Enhancing the Business Process Lifecycle Using Twitter Mining

Hend Montasser[1], Hanan Fahmy[2], Sayed Abdelgaber[2], Yehia Helmy[1]
[1] *Faculty of Commerce and Business Administration, Helwan University, Egypt*
Business Information Systems Department
[2] *Faculty of Computers and Artificial Intelligence, Helwan University, Egypt*
Hend.montasser@fci.helwan.edu.eg (correspondence author)

Abstract— Nowadays, reviews on different items can be found in blogs, social networking sites, e-commerce sites, and web forums. Whenever a consumer wishes to decide on the buying of some product, the opinion plays an important part. Before buying any new product, the buyer must refer to the thoughts or views of the people currently using the product. Customers who purchased the items will share their experience with the product online in the form of textual feedback. Reviews are also useful for purchasers of products. The sentiment contained in the feedback on the characteristics of the product offer a detailed interpretation of the reviews and also aid in the examination of the characteristics of the product. However, it is difficult for consumers to derive the necessary details from the opinion reviews. Therefore, this paper aims to collect customer reviews from Twitter to support the Business Process Management lifecycle. The collected tweets will be cleaned, preprocessed, and organized using Natural Language Processing techniques. Besides, a sentiment analysis will be performed to derive meaningful knowledge from the collected tweets. Finally, the extracted reviews will be automatically stored in the generated ontology to be available for query by process owners. The proposed framework aims to enhance the knowledge creation, acquisition and dissemination throughout the organization.

Index Terms— business process, business process lifecycle, knowledge management, ontology, twitter mining

1 Introduction

The Business Process Management (BPM) model has been more socially focused in recent years. The socialization of systems has become an unavoidable means of integrating agile methods, including enhancing collaboration, knowledge sharing, and collective decision-making. The social element is an inevitable way to improve collaboration, knowledge sharing, and collective decision-making at BPM. However, this social dimension is effective and efficient, as well as well defined, correctly applied, and independent of any technology.

Traditional BPM is rooted in the principle of empirical management was to function effectively in a well-structured, secure, and consistent economic climate. Because today's business performance in diverse, global, and challenging markets depends primarily on the ability to evolve, companies need to reshape themselves. Networks inside and between organizations lead to increased stability, agility, and creativity in the way organizations address consumer demands and requirements. It applies to goods, resources, and processes. As a result, there is a continuous change from precisely organized and manageable organizational types with well-defined boundaries (information society) to highly complex, self-organizing network organizations with fuzzy boundaries (knowledge society) [1].

However, much of the current approaches remain confined to the usage of social media without providing a coherent, Scholarly, and reusable sense, irrespective of such technology. This paper precisely explains the social dimension to be taken into account in BPM and suggests a set of models for the design and development of Social BPM. A social framework is proposed to collect customer reviews from Twitter to support the business process life cycle. Then, Natural language process techniques were used to clean and analyse the data. A social ontology with Protege 5.0 has already been developed, and SWRL and SPARQL have implemented rules for the inferring of knowledge and manipulation queries.

The rest of the paper is structured as follows: Section 2 will provide a background of the state-of-art for this area of study. Section 3 will introduce and discuss the proposed conceptual framework. Finally, section 4 will conclude the paper and present the possible future work.

2 Background

2.1 Business Process Optimization

Business Process Optimization is the act of taking your old business process and optimizing them for efficiency [2]. Business Process Optimization is one of the final steps for BPM a methodology that advocates for constant process re-evaluation and improvement [3]. So, to make it work, you should have already carried out the first three steps critical for any BPM initiative. Optimizing business processes offers many benefits that can help businesses stay afloat in the tidal waves of change:

- Market compliance
- Streamlined operations
- Reduced risks
- Well-used resources
- Consistency
- Assured quality
- End-to-end visibility

2.2 Knowledge Management

Knowledge Management (KM) refers to any intentional effort to manage the knowledge of the workforce of the organization [4]. Such activities can be defined as a mechanism for developing, capturing, assimilating, adapting, using, and exchanging information using direct or indirect approaches, such as unique organizational culture and social processes [5]. KM can be defined as a discipline for the utilization of knowledge, people, systems, thought, and collective expertise that will help or lead to the development of new skills, creativity, and understanding that in turn enhances market performance or business requirements.

