Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/120921

Sentiment Analysis and Classifying Hashtags in Social Media Using Data Mining Techniques

G. Kassem[1], E. Asfoura[2,], B. Alhuthaifi[2], J. J. C. Gallego[3] and F. Balhaj[2]*

[1]Department of Business Information System, College Business, German University in Cairo, Cairo, Egypt
[2]Department of System, Faculty of Business Studies, Arab Open University, Riyadh, Kingdom of Saudi Arabia
[3]Department of Software Engineering, Faculty of Computer Science, Alcala Madrid University, Madrid, Spain

Received: 17 Feb. 2023, Revised: 30 Jun. 2023, Accepted: 07 Jul. 2023.
Published online: 1 Sep. 2023.

Abstract: Big data is one of the important topics which is still open for a wide range of applications for extracting useful information and knowledge for supporting organizations by planning and decision-making. Social media as a technology is an important resource of data, especially because it has been widely used in the last years. A Hashtag is recently one of the most popular features provided by Social media and is used by most social media users to express, share, and retrieve opinions and feelings regarding a specific theme. Hashtag features in social media are used more and more in recent years to discuss and debate important current events by public audience. This paper sheds light on how business can use such sources of information and how needed technical processes can be implemented accordingly. The paper demonstrates sentiment analysis as a scenario for such implementation. The main innovation in this paper is not limited to the technical method used, but rather to focus on the idea of using hashtags as information source in business, which is still rarely addressed in science. This paper will provide a novel model based on text mining techniques to provide a sentiment analysis for classifying business-related Hashtags posted on social media from the customers. The results will be presented and verified through samples of positive, and negative classified comments extracted from the Hashtags for supporting the organization by planning and decision making for generating completive advantages.

Keywords: Social Media, Hashtag, Data Mining, sentiment analysis, Decision Making.

1 Introduction

In the last few years, there was an increased focus on social media as an important source of data that can be analyzed for many purposes because of the huge amount of data that is generated through interaction between people and posting information daily. Many researchers published articles about social media from different aspects like big social media as an important source of big data for analytical purposes [1-2-3], the related privacy issues [4-5], recent achievements and new challenges [6-7] representativeness, validity and the quality of extracted data from social media [8-9] in addition to many another aspect. A type of the expressed and posted data daily through social media is Hashtags, which have not been considered significantly especially in business despite their relevancy.

Hashtags have become a part of pop culture in the last few years. People see it as a way to express the thought of the moment to post a hashtag on social media. In addition, they use hashtags to recognize the content or the theme of a group of messages.

Users describe their sentiments in a social media text. Therefore, both public and private sectors can gain benefits from analyzing their stakeholders' attitudes toward their products or services. For example, companies can examine the level of satisfaction of their customers, governments can recognize their public emotions toward an event or decision.

Hashtags might occur in a single word like #angry or might contain multiple words like #cantwait4tmrw. At the same time, hashtags like #angryattheworld, and #happyasalways might describe an emotional state. Therefore, it will not be easy for organizations to manually analyze hashtags in social media that match their interests.

The objective of this paper is to analyze and evaluate hashtags to improve decision-making in organizations. Sometimes hashtags are not clear from the customer's opinion; for example, the hashtag (#good); for businesses they cannot classify exactly the customer's opinion through hashtags whether it is a positive, negative, or neutral hashtag. For instance, as a use case, those hashtags can be classified by using unsupervised data mining techniques by comparing the car companies through social media.

*Corresponding author e-mail: e.asfoura@arabou.edu.sa

Next, we will provide an overview of different aspects of social media such as platforms, analysis, types, and gained benefits.

2 Background

In the last years, the popularity of social media technologies has increased and has been growing rapidly, changing how people interact, communicate and even conduct business [10-11]. Web 2.0 is the foundation of social media applications which permits the exchange of contents generated by their users [12-1]. The most common social media categories are blogs, microblogs, opinions, photo and video sharing, social bookmarking, instant messaging, social networking sites, social news, and wikis [14]. IT enables users to create public or semi-public profiles, generate and connect to several other users forming virtual communities, and visualize and go through the connections of users [15-16]. Businesses can use social media to allow customers to build up loyalty toward their brands. By creating special fan pages, customers can visit the sites and the brands all day. This is crucial and essential to any company because it is now the easiest way to give out information to customers and collect personal data about customers [17-18]. Data obtained from social media is particularly beneficial for mapping relationships that include connections as well as interests [19]. On social media platforms, one of the most unique tools available is the hashtag. Little research exists on how social work and advocacy organizations use hashtags, much less on how such use can be effective [20].

Users of Social media platforms such as Twitter and Facebook are rapidly sharing their news and ideas with otsers [21-22]. Twitter has become one of the largest and most popular microblogging websites. To deal with the volume of information shared daily, Twitter has introduced hashtags. Other social media platforms like Facebook and Instagram copied the idea of hashtags invented by Twitter and it are largely and regularly employed by millions of social media users [23]. Facebook comes first among all social media platforms with more than 800 million users. Next comes Twitter with 280 million users. Facebook grew to be the second most opened internet website every month, about 152 million people from the Americans went online in only May 2012, which was estimated to be about 72%. Every Facebook user spends at least 7 hours every month on Facebook, and about 70.15% of the United States had a Facebook personal account by October 2012. Facebook is also used as a reference for some people to search for information [24]. In social media, models can be built to predict future trends by understanding targeted customers' behavior. in addition, companies can utilize the gathered information from social media platforms to plan their marketing activities [25].

