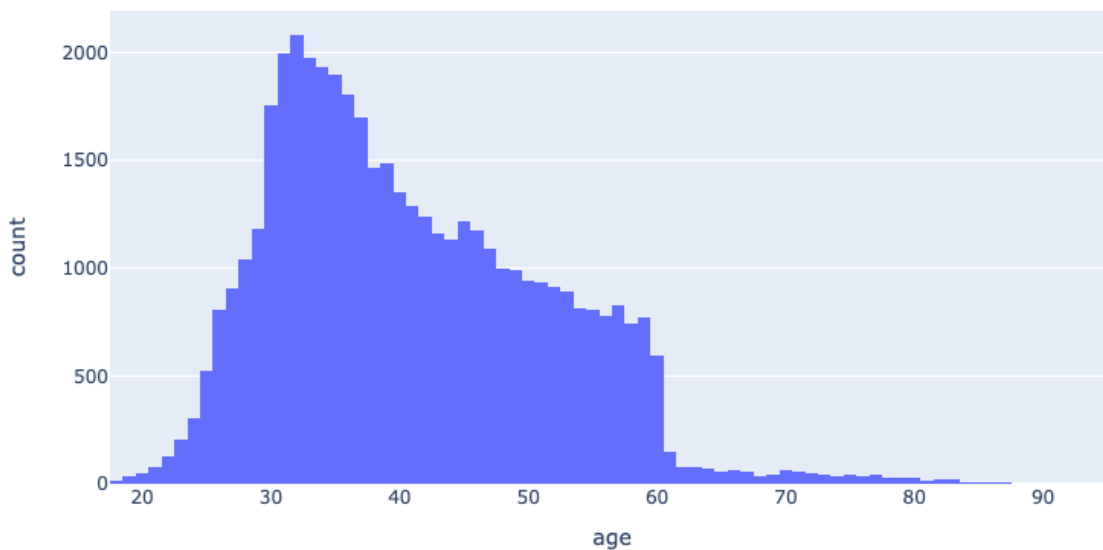
[¿Qué variables predicen el éxito de la campaña?](#)

[¿Qué valores de estas variables predicen el éxito?](#)

[Conclusión](#)

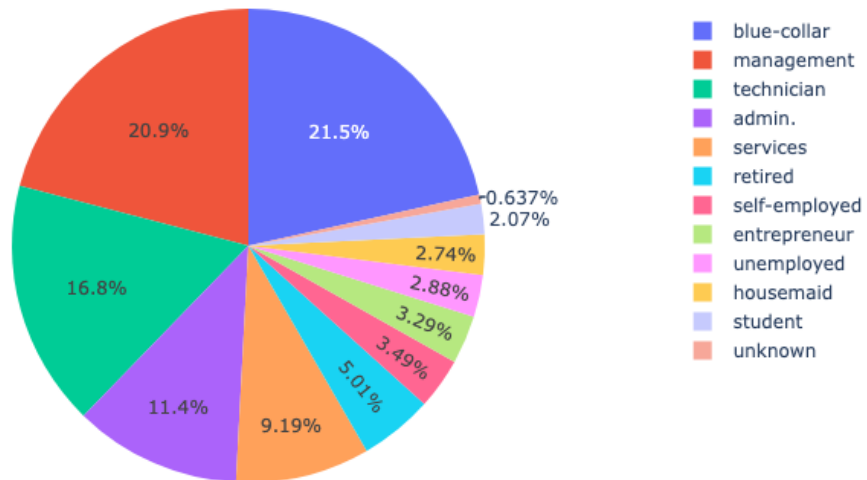
## ¿A quienes hemos preguntado?

