

# MODELO INTELIGENTE PARA LA PREDICCIÓN DE LA EVOLUCIÓN EN REDES SOCIALES.

Marina Dáder Suárez  
Grado en Ciencia e Ingeniería de Datos



## Cátedra TotalEnergies de Analítica de Datos e Inteligencia Artificial

### INTRODUCCIÓN

El departamento de comunicación es de vital importancia para el Desarrollo de una empresa. La toma de decisiones en este ámbito puede ser determinante a la hora de alcanzar los objetivos fijados.

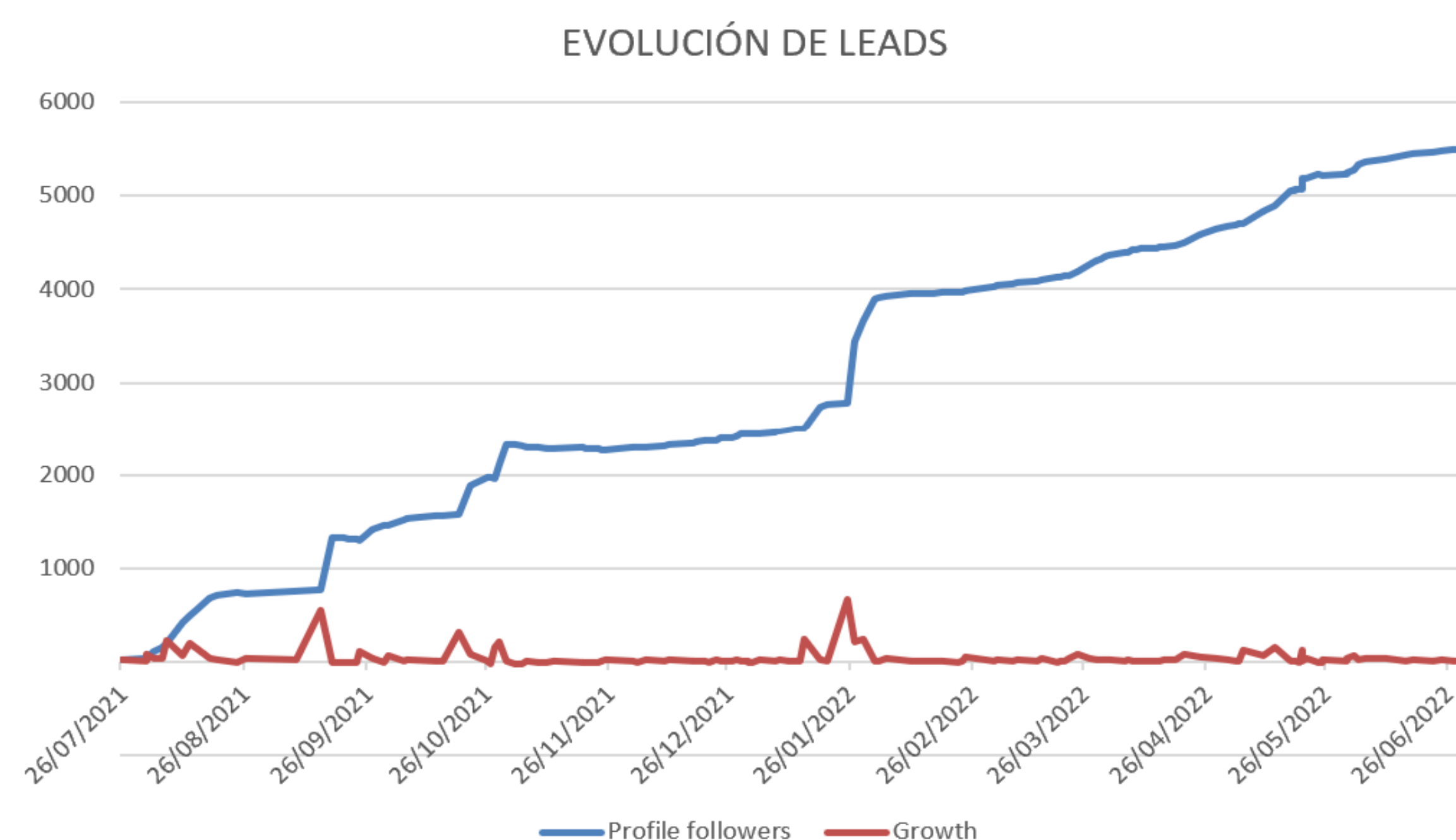
El éxito de la comunicación empresarial depende de diversos factores entre los que destacan una correcta aplicación de técnicas de *copywriting* y la elección de un contenido de calidad para lograr la imagen deseada entre otras.