Social media analysis is basically mining data in social media in order to discover concealed information and knowledge that could be useful for businesses, users, and consumers, due to the effortlessly reachable platform of users who share information that social media offers [26]. Data mining is recognizing data patterns, which assist people to understand huge sets of data. Data mining is correlated to machine learning, information retrieval, statistics, databases, and even data visualization as well as it has techniques, the most common are classification, clustering, association rules, text mining, link analysis, and multi-relational data mining. In addition, social media analysis is considered a type of social computing, which is a computing application where the software servers play a role in improving social relations [27-28].

The main challenges that are faced while mining social media are the enormous data that is generated by the users on social media sites, which is noisy, distributed, unstructured, and dynamic and the privacy concerns due to the sensitivity of the information on the social relationship during the information acquisition process for the data mining. Another challenge is identifying implicit relationships in a social network, which is the relationship between users who have common interests or activities and interact directly with each other, thus normally this relationship could not be tracked by public records. While determining implicit relationships offers important information in marketing in general as well as for suggesting products/services to users who are likely to be interested in them, which conquers the broad links due to identifying only the explicit relationships in a social network.

Traditional analysis is analyzing the structural properties of an online network by determining values of actors and relations calculations, like degree, in-degree, and out-degree, as well as identifying patterns in a dataset that has a set of independent instances of single relation, by applying data mining techniques like rule mining, market basket analysis, and cluster analysis. In other words, it is creating a model out of a haphazard sample from a common fundamental distribution [29].

Link mining analysis

Link mining is a recently evolving research area that emerged long after the traditional social network analysis tasks as they are depending on new advances in data mining techniques. It is an instance of multi-relational data, but the term link mining is used in order to add more focus on the links themselves. Link mining could be seen as data mining on social networks, in which the conventional classification and clustering techniques of data mining are regenerated in a way in order to work with network data**.** Link mining techniques are either classified as predictive - which is predicting the values of a specific attribute according to the values of other attributes - or descriptive – which is identifying patterns that sum

up the underlying relationships in data [30].

Two main tasks in data mining. One is descriptive where the user can describe the data relationships and determine their patterns. The second is predictive where the user can predict unknown values based on previously collected data.

Summarization, clustering, sequential patterns, and association learning are considered as descriptive tasks. Summarization is the practice of describing a subset of data in a simple way. Clustering is the practice of grouping things into their categories. Sequential patterns help in finding the root of a set of sequences. Association learning is the practice of identifying relationships between data by the use of if/then statements.

Whereas predictive tasks include regression, classification, and deviation detection. Regression is used to predict a value by mapping the data to the prediction variable. Classification is the practice of grouping data objects that share similar attributes. Deviation detection is the practice of detecting variation from usual behavior [31].

Visual analysis

Visual analysis is common in social network analysis. It is mainly about visualizing information in social networks. In fact, visual analysis allows social network analysis experts to easily obtain conclusions about social networks that would have stayed hidden even after applying other analysis techniques; for that reason, visual analysis is crucial to social network analysis [29].

Social media analysis helps in improving the understanding of people's insights and points of view on a certain matter, spotting out implicit or hidden groups of people from between loads of population, recognizing the long term dynamic alters of the group, and locates influential people as well as suggests a product or activity to user systems in order to sustain the present customers, obtaining more new customers and maintaining high trust between customers themselves or between customers and stakeholders.

Hashtags

Hashtags contain the # symbol followed by single or multiple words without spaces or punctuation. They are grouping users' posts for the same topic. Such hashtags could be used to search for a specific topic. Simply by clicking on the Hashtag, you will reach all the posts tagging it on a single page.

It is not necessary to have the hashtag at the beginning or the end of your microblog, it could occur anywhere. With the wide use of social media, microblogs have become an essential source of data to predict public opinions. According to [32] multiple approaches for diverse applications have also confirmed the usefulness of hashtags. for example, they are useful in sentiment analysis, query expansion, and microblog retrieval.

Hashtags are everywhere in the social networking world. Users use hashtags for multiple reasons, many of these reasons can be considered as labels for classical NLP tasks like sentiment (#dislike), subject annotation (#yoga), or named entities identification (#sf49ers) [33].

A new way of public communication and interaction has occurred because of social media invention. One of the most important tools that cause such changes is the hashtag. In 2007, the first post with a hashtag was created by a Twitter employee called Chris Messina. Since then, hashtags have become widely used not only on Twitter platforms but also on other social media platforms. Since hashtags group related posts and make it easier to reach these posts, users can share their ideas and opinions with the community using these hashtags [20].

