

Data Analysis Of Hashtag Mining : A Survey

¹Mr. Girish Regunta, ²Prof.Lavina Jadhav
^{1,2} MET Institute of Computer Science,
Bandra West, Mumbai, India

Abstract: *Social media platforms rapid growth has enhanced its significance and usage. Social media platforms are powerful communication and records sharing gear which are used by many human beings ensuing in manufacturing of heaps and hundreds of random records. With this sort of speedy increase of data each day, IT communities have always tailored new ways to absorb this random data, finding patterns to provide higher conversation capabilities in the social media. The technology behind locating styles among random information is data mining and text mining is one of the branches implemented in those social media structures enhancing pleasant of usage of those platforms making them play a distinguished position in an average person's lifestyles. One such data mining approach is Hashtag. Hashtag is a social media term used on various social media systems with posts with the aid of the users themselves. It is used with the (#) sign following by keywords with no spaces, e.g. #helloworld in various social media platforms such as Twitter, Instagram, Facebook, etc. Moreover, Hashtag is a data collecting tool which is analyzed to create AI models, Neuron network models and numerous different data mining strategies that's similarly applied in exceptional fields as a powerful business device for different businesses, establishments and social media platforms. In this paper, we are briefly discussing about origin of hashtag, its growth throughout the years, its usage over different social media platforms and its applications as data mining models and in real life*

Keywords: *Data mining, hashtag, social media, text mining techniques.*

1. INTRODUCTION

Social media is a place, network or a platform wherein humans can engage using their critiques, photos, emotions, experience and debate on any subject matter with their social contacts and additionally to boom their social contacts. Text mining has been in Life even earlier than the social media has walked into an Average person's lifestyles. For some time, data mining turned into use to find numerous hidden styles among the random records till It made a step forward of producing the hundreds of statistics already in a pattern. Even though it has become a random idea to implement facts clustering the usage of a described symbol observed through key phrases. It gave delivery to whole new methods of patterning information before generating them [1].

Boyd, who published the primary recognized use of the word "hashtag" told that the call "hashtag" comes from programmer subculture due to the fact he and his friends would refer to the symbol as the hash, not the pound sign. In the pre-internet days Americans and Canadians called it as pound sign or the number sign, while irish members referred to as it a hash. However, within the context of the internet, we all name the # a hashtag. Nowadays, the symbol is so pervasive in social media that almost each essential social media platform supports some functionality for the hashtag.

Hashtag is a metadata tag used preceded by way of pound (#) signal On social media systems which include twitter, instagram, Youtube, pinterest, and so on. The overall concept behind Creation of hashtag become allow users to apply Dynamic, person-generated tagging which allows customers to Without difficulty identify posts associated with the respective terms. Hashtag is generated by using the person the usage of the (#) signal observed with the aid of Phrases or words particular through the person. The styling of hashtag includes hash (#) sign followed by words with no spaces and numeric, e.g. #helloworld, #metoo. It should to be Mentioned that these terms are not case-sensitive and cannot be Started out with numeric (#123cannotuse), while numeric May be followed via characters (#canuse123) [6]. The Keywords may be of any language, capitalized, mixed, lowercase or uppercase.

This option of hashtag of allowing users to create their personal specific hashtags has involved customers so much utilization of hashtags. Usage became so comfortable that it made others Agencies or social media platforms to adapt them as soon As possible. The usage of hashtags has encouraged various Social troubles and areas which include political events, Leisure, corporate advertising or virtual advertising, a Social reason, conversation among other social contacts.

2. EVOLUTION

In 1988, the primary hash symbol used on Internet relay chat (IRC) to label agencies and subjects that had been available across the whole community. They had been used for grouping similar messages and content material to make it smooth for users to find the statistics they had been searching out. In its early days Twitter did not have a hashtag characteristic. Chris Messina took inspiration from IRC and unleashed the first Twitter hashtag on August 23, 2007.