### OBJETIVOS DEL ESTUDIO

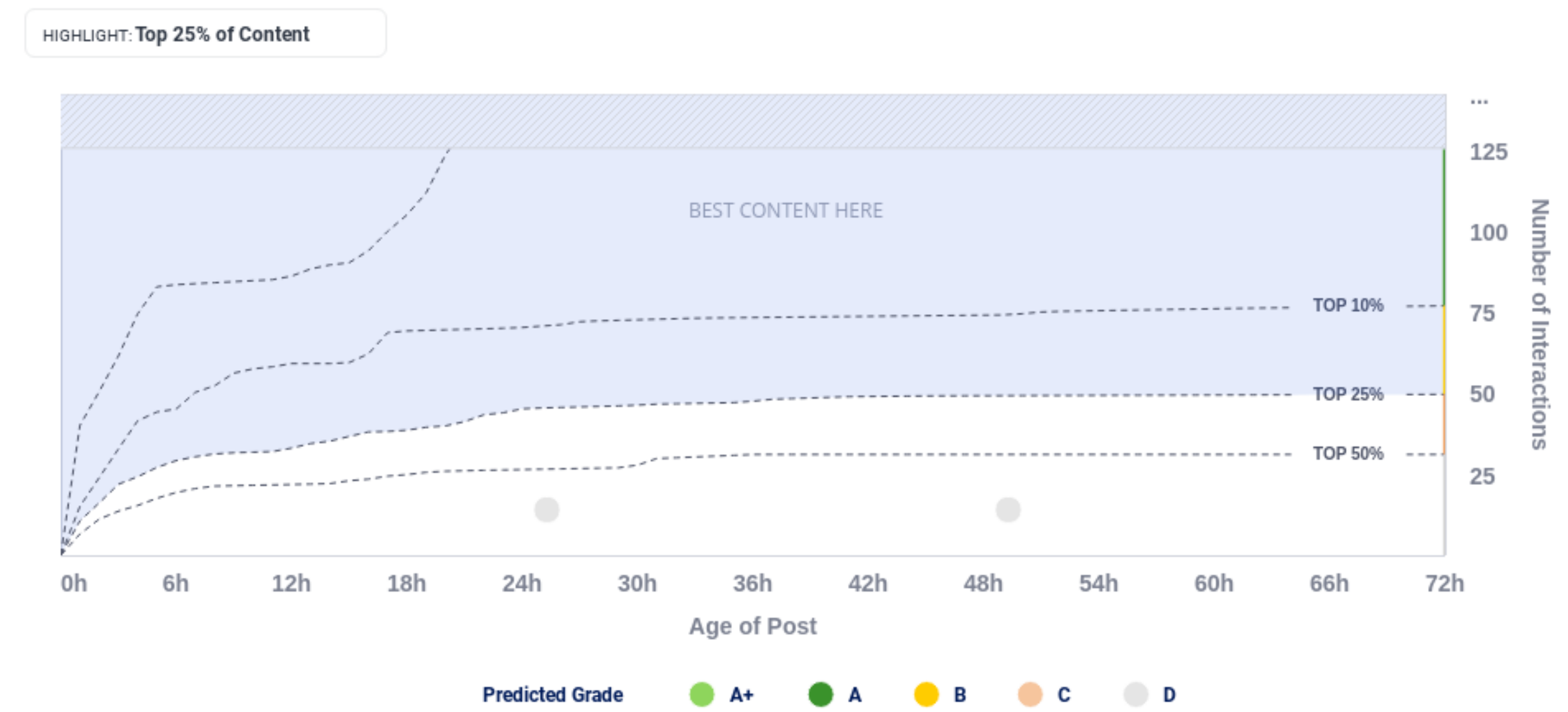
En este contexto, la predicción de las reacciones de los usuarios interesados en el perfil empresarial en relación al contenido publicado es importante porque permite anticiparse la futura creación de contenido. Aporta una organización necesaria y en esencia, permite ajustar de forma eficaz el número de usuarios que se alcanzará para tener una prevision de posibles contratos.

### MÉTODOS

En este proyecto se desea desarrollar un módulo predictivo con técnicas de Inteligencia Artificial para lograr un incremento significativo de la tasa de conversión obtenido a partir de las redes sociales de la empresa aumentando el alcance y el número de *leads* de éstas.



### FIGURAS Y RESULTADOS



Tomando como referencia los datos obtenidos durante la medición de las estadísticas de las diversas redes sociales desde la Fundación de la empresa hasta la actualidad, se buscó organizar la información. Proceso costoso pues, recabar la información y acondicionarla es una ardua tarea.

Se aplicó la mejor estrategia posible para la integración de estos llevándolos a un algoritmo modelo de predicción.

La ciencia del Aprendizaje automático o Deep Learning llevará al computador a reconocer las diversas imágenes propuestas por los empleados para ser publicadas. El programa indicará de forma certera el impacto que provocará en los *posibles* clientes en función de los datos a partir de los cuales se ha entrenado.

Se precisará dividir el data set en dos conjuntos: uno de entrenamiento y otro de test de forma y manera que el algoritmo será capaz de reconocer y asociar el contenido de la imagen con el *copywriting* correspondiente. Dará así una predicción acertada.