En los siguientes gráficos podemos ver la distribución de la edad de los encuestados, su trabajo, estado marital y su educación



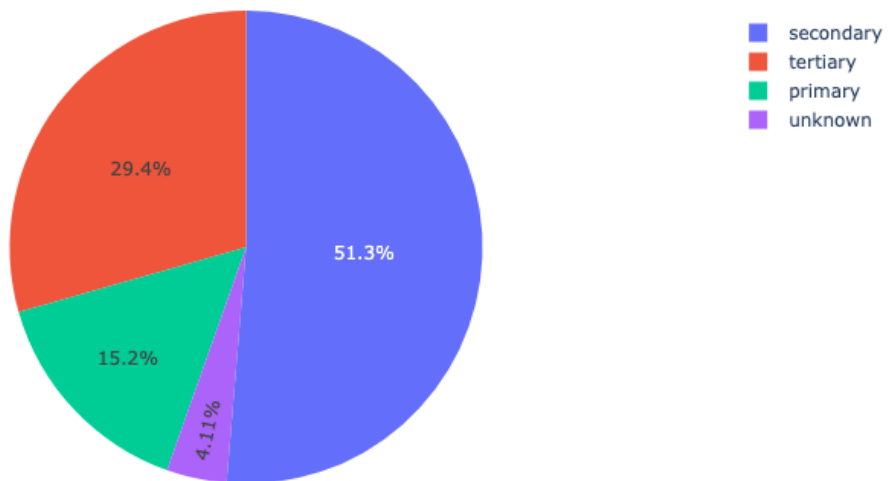
Distribución de la edad de los encuestados

La mayoría de las personas encuestadas se sitúan entre los 30 y 50 años, seguido de las personas entre 50 y 60 años.



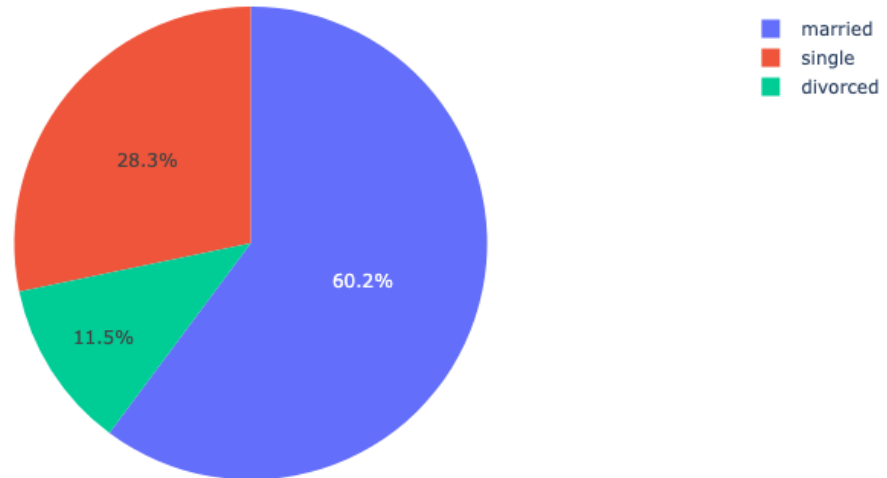
Distribución de trabajos

Las personas encuestadas son mayoritariamente, *blue-collar*, dirección y técnicos, cubriendo entre esto tres grupos más del 50% de las encuestas.



Distribución de su educación

A su vez, la mayoría de encuestados tiene la educación secundaria, seguidos un 30% con educación terciaria y un 15% con educación primaria.



