

Defining and Measuring Microblogging Sentiment Investors on Stock Market: A Literature Review

Putri Fariska^{1,*}, Nugraha², Mochamad Malik Akbar Rohandi³

¹ Universitas Pendidikan Indonesia

² Universitas Pendidikan Indonesia

³ Universitas Islam Bandung

*Corresponding author. Email: putri.fariska@upi.edu

ABSTRACT

The purpose of this paper is to review microblogging sentiment investors on a literature-based. Focusing on defining and measuring online sentiment investors and understanding their impact on financial market behavior—using scholarly articles to analyze, define, and measure microblogging sentiment investors. All the articles have been classified to measure and underpinnings online investor sentiment theory and other determinants in psychological measure. We have used citation and bibliographic coupling analysis to provide a structure of microblogging sentiment investors. This research also investigates how microblogging forums have become an online platform for exchanging stock information among investors. This paper's result defines investor sentiment research development will further reveal new ways of measuring investor sentiment through big data social media on microblogging sentiment investors construct by identifying authors, cite, references, or keywords.

Keywords: *Microblogging, Sentiment, Investors, Stock Market.*

1. INTRODUCTION

Investing in speculative assets is a social activity influenced by social movements [1]. Social dynamic provides open information on whether the truth is known or not. The information can have an impact on investors who assume that the information is accurate. This tends to bias every decision and is considered a cognitive bias [2].

Research that discusses the relationship between social interaction and investment is in sentiment investors [3]. The sentiment is considered a way for investors to form trust [4]. When investors decide, their sentiment or state of mind influences it [5], [6], sentiment can be in the form of optimism or pessimism about the future stock activity [7].

Sentiment investor formed by three things: First is news media content [8]. Second is data search or a query on the internet are proven to be related and even predict market fluctuations [9], and the third is posting on social media becomes an important data source to measure sentiment investor and social mood extraction [10], [11].

In some conversations, everyone can quickly exchange information, opinions, and emotions, which is considered necessary in the stock and investment markets [12], and information about investments will spread through the environment from word of mouth effects [13].

Information technology is a significant factor for investor sentiment in financial markets. Sentiment appears in robust information and captures information inherent in market variables related to sentiment, such as volatility [14]. As the development of information technology, the internet and social media provide a platform to express emotions to the public and have a significant influence on overall public opinion [15].

Through social media, online sentiments have a predictive value on stock prices using Google Insights [5], [15]. Yahoo Finance also helps predict market volatility [10]. Twitter can predict the stock market, investor alpha impact on followers, and casualty relationships among them [5], [16], [17], [18], [19], [20].

As in [3] investor sentiment's theoretical structure by identifying the research area's foundations and leading

journals, references, authors, or keywords. Due to the increase of publications on investor sentiment, microblog sentiment investor literature lacks cohesion.

Through microblogging (e.g., Twitter), which has become a dynamic online platform for exchanging stock-related information and by using social media data mining methods for the analysis of public sentiment, investor sentiment research with a bottom-up approach can be developed in order to be able to show decisions taken by investors affected by psychological biases.

This study aims to underpinning online investor sentiment theory and other determinants in a psychological measure of behavioral finance's theoretical field. We collected information from the Scopus database using Harzing's Publish or Perish about all the articles with keywords sentiment investor online or microblogging that publish between 1980 – 2020.

We used bibliometric analysis as a technique that used by [3] and [21] in the investor sentiment research area using VOSviewer that provides three visualizations, referred to as the network visualization, the overlay visualization, and the density visualization.

One of the essential things that form sentiment investors is news media content [7]. News that appears in the media impacts investor behavior and stock prices even when new information is not available [12].

The human mind tends to link social results with individual actions. One example of investment socialization is the investment group. Through this group, individual investors discuss investment stocks that have the potential to make a profit. It is known that the investment behavior in the group shows the same psychological bias as individual investors [12].

Individual investors react to the news that contains old information that has been previously released. Usually, shares dominated by individual investors experience a significant return on the day investors act on news that contains old information and drop the next day significantly. This is caused by investors' failure to understand new or old news, especially financial media, which generally publish old news so that the media exacerbates investor bias [8].

Conversations make it possible to exchange information, opinions, and emotions very quickly. This process is considered to be very important for the stock and investment markets. Investors can seek information and expert opinions from social media such as Twitter or Facebook, or institutional investors can form groups to share investments, or analysts communicate with executives, managers, and local groups and associations to interact with each other.

