European Journal of Business & Social Sciences
Available at https://ejbss.org/
ISSN: 2235-767X
Volume 06 Issue 09
September 2018

Social Media & Big Data Analytics – Literary Review

Anu Dahiya

anudahiya19@gmail.com

Abstract

With the immense use of online services and mobile technologies, bulk of data is generating in huge quantity and day by day it is increasing and that is why today's era is known as big data era. Big data refers to data sets that are not only big, but also high in variety and velocity, which makes difficulty in handling such a big data using traditional tools and techniques. Due to the rapid growth of data, solutions need to be studied and provided in order to handle and extract value and knowledge from these data sets. Furthermore, decision makers need to be able to gain valuable insights from such varied and rapidly changing data, ranging from daily transactions to customer interactions and social network data. Such value can be provided using big data analytics, which is the application of advanced analytics techniques on big data. This paper aims to analyze some of the different analytics methods and tools which can be applied to big data, as well as the applications of big data analytics in social media.

Keywords: Social media, Big data Analytics, review, applications

1. Introduction

In the past few years, the widespread use of multimedia data, including videos, audios and text etc. from the social media sites have resulted in big data revolution in social media analytics system. Currently various social media platforms like You Tube, Facebook, Whatsapp, Instagram and Twitter are considered as inimitable and even valuable sources of multimedia data (Zafarani et al. 2014). For example, 255 million users of Twitter send approx 500 million tweets every minute in a day (Smith 2015). It is predicted that by 2020 the digital data rate will exceed 40ZB which means almost 5200GB of data will be produced by every person (Mearian 2012). Due to the rapid growth of data, solutions need to be studied and provided in order to handle and extract value and knowledge from these data sets. Furthermore, decision makers need to be able to gain valuable insights from such varied and rapidly changing data, ranging from daily

ISSN: 2235-767X
Volume 06 Issue 09
September 2018

transactions to customer interactions and social network data. Such value can be provided using big data analytics, which is the application of advanced analytics techniques on big data.

The big data concept is used to describe extremely large datasets. However, different scientists and authors have various definitions for this term. For example, it is first used in 1997 article that was published by NASA scientists explaining the computer systems challenges (Cox and Ellsworth 1997). After then, (Bryant et al. 2008) used the "Big Data Computing" term in 2008. And finally Apache Hadoop defined big data as "datasets which could not be captured, managed, and processed by general computers within an acceptable scope" (Chen et al. 2014).

In the current big data era, new opportunities as well as challenges appear with high-diversity and voluminous social media data. Hence, big data analytics has attracted a lot of attention in both industries and academia in recent few years. It is considered an emerging topic due to its critical and valuable insights (Chen 2010; Smith 2013; Tian et al. 2015).

2. Applications of big data in social media analytics

Today, social media levered with big data analytical techniques help in sensing the big data in social media in cost effective way. Following are the most popular big data applications in social media to show the significant role of big data analytics in social media era:

a. Social networking sites:

There are lots of studies carried out on social media and big data analytics (Cheung et al. 2015). For example, Tufekci (2014) carried out a research on the current challenges in analyzing social activities and behaviors of humans, specially based on twitter hastags. Wilson et al. 2012 conducted a study on Facebook as a valuable source of social media data.

b. Smart cellphones

In this new era of technology, smart phones have taken over other electronic devices, such as laptops and computers. Due to advancement in smart phones like internet connectivity, Bluetooth, GPS navigation, cameras etc has generated lots of meaning c

ISSN: 2235-767X
Volume 06 Issue 09
September 2018

c. Surveillance videos

These are considered as the large source of social media data. Research using surveillance data got valuable insights with the advent of innovative big data solutions. There is one remarkable application of surveillance videos is how to automatically detect semantic concepts from videos (Chen et al. 2001b, 2003; Shyu et al. 2008).

d. Some other applications

There are other applications of bug data in social media that can be categorized into Smart TVs (Fleites et al. 2015a, 2015b), disaster management systems (Yang et al. 2012, 2015) and IOTs (Zhou and Chao 2011).