On the Twitter platform, the use of the symbol (#) means it is a hashtag and this hashtag is clickable. Once it is clicked, you will navigate to a new page that displays all posts using the same hashtag. This will improve the reachability of the interesting topics. And also, will allow participation, interaction, and sharing of ideas. This was very difficult before the invention of hashtags. Moreover, on the Twitter platform, there is no limit to the users you can follow. Whatever those followed users post, it will appear on their Twitter pages, and you can see it easily. However, in order to see the posts of Twitter users that you don't follow, hashtags are the best way.

3 Methodology

This paper will provide a sentiment analysis and classification model based on data mining techniques for extracting business-related Hashtags which will be useful for supporting better planning and decision making for developing the business. Designing the proposed model will be achieved through sequential steps from determining the business-related Hashtags to conceptualizing the proposed model as shown in figure 1.

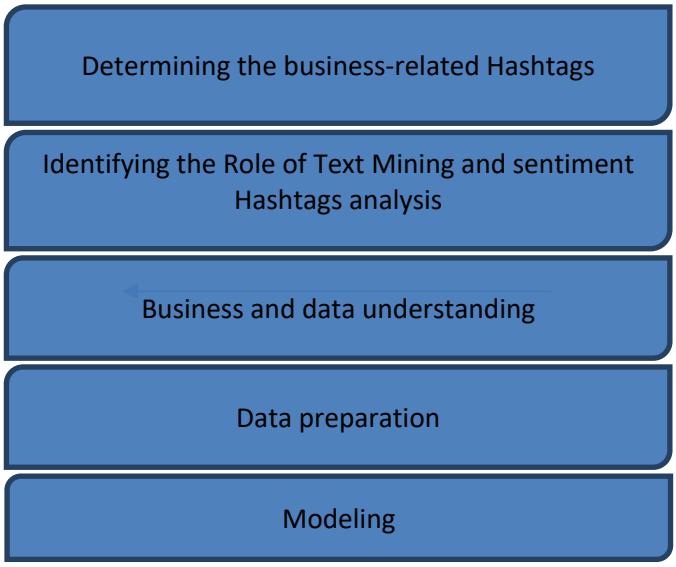


Fig. 1: required steps for the proposed model for Sentiment Analysis and classifying Hashtags

3.1 Types of Hashtags related to Business

The use of hashtags within social media creates a unique opportunity for businesses to connect with their fans and implement new branding strategies. Businesses can use hashtags to increase their social reach and improve brand awareness by seeking out popular hashtags related to their services or relevant to their target audience. Most social networks include a list of trending hashtags that allow users to see the latest news and popular topics in real time. Businesses can use these trending hashtags to their advantage. There are 4 main types of hashtags that business use to reach more customers in social media, which are branding hashtags, campaign hashtags, product hashtags, and event hashtags.

Branding hashtags

The goal of branding is to create an emotional connection between a company and its consumers. Branding results from the sum of many different parts, including the brand name, logo, colors, and more. It allows companies to differentiate themselves in an increasingly crowded market. The rise of new technology is changing how companies communicate with consumers and in turn how they use branding.

For example, KitKat uses a brand hashtag. They use their tagline: #HaveABreak. They consistently use it on all their social sites from Twitter to Google+. Their consumers know their brand hashtag, and they use it to engage with the social KitKat community.

Campaign hashtags

Campaign hashtags are created specifically for a particular marketing campaign, intended to reinforce campaigns and brands. Campaign hashtags are often used to promote but are also occasionally a way for consumers to unlock exclusive discounts or promotions just for being on Twitter. For example, Home Depot is using the hashtag #HDgameday. They are running a photo contest (using Wishpond's Facebook Contest Apps) of bucket drumming, to celebrate the start of college football. The campaign is being run through both Twitter and Instagram. They're also promoting photo winners on Facebook and their website.

Product hashtags

A product hashtag is to build up a product's individual brand underneath the umbrella of the brand. People search for common product tags. They use them too. When a business posts their products, they think like their customers. Using hashtags that both connect their product and their market. For example, coffee shops post the best images of their lattes using the #latte hashtag.

Event hashtags

An event could be anything from a local community fair, to a well-known global celebration, to a live product launch or a live webinar you're hosting online. Williams Sonoma tells their Facebook Fans about a local charity event they're attending. They use the event hashtag #FeastPDX to support the cause and promote it with partners through social sites.

3.2 Role of Text Mining and sentiment Hashtags analysis

Text Mining has become a very popular and unique topic since text mining applications are used in opinion mining and sentiment analysis and explore how it can be beneficial in analyzing customers' comments and complaints. However, to our knowledge, too little research focused on the customer's sentiment hashtags and how the hashtags are classified and analyzed. The purpose of this study is to analyze the research gap "Analyzing and classifying sentiment analysis hashtags from social media."

The text mining process starts with a document collection from various resources. A text mining tool would retrieve a particular document and pre-process it by checking the format and character sets. Then the document would go through a text analysis phase. Text analysis is semantic analysis to derive high quality information from text. Many text analysis techniques are available; depending on the goal of the organization combinations of techniques could be used. Sometimes text analysis techniques are repeated until the information is extracted. The resulting information can be placed in a management information system, yielding an abundant amount of knowledge for the user of that system.