Reference [22] examines stocks that are most actively discussed by online posters and see if the messages posted about these stocks have information or if they are

just noise. The result provides evidence that online talk affects stock prices and suggests that corporate managers, tiny firms, should monitor the stock message boards, e.g., yahoo finance.

Web blog also provides potential information. [23] extracted investor sentiment from weblog text using a semantic analysis that starts on the word and sentence level and employs ontology-guided and finds that sentiments from weblogs provide evidence for investor's economic utility.

Defining bottom-up approach on microblogging sentiment investors among of empirical evidence that concluded individual emotion in online perspective. Using lexical-semantic extension and correlation analysis method to extend the scale of emotion words, which can capture strong emotion, find that this model classifying opinions in the stock message board [24]. [25] explores individual emotion from an online perspective through 10 million stock and finds that five attributes of the stock market can completely predict by various online emotions like disgust, joy, sadness, and fear.

Another empirical evidence used virtual investment communities (VICs) to study the propensity of investors. Analyze 682,781 messages on Yahoo! Finance message boards for 29 Dow Jones stocks and assess how investors select a particular thread to respond. Results revealed that, despite the benefits from heterophily, investors are not immune to the allure of homophily in interactions in VICs [26]. Homophily refers to the propensity to seek interactions with others of similar status (e.g., religion, education, income, occupation) or values (e.g., attitudes, beliefs, and aspirations).

Another study examines stock message boards to analyze. [27], using data from 502 investor responses analyses, revealed that investors exhibit confirmation bias and a negative interaction effect between the perceived knowledge and the strength of prior belief on confirmation bias.

A recent study also examines the relationship between local daily happiness sentiment extracted from Twitter and stock returns [28], [29], [30]. All these findings indicate interdependencies between online activities and stock markets. Online stock investment message boards play a more significant role in social, political, and economic interactions through information published that many investors consider highly valuable.

As in [31], proposes a general framework for identifying expert investors and uses it as a basis for several models that predict stock rise from stock microblogging messages (stock tweets). This indicates that investor alpha, or we can say as an expert user followed by another user, influences the diffusion of information in microblogging forums and provides a mechanism for the reader to weigh information.

The top-down approach on microblogging sentiment investors is used to see the influence of investor sentiment on the capital market to show that investors are not as rational as they are. Reference [11] and [33] predict the market through online sentiment and find that investor sentiment on Twitter is considered as a very significant predictor in market daily returns. However, sentiment indicators do not affect financial market values after controlling mood indicators of emotional tweets that are significantly negatively correlated with Dow Jones, NASDAQ, and S&P 500 but show significant positive correlations with VIX.

In a top-down approach, among of empirical evidence that concluded sentiment investor as a determinant in a stock market [18], [22], [33], [34], [35], [36], [37], [38], [39], [40], [41] analyzed financial blogs and online news articles to develop a public mood dynamic prediction model for stock markets. The result assesses the emotional content of commentary on current stock or financial issues that can effectively forecast stock price movement [42].

Other empirical evidences predicting the effect of news and social media on stock market prediction are [43], [44], [45].

In general, investor emotions have traditionally been measured using surveys. However, with social media platforms such as Twitter, Yahoo Finance, and Google Insight, many researchers have used behavioral finance research in both the macro and micro approaches.

Researcher measuring microblogging sentiment investors with text classifier technology [32], lexical-semantic extension and correlation analysis method [24], naive Bayesian classification method [10], [18], Sprenger et al., 2014), text mining [46], [47], combines text-mining, feature selection and Bayesian networks [48], IBM natural language processing [49], machine learning algorithms to infer the relationship between the general public view regarding a stock and its evolution within the stock market [50].

2. METHODS

We used the Scopus database using Harzing's Publish or Perish and Sefora to find all the articles with keywords sentiment investor online or microblogging that publish between 1980 – 2020.

From the Scopus search method, the Scopus database's search system uses authors, affiliations, publication names, title words, and keywords by entry the years. This procedure ensures that all publications on sentiment investor online are included.

As a result, we downloaded about 1.272 papers on Scopus using Harzing's Publish or Perish. We used Microsoft Excel and VOSViewer 1.6.13 to process and analyze the data. The VOS approach provided a visual

outline of the relationships among authors, publications, journals, and keywords [3] and applied it to investor sentiment research analysis.