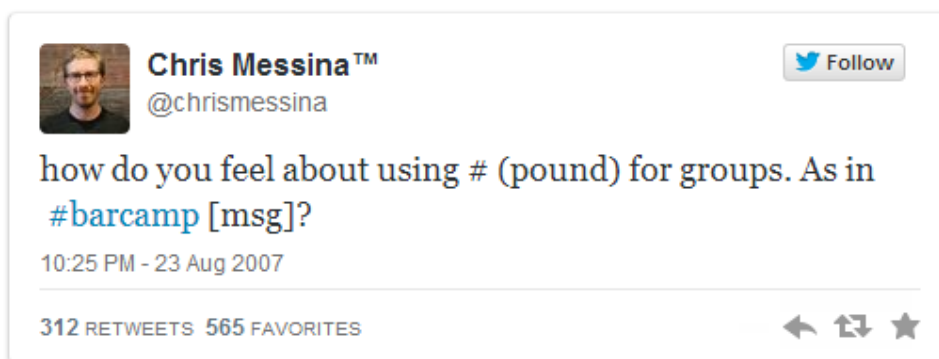


Fig 1.1 : First tweet about hashtag

Even twitter did not support the hashtag in 2007, chris messina saw an possibility to get the hashtag noticed. California resident nate ritter happened to be prolifically tweeting about the wildfires. On the equal time, messina noticed that sandiegofire become getting used as a tag on flickr. This inspired messina to reach out to ritter and endorse the usage of #sandiegofire on all relevant tweets. Ritter's tweets have become so well known that twitter users started out the use of hashtags to group relevant content. In 2009, twitter finally embraced them and brought a search device, in order that customers ought to see who else changed into the usage of a specific hashtag. From there, hashtags were followed through other platforms and have become part of the internet lexicon. Instagram, which was launched in 2010, has used them from day one. Facebook added them in 2013. Google+ and Pinterest also let users group content by simply using the #. [Ref 1]

3. DATA MINING APPLICATIONS OF HASHTAG

Text mining is a branch of data mining which is used to discover knowledge or patterns for better or further use. The text mining process undergoes various stages that is, collection of data, then preparing the data in order to analyse further, then applying clustering and classification methods to analyse text, forming a model which defines a pattern resulting in final stage that is knowledge discovery.

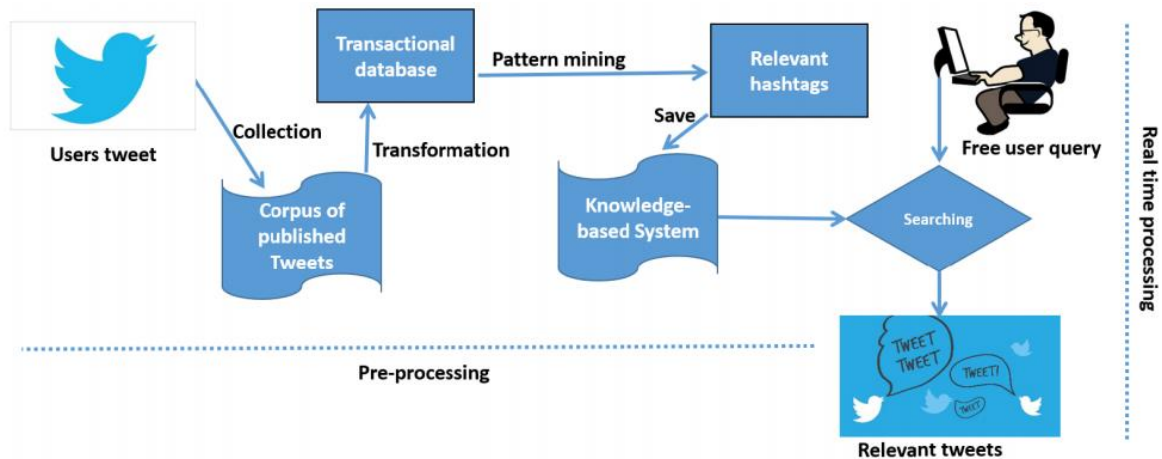


Fig 2.1 : Mining Process[9]

As discussed above that Hashtag is a data collecting tool in a patterned manner it is evident it has been influencing many works for a long time. Following are related works of hashtag explained:

- Bidisha Samanta, Abir De, Abhijnan Chakraborty, Niloy Ganguly [1] proposed Large Margin Point Process (LMPP), a probabilistic framework for Hashtag propagation integrating two main factors: hashtag-tweet influence and hashtag-hashtag competitions. The proposed model can also accurately predict the relative rankings of competing hashtags providing valuable data for further analysis.
- Jie Liu, Zhicheng He and Yalou Huang [2] devised an Embedding model, Hashtag2Vec, using multiple relations among hashtag-hashtag, hashtagtweet, tweet-word, word-word based on hierarchical heterogeneous network. It is capable of embedding hashtags and short social texts as well.
- Ramzan Talib, Muhammad Kashif Hanif, Shaeela Ayesha, and Fakeeha Fatima [3] discussed and analysed different text mining techniques, their applications in real life and also their adverse effects

4. REAL LIFE APPLICATIONS OF HASHTAG

As said before, with the increasing use of hashtag as style of posting on various social media platforms it has greatly impacted or also can be said to have changed the usage of social media in present days. Following are some of the factors that point the impact caused by this Meta tagging tool.