Use Case for Sentiment Analysis and classifying Hashtags in Social Media using Text Mining Process.

In this use case, Hashtags were collected from different social media sources like Facebook, Twitter, and Instagram, the dataset of hashtags is related to the reviews of customers about Automobile companies. Thereby, the analytics tools of RapidMiner1 and Waston IBM2 are used to analyze and classify the sentiment analysis hashtags. The CRISP-DM (Cross Industry Standard Process for Data Mining) outlines a process model by providing a framework for supporting data mining projects. Also, it helps in resolving issues that are faced during data mining use case developments and helps in assuring the quality of the results. Even though CRISP-DM is designed mainly for data mining projects we have to customize the sequential phases of the CRISP-DM as follows:

3.3 Business and data understanding

Business understanding

In this phase, it is essential to concentrate on understanding the objective of the use case and its requirements. In the use case, the objective is to find how to use text mining in order to classify and analyze sentiment hashtags. Customer complaints and comments were analyzed and classified as texts only, also classifying and analyzing hashtags is important in the social media platforms as people write their reviews or complaints using the hashtags. Since the hashtags are in the form of text; text mining analysis can be used to assist in this process. Using rapid miner, we needed to create a model to classify the sentiment hashtags instead of trying to analyze every hashtag alone and trying to classify if it is positive, negative, or neutral.

Data understanding

The data understanding starts with the initial data collection of hashtags. 50 hashtags were collected from car organizations on social media platforms. The dataset was divided into three parts which are the company, car model, and sentiment (the hashtags) as shown in figure 2.

ExampleSet (49 examples, 0 special attributes, 3 regular attributes)

Row No.	Company	Car Model	sentiment
1	Nissan	Z350	power
2	Nissan	Z350	badexperince
3	Nissan	Z350	cool
4	Nissan	Z350	notgood
5	Nissan	versa	cheapprice
6	Nissan	versa	awful
7	Nissan	versa	comfortable
8	Nissan	versa	horrible
9	Kia	Rio	awesome
10	Kia	Rio	amazing
11	Kia	Rio	poordesign
12	kia	sportage	expensive
13	Kia	sportage	badexperince
14	Renault	stepway	uniquefeatures
15	Renault	stepway	poordesign
16	BMW	4 series conv...	convertiblepo...
17	BMW	4 series conv...	terrible

Fig. 2: Sample of the dataset

[1] https://rapidminer.com/

[2] https://www.ibm.com/watson/

3.4 Data preparation

The data preparation/pre-processing phase covers all tasks required to reach the final dataset (data that will be fed into a rapid miner for training the model) from the raw dataset. To analyze the data, the Rostte Text ToolKit is used, it is an easy tool to classify and analyze the dataset. The column sentiment was rolled as a label as shown in figure 3.

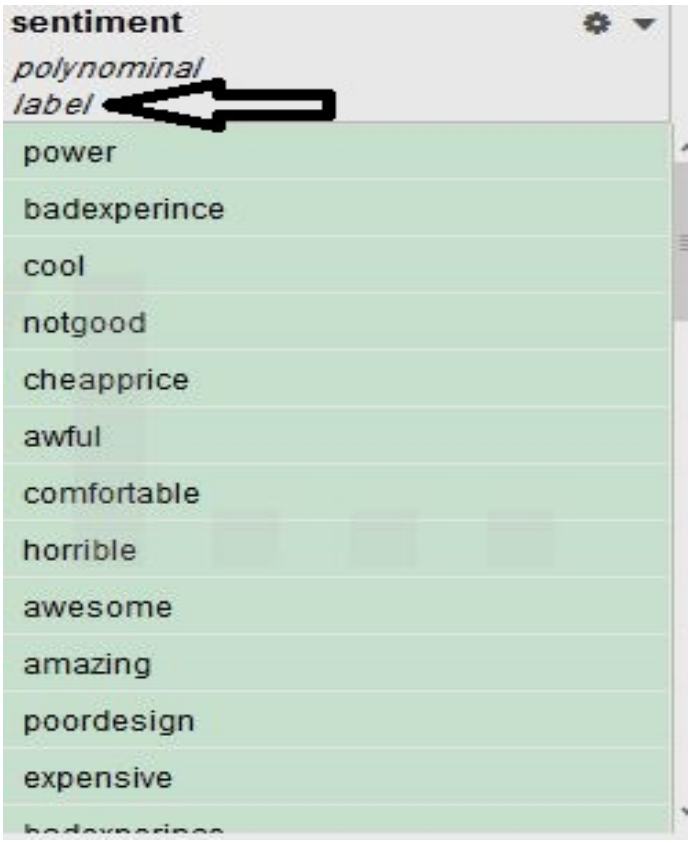


Fig. 3: Sentiment column labeled

Two tools are applied to analyze the hashtags which are Aylien Text Analysis and Rostte Text ToolKit.

Alyien Text Analysis tool is used to analyze sentiment hashtags.