We create a map based on text data through API in a specific search query with a DOI file. Also, we create a map based on bibliographic data, read from reference manager file using RIS and Endnote file to analyze co-authorship and co-occurrence. Reference [51] concluded that co-citation measures the number of common forward citations of two studies or journals, and bibliographic coupling measures the number of common backward citations. In comparison, [52] assumes that co-citation is the most appropriate technique to identify the primary authors and studies published on investor sentiment applied to authors and references.

3. RESULTS AND DISCUSSION

Research that discusses investor sentiment began in 1989 on Bullish or bearish: A dichotomous Bayesian model for forecast turning points in the foreign exchange market [53], over a year rapid development occurred in this research area in the figure below.

From figure 1, we can see that the most extensive keyword in publication articles base on keyword co-occurrence analysis is sentiment analysis.

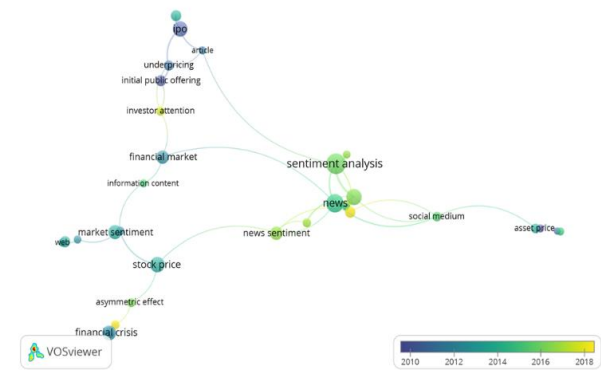


Figure 1 Bibliographic keywords co-occurrence analysis

Sentiment analysis in green color is the center of the map, its shows that the largest average publications year is between 2014 – 2016. Before that, the publications area is more concentrated on the top-down approach with the area financial market, IPO, underpricing, volatility, and price in the stock market. In the financial market, there is a link between information content and trade.

Furthermore, sentiment in the financial market is influenced by news that causes news sentiment. Besides, in predicting the state of the capital market, sentiment analysis measures the emotional relationship in social media.

Twitter is a social media platform used in analyzing sentiments in the capital market, and it can be seen that

there is a link between sentiment analysis and Twitter. In 2017, a growing area of research was the relationship between "sentiment analysis", "stock market prediction", "social media", and "news sentiment", this is due to changes in investor behavior using social media as a source of information in taking the decision.

Investor sentiment research is divided into eight clusters, where the first cluster shows the link between asset prices, closes end funds, investor confidence, equity, and social media. The second cluster shows links between financial markets, information content, IPOs, and underpricing. In cluster three, investor sentiment is related to asymmetric effects, economic policy, a financial crisis in the capital market. Furthermore, in the fourth cluster, stock market volatility is related to excess return, trade, web information, and market sentiment.

The fifth cluster shows the relationship between emotions, news sentiment for stock market prediction. Cluster six illustrates the relationship between IPO, article, and pricing. Cluster seven illustrates the relationship between news, news strategy, and strong investor sentiment. The final cluster illustrates the relationship between sentiment analysis using the Twitter platform to predict the stock market.

In a recent study, from the top-down approaches, a growing area of research links the financial market and investor attention through initial public attention that caused in the underpricing area. Besides, some studies discuss the asymmetric effect, economic policy, and financial crisis that affect investor sentiment.

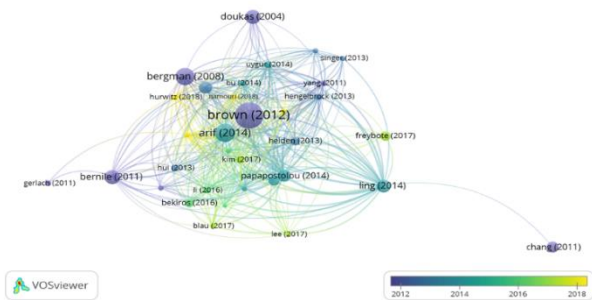


Figure 2 Bibliographic coupling document analysis.

It can be concluded that the most extensive total strength of links between the author and the investor sentiment research area is Brown (2012) with a total link strength of 101, followed by Ari (2014) with a total link strength of 108, Bergman (2008) with a total link strength of 60, Bernile (2011) with a total link strength of 58 and Doukas (2004) with a total link strength of 32.