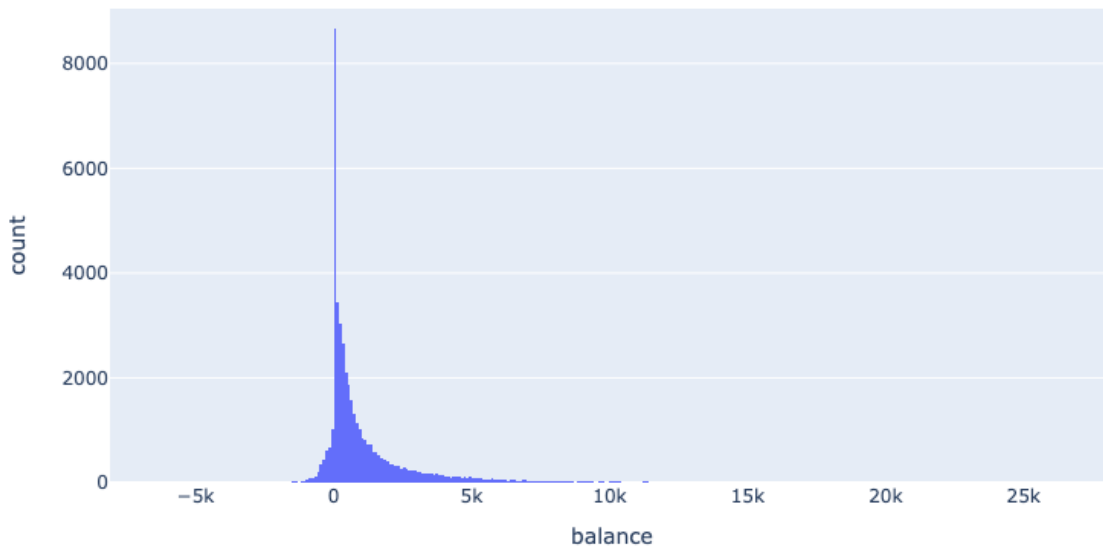
Distribución de su estado marital

Finalmente, la mayoría de encuestados está casado (lo que guarda cierta relación con la edad media del encuestado).

Por tanto, podríamos decir que el perfil medio encuestado es una persona de 35 años, *blue-collar*, con educación secundaria y casado.

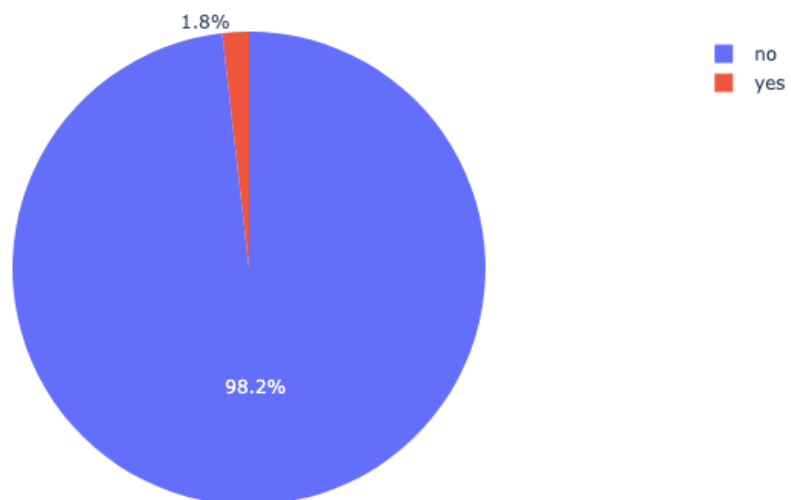
## ¿Cómo están las finanzas de los encuestados?

A continuación evaluaremos la situación económica de las personas encuestadas.



Distribución del balance

Podemos ver por el histograma del balance, que la mayoría de encuestados tienen un balance medio anual de entre 0 y 2500€.



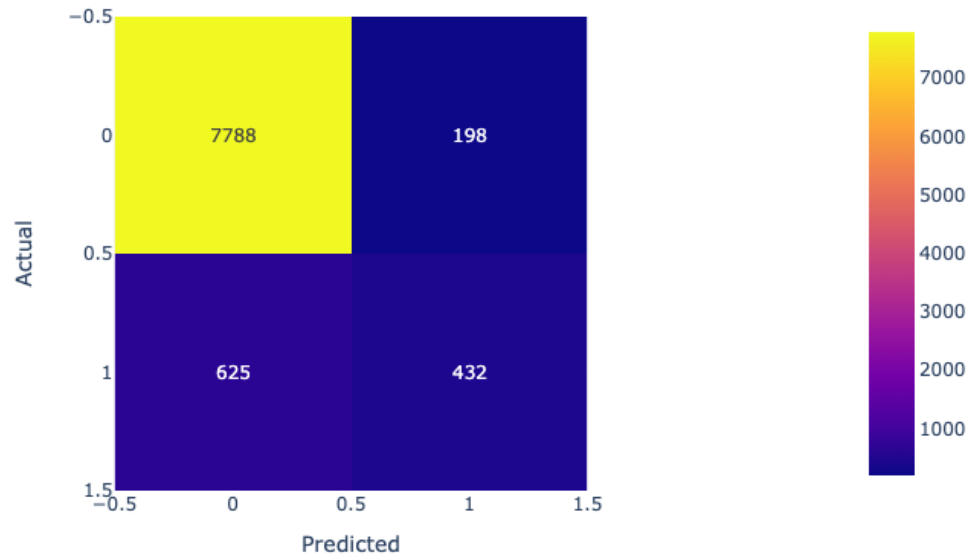
% de encuestados que no han pagado sus deudas