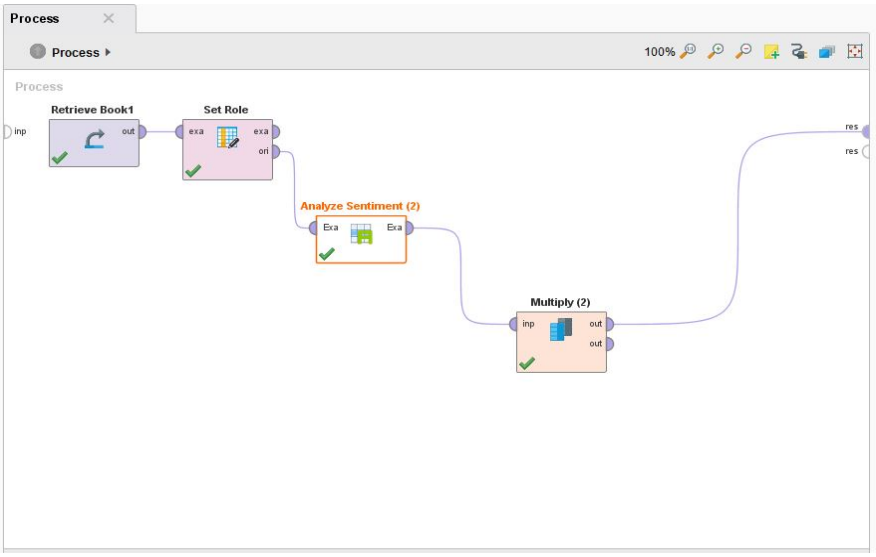


Fig. 4: Analyzing sentiment

In figure 4, first, the data is processed to make sure that it's running without any errors, and then it is connected to the set

role operator to make sure that the column should be classified is labeled. The Alyien text analysis tool had an operator called analyze sentiment. The input attribute as shown in figure 4 was column sentiment as that the column should be classified as this column contains the hashtags that would be classified.

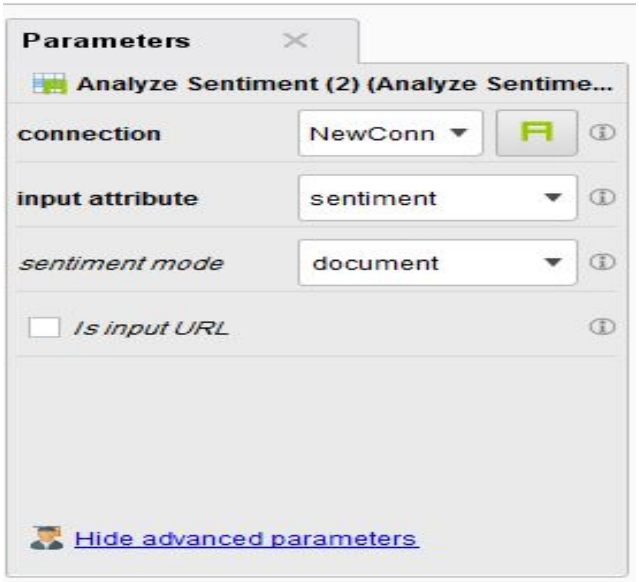


Fig. 5: Input Attribute (sentiment)

Finally, the multiply operator is connected to copy the objects at its input port to the output ports multiple number of times. As more ports are connected, more copies are generated.

3.5 Modeling

During this phase, various modeling techniques are selected and applied to choose the best model. The Rosette Text Toolkit as seen in figure 6 the final dataset is going to be used to classify the model. It is read by rapid miner where we set the label attribute (sentiment) to represent the target class for classification using a set role operator. Then it is connected to the analyze sentiment operator which analyzes the sentiment (subjective attitude) of the input as either positive or negative. Before analyzing the text, Rosette filters out some stop words and punctuation, such as “the” “?”, “a” and “it”, to increase the accuracy of the assessment. The output contains the Sentiment attribute populated with the collected data.

