

Sentimiento de los medios de comunicación españoles en formato digital sobre el Ibex 35

News sentiment from Spanish digital media to the Ibex 35

Carmen Orden-Cruz
Doctora en Ciencias Económicas y Empresariales,
Profesora de la Universidad Rey Juan Carlos
carmen.delaorden@urjc.es

Raúl Gómez-Martínez
Doctor en Economía de la Empresa y Finanzas,
Profesor de la Universidad Rey Juan Carlos
raul.gomez.martinez@urjc.es

Jessica Paule-Vianez
Doctoranda en Economía de la Empresa en el área de
Finanzas. Profesora de la Universidad Rey Juan Carlos
jessica.paule@urjc.es

Orden-Cruz, C., Gómez-Martínez, R. y
Paule-Vianez, J. (2019)
Sentimiento de los medios de comunicación españoles
en formato digital sobre el Ibex 35
Revista Internacional de Investigación en Comunicación
aDResearch ESIC. Nº 19 Vol 19
Primer semestre, enero-junio 2019 · Págs. 56 a 67
<https://doi.org/10.7263/adresic-019-03>

RESUMEN

Clasificación JEL:
G10, G40, C88

Palabras clave:
Sentimiento del
Inversor,
Transmisión de
sentimiento,
Medios de
Comunicación,
Información Bursátil,
Ibex 35

En este artículo se analiza en qué medida y sentido los medios de comunicación españoles en soporte digital transmiten sentimiento positivo o negativo en la información que publican sobre la bolsa. Concretamente, se evalúa el tratamiento que dan estos medios a las noticias que publican sobre el comportamiento del índice Ibex 35 con una transmisión de sentimiento positivo o negativo. Para captar y analizar el sentimiento contenido en las noticias, se ha utilizado una muestra de publicaciones con contenidos relativos a la evolución del índice bursátil desde el 11 de septiembre de 2015 hasta el 19 de noviembre de 2018, analizando un total de 1.165 días de publicaciones en 18 medios de comunicación españoles en el medio digital. Para llevar a cabo el objetivo de este trabajo, se han aplicado algoritmos de Procesamiento de Lenguaje Natural (PLN) sobre título y resumen de la noticia, con los cuales se ha extraído el sentimiento implícito de las noticias. Los resultados obtenidos muestran que los medios de comunicación españoles en formato digital sobre el Ibex 35 transmiten sentimiento en sus publicaciones, encontrándose diferencias entre los medios especializados y generalistas. Se ha observado, a su vez, cómo los medios que transmiten menos sentimiento, principalmente este es negativo. Este trabajo contribuye en el debate del impacto que el sentimiento generado por los medios de comunicación puede presentar en los mercados de valores.

ABSTRACT

JEL Classification:
G10, G40, C88

Key words:
investor sentiment,
sentiment
transmission,
media,
stock information,
Ibex 35

This article analyses the extent to which the Spanish digital media transmits positive or negative sentiment when they publish information about the stock market. Specifically, the evaluation was performed on news published about the performance of the Ibex 35 index with the objective of determining positive or negative sentiment transmission. The sample set, utilized publications with content related to the evolution of the stock index, covered the period from September 11, 2015 to November 19, 2018 and employed a total of 1,165 days of publications in 18 Spanish digital media. The publications were analysed with Natural Language Processing (PLN) algorithms to determine the extent to which sentiment was implied and transmitted by the news. The results showed that, the Spanish digital media transmitted sentiment in their publications about the Ibex35; there were differences between specialized and generalist media; the media that transmitted less sentiment were mainly negative. This work contributes to the discussion on the impact of sentiment generated by the media in the stock markets.

Introducción

La forma de procesar la información sobre los hechos acontecidos por parte de los medios de comunicación constituye uno de los temas que genera más debate. Existen múltiples estudios que corroboran la influencia de los medios de comunicación en la opinión de los receptores de la información en diversos ámbitos tales como la política (Entman, 1989), la economía y las finanzas (Dougal, Engelberg, García y Parsons, 2012; Balduzzi, Elton y Green, 2001) y la sociedad (Slater, 2007), entre otros.