2.3 Ontology

Ontologies are state-of-the-art constructions that represent a rich and complex awareness of objects, their properties, the class of entities, and the relationship between things. A great deal of focus has been paid to the use of web-based ontologies and their contribution to business advancement in recent years. Ontologies provide the means to freely describe the different dimensions of a business domain, to provide semantics, and to define both the semantics of language constructs and the semantics of model instances. With a web-based semantic schema such as the Web Ontology Language (OWL), the creation and usage of basic models can be improved and implicit semantics contained in the models can be partially articulated and used for processing purposes [6].

2.4 Twitter Mining

Twitter mining is a new field of research under the umbrella of data mining and machine learning. Twitter is the preferred route of user reviews for the following reasons: a) Twitter is a popular platform for academics in a range of fields, allowing the use of a vast volume of content. b) Twitter facilitates data collection by offering easy access to the Twitter APIs. c) It is impossible to evaluate the exact domain(s) to which the user's tweet refers due to the economy and the uncertainty and brevity of the content of the tweet.

Some researchers have previously discussed this area like [7] who used a Twitter mining approach to the domain-based classification of users and their textual content. The suggested solution involves machine learning units. The method consists of two steps of analysis: a time-aware semantic study of the user's historical aspects, including five widely used machine learning classifiers. This system classifies consumers into two main categories: policy-related and non-political groups. This research is restricted to an on / off domain classification for OSN material. While [8] suggested a way to derive valuable data from a tweet object and to represent it in the Resource Description System (RDF) data model. The Streaming API used to gather twitter data by filtering the information using unique keywords.

3 CONCEPTUAL FRAMEWORK

The proposed conceptual framework, shown in fig. 1, is used to extract the customer reviews from Twitter and perform some pre-processing analysis to be dynamically stored in the ontology. This ontological approach helps in disseminating knowledge throughout the Business Process Management lifecycle. The focus of the work is on Twitter data because: (I) it is readily accessible, (ii) unique domains can be easily identified using "# tags" and (iii) the desired ontologies are minimal (comprising a small number of entities and relationships). Following, the steps for this framework will be discussed.

3.1 Data Collection

After the introduction of Twitter in 2006, Twitter generated a rich data set with over 500 million messages per day, some 200 billion tweets per year [9]. Twitter is a microblogging site which has evolved exponentially over the past three years. It is recognized as an anonymous social networking site, with about 336 million active monthly users at the beginning of 2018 [8]. Twitter allows people to send and read tweets which are limited to 140 characters per tweet. People may engage and share their views on products or films and future events, such as sports or political elections, etc. Owing to the exponential development of Twitter, academics are exploring the content features of tweets so that they can help to retrieve details such as consumer problems, topic views.

The proposed scheme used the Twitter Streaming API, a gateway that lets people connect to their computer programs and online resources. The Tweepy-based Python library uses the Streaming API to retrieve Tweets streams in real-time; the first step in accessing Twitter data is to create a Twitter application to allow access to the Twitter API. Data collection was accomplished by logging in to the Twitter profile account and switching to https:/apps.twitter.com/application server that helps the researcher to create an application.

3.2 Data Cleaning and Preprocessing

The cleaning process of each tweet was carried out as in the following steps • Remove repeats in characters • Eliminate white spaces. • Delete the URI. • Delete symbols such as RT, *, /? • Erase markers such as semicolon (;), column (,), circles (.), quotations ('/"), etc. • Eliminate arcs (, [.,]). • Delete the numbers. • Delete words like (to, to, and to). • Switch capital letters to small letters.

At this point, multiple operations are carried out on text resulting from the previous step of removing important entities, as in the following steps: • Tokenize text to individual words. • To delete all suffixes from the terms, use the Porter Stemmer API in the Java language. • Add POS for each word using the Ark-tweet NLP API, a Part-of-Speech tag deemed to be a crucial problem in natural language processing and a critical step in deciding the part of each word in the text.

