

Social Media Mining

13**Roberto Marmo***Universita' Pavia, Italy*

INTRODUCTION

Today, Social Media is a relevant part of human life, because it provides easily an accessible platform for users to share information, and it is truly the reflection of today's society, so the pervasive use has generated unprecedented amounts of social data. The result is large quantity of social data and user-generated content that, if properly mined and analyzed, could help the public and private sectors improve the quality of their products and services while reducing costs.

The aim of this contribution is to describe some technologies and methodologies to compute Social Media Mining, in order to collect information about individuals and entities, measure their interactions, and discover patterns to understand human behavior. The chapter also discusses background, knowledge, challenges and critical factors necessary for successful analysis.

BACKGROUND

Social media is the collective of online communications channels dedicated to community-based input, interaction, content-sharing and collaboration. social media can be considered a world of social atoms (i.e., individuals), entities (e.g., content, sites, networks, etc.), and interactions between individuals and entities (Zafarani, 2014). Social network is a social structure composed of individuals, organizations, company etc. which are connected by relationships and interactions.

Social media are used intensely to communicate, to publish resources (personal data, photo, video, blog), to establish relationships of a different type share information, make decisions, and do business in many ways, with hundreds web platforms in the world collecting the information of more than one billion registered users. Categories of social media regards: blogs, discussion forums, blogs, social networking, news discussions, media sharing sites, collaborative websites. Well-known online platforms are: Facebook as generic social network, LinkedIn as business social network, Twitter as micro-blogging. It is possible to seek social media as the tool for marketing (Goyal, 2018).

SOCIAL MEDIA MINING

Social Media Mining (SMM) is the process of representing, analyzing, and extracting actionable patterns and trends from massive raw social data, that are publicly available on social media web platforms (Bonzanini, 2016; Ravindran, 2015; Zafarani, 2014).

SMM discusses theories and methodologies from different disciplines such as computer science, data mining, machine learning, social network analysis, network science, sociology, ethnography, statistics, optimization, and mathematics.

DOI: 10.4018/978-1-7998-3473-1.ch149

The term mining is an analogy to the resource extraction process of mining for rare minerals, as stated in typical data mining approaches that analyze the huge volumes of data and seeks out patterns, trends and clusters. The goal of work (Injadat, 2016) is to analyze the data mining techniques that were utilized by social media between 2003 and 2015, conclusions suggest that more research be conducted by both the academia and the industry since the studies done so far are not sufficiently exhaustive of data mining techniques.

Data generated on social media sites are different from conventional attribute-value data for classic data mining. Social media data are largely user-generated content on social media sites. Social media data are vast, noisy, distributed, unstructured, and dynamic. These characteristics pose challenges to data mining tasks to invent new efficient techniques and algorithms.

SMM requires human data analysts and automated software programs to sift through massive amounts of raw social media data in order to discern patterns and trends. SMM cultivates a new kind of data scientist who is well versed in social and computational theories, specialized to analyze recalcitrant social media data, and skilled to help bridge the gap from what we know (social and computational theories) to what we want to know about the vast social media world with computational tools (Zafarani, 2014). The key questions for researchers are:

- How to derive information from unstructured data as textual, image, video?
- How to form analysis to assist in decision making based on the information derived from unstructured data?
- How to grasp opportunities for business success based on the generated report?
- Which source of social information should a user use?
- How can to identify communities in a social network?
- Which pieces of information are popular and receive a lot of attention?
- How quickly to uncover insights about customers? Who is customer talking to? Who are his friends? Who is influencing him?

Nowadays, sentiment analysis and influencer detection are relevant research areas. Social media data are vast, noisy, unstructured, and dynamic in nature, thus novel research areas arise. SMM is a burgeoning multidisciplinary area, where researchers of different backgrounds can make important contributions that matter for social media development.

APPROACHES

SMM integrates many approaches to provide a coherent platform to understand the basics and potentials of social media platforms. A methodology to integrate several techniques for analyzing and comparing social media content from business competitors is detailed on (Wu et al, 2015). The work of Adedoyin-Olowe et al (2014) discusses different data mining techniques used in mining diverse aspects of the social network over decades going from the historical techniques to the up-to-date models.

The book (Ravindran, 2015) provides a theoretical background, comprehensive instructions, and state-of-the-art techniques through explanations of implementation of various use cases using R programming language. Tools for data analysis in Python programming language are discusses in Bonzanini (2016).

In this section, some representative research issues are discussed.

11 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/social-media-mining/263682

Related Content

Organization, Information, and Human Capital: Troika Requisites for HEI Leadership and Organization Performance

Teay Shawyun (2022). *Key Factors and Use Cases of Servant Leadership Driving Organizational Performance* (pp. 225-250).

www.irma-international.org/chapter/organization-information-and-human-capital/294791

Student Nurses' Perception on the Impact of Information Technology on Teaching and Learning

Nahed Kandeel and Youssef Ibrahim (2018). *Nursing Education, Administration, and Informatics: Breakthroughs in Research and Practice* (pp. 53-66).

www.irma-international.org/chapter/student-nurses-perception-on-the-impact-of-information-technology-on-teaching-and-learning/202157

A Survey on Forensic Accounting

Roseline Oluwaseun Ogundokun and Opeyemi Eyitayo Ogundokun (2021). *Encyclopedia of Organizational Knowledge, Administration, and Technology* (pp. 35-44).

www.irma-international.org/chapter/a-survey-on-forensic-accounting/263525

When Do Working Consumers Become Prosumers?: Exploring Prosumer Characteristics for Organizational Value Creation Strategies

Sabina Alina Potra and Adrian Pugna (2021). *Encyclopedia of Organizational Knowledge, Administration, and Technology* (pp. 2261-2277).

www.irma-international.org/chapter/when-do-working-consumers-become-prosumers/263689

Manipulating Business Performances by Creative Accounting

Ionica Oncioiu, Traian Ovidiu Calot and Alin Eliodor Tnase (2021). *Encyclopedia of Organizational Knowledge, Administration, and Technology* (pp. 24-34).

www.irma-international.org/chapter/manipulating-business-performances-by-creative-accounting/263524