Y en este gráfico podemos ver que la mayoría de encuestados han pagado sus deudas (no han *defaulted*).

## ¿Qué variables predicen el éxito de la campaña?

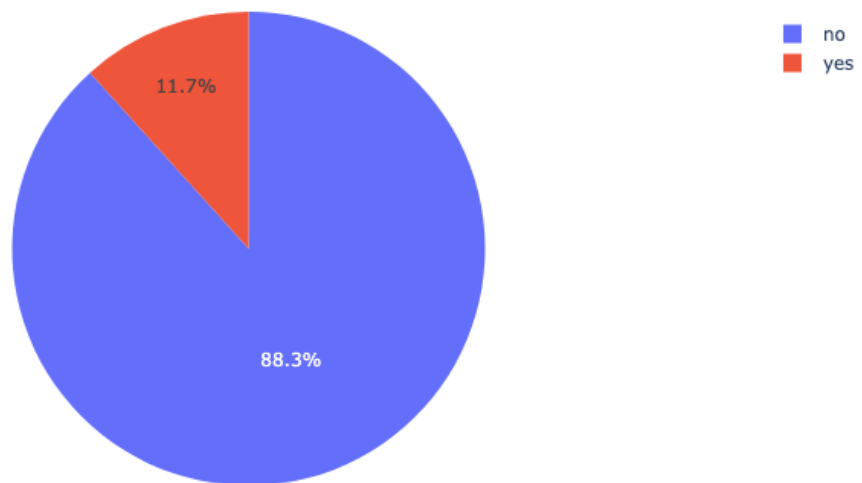
Más importante que conocer el perfil de los encuestados es saber qué variables predicen el éxito de la campaña. Para extraer esta información se ha entrenado un *random forest* y se ha extraído la importancia de cada variable.

Antes de ver qué variables son más importantes, vamos a evaluar la predicción del clasificador.



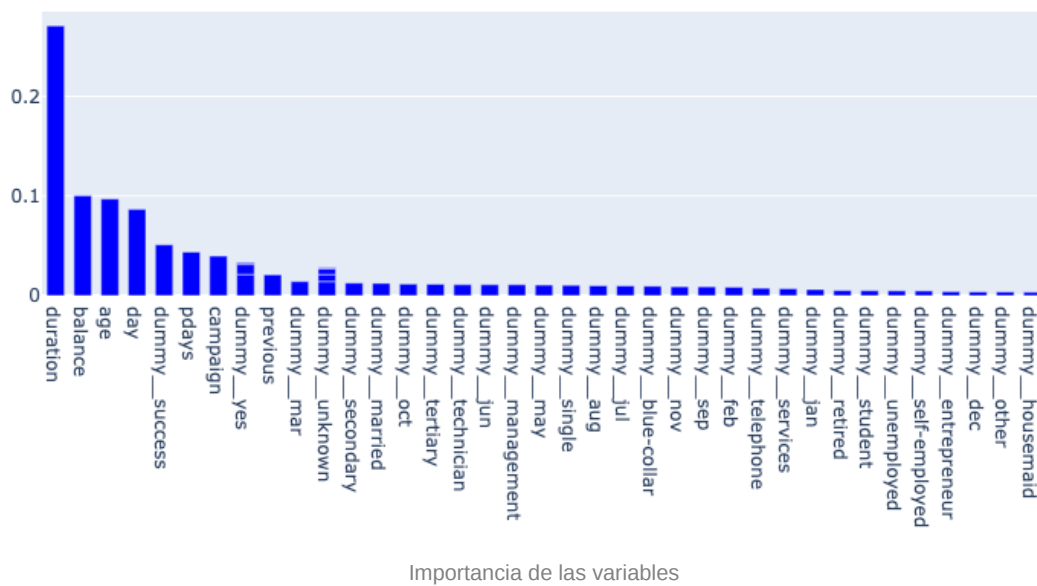
Matriz de confusión sobre el set de test