There are six different clusters in Total. The first cluster discusses the stock returns and volatility with the investor sentiment index. The second cluster examines the market response to investor sentiment. The third cluster discusses aggregate sentiment and investor

sentiment. The fourth cluster discusses how investor sentiment can predict market returns, the fifth cluster discussed how to measure investor sentiment, and the last cluster is measuring investor sentiment towards private market return.

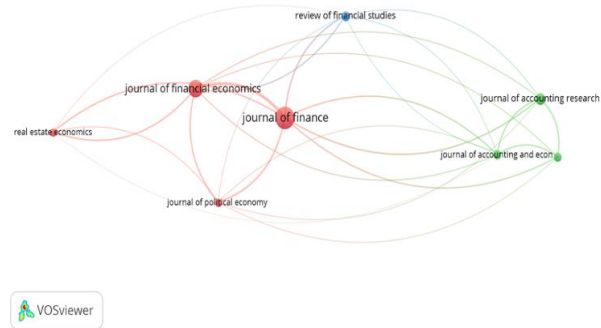


Figure 3 Bibliographic co-citation analysis.

It can be concluded that journal finance is the center of a map. In other words, journals that discuss investor sentiment are published in the journal of finance with a total citation of 132. These results show a strong relationship between the publications of this journal and investor sentiment.

Divided into three clusters, in the first cluster the link between "journal of finance", "journal of financial economics" and "journal of political economy" with a total link strength of 1781, while in the second cluster, there is a link "journal of accounting and economics", "The journal of accounting research" and "the accounting review" with a total link strength of 709, the last being a review of financial study cluster with a total link strength of 573.

4. CONCLUSION

The bibliometric analysis shows that the distribution of research in the area of sentiment analysis is divided into two phases. The first phase concentrated on the top-down analysis approach. The distribution of maps is connected to the financial market, IPO, underpricing, volatility, price, and a link between information content and trade in the stock market. The second phase is the most significant phase of publication in the investor sentiment area. This phase has begun to develop research from the bottom-up approach by examining emotions on investor sentiment. When investors make a decision, their mental and emotional basis of human interaction, including economic interactions that focus on elements such as self-esteem, shame, insecurity, and selfishness [54], with technology, the development of investor sentiment research will further reveal new ways of measuring investor sentiment through big data social media. With the development of information technology, social media (Twitter) is used to capture emotional signals over information and news that develops among

investors. It is also used to measure investor confidence in its relationship with asset price.