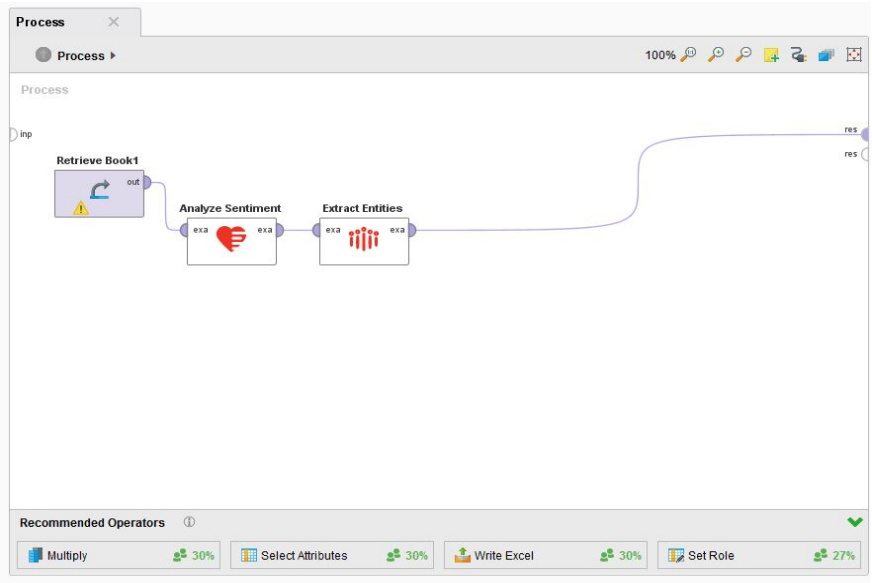


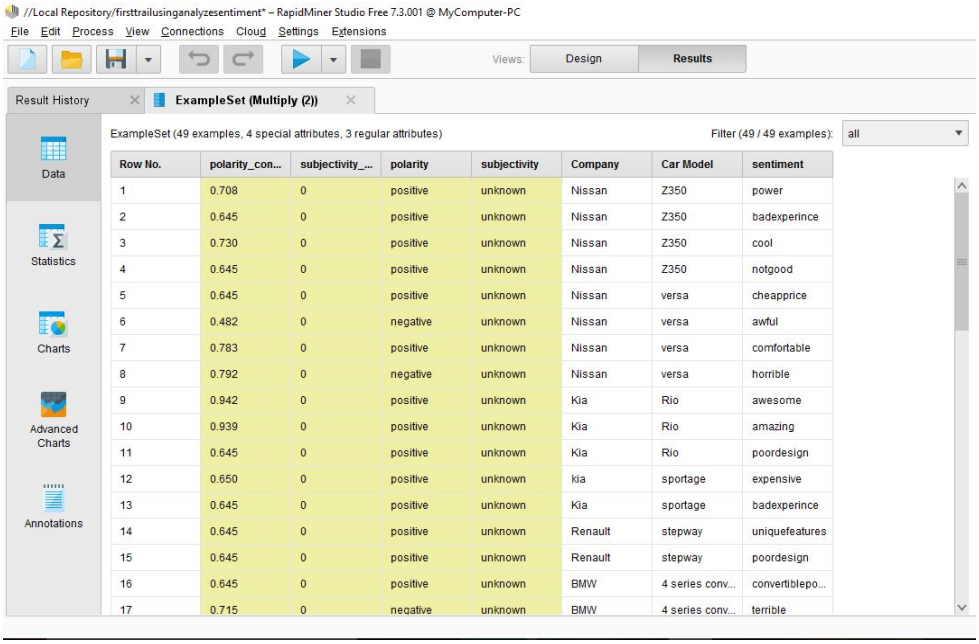
Fig. 6: Modeling

Finally, the analyze sentiment operator is connected to the extract entities operator refers to an object of interest such as

a person, organization, location, or email address. Identifying entities can help you classify the text and determine what kinds of data it contains.

4 Deployment: Results

After creating the model using the Rosette Text Toolkit the results were classified accurately whether the hashtag is positive, negative, or neutral. While the results using the Aylien Text Analysis are as shown in figure 7 the polarity confidence was calculated and also the data was classified as whether the hashtags are positive, negative or neutral.



Row No.	polarity_con...	subjectivity_...	polarity	subjectivity	Company	Car Model	sentiment
1	0.708	0	positive	unknown	Nissan	Z350	power
2	0.645	0	positive	unknown	Nissan	Z350	badexperince
3	0.730	0	positive	unknown	Nissan	Z350	cool
4	0.645	0	positive	unknown	Nissan	Z350	notgood
5	0.645	0	positive	unknown	Nissan	versa	cheapprice
6	0.482	0	negative	unknown	Nissan	versa	awful
7	0.783	0	positive	unknown	Nissan	versa	comfortable
8	0.792	0	negative	unknown	Nissan	versa	horrible
9	0.942	0	positive	unknown	Kia	Rio	awesome
10	0.939	0	positive	unknown	Kia	Rio	amazing
11	0.645	0	positive	unknown	Kia	Rio	poordesign
12	0.650	0	positive	unknown	kia	sportage	expensive
13	0.645	0	positive	unknown	Kia	sportage	badexperince
14	0.645	0	positive	unknown	Renault	stepway	uniquefeatures
15	0.645	0	positive	unknown	Renault	stepway	poordesign
16	0.645	0	positive	unknown	BMW	4 series conv...	convertiblepo...
17	0.715	0	negative	unknown	BMW	4 series conv...	terrible

Fig. 7: Results using Alyien tool

After analyzing those sentiment hashtags, it was easy to classify them by using two software which are IBM Waston analytics and RapidMiner. They are different tools and each of them analyzes the hashtags in different ways. The IBM Waston analytics suggested questions or what the users want to predict or analyze; the results were classified by counting the negative hashtags and positive hashtags separately. While in RapidMiner software two tools were used which are Rosette Text Toolkit which classified the data into three sentiments which are positive, negative, and neutral. the data was classified accurately. In the other tool, which is Alyien analysis tool, the confidence was calculated, and the data was classified whether it is positive, negative, or neutral.

5 Discussions

The summary of the related work to sentiment analysis and classification for data extracted from social media in literature will be presented inf table 1 for an explanation of the cross different aspects of the proposed subject of this paper.