Organic interactions	Labels (Temas)	Organic interactions	Valor	Alcance	Impresiones
2,5449918	SR - TotalEnergie	14	D	588	619
2,5482344	SR - TotalEnergie	14	D	639	819
112,0642453	SR - Sponsoring	614	A+	3391	3852
2,7412281	PAT_SP;SR - Spor	15	D	860	959
7,8639356	SR - CSR content;	43	C	1220	1328
2,5683361	Benchmark Soste	14	D	749	1000
3,4939316	SR - Sponsoring	19	D	907	1006
4,8165987	PAT_SP;SR - Spor	26	D	1029	1162
5,7825033	SR - TotalEnergie	31	C	1169	1562
48,5109571	PAT_SP;SR - Com	259	A+	1693	1923
7,9666161	PAT_SP;SR - Spor	42	C	1510	1718
5,3435114504	SR - Sponsoring	28	D	1138	1282
1,7178850926	SR - Sponsoring	9	D	0	0
19,7280214518	Benchmark Soste	103	A	1682	2129
0,5747126437	Benchmark Soste	3	D		
19,7280214518	SR - Sponsoring	103	A	1169	1309

La tabla muestra una pequeña parte del conjunto de datos (*dataset*) sobre los que se ha efectuado el estudio. Como se puede comprobar, las imágenes publicadas en las redes sociales fueron clasificadas en función de un criterio de "Valor". Esta tabla es más compleja en la práctica pues, la figura adjunta refleja tan solo cinco atributos más la llamada Clase.

### CONCLUSIONES

Las conclusiones obtenidas son realmente interesantes. Tomando los datos ya preparados se buscó preparar el Data Set para en el futuro, utilizar los algoritmos objetivo de forma óptima.

Para el Departamento de Comunicación y Marca es muy importante visualizar de forma correcta y simple los datos. Al expresar los datos como un Histograma se necesitó calcular la correlación de estos.

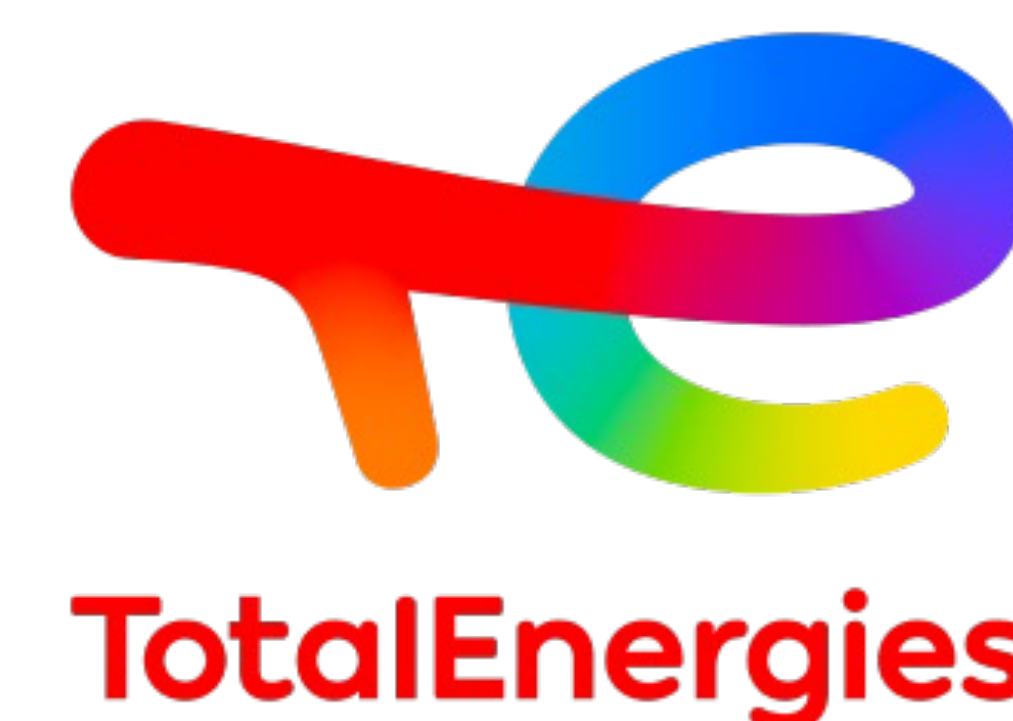
Como conclusión se obtuvo que la correlación es bastante baja. Esto es debido a que las variables son demasiado independientes por ello, no es posible determinar un único factor de las métricas seleccionadas como única causa de un incremento de seguidores.

Por otro lado, se identificaron las mejores horas de publicación, un dato muy relevante para la organización además del público al que deben de estar dirigidas

### TRABAJO FUTURO

El estudio se ha realizado con los datos de sólo una red social. El trabajo a futuro se basa en expandir esta actividad a las demás redes.

Por otro lado, se deberá concluir el trabajo de predicción utilizando técnicas de Clustering.



Universidad de Oviedo