REFERENCES

- [1] Shiller, J. R. 1984. *Stock prices dan Social Dynamics*. Brookings Papers on Economic Activity. 2(1984):457-510.
- [2] Pompian, Michael, M. 2006. Behavioral Finance and Wealth Management: How to Build Optimal Portfolios That Account for Investor Biases. John Wiley & Sons, Inc., Hoboken, New Jersey.
- [3] M. Ángeles López-Cabarcos, Ada M. Pérez-Pico, Paula Vázquez-Rodríguez & M. Luisa López-Pérez. 2019. *Investor sentiment in the theoretical field of behavioural finance*. Economic Research-Ekonomska Istraživanja.
- [4] Barberis, N., Shleifer, A. dan Vishny, R. 1998. *A model of investor sentiment*. Journal of Financial Economics, 49(1998): 307-343.
- [5] Blajer-GoleRbiewska, A., Wach, D., & Kos, M. 2018. *Financial risk information avoidance*. Economic Research, 31(1), 521–536.
- [6] Lucey, B. M., & Dowling, M. 2005. *The role of feelings in investor decision-making*. Journal of Economic Surveys, 19(2), 211–237.
- [7] Baker, Malcolm dan Jeffrey Wurgler. 2006. *Investor Sentiment dan the Cross-Section of Stock Returns*. Journal of Finance 61: 1645–1680.
- [8] Tetlock, P. C. 2007. Giving content to investor sentiment: The role of media in the stock market. The Journal of Finance, 62, 1139e1168.
- [9] Da, Zhi, Joseph Engelberg dan Pengjie Gao. 2014. *The Sum of All FEARS Investor Sentiment and Asset Price*. The Review of Financial Studies, Volume 28, Issue 1, January 2015, Pages 1–32.
- [10] Antweiler, W. dan Frank, M.Z. 2004. Is all that talk just noise? The information content of internet stock message boards. Journal of Finance, Vol. 59, pp. 1259–94.
- [11] Bollen, J., Mao, H. dan Zeng, X-J. 2011. *Twitter mood predicts the stock market*. Journal of Computational Science, Vol. 2, pp. 1–8.
- [12] Nofsinger, R. J. 2005. *Social Mood dan Financial Economics*. The Journal of Behavioral Finance. 6(3):144-160.
- [13] Ivković Zoran, Scott Weisbenner. 2005. Local Does as Local Is: Information Content of the Geography of Individual Investors' Common Stock Investments. Journal of Finance 60: 267–306.
- [14] Petit, Juan Jose García, Esther Vaquero Lafuente, Antonio Ru'a Vieites. 2019. *How information technologies shape investor sentiment: A web-based investor sentiment index*. Borsa _Istanbul Review 19-2 (2019) 95e105.
- [15] Sahana, T.P. & J. Anuradha. 2019. *Analysis and Prediction of Stock Market Using Twitter Sentiment and DNN*. Advance in Intelligent Systems and Computing.
- [16] Zhang, X., Fuehres, H., & Gloor, P. A. 2011. *Predicting stock market indicators through twitter "I hope it is not as bad as I fear*. Procedia—Social and Behavioral Sciences, 26, 55–62.
- [17] Makrehchi, M., Shah, S., & Liao, W. 2013. *Stock prediction using event-based sentiment analysis*. In Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013 IEEE/WIC/ACM International Joint Conferences (Vol. 1, pp. 337–342).
- [18] Sprenger, O. T., Tumasjan, A., Sdanner, G. P. dan Welpe M. I. 2014. *Tweets and trades: the information content of stock microblogs*. European Financial Management. 20(5): 926-957.
- [19] Pagolu, V.S., Reddy, K.N., Panda, G., Majhi, B. 2016. *Sentiment analysis of Twitter data for predicting stock market movements*. International Conference on Signal Processing, Communication, Power and Embedded System (SCOPEs), Paralakhemundi, pp. 1345– 1350.
- [20] Sul, H. K., Dennis, A. R., & Yuan, L. 2016. *Trading on Twitter: Using social media sentiment to predict stock returns*. Decision Sciences, 48(3), 454–488.
- [21] Roig-Tierno, N., Gonzalez-Cruz, T. F., & Llopis-Martinez, J. 2017. *An overview of qualitative comparative analysis: A bibliometric analysis*. Journal of Innovation and Knowledge, 2(1), 15–23.
- [22] Sabherwal, S. 2011. Do internet stock message boards influence trading? Evidence from heavily discussed stocks with no fundamental news. Journal of Business Finance and Accounting, 389, 1209-1237, ISSN 0306-686X.
- [23] Klein, A., Altuntas, O., Häusser, T., Kessler, W. 2011. *Extracting investor sentiment from weblog texts: A knowledge-based approach*. Proceedings - 13th IEEE International Conference on Commerce and Enterprise Computing, CEC 2011, Article number 6046948, Pages 1-9.
- [24] Luo, B. 2016. *Emotion space model for classifying opinions in stock message board*. Expert Systems with Applications, 44, 138-146, ISSN 0957-4174.

