

Accepted Manuscript

A survey on Opinion Summarization Techniques for Social Media

Mohammed Elsaid Moussa, Ensaf Hussein Mohamed, Mohamed Hassan Haggag

PII: S2314-7288(17)30058-2

DOI: [10.1016/j.fcij.2017.12.002](https://doi.org/10.1016/j.fcij.2017.12.002)

Reference: FCIJ 33

To appear in: *Future Computing and Informatics Journal*

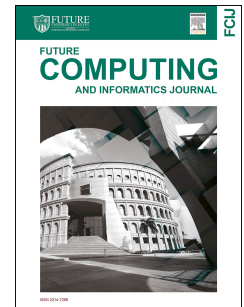
Received Date: 13 October 2017

Revised Date: 23 November 2017

Accepted Date: 19 December 2017

Please cite this article as: Moussa ME, Mohamed EH, Haggag MH, A survey on Opinion Summarization Techniques for Social Media, *Future Computing and Informatics Journal* (2018), doi: 10.1016/j.fcij.2017.12.002.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



A survey on Opinion Summarization Techniques for Social Media

Mohammed Elsaid Moussa ^{a,*}, Ensaf Hussein Mohamed ^b, Mohamed Hassan Haggag ^c

Computer Science Dept., Faculty of Computers and Information, Helwan University, Cairo, Egypt.

*Corresponding author. E-mail addresses: ^a mohammed.elsaid@fci.helwan.edu.eg ,

^b ensaf_hussein@fci.helwan.edu.eg , ^c mohamed.haggag@fci.helwan.edu.eg

Abstract

The volume of data on the social media is huge and even keeps increasing. The need for efficient processing of this extensive information resulted in increasing research interest in knowledge engineering tasks such as Opinion Summarization. This survey shows the current opinion summarization challenges for social media, then the necessary pre-summarization steps like preprocessing, features extraction, noise elimination, and handling of synonym features. Next, it