Table 1: Summary of the related work to sentiment analysis and classification

Reference	Type of processed data	Used techniques	the objective
[37]	Tweets (terrorism-related content)	Deep learning techniques	Classification of tweets into extremist and non-extremist
[34]	Feedbacks and comments posted on social media	Fuzzy system	Processing and Integration of non-English languages for better understanding
[38]	Data about Costumer reactions on Posted (photos and videos) in social media from companies	various statistical tools and the sentiment analysis technique applied to big sets of data	Understanding of customer preferences related different social media posts
[35]	Facebook and twitter comments about products (drugs and cosmetics)	machine learning, text mining and sentiment analysis	Developing product safety lexicon

[36]	tweets for the #Startups hashtag on Twitter	Latent Dirichlet Allocation (LDA) model which is modeling tool works in Python. And text mining techniques	Identifying the success factors for startup business
Current work	Automobile companies' hashtags on Facebook, Twitter, and Instagram	Sentiment analysis & classification by utilizing RapidMiner3, Waston IBM4, and CRISP-DM Text mining by utilizing Aylien Text Analysis and Rostte Text ToolKit	Develop a model that support decision making

The main contribution of this work is using sentiment analysis for Hashtags because they are important for many business sectors that need real-time information of current events like stocks trade markets, supply chain, marketing or even public authorities, Hashtag has not yet been sufficiently exhausted for business and has received less academic attention than it deserves. This article is intended to emphasize the importance of hashtags and technical implementation to use them for companies.

6 Conclusion

Hashtags are a great way for businesses to encourage interaction with their customers through social media. The hashtags that customers commented on through social media were very huge. The companies were not able to analyze those hashtags but by using the data mining techniques and tools that became much easier for the companies to analyze the data. The big challenge of social media analysis is the huge amount of data that has to be processed, especially if a company wants to scan all comments shared to the entire social media and related to its products. Although Big Data technology can solve this problem (all or partial), many SMEs don't have the resources to deploy it. Hashtags can be considered per definition as pre-filtering of data by grouping all comments related to a specific topic. Furthermore, Hashtags indicate the recent importance of specific topics and show new trends that could be helpful for companies to respond to customer needs and complaints, or to consider for planning of marketing strategies. The last aspect that characterizes Hashtags is the addressed group themselves who are engaged and active in specific Hashtags. These Groups can be directly addressed for marketing or crowd mining for example. In this paper, a simple use case is presented to demonstrate the feasibility of sentiment analysis and classifying hashtags in social media using data mining techniques and how it can be beneficial for business. Future work may develop a framework for hashtag-based analytics (business objectives, data mining methods, integrations, extracting processes, etc.) to develop generic hashtags analytics landscape and roadmaps, including the sentiment analysis presented in this paper.

Acknowledgment

The authors extend their appreciation to the Arab Open University for funding this work through AOU research fund No. (AOURG-2023-013).

Conflict of interest

The authors declare that there is no conflict regarding the publication of this paper.

References

[1] Gole, S. (2015). A survey of big data in social media using data mining, in proceeding of IEEE 2015 International Conference on Advanced Computing and Communication Systems, Coimbatore, India

[2] Abdul Ghan, N. Hamid Ibrahi m, S. Targio Hashem, A & Ahmedc, E (2019), Social media big data analytics: A survey, Computers in Human Behavior, vol 101, pp 417-428.

[3] Abkenar, S. B., Kashani, M. H., Mahdipour, E., & Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. Telematics and informatics, 57, 101517.

[4] Smith. M., Szongott. C, Hanne. B and& von Voigt. G (2012), Big data privacy issues in public social media, in

[3] https://rapidminer.com/
[4] https://www.ibm.com/watson/

proceeding of 2012 6th IEEE International Conference on Digital Ecosystems and Technologies (DEST), 18-20 June 201, Campione d'Italia, Italy.

[5] Sriram, G. S., & Sriram, G. S. (2022). Security challenges of big data computing. International Research Journal of Modernization in Engineering Technology and Science, 4(1), 1164-1171.

[6] Bello-Orgaz, G. Jason J. Jung and David Camachoa (2016), Social big data: Recent achievements and new challenges, Information Fusion, Vol 28, pp 45-59.

[7] Abkenar, S. B., Kashani, M. H., Mahdipour, E., & Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. Telematics and informatics, 57, 101517.

[8] Immonen, A., Paakkonen, p. & Ovaska, E. (2015). Evaluating the Quality of Social Media Data in Big Data Architecture, IEE Access, Vol 3, pp 2028 – 2043.

[9] Tufekci, Z. (2014). Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls, in Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, ne 1-4, 2014 in Ann Arbor, Michigan USA.

[10] Venkatraman, S. S. (2010), Social Networking Technology as a Busines7-8 Jan 2015, s Tool: Vol. 33, No. 3, pp. 476-498

[11] Nicolaou, C. (2021). Development of business through the internet and social media: The professional use of audiovisual media technologies through strategic tactics and practices. In Handbook of research on IoT, digital transformation, and the future of global marketing (pp. 193-211). IGI Global.

[12] Kaplan, A. M.; Haenlein, M. Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons, v. 53, p. 59-68, 2010.