- **Politics:**

In Politics, There has been various instances in history where social media has greatly impacted politics. a recent political event in Indian Lok Sabha elections 2019 where a snide comment made by opposing party on current ruling party, resulted in a controversial hashtag, #meinbhichowkidar increased the attention towards ruling party winning them the elections with a clean sweep

- **Business:**

With the gaining popularity of social media, now there are many companies that are run using social media. Hence, hashtag plays a key role in branding of these companies. Usage of hashtags expands one's content reach, gradually increases the target audience, gathers attention required to the company's content The usage of hashtags are based upon three factors that are Brand and Campaign specific, Trending and Content. In the first kind, Companies are required to use hashtags that represent their motto or as said their tagline. For example, KitKat using #takeabreak, Nike using #justdoit etc.

- **Social Campaigns :**

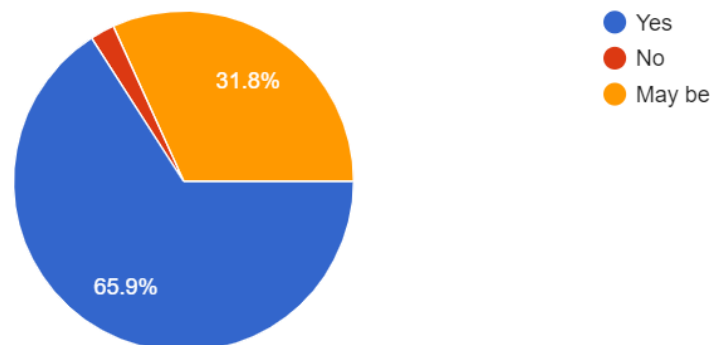
Social media has played a very important role in providing more than the required attention to the campaigns be it in any given field such as in health awareness there has been various campaigns about HIV / AIDS, cancers, Diabetes, Migraines, Vaccinations, self-care, diets and many more. In Entertainment field, there has been #metoo campaign that has run for a long time in Indian entertainment industry where women shared their disturbing experiences while pursuing their career or in their career span

5. RESULTS AND DISCUSSION

A survey was undertaken to demonstrate the usage of hashtag across various social media platforms and their idea of usage by a normal person's life

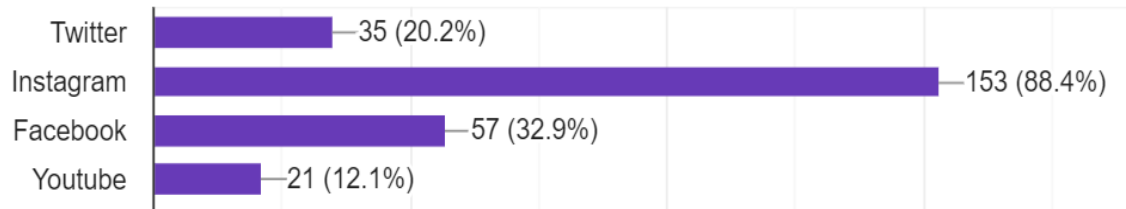
Q. Is hashtag allow people to find related posts of their own Interests ?

The graph below illustrates that 65.9 % of the survey respondents have positive view towards finding their interests using hashtag. 31.8% feels that chances are equal and remaining don't think it is useful for related post findings



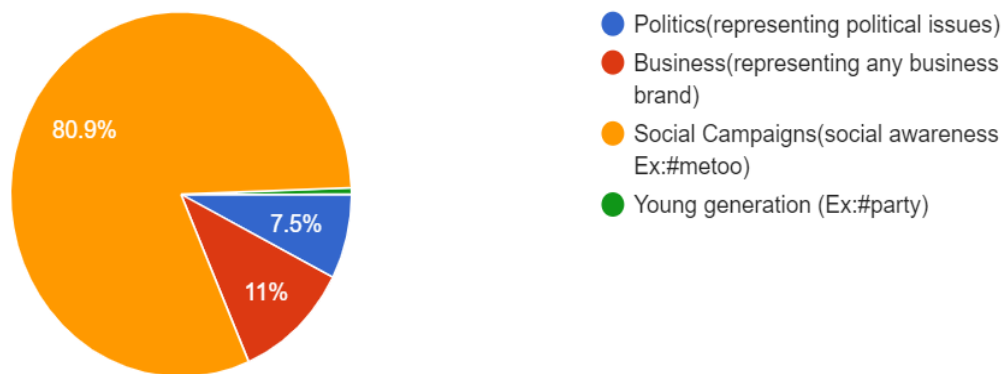
Q. On which platforms do you frequently use hashtags ?