Podemos ver que el clasificador tiene una precisión del 91% y un *weighted recall* del 91% también. Además cabe destacar que hay más verdaderos positivos que ningún otro valor. Esto se podría deber a que hay más noes que síes y por tanto los datos pueden estar sesgados hacia ese valor.



La precisión del modelo nos indica que este (aparentemente) es un buen modelo para predecir la respuesta de la encuesta.

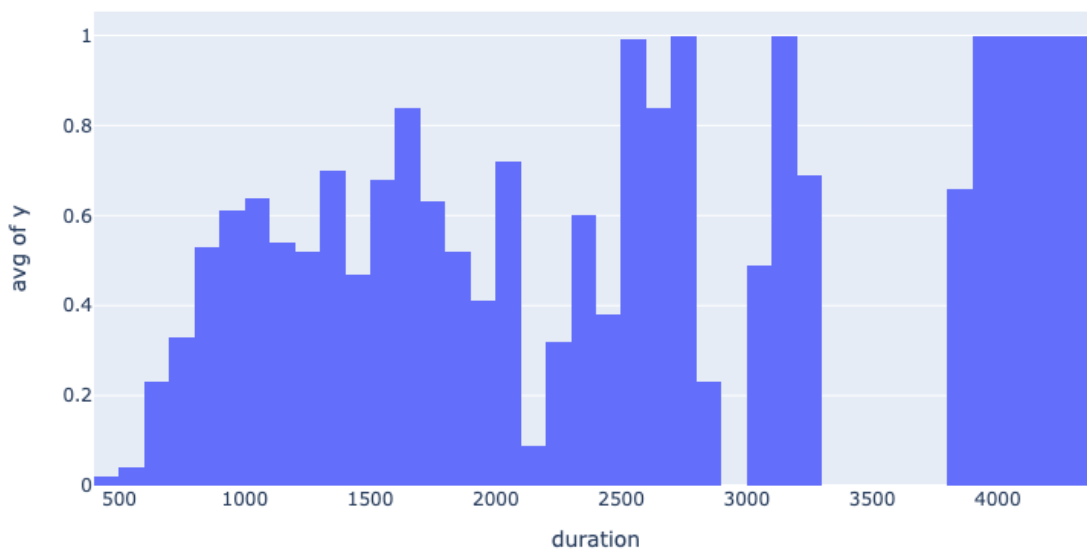
Sabiendo que el modelo, funciona correctamente, podemos evaluar qué variables son más significativas. La siguiente gráfica muestra la importancia de las variables ordenadas de mayor a menor.



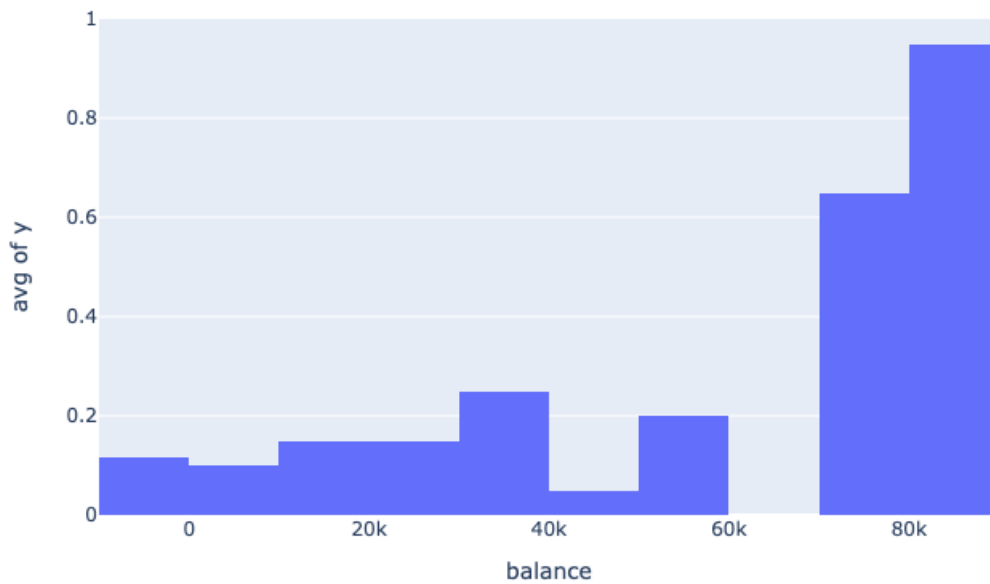
Podemos ver en el gráfico que la variable más importante a la hora de predecir es la duración, seguido del balance anual, la edad y el día. Estas son las variables que más predicen el éxito de la campaña: **duración, balance y edad**.