En concreto, en el ámbito de la información financiera, la literatura presenta estudios sobre los factores que pueden afectar al sentimiento de los inversores y el efecto que dicho sentimiento puede tener en los mercados financieros. Así, el sentimiento del inversor ha sido un factor clave en la explicación de los continuos desequilibrios observados en los mercados de valores, destacando cómo impulsor de dicha medida el trabajo de Baker y Wurgler (2006). En esta línea, existen estudios que evalúan el impacto del estado ánimo generado por diversas variables sobre los mercados bursátiles, cómo son el clima (Kamstra, Kramer y Levi, 2003; Hirshleifer y Shumway, 2003; Jacobsen y Marquering, 2008), la estacionalidad (Seyhun, 1988), la calidad del aire (Li y Peng, 2016), las vacaciones (Kaplanski y Levy, 2012), los eventos deportivos (Gómez y Prado, 2014), las guerras o desastres (Kaplanski y Levy, 2012) y los ciclos lunares (Wang, Lin y Chen, 2010), entre otros.

Los medios de comunicación son otro factor estudiado en cuanto a su impacto en el sentimiento y, por ende, en los movimientos bursátiles. Diversos trabajos relacionan el sentimiento transmitido por distintos medios de comunicación y difusión con el comportamiento de los mercados de valores. Se encuentran estudios que ana-

lizan el impacto de distintas publicaciones en redes sociales cómo Twitter en el comportamiento de estos mercados. En esta línea, Oliveira, Cortez y Areal (2013) utilizan diversos indicadores de sentimiento y atención extraídos de Twitter y descubren que la opinión de Twitter y el volumen de publicación son factores relevantes para la previsión del rendimiento. Sprenger, Tumasjan, Sandner y Welppe (2014) analizan con métodos de lingüística computacional aproximadamente 250.000 mensajes relacionados con las acciones a diario, y encuentran una asociación entre el sentimiento del tweet y el rendimiento de las acciones, el volumen de mensajes y el volumen de transacciones, así como el desacuerdo y la volatilidad. Zhang, Li, Shen y Teglio (2016) examinan las relaciones entre el sentimiento de felicidad diaria extraído de Twitter y el desempeño de los mercados de valores de 11 mercados internacionales. Estos autores encuentran que la felicidad diaria extraída de Twitter tiene poder explicativo sobre la rentabilidad y la volatilidad de estos mercados.

Respecto al impacto que el sentimiento de la información publicada a través de noticias de prensa puede tener en estos mercados, se encuentran trabajos cómo el de Gidofalvi y Elkan (2001) en el que muestran como los movimientos a corto plazo del precio de las acciones se pueden predecir utilizando artículos de noticias financieras. Tetlock (2007) mide cuantitativamente las interacciones entre los medios y el mercado de valores utilizando el contenido diario de una columna de Wall Street Journal. Este autor descubre que el pesimismo de los medios de comunicación predice una presión a la baja sobre los precios del mercado seguida de una reversión a los fundamentos. García (2013) utiliza la fracción de palabras positivas y negativas de dos columnas de noticias financieras del The

New York Times para medir el sentimiento del inversor y llega a que esta medida permite predecir el rendimiento de las acciones, especialmente durante recesiones. Alanyali, Moat y Preis (2013) cuantifican la relación entre las decisiones tomadas en los mercados financieros y la evolución de las noticias financieras del Financial Times. Estos autores encuentran una correlación positiva entre la cantidad diaria de menciones de una compañía en el Financial Times y el volumen de transacciones diarias de las acciones de una compañía. Hillert, Jacobs y Müller (2014) descubren que el tono de artículos específicos pronostica negativamente el rendimiento del mercado, especialmente durante recesiones. García-Medina, Junior, Bañuelos y Martínez-Argüello (2018) analizan las correlaciones y el flujo de información entre 64.939 noticias en el The New York Times y 40 índices financieros mundiales durante 10 meses, sugiriendo que existe una relación profunda entre las noticias y los índices mundiales, impulsando las noticias los movimientos del mercado mundial.

