



Guru Nanak Institute of Technology
Department of Information Technology

A Systematic Review Towards Big Data Analytics in Social Media

Under Guidance of :
Mr. A. RAJKUMAR
(Assistant Professor)

PRESENTED BY :

S. THANVI YADAV (19831A1245)

T. YUGANDHAR (19831A1249)

K. ESHWAR (19831A1225)

S. SAI KUMAR (19831A1244)

LIST OF CONTENTS:

- ❑ INTRODUCTION
- ❑ ABSTRACT
- ❑ LITERATURE SURVEY
- ❑ EXISTING SYSTEM
- ❑ PROPOSED SYSTEM
- ❑ MODULES
- ❑ CONCLUSION

INTRODUCTION TO DOMAIN:

- Big data analytics in social media is an emerging field of research that is becoming increasingly important as the world becomes increasingly connected through digital platforms. This project will examine the state of the art in this domain, exploring the challenges and opportunities with the use of big data analytics in social media. The review will focus on how big data analytics can be used to gain insights from vast amounts of data generated by social media users, and how this data can be used for business and marketing purposes. It will also examine the ethical considerations that need to be taken into account when using big data analytics in social media. Finally, the review will look into the current trends and future prospects of this field.

ABSTRACT :

- Social media have become an integral part of our lives, and the increasing availability of data has made it possible to gain valuable insights from such data through big data analytics. This project studies the current state-of-the-art research in the field of big data analytics in social media, covering topics such as sentiment analysis, opinion mining, text and data mining, and machine learning. Examining the literature, this project highlights the challenges and opportunities associated with the new field of big data analytics in social media, and identifies open research challenges in the areas of data collection, storage, retrieval, and analysis. Furthermore, this project provides a comprehensive overview of the current research trends in the field, discussing the strengths and limitations of existing approaches.

LITERATURE SURVEY :

- ▶ **TITLE :** A Systematic Review Towards Big Data Analytics in Social Media
- ▶ **AUTHOR:** Md. Saifur Rahman And Hassan Reza
- ▶ **YEAR :** June, 2022
- ▶ **DESCRIPTION:** The recent advancement in internet 2.0 creates a scope to connect people worldwide using society 2.0 and web 2.0 technologies. This new era allows the consumer to directly connect with other individuals, business corporations, and the government. People are open to sharing opinions, views, and ideas on any topic in different formats out loud. This creates the opportunity to make the “Big Social Data” handy by implementing machine learning approaches and social data analytics. This study offers an overview of recent works in social media, data science, and machine learning to gain a wide perspective on social media big data analytics.

- ▶ **TITLE :** Big data analytics meets social media: A systematic review of techniques, open issues, and future directions
- ▶ **AUTHOR:** Sepideh Bazzaz Abkenar, Mostafa Haghi Kashani, Ebrahim Mahdipour, and Seyed Mahdi Jameii
- ▶ **YEAR :** 2020, Oct 14
- ▶ **DESCRIPTION:** This paper presents a systematic review of big data analytics in social networks. We explained the research methodology, paper selection process, and selected 74 papers between 2013 and August 2020, from among 785 papers in our search query. A significant number of the studied papers were related to IEEE, Springer, and ScienceDirect journals, with 37%, 27%, and 19%, respectively. On the other hand, each of Taylor&Francis and Emerald publishers with 1% had the lowest number of published papers. From these studies, 74 papers were categorized into two approaches: Content-oriented approaches (51%) and network-oriented approaches (49%). Besides, the main ideas, advantages, disadvantages, evaluation methods, tools, and evaluation parameters of each studied paper were discussed. It was found that the most widely considered evaluation parameters were accuracy (20%), time (16%), and scalability (12%), but privacy, reliability, and security measures were somewhat neglected.