3. Social media Big data tools

TOOL	PURPOSE	BIG DATA TECHNIQUES	LANGUAGE
Apache Mahout	- Data stochastic - Clustering - Classification - Collaborative filtering	- Hadoop - Spark - H2O - Flink	- Java - scala
MOA/ Weka	- stream classification, clustering, and regression - graph mining - visualization	- stream mining - large scale machine learning	- java
R	- Statistical modeling - Graphing - Machine learning - Time series	- Oracle R - sparkR - RHadoop	- C Fortran
Rapidminer	- Business	- Radoop	- java

ISSN: 2235-767X
Volume 06 Issue 09
September 2018

	processes - Predictive analysis - Text mining - Social media analysis - Vizualization		
KNIME	- statistical modeling - preprocessing - machine learning - image processing - time series - network analysis	- Hadoop - Spark	- Java (Eclipse-based)

Source: Samira Pouyanfar et al. (2018). Multimedia Big Data Analytics: A Survey

4. Social media big data challenges

Social media big data or we say the "biggest big data", not only exhibits all the 5Vs of big data but it also escalates the difficulties into another level. As we know social media big data provides the industries and professionals a valuable insight, at the same time it also brings new challenges, such as scalability, storage and security.

Processing of heterogeneous social media data in an effective way is also a major challenge now days.

Retrieving valuable information from large data is yet another big challenge. Recent studies have shown the importance of hashing techniques to extract the correlations between different modalities from various sources and also to tackle the problem of scalability (Song et al. 2013). The security of data is the major issue of social media big data analytics.

Overall, although new advances in big data turn social media challenges into new new opportunities, there still remain many concerns that need to be addressed.

5. Opportunities in social media big data analytics

ISSN: 2235-767X
Volume 06 Issue 09
September 2018

In spite of all the challenges, ever increasing social media big data provides great and valuable insights into human behaviors and sentiments, which again turn into new opportunities to making great progress in many fields (Michael and Miller 2013; Chen and Zhang 2014). Following are some areas of new opportunities, which social media big data analytics provides:

- Large number of mobile users could provide data of classifying or categorizing behavior of customers and it helps in decision making process. It also helps in product improvement and target marketing (Micheal and Clarke 2013).
- Analytic of patient data could lead to detection of devastating diseases at an early stage and also provide timely cures (Strickland 2013).
- Many of the organizations could use big data to learn more about their workforce, increase productivity and also helps in improving business processes. (Michael and Miller 2013).
- Climate science also utilize social media spatial-temporal data to analyze and monitor the temperature of the earth (Yang et al. 2017)

To summarize, the proper utilization of social media big data helps a lot in providing new insights to the world (Chen and Zhang 2014).

6. Summary

This paper has presented various applications of social media big data analytics. It also covered various tools some of which are provided to us as free access and some are partially free to use. These tools help us in classification, clustering, prediction, time series and many more using the social media big data. Further the paper covered some challenges of social media big data and also highlights some new insights in form of opportunities.

7. References

Reza Zafarani, Mohammad Ali Abbasi, and Huan Liu. 2014. *Social Media Mining: An Introduction*. Cambridge University Press.

John R. Smith. 2013. Riding the multimedia big data wave. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, Dublin, Ireland, 1–2.

Thurston Smith. 2015. Social media and big data statistics. https://www.youtube.com/watch? v=hoZ7YhfRchw.

Lucas Mearian. 2012. By 2020, there will be 5,200 GB of data for every person on the Earth. http://www.computerworld.com/article/2493701/data-center/by-2020-there-will-be-5-200-gb-of-datafor-every-person-on-earth.html.