On account that hashtag has been added on twitter platform but as consistent with survey facts instagram has been very popular platform for utilization of hashtag with a response of 88.4% whereas facebook with 32.9%, twitter with 20.2% and youtube with 12.1% utilization has been recorded



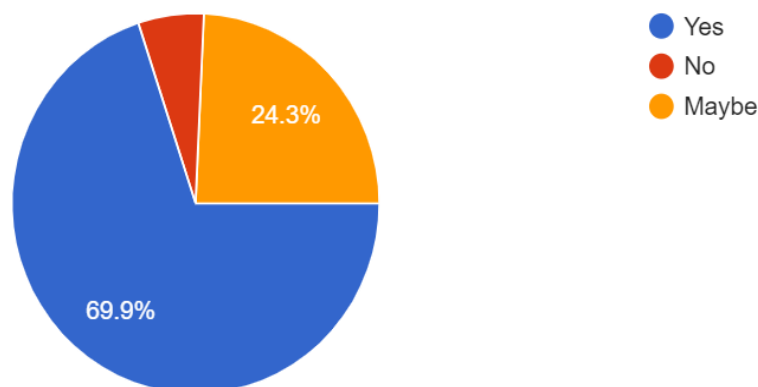
Q. According to you which sector has more advantages of using hashtag ?

Hashtags are useful to spread Awareness, Social campaigns have dominating usage with 80.9%, observed by using commercial enterprise sector with 11% response and political components with 7.5%



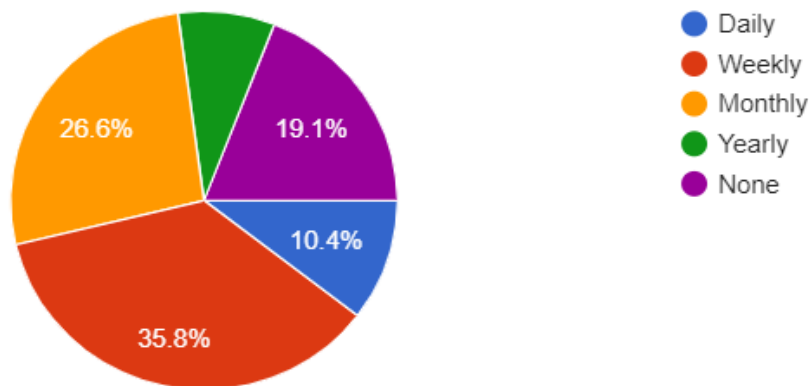
Q. Do you think hashtag will help you to know latest trends on social media ?

The underneath graph depicts hashtag outcomes in locating trendy tendencies, in which majority of the respondents have high quality view



Q. Hashtag usage frequency ?

As per Hashtag survey the frequency chart of hashtag usage is listed in following graph



6. CONCLUSION

As discussed in this paper, we can conclude that the maximum charming reality approximately hashtag is that it was brought to the world as a easy information collecting tool which especially revolutionized text mining techniques and affected people undoubtedly with its simplicity in use and making an effect in huge and additionally has been in lifestyles for extra than a decade now.

REFERENCES

- [1]. Bidisha Samanta, Abir De, Abhijnan Chakraborty, Niloy Ganguly: \ LMPP: A Large Margin Point Process Combining Reinforcement and Competition for Modeling Hashtag Popularity\, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)*.
- [2]. Jie Liu, Zhicheng He and Yalou Huang: \ Hashtag2Vec: Learning Hashtag Representation with Relational Hierarchical Embedding Model\, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*.
- [3]. Ramzan Talib, Muhammad Kashif Hanif, Shaeela Ayesha, and Fakeeha Fatima: \ Text Mining: Techniques, Applications and Issues\, (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, Vol. 7 No. 11, 2016
- [4]. article: George N, Britto DR, Krishnan V, Dass LM, Prasant HA, Aravindhan V. Assessment of hashtag (#) campaigns aimed at health awareness in social media. *J Edu Health Promot* 2018;7:114
- [5]. *The Surprising History of Twitter's Hashtag Origin and 4 Ways to Get the Most out of Them* by Belle Beth Cooper <https://buffer.com/resources/a-concise-history-of-twitter-hashtags-and-how-you-should-use-them-properly>
- [6]. Andrea Sandre , —Twitter didn't invent the hashtag... Chris Messina did!" , <https://hackernoon.com/twitter-didnt-invent-the-hashtag-chris-messina-did1020969abfcd>.
- [7]. *History of Hashtags: How a Symbol Changed the Way We Search & Share* <https://socialmediaweek.org/blog/2018/02/history-hashtags-symbol-changed-way-search-share/>
- [8]. *Hashtag* <https://en.wikipedia.org/wiki/Hashtag>
- [9]. *Exploring Pattern Mining Algorithms for Hashtag Retrieval Problem* <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8951154>