Los trabajos que han investigado el sentimiento generado por los medios de comunicación en los inversores han aportado conclusiones valiosas en cuanto al impacto de ese sentimiento en los mercados de valores, sin embargo, estos estudios no ofrecen una radiografía del sentimiento de los medios de comunicación. De este modo, el objetivo de este trabajo se centra en estudiar el sentimiento que los medios de comunicación, en este caso españoles, transmiten sobre la evolución del Ibex 35, índice representativo del mercado bursátil español. Para conseguir este objetivo, la información publicada por los medios de comunicación en formato digital se configuró como la mejor opción para tener información disponible en texto plano y poder ser tratada a través de Procesamiento de Lenguaje Natural

(PNL). Esta técnica presenta la ventaja de que permite extraer información del lenguaje utilizado en textos o el habla (Chowdhury, 2003) y que permite procesar un alto volumen de información con bastante eficacia, encontrándose entre sus múltiples aplicaciones, su función de motor de búsqueda web (Vallez y Pedraza, 2007).

Los resultados obtenidos en el estudio ponen de manifiesto cuánto y en qué medida los medios de comunicación españoles en sus publicaciones transmiten sentimiento, encontrándose diferencias entre los medios especializados y los generalistas.

Tras la introducción, en el segundo epígrafe, se plantean las hipótesis del estudio, continuando en el tercer epígrafe con la explicación de la metodología y la muestra utilizada; en el tercero se hace un análisis de los datos y se extraen los resultados, y, por último, en el cuarto epígrafe, se finaliza con las conclusiones obtenidas.

Hipótesis

Debido a que en este trabajo se persigue estudiar cuánto y en qué medida los medios de comunicación manifiestan sentimientos en sus publicaciones, y si dichos sentimientos difieren entre unos medios de comunicación u otros, las hipótesis que se plantean son las siguientes:

H1: Los medios de comunicación digitales españoles transmiten sentimiento en la información sobre el Ibex 35.

En función de la literatura expuesta que demuestra la existencia de transmisión de sentimiento en los medios de comunicación, en este trabajo, se parte de que los medios de comunicación digitales españoles transmiten sentimientos positivos o negativos respecto al comportamiento del Ibex 35.

Si el análisis del contenido de las informaciones arroja sentimiento positivo o negativo, quedaría validada la H1.

H2: El sentimiento que transmiten los medios de comunicación digitales españoles sobre la información del Ibex 35 es diferente en unos y otros.

Las diferencias en la forma de transmitir la información bursátil en publicaciones de distintos medios de comunicación ha sido un área poco explorada. Así, en este trabajo, se sugiere que cada medio puede transmitir información en un sentido u otro en función de sus sentimientos y percepciones individuales, por lo que deberían de encontrarse diferencias entre los sentimientos transmitidos por unos medios u otros respecto al Ibex 35.

Si el porcentaje de noticias que transmiten sentimiento positivo o negativo es diferente dependiendo de cada medio de comunicación, quedaría validada la H2.

H3: Los medios que transmiten con menos frecuencia un sentimiento en la información sobre el Ibex 35, cuando lo hacen es en negativo.

Hasta donde nuestro conocimiento llega, no se ha encontrado literatura que relacione la frecuencia de la transmisión de sentimientos de los medios con el signo de dicho sentimiento. En este trabajo se parte de que, en los medios que transmiten con menor frecuencia un sentimiento respecto al comportamiento del Ibex 35, el sentimiento transmitido será mayoritariamente negativo.

Si los medios que transmiten con menos frecuencia un sentimiento, cuando lo hacen es con un sentimiento negativo superior al 50%, quedaría validada la H3.

Metodología y datos

Para poder confirmar las hipótesis expuestas se han analizado los sentimientos extraídos de dis-

tintos medios de comunicación en formato digital españoles que han transmitido noticias respecto al Ibex 35. La metodología seleccionada para la extracción del sentimiento de las publicaciones de estos medios ha sido la técnica de Procesamiento del Lenguaje Natural (PNL), concretamente, la herramienta que ofrece Sigma Technologies¹.

El PNL es una disciplina que tiene su base en áreas como la Informática, la Lingüística, las Matemáticas, la Inteligencia Artificial y la Psicología, entre otras, y que explora cómo se pueden usar las nuevas tecnologías para «comprender y manipular el texto o el habla en lenguaje natural para hacer cosas útiles» (Chowdhury, 2003, p.1).