- [25] Zhou, Z. 2016. *Can online emotions predict the stock market in China?*. Lecture Notes in Computer Science including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, 10041, 328-342, Allen, D. 2015. Machine News and Volatility: The Dow Jones Industrial Average and the TRNA Real-Time High-Frequency Sentiment Series. Handbook of High-Frequency Trading, 327-344.
- [26] Gu, B. 2014. The allure of homophily in social media: Evidence from investor responses on virtual communities. *Information Systems Research*, 253, 604-617.
- [27] Park, J. 2013. Information valuation and confirmation bias in virtual communities: Evidence from stock message boards. *Information Systems Research*, 244, 1050-1067.
- [28] Li, X. 2017. Daily happiness and stock returns: The case of Chinese company listed in the United States. *Economic Modelling*, 64, 496-501, ISSN 0264-9993.
- [29] Zhang, X. 2018. Improving stock market prediction via heterogeneous information fusion. *Knowledge-Based Systems*, 143, 236-247, ISSN 0950-7051.
- [30] Zhao, R. 2020. Quantifying the cross sectional relation of daily happiness sentiment and stock return: Evidence from US. *Physica A: Statistical Mechanics and its Applications*, 538.
- [31] Bar-Haim, R. 2011. *Identifying and following expert investors in stock microblogs*. EMNLP 2011 - Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, 1310-1319.
- [32] Zhang, Y. 2010. *Are day traders bias free?-evidence from internet stock message boards*. *Journal of Economics and Finance*, 341, 96-112, ISSN 1055-0925.
- [33] Joseph, K. 2011. Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting*, 274, 1116-1127, ISSN 0169-2070.
- [34] Takeda, F. 2014. Google search intensity and its relationship with returns and trading volume of Japanese stocks. *Pacific Basin Finance Journal*, 271, 1-18, ISSN 0927-538X.
- [35] Renault, T. 2017. Intraday online investor sentiment and return patterns in the U.S. stock market. *Journal of Banking and Finance*, 84, 25-40.
- [36] Rechenhain, M. 2013. *Stock chatter: Using stock sentiment to predict price direction*. *Algorithmic Finance*, 23, 169-196, ISSN 2158-5571.
- [37] Coyne, S. 2018. *Forecasting stock prices using social media analysis*. Proceedings - 2017 IEEE 15th International Conference on Dependable, Autonomic and Secure Computing.
- [38] Nooijen, S.J. 2016. Predicting Equity Markets with Digital Online Media Sentiment: Evidence from Markov-switching Models. *Journal of Behavioral Finance*, 174, 321-335, ISSN 1542-7560.
- [39] Yao, C. 2017. A study of correlation between investor sentiment and stock market based on Copula model. *Kybernetes*, 463, 550-571, ISSN 0368-492X.
- [40] Xie, C. 2017. Does Online Investor Sentiment Affect the Asset Price Movement? Evidence from the Chinese Stock Market. *Mathematical Problems in Engineering*, 2017, ISSN 1024-123X.
- [41] Coelho, J. 2019. *Social media and forecasting stock price change*. Proceedings - International Computer Software and Applications Conference, 2, 195-200.
- [42] Chen, M. 2019. Modeling public mood and emotion: Blog and news sentiment and socio-economic phenomena. *Future Generation Computer Systems*, 96, 692-699.
- [43] Liu, J. 2013. *Looking for gold in the sands: Stock prediction using financial news and social media*. Proceedings - Pacific Asia Conference on Information Systems, PACIS 2013.
- [44] Allen, D. 2019. *Daily Market News Sentiment and Stock Prices*. *Applied Economics*, 5130, 3212-3235.
- [45] Shah, D. 2019. *Predicting the Effects of News Sentiments on the Stock Market*. Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018, 4705-4708.
- [46] Shi, Y. 2018. *A text mining based study of investor sentiment .and its influence on stock returns*. *Economic Computation and Economic Cybernetics Studies and Research*, 521, 183-199, ISSN 0424-267X.
- [47] Huang, J.Y. 2020. Using social media mining technology to improve stock price forecast accuracy. *Journal of Forecasting*, 391, 104-116, ISSN 0277-6693.
- [48] Nasser, A. Al 2014. *Big data analysis of stocktwits to predict sentiments in the stock market*. Lecture Notes in Computer Science including subseries Lecture Notes in Artificial Intelligence and Lecture

Notes in Bioinformatics, 8777, 13-24, ISSN 0302-9743.

- [49] Carvalho, A. 2019. Off-the-shelf artificial intelligence technologies for sentiment and emotion analysis: A tutorial on using IBM natural language processing. *Communications of the Association for Information Systems*, 441, 918-943, ISSN 1529-3181.
- [50] Bourezk, H. 2020. *Analyzing Moroccan Stock Market using Machine Learning and Sentiment Analysis*. 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology.
- [51] Kovacs, A., Van Looy, B., & Cassiman, B. 2015. Exploring the scope of open innovation: A bibliometric review of a decade of research. *Scientometrics*, 104(3), 951–983.
- [52] Boyack, K. W., & Klavans, R. 2010. Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology*, 61(12), 2389–2404.
- [53] Shyy G. 1989. Bullish or bearish: A Bayesian dichotomous model to forecast turning points in the foreign exchange market. *Journal of Economics and Business*, 41 (1) , pp. 49-60.
- [54] Smith, Adam. 1759. *The Theory of Moral Sentiments*.