ISSN: 2235-767X
Volume 06 Issue 09
September 2018

Michael Cox and David Ellsworth. 1997. Application-controlled demand paging for out-of-core visualization. In *Proceedings of the 8th Conference on Visualization'97*. IEEE Computer Society Press, 235–244.
Samira Pouyanfar, Yimin Yang, Shu-Ching Chen, Mei-Ling Shyu, and S.S. Iyenger. 2018. Multimedia Big Data Analytics: A Survey. *ACM Comput. Surv. 51, 1, Article 10(january 2018), 34 pages.*
Randal Bryant, Randy H. Katz, and Edward D. Lazowska. 2008. Big-data computing: Creating revolutionary breakthroughs in commerce, science and society. Retreived from https://pdfs.semanticscholar.org/65a8/b00f712ffd5c230bf0de6b9bd13923d20078.pdf .
CL Philip Chen and Chun-Yang Zhang. 2014. Data-intensive applications, challenges, techniques and technologies: A survey on big data. *Inf. Sci.* 275 (2014), 314–347.
Min Chen. 2014. A hierarchical security model for multimedia big data. *Int. J. Multimedia Data Eng. Manage.* 5, 1 (2014), 1–13.
Min Chen, Shiwen Mao, and Yunhao Liu. 2014. Big data: A survey. *Mobile Netw. Appl.* 19, 2 (2014), 171–209.
Shu-Ching Chen. 2010. Multimedia databases and data management: A survey. *Int. J. Multimedia Data Eng. Manage.* 1, 1
(2010), 1–11.
Shu-Ching Chen and Rangasami L. Kashyap. 2001. A spatio-temporal semantic model for multimedia database systems and multimedia information systems. *IEEE Trans. Knowl. Data Eng.* 13, 4 (2001), 607–622.
Shu-Ching Chen, Mei-Ling Shyu, Chengcui Zhang, and Jeff Strickrott. 2001b. Multimedia data mining for traffic video sequences. In *Proceedings of the 2nd International Workshop on Multimedia Data Mining*. 78–86.
Ming Cheung, James She, and Zhanming Jie. 2015. Connection discovery using big data of user-shared images in social media. *IEEE Trans. Multimedia* 17, 9 (2015), 1417–1428.
Fausto Fleites, Haohong Wang, and Shu-Ching Chen. 2015a. TV shopping via multi-cue product detection. *IEEE Trans. Emerg. Top. Comput.* 3, 2 (2015), 161–171.
Fausto C. Fleites, Haohong Wang, and Shu-Ching Chen. 2015b. Enabling enriched TV shopping experience via computational and temporal aware view-centric multimedia abstraction. *IEEE Trans. Multimedia* 17, 7 (2015), 1068–1080.
Katina Michael and Roger Clarke. 2013. Location and tracking of mobile devices: Uberveillance stalks the streets. *Comput. Law Secur. Rev.* 29, 3 (2013), 216–228.
Katina Michael and Keith Miller. 2013. Big data: New opportunities and new challenges [guest editors' introduction]. *Computer* 46, 6 (2013), 22–24.
Mei-Ling Shyu, Zongxing Xie, Min Chen, and Shu-Ching Chen. 2008. Video semantic event/concept detection using a subspace-based multimedia data mining framework. *IEEE Trans. Multimedia* 10, 2 (2008), 252–259.
Jingkuan Song, Yang Yang, Yi Yang, Zi Huang, and Heng Tao Shen. 2013. Inter-media hashing for large-scale retrieval from heterogeneous data sources. In *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*. 785–796.
Eliza Strickland. 2013. The gene machine and me. *IEEE Spectr.* 50, 3 (2013), 30–59.
Robert E. Wilson, Samuel D. Gosling, and Lindsay T. Graham. 2012. A review of Facebook research in the social sciences. *Perspect. Psychol. Sci.* 7, 3 (2012), 203–220.

ISSN: 2235-767X
Volume 06 Issue 09
September 2018

Chaowei Yang, Qunying Huang, Zhenlong Li, Kai Liu, and Fei Hu. 2017. Big data and cloud computing: Innovation opportunities and challenges. *Int. J. Digit. Earth* 10, 1 (2017), 13–53.

Yimin Yang. 2015. *Exploring Hidden Coherent Feature Groups and Temporal Semantics for Multimedia Big Data Analysis*. Ph.D. Dissertation. Florida International University.

Yimin Yang, Wenting Lu, Jesse Domack, Tao Li, Shu-Ching Chen, Steve Luis, and Jainendra K. Navlakha. 2012. Madis: A multimedia-aided disaster information integration system for emergency management. In *Proceedings of the 8th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*. ICST/IEEE, 233–241.

Liang Zhou and Han-Chieh Chao. 2011. Multimedia traffic security architecture for the internet of things. *IEEE Netw.* 25, 3 (2011), 35–40. Qiusha Zhu, Zhao