El tratamiento de la información obtenida se hizo siguiendo la metodología de InvestMood², un sistema de negociación bursátil basado en algoritmos que tratan la información extraída a través de PNL. La información la extrae mediante un proceso de rastreo diario de todas las noticias publicadas en medios digitales relativas a un término. Sus sistemas recogen la información desestructurada y clasifican cada noticia mediante algoritmos automáticos de PNL en los grupos que se deseen según las condiciones establecidas.

De este modo, las noticias que mencionaban la palabra «Ibex» en medios de comunicación digitales especializados en Economía y Finanzas se clasificaron en tres grupos según el sentimiento transmitido. Para ello, se parametrizó como «sentimiento positivo» aquellas noticias que transmitían un sentimiento favorable sobre el Ibex 35, como «sentimiento negativo» a las publicaciones que transmitían un sentimiento desfavorable sobre el Ibex 35, y como «sentimiento

1 Véase <http://www.sigmatechnologiesglobal.com/>

2 Véase <http://www.investmood.com/>

neutro» a aquellas noticias que no transmitían ningún sentimiento o en las que no había noticia alguna ese día sobre dicho índice. Dos ejemplos del análisis automatizado serían los siguientes. La noticia del 29 de enero de 2016 a las 09:44 de Europa Press: «El Ibex 35 se dispara un 1,6% en la apertura, animado por Repsol y los bancos»³ fue calificada como «positiva», y la noticia del 15 de agosto de 2018 a las 18:59 de El País: «Las Bolsas europeas caen y el Ibex marca su nivel más bajo desde marzo»⁴ fue calificada por el algoritmo como «negativa».

En cuanto a la muestra, dada la especialización del término «Ibex», para que esta fuera representativa se decidió considerar como fuente aquellos medios con diversos indicadores de relevancia, tomando como fuente de información el informe «Estudio general de medios de comunicación. Año móvil abril 2017 – marzo 2018» que elabora periódicamente la Asociación para la Investigación de Medios de Comunicación (2018). En el caso de los diarios, se consideró el número de lectores y para las cadenas de televisión, se utilizó el porcentaje de cuota de audiencia ya que las noticias publicadas en su página web corresponden a las emitidas en pantalla. Como agencias de noticias, se eligieron dos relevantes, una a nivel nacional⁵ (Europa Press) y otra internacional⁶ (Reuters). En total, se eligieron 18 medios de comunicación, siendo los siguientes (por orden alfabético):

3 Véase <https://www.europapress.es/economia/bolsa-00348/noticia-ibex-35-dispara-16-apertura-animado-repsol-bancos-20160129091815.html>

4 Véase https://elpais.com/economia/2018/08/15/actualidad/1534344992_597270.html

5 Véase Ranking sectorial de empresas. Sector CNAE ((6391) Actividades de las agencias de noticias. Disponible en: <http://ranking-empresas.economista.es/sector-6391.html>

6 Véase MMC208, Top 10 news agencies in the world. Disponible en: <http://mmc208.aue.ae/?p=15171>

co): 20 Minutos, ABC, Antena 3, Cinco Días, El Confidencial, eldiario.es, El Mundo, El País, El Periódico, Europa Press, Expansión, La Razón, La Sexta, La Vanguardia, Reuters, RTVE, Telecinco y Telemadrid.

Los textos originales se obtuvieron de las noticias publicadas en soporte digital lo que permitía disponer de los textos de las noticias. Según la AIM (2017), la prensa digital se ha impuesto al resto de los medios pues cuenta con el 84,3% del consumo semanal frente al 56,2% de la prensa en papel. Además, según el AIMC (2018), en España, el ratio de penetración de la prensa tradicional general ha ido cayendo en la última década, situándose actualmente en el 24,2%. La prensa económica especializada tiene unos 140.000 lectores, muy alejado de la prensa general que se sitúa en torno al millón de lectores diarios en el mejor de los casos. De este modo, la utilización de este soporte digital se configura como mejor opción frente a la de papel.

Bajo estas condiciones, la muestra de las noticias publicadas se obtuvo en el periodo que abarca desde el 11 de septiembre de 2015 hasta el 19 de noviembre de 2018 (1.165 días). Dado que hay medios que en el mismo día publican varias noticias sobre el término «Ibex», más de 50.000 noticias fueron analizadas en este proceso.