[13] Thao, T., & Shurong, T. (2020). Is it possible for" electronic word-of-mouth" and" user-generated content" to be used interchangeably. Journal of Marketing and Consumer Research, 65, 41-48.

[14] Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P., & Bhattacharjee, B. (2007, October). Measurement and analysis of online social networks. In Proceedings of the 7th ACM SIGCOMM conference on Internet measurement (pp. 29-42).

[15] Pinho, J. C., Soares, A. M., & Nobre, H. (2012). From social to marketing interactions: the role of social networks. Journal of Transnational Management , 17 (1), 45-62.

[16] Cheung, M. L., Pires, G. D., & Rosenberger III, P. J. (2019). Developing a conceptual model for examining social media marketing effects on brand awareness and brand image. International Journal of Economics and Business Research, 17(3), 243-261.

[17] Ang, L. (2011). Community relationship management and social media. Journal Of Database Marketing & Customer Strategy Management, 18(1), 31-38.

[18] Wongsansukcharoen, J. (2022). Effect of community relationship management, relationship marketing orientation, customer engagement, and brand trust on brand loyalty: The case of a commercial bank in Thailand. Journal of Retailing and Consumer Services, 64, 102826.

[19] Cooper, G. (2012). Data Mining and Social Media. 1-31. Retrieved from http://blog.geoffcooper.webfactional.com/sites/default/files/gcooper_datamining_04192012.pdf

[20] Saxton, G. D., Niyirora, J. N., Guo, C., & Waters, R. D. (2015). # AdvocatingForChange: The Strategic Use of Hashtags in Social Media Advocacy. Advances in Social Work, 16(1), 154-169.

[21] Oeldorf-Hirsch, A., Sundar, S. S. (2015). Posting, commenting, and tagging: Effects of sharing news stories on Facebook. Computers in Human Behavior, 44, 240-249.

[22] Kitchens, B., Johnson, S. L., & Gray, P. (2020). Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption. MIS quarterly, 44(4).

[23] Caleffi, P. M. (2015). The 'hashtag': A new word or a new rule. Skase Journal of Theoretical Linguistics, 12(2).

[24] Delaney, J. (2013). Employer use of Facebook and online social networks to discriminate against applicants for employment and employees: An analysis balancing the risks of having a Facebook account and the need for protective legislation. Labor Law Journal, 64(2), 86.

[25] Sheshasaayee, A., & Jayanthi, R. (2014). Exploring the potential of social media data using text mining to augment business intelligence. Compusoft, 3(4), 738.

[26] Tundjungsari, V. (2013). Business Intelligence with Social Media and Data Mining to Support Customer Satisfaction in Telecommunication Industry. International Journal of Computer Science and Electronics Engineering (IJCSEE) Volume, 1.

[27] Barbier, G., & Liu, H. (2011). Data mining in social media. In Social network data analytics (pp. 327-352). Springer US.

[28] Xiong, F., Pan, S., & Zhu, X. (2022). Collective Behavior Analysis and Graph Mining in Social Networks 2021. Complexity, 2022..

[29] Getoor, L. (2003). Link mining: a new data mining challenge. ACM SIGKDD Explorations Newsletter, 5(1), 84-89.

[30] Zhang, C., Chen, T., Li, Z., Liu, A., Xu, Y., Gao, Y., & Xu, D. (2021). Depiction of tumor stemlike features and underlying relationships with hazard immune infiltrations based on large prostate cancer cohorts. Briefings in bioinformatics, 22(3), bbaa211.

[31] Pippal, S., Batra, L., Krishna, A., Gupta, H., & Arora, K. (2014). Data mining in social networking sites: A social media mining approach to generate effective business strategies. International Journal of Innovations & Advancement in Computer Science (IJIACS), 3(2).

[32] Gong, Y., & Zhang, Q. (2016, July). Hashtag recommendation using attention-based convolutional neural network. In IJCAI 2016, Proceedings of the 26rd International Joint Conference on Artificial Intelligence, New York City, USA.

[33] Weston, J., Chopra, S., & Adams, K. (2014). # TagSpace: Semantic embeddings from hashtags.

[34] Wang, Z., Joo, V., Tong, C., & Chan, D. (2014, December). Issues of social data analytics with a new method for sentiment analysis of social media data. In 2014 IEEE 6th International Conference on Cloud Computing Technology and Science (pp. 899-904). IEEE.

[35] Isah, H., Trundle, P., & Neagu, D. (2014, September). Social media analysis for product safety using text mining and sentiment analysis. In 2014 14th UK workshop on computational intelligence (UKCI) (pp. 1-7). IEEE

[36] Saura, J. R., Palos-Sanchez, P., & Grilo, A. (2019). Detecting indicators for startup business success: Sentiment analysis using text data mining. Sustainability, 11(3), 917.

[37] Ahmad, S., Asghar, M. Z., Alotaibi, F. M., & Awan, I. (2019). Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. Human-centric Computing and Information Sciences, 9(1), 24.

[38] Păvăloaia, V. D., Teodor, E. M., Fotache, D., & Danileț, M. (2019). Opinion mining on social media data: sentiment analysis of user preferences. Sustainability, 11(16), 4459.