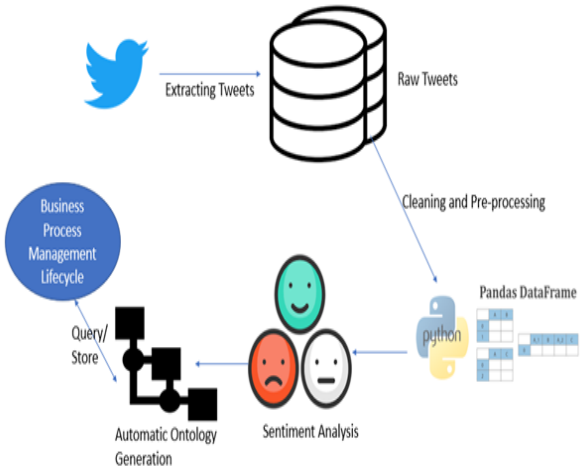

Figure 1: Twitter Mining framework

3.3 Twitter Sentiment Analysis

The ontology model that was developed earlier was used to classify the services/products for which consumers published derogatory tweets. The Ontology model was searched using SPARQL with feedback as various combinations of nouns and verbs derived from these negative tweets. Results can be enhanced by refining the interaction between objects in the ontology model. Often, adjectives and adverbs that modify the terms of the product are called sentimental words. On this basis, the sentimental polarity of the established sentimental terms is examined by the method of the classification of opinions, which primarily promotes fuzzy logic and SentiWordNet to classify the opinion words into five degrees of polarity, i.e., strong-weak negative, neutral, positive, and strong positive. Specifically, a fuzzy OWL plugin in Protégé can be used to add a polarity attribute to any fuzzy opinion vector in which the min-max normalization is used to normalize the polar spectrum [0, 1]. The difference between each type of sentiment can be rendered as follows: strong negative spectrum is [0.0–0.25], negative is [0.25–0.5], neutral is [0.5], positive is [0.5–0.75] and strong positive is [0.75–1].

3.4 Automatic Ontology Generation

Ontology generation is a resource-intensive undertaking; the primary challenge is to identify and define the various entities and relationships that form the target domain and need to be included in the intended ontology. We strive to build a Product Review Ontology that is capable of collecting consumer feedback from a range of sites to be used by the manufacturer or engineer at the design stage of the product. Below are the potential motivational questions that should be posted by the designer about the product based on the consumer review: 1. List all the consumer feedback. 2. What part of the structure is alluded to in the reviews? 3. What aspects of the segment do the comments refer to? 4. Compare the ratings of the unique aspects of the product of another brand. 5. What are the customer's suggestions?

4 Conclusion

Social networking services have arisen as essential forums for communicating and sharing about news, activities unfolding around the world. Such user-generated knowledge can be a significant source of evidence for study in a range of areas, such as data science, sociology, psychology, and historical studies, which can be used by researchers to explain actions, patterns, and opinions. Twitter is a microblogging site that has evolved exponentially in the past three years. It is regarded to be an online social network site that numbers about 336 million active monthly users at the beginning of 2018. This paper proposed a conceptual framework for business process optimization through Twitter mining. It starts with collecting tweet posts containing customer reviews through Twitter API. Then, a series of Natural Language Processing techniques is used to clean, pre-process, and analyze the data. The sentiment analysis is used to derive knowledge from the collected reviews to be automatically stored in the generated ontology. Process owners are able to query the ontology using SPARQL queries throughout the business process lifecycle. The conceptual framework aimed to optimize the business process by facilitating knowledge acquisition and transfer. In the future, a case study will be applied to evaluate the efficiency of the framework.

References

[1] Kocbek, M., et al., Business process model and notation: The current state of affairs. Computer Science and Information Systems, 2015. 12(2): p. 509-539.

[2] Scheer, A.-W., ARIS-business process frameworks. 2012: Springer Science & Business Media.

[3] Mincheva, A. Goals and Benefits of Implementing Business Process Management Systems. in Proceedings of International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE). 2018. International Conference on Application of Information and Communication ….

[4] Omotayo, F.O., Knowledge Management as an important tool in Organisational Management: A Review of Literature. Library Philosophy and Practice, 2015. 1(2015): p. 1-23.

[5] Jafari, M. and M. Maleki, A Review on Knowledge Management Process Models in Former Two Decades. Business and Management, 2013. 5(2).

[6] Ternai, K., M. Khobreh, and F. Ansari, An Ontology Matching Approach for Improvement of Business Process Management, in Integrated Systems: Innovations and Applications, M. Fathi, Editor. 2015, Springer International Publishing: Cham. p. 111-130.

[7] Bilal Abu-Salih, P.W.a.K.Y.C., Twitter mining for ontology-based domain discovery incorporating machine learning. Journal of Knowledge Management., 2018.

[8] Hadi, A.S. and E.M. Abdulshaheed, Twitter Information Representation using Resource Description Framework. International Journal of Engineering & Technology, 2019. 8(1.5): p. 497-502.

[9] Sayce, D. Number of tweets per day?. 2016 15 October 2020]; Available from: https://www.dsayce.com/social-media/tweets-day/.