Una vez obtenidos los datos, se procedió a la extracción de datos estadísticos sobre la repercusión y el sentimiento de las noticias publicadas.

Resultados

Con los datos recopilados aplicando la metodología planteada, en la tabla 1 se muestra un resumen del análisis de sentimiento captado para los medios de comunicación seleccionados:

Tabla 1 · Clasificación del Sentimiento transmitido por los 18 medios de comunicación digitales españoles sobre el Ibex 35 en el período 11/9/2015-19/11/2018

Medio	Sentimiento Neutro	Sentimiento Positivo	Sentimiento Negativo	Transmisión %	Sentimiento Positivo %	Sentimiento Negativo %	Sentimiento Predominante
20 minutos	789	13	20	4.01%	39.39%	60.61%	Negativo
Abc	682	52	88	17.03%	37.14%	62.86%	Negativo
Antena 3	818	0	4	0.49%	0.00%	100.00%	Negativo
Cinco días	183	265	374	77.74%	41.47%	58.53%	Negativo
El Confidencial	589	87	146	28.35%	37.34%	62.66%	Negativo
El Mundo	816	0	6	0.73%	0.00%	100.00%	Negativo
El País	773	16	33	5.96%	32.65%	67.35%	Negativo
El Periódico	745	28	49	9.37%	36.36%	63.64%	Negativo
Eldiario.es	799	9	14	2.80%	39.13%	60.87%	Negativo
Europa Press	114	405	303	86.13%	57.20%	42.80%	Positivo
Expansión	373	241	208	54.62%	53.67%	46.33%	Positivo
La Razón	270	284	268	67.15%	51.45%	48.55%	Positivo
La Sexta	814	1	7	0.97%	12.50%	87.50%	Negativo
La Vanguardia	243	308	271	70.44%	53.20%	46.80%	Positivo
Reuters	775	27	20	5.72%	57.45%	42.55%	Positivo
Rtve	632	80	110	23.11%	42.11%	57.89%	Negativo
Telecinco	789	17	16	4.01%	51.52%	48.48%	Positivo
Telemadrid	775	29	18	5.72%	61.70%	38.30%	Positivo

Fuente: Elaboración propia

En la tabla 1, se observa en las tres primeras columnas el número de noticias sobre el Ibex 35 que se han encontrado en cada medio que presentan un sentimiento neutro, un sentimiento positivo y un sentimiento negativo. En la cuarta columna se observa la transmisión de cada medio, entendiéndose la misma por el porcentaje de noticias con sentimiento respecto al total de noticias emitidas

sobre el Ibex 35. La quinta y sexta columna muestran, dentro de las noticias con sentimiento, el % de noticias que lo incorporan en sentido positivo y negativo, respectivamente. Y, por último, la séptima columna muestra el sentimiento predominante transmitido por cada medio de comunicación.

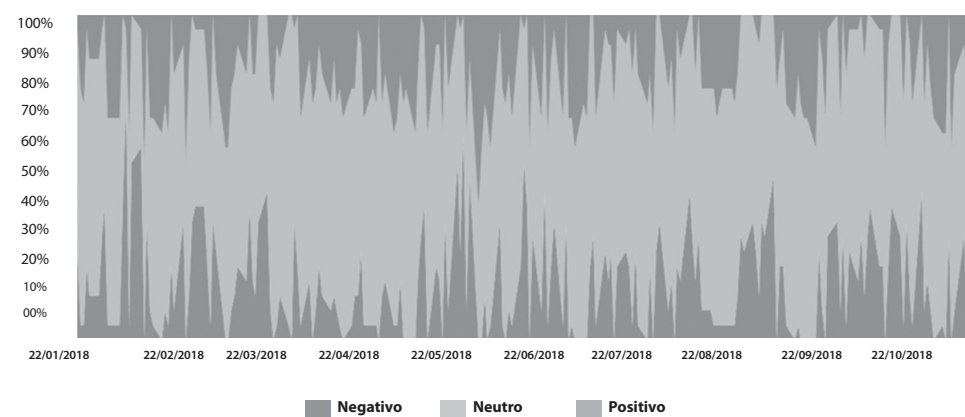
Con todo esto, se procede a validar o rechazar las hipótesis planteadas.