## ¿Qué valores de estas variables predicen el éxito?

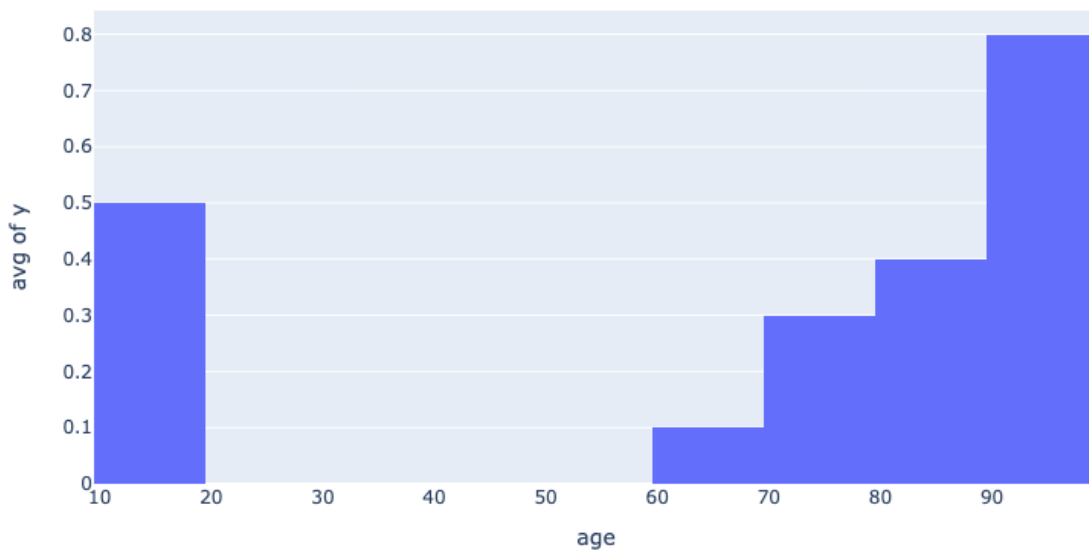
La última pregunta a responder es, ¿qué valores de estas variables predicen el éxito? Con esta información podemos optimizar las ventas para maximizar el número de conversiones.



Este gráfico muestra cómo responde la persona para diferentes duraciones. Cuando la y vale 0, el cliente ha comprado el depósito, por tanto, **cuanto menor, mejor**. Podemos ver entonces que, para duraciones entre 0 y 500 segundos, y entre 3200 y 3800, la conversión es mayor. A partir de los 500 segundos, es más improbable que el cliente compre el depósito a no ser que el cliente se mantenga en línea hasta los 3200. A partir de los 3800, el cliente se puede dar por perdido.



En cuanto al balance, personas con menos de 70.000€ medios anuales tienen altas probabilidades de contratar el depósito, sobre todo las personas en el rango entre 60-70k€.



Finalmente, los grupos de edad que más contratan el depósito son las personas entre 20 y 60 años. A partir de los 60 años cada vez es menos probable que contraten el depósito.



## **Conclusión**

Para maximizar las ventas hay que centrarse en conversaciones cortas (menos de 500 segundos), con personas de menos de 70.000€ anuales y que estén entre los 20 y 60 años.