- ▶ **TITLE :** Accurate Content Push for Content-Centric Social Networks: A Big Data Support Online Learning Approach
- ▶ **AUTHOR:** Yinan Feng; Pan Zhou; Dapeng Wu; Yuchong Hu
- ▶ **YEAR :** 2020, Jan
- ▶ **DESCRIPTION:** In this paper, we model this issue as a novel contextual multiarmed bandit based Monte Carlo tree search problem and propose a big data support online learning algorithm to meet the demand of content push with low cost. To avoid destroying CCN's energy efficient feature, the energy consumption is considered into our module. Then, we theoretically prove that our online learning algorithm achieves sublinear regret bound and sublinear storage, which is very efficient in the big data context and do not increase the network burden. Experiments in an offline collected dataset show that our approach significantly increases the accuracy and convergence speed against other state-of-the-art bandit algorithms and can overcome the cold start problem as well.

- ▶ **TITLE:** Twitter sentiment analysis with a deep neural network
- ▶ **AUTHOR:** A. Ahmed and D. Elise
- ▶ **YEAR:** 2019.
- ▶ **DESCRIPTION:** Sentiment analysis on social media such as Twitter has become a very important and challenging task. Due to the characteristics of such data—tweet length, spelling errors, abbreviations, and special characters—the sentiment analysis task in such an environment requires a non-traditional approach. Moreover, social media sentiment analysis is a fundamental problem with many interesting applications. Most current social media sentiment classification methods judge the sentiment polarity primarily according to textual content and neglect other information on these platforms. In this paper, we propose a neural network model that also incorporates user behavioral information within a given document (tweet). The neural network used in this paper is a Convolutional Neural Network (CNN). The system is evaluated on two datasets provided by the SemEval-2016 Workshop.

EXISTING SYSTEM:

- ▶ The existing system for big data analytics in social media is based mainly on supervised and unsupervised machine learning algorithms, such as support vector machines, decision trees, and topic modeling. While these algorithms have been successfully used for various tasks, they have several limitations. For example, supervised learning algorithms require labeled data for training, which can be difficult and time-consuming to obtain. In addition, unsupervised learning algorithms may suffer from the curse of dimensionality, where the performance of the algorithm decreases with increasing number of features. Furthermore, these algorithms are limited in their ability to deal with non-structured data, such as text and images.
- ▶ The main drawback of the current system for big data analytics in social media is its limited ability to deal with non-structured data, such as text and images. Additionally, current algorithms are limited in their ability to extract meaningful insights from large volumes of data in a timely and efficient manner. Finally, current algorithms are often not designed to capture the dynamic nature of social media data, such as user interactions, trends, and sentiment.

PROPOSED SYSTEM :

- ❑ The proposed system for big data analytics in social media uses a combination of supervised and unsupervised learning algorithms. This approach has several advantages over existing algorithms.
- ❑ The Sunflower Model is a proposed system for big data analytics in social media that uses a combination of supervised and unsupervised learning algorithms. The Sunflower Model is designed to help organizations extract insights from large amounts of social media data. It is based on a hierarchical structure that consists of three layers: the data layer, the feature extraction layer, and the learning layer.
- ❑ Furthermore, it is designed to be scalable, so that it can handle large amounts of data.

MODULES :

- ▶ **MODULES NAME:**
- ▶ **1. Admin :** In this module we design the windows for the project. These windows are used for secure login for all users. To connect with server user must give their username and password then only they can able to connect the server.
- ▶ **2. Selection Criteria :** This is the first module Data User can register and Login. After login Data User have an option of searching the files as a file name. Data user can also have a download file it will show an encrypted data. Data user can also send a trapdoor request to the server.
- ▶ **3. Social Media Analytics :** This is the Second module of this project. In this module Data Owner should register and Login. Data Owner will Uploads the files into the database. Data owner can also send request to the data user.
- ▶ **4. Sunflower Media :** This is the third module of this project. In this module Cloud Server can login. After login it will see all data owners' information. Cloud server can see all users' information. Cloud server can see an all stored data files. Cloud server can give keys request to the user. Cloud server can also see an attacker information of file.

CONCLUSION :

- This project provides a comprehensive overview of the current state-of-the-art research in the field of big data analytics in social media. It discusses the challenges and opportunities associated with the new field, and identifies open research challenges in the areas of data collection, storage, retrieval, and analysis. Furthermore, the project highlights the strengths and limitations of existing approaches, and provides an overview of the current research trends in the field. Finally, it proposes the Sunflower Model as a proposed system for big data analytics in social media.

**Thank
you**