Comenzando por la H1: «Los medios de comunicación digitales españoles transmiten sentimiento en la información sobre el Ibex 35», esta quedó validada. En la Tabla 1, se observa claramente cómo todos los medios, en mayor o menor medida, han transmitido sentimiento en algunas de sus publicaciones.

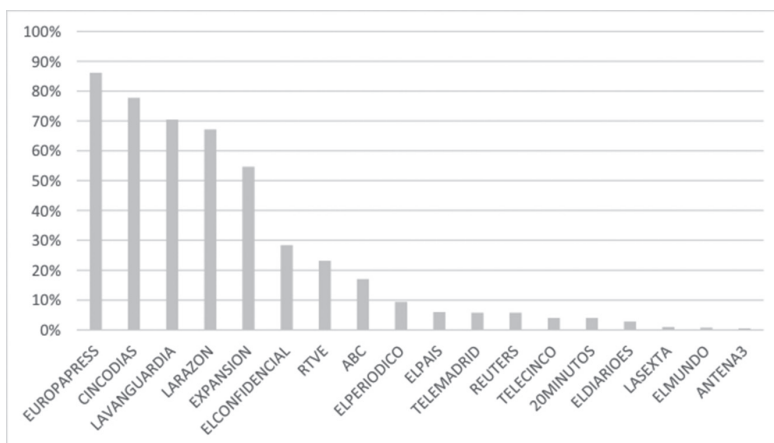
En el Gráfico 1 se puede visualizar el sentimiento medio transmitido por los medios de comunicación digitales seleccionados diariamente sobre el Ibex 35 en el periodo seleccionado. Aunque, en general, el sentimiento neutro ha sido el predominante, se observa como los medios, en su conjunto, efectivamente transmitieron sentimiento, ya fuera positivo o negativo en la mayoría de los días. En término medio, el sentimiento positivo transmitido ha sido de un 17,05%, el sentimiento neutro de un 65,70% y el sentimiento negativo de 17,26%.

Respecto a la H2: «El sentimiento que transmiten los medios de comunicación digitales es-

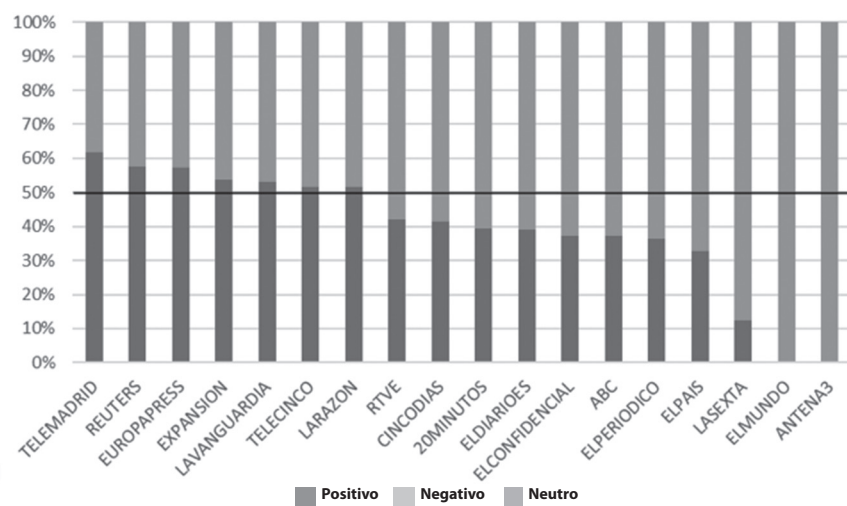
pañoles sobre la información del Ibex 35 es diferente en unos y otros», también quedó validada. El estudio revela diferente activismo según el tipo de medio del que se trate. En concreto, Europa Press es el medio de comunicación más activo ya que transmite sentimiento en un 86,13% de los casos, seguido de Cinco Días con un 77,74% y La Vanguardia con un 70,44% cómo se observa en la Tabla 1 y en el Gráfico 2. Esto demuestra, a su vez, que los medios que más han transmitido sentimiento han sido estos tres. En el otro extremo, los medios que menos lo han hecho han sido Antena 3 con un 0,49%, El Mundo con un 0,73% y La Sexta 0,97%, medios en su conjunto generalistas. Estos resultados, además de demostrar que existen diferencias en la transmisión de sentimiento por los distintos medios, parecen apuntar a la presentación de información con más sentimiento por parte de medios especializados que por parte de medios generalistas.

Gráfico 1 · Transmisión del sentimiento medio diario sobre el Ibex 35 en el período 11/9/2015-19/11/2018

Fuente: Elaboración Propia

Gráfico 2 · Grado de transmisión del sentimiento, en término medio, por cada medio de comunicación

Fuente: Elaboración Propia

Gráfico 3: Sentimiento transmitido por cada medio de comunicación

Fuente: Elaboración Propia

Por último, respecto a la H3: «Los medios que transmiten con menos frecuencia un sentimiento en la información sobre el Ibex 35, cuando lo hacen es en negativo», del mismo modo que las anteriores, quedó validada. La Tabla 1 muestra cómo los medios de comunicación que apenas presentan sentimiento, como Antena 3, El Mundo y La Sexta, suponen, respecto al total de sentimiento transmitido, un 100% de sentimiento negativo Antena 3 y el Mundo, y un 87,50% de sentimiento negativo La Sexta. El Gráfico 3 nos muestra el sentimiento positivo y negativo transmitido por cada medio de comunicación en proporción, ordenados de menor sentimiento negativo a mayor. Se aprecia claramente una mayor transmisión de sentimiento negativo respecto al positivo.

Conclusiones

En este trabajo se ha realizado un estudio sobre la transmisión de sentimiento de la información publicada por los medios de comunicación españoles en formato digital respecto al Ibex 35. Para ello, se ha aplicado un algoritmo de PNL sobre la información de las publicaciones de 18 medios de comunicación españoles durante 1.165 días sobre el comportamiento del Ibex 35.

El estudio demuestra la existencia de transmisión de sentimiento de la información por parte de estos medios, encontrándose, por término medio, un 34,31% de sentimiento emitido. Si bien, cabe precisar que se han encontrado diferencias entre el sentimiento transmitido sobre el comportamiento del Ibex 35 por unos medios de comunicación y otros. Los resultados apuntan que los sentimientos son transmitidos principalmente por los medios especializados en la materia y lo hacen de forma bastante equilibrada tan-

to en positivo como negativo. Sin embargo, en el caso de los medios más generalistas, se ha encontrado que los medios que menos sentimiento transmiten, cuando lo hacen, lo transmiten con sentido negativo. Así, parece que los medios generalistas se enfocan más en transmitir la información negativa con sentimiento negativo que la positiva.

De este modo, el estudio realizado constata que cuando los medios de comunicación digitales abordan una noticia sobre el Ibex, en un elevado porcentaje de las ocasiones imprimen una valoración personal del redactor tanto en el titular como en el resumen. Estas connotaciones, positivas o negativas, pueden afectar al sentimiento de los inversores y, por tanto, a la evolución del mercado, tal y como demostraron Hardling y He (2016). Según estos autores, las decisiones de inversión tras un sentimiento negativo eran mucho más conservadoras que tras un sentimiento positivo. Así, tal y como demuestra este estudio, los medios de comunicación españoles en formato digital imprimen en sus noticias con más frecuencia un sentimiento negativo que positivo. Este aspecto podría afectar al mercado provocando una menor tolerancia al riesgo y un menor apetito por la inversión, relación que podría ser testeada en el caso español en un próximo trabajo.

Por último, cabe señalar que este estudio presenta como principal limitación la utilización de un sólo término «Ibex» lo que acota los resultados en la detección de noticias que se refieran al comportamiento de la bolsa. En próximas investigaciones se utilizarán otros términos que hagan referencia al mismo aspecto como «mercado bursátil» o «bolsa» condicionado que en el texto se haga referencia al mercado español.

Bibliografía

- AIM (Asociación de Medios de Información) (2017): *Libro Blanco de la Información 2017*. Disponible en: <https://www.ami.info/estudios-informes/libro-blanco>
- AIMC (Asociación para la Investigación de Medios de Comunicación) (2018): *Estudio general de medios de comunicación*. Año móvil abril 2017 – marzo 2018. Disponible en: <https://www.aimc.es/aimc-c0nt3nt/uploads/2018/04/resumegml18.pdf>
- Alanyali, M., Moat, H. S., & Preis, T. (2013). Quantifying the relationship between financial news and the stock market. *Scientific reports*, 3, 3578.
- Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The journal of Finance*, 61(4), 1645-1680.
- Balduzzi, P., Elton, E. J., & Green, T. C. (2001). Economic news and bond prices: Evidence from the US Treasury market. *Journal of financial and Quantitative analysis*, 36(4), 523-543.
- Chowdhury, G. G. (2003). Natural language processing. *Annual review of information science and technology*, 37(1), 51-89.
- Dougal, C., Engelberg, J., Garcia, D., & Parsons, C. A. (2012). Journalists and the stock market. *The Review of Financial Studies*, 25(3), 639-679.
- Entman, R. M. (1989). How the media affect what people think: An information processing approach. *The journal of Politics*, 51(2), 347-370.
- García, D. (2013). Sentiment during recessions. *The Journal of Finance*, 68(3), 1267-1300.
- García-Medina, A., Junior, L. S., Bañuelos, E. U., & Martínez-Argüello, A. M. (2018). Correlations and flow of information between the New York Times and stock markets. *Physica A: Statistical Mechanics and its Applications*, 502, 403-415.
- Gidofalvi, G., & Elkan, C. (2001). Using news articles to predict stock price movements. *Department of Computer Science and Engineering, University of California, San Diego*.
- Gómez, R., y Prado, C. (2014). Sentimiento del inversor, selecciones nacionales de fútbol y su influencia sobre sus índices nacionales. *Revista Europea de Dirección y Economía de la Empresa*, 23(3), 99-114.
- Harding, N. & He, W. (2016). Investor Mood and the Determinants of Stock Prices: An Experimental Analysis. *Accounting & Finance*, Vol. 56 (2), 445-478, 2016.
- Hillert, A., Jacobs, H., & Müller, S. (2014). Media makes momentum. *The Review of Financial Studies*, 27(12), 3467-3501.
- Hirshleifer, D., & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance*, 58(3), 1009-1032.
- Jacobsen, B., & Marquering, W. (2008). Is it the weather?. *Journal of Banking & Finance*, 32(4), 526-540.
- Kamstra, M. J., Kramer, L. A., & Levi, M. D. (2003). Winter blues: A SAD stock market cycle. *American Economic Review*, 93(1), 324-343.
- Kaplanski, G., & Levy, H. (2012). The holiday and yom kippur war sentiment effects: the Tel Aviv Stock Exchange (TASE). *Quantitative Finance*, 12(8), 1283-1298.
- Li, Q., & Peng, C. H. (2016). The stock market effect of air pollution: evidence from China. *Applied Economics*, 48(36), 3442-3461.
- Oliveira, N., Cortez, P., & Areal, N. (2013, September). On the predictability of stock market behavior using stocktwits sentiment and posting volume. In *Portuguese Conference on Artificial Intelligence* (pp. 355-365). Springer, Berlin, Heidelberg.
- Seyhun, H. N. (1988). The January effect and aggregate insider trading. *The Journal of Finance*, 43(1), 129-141.
- Slater, M. D. (2007). Reinforcing spirals: The mutual influence of media selectivity and media effects and their impact on individual behavior and social identity. *Communication theory*, 17(3), 281-303.
- Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and trades: The information content of stock microblogs. *European Financial Management*, 20(5), 926-957.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139-1168.
- Vallez, M. y Pedraza, R. (2007). El Procesamiento del Lenguaje Natural en la Recuperación de Información Textual y áreas afines. *Hipertext.net*, 5.
- Wang, Y. H., Lin, C. T., & Chen, W. L. (2010). Does lunar cycle effect exist? Lunar phases and stock return volatilities. *African Journal of Business Management*, 4(18), 3892-3897.
- Zhang, W., Li, X., Shen, D., & Teglio, A. (2016). Daily happiness and stock returns: Some international evidence. *Physica A: Statistical Mechanics and its Applications*, 460, 201-209.
- Zhang, W., Li, Y., Zhang, Z., & Shen, D. (2018). The dynamic cross-correlations between foreign news, local news and stock returns. *Physica A: Statistical Mechanics and its Applications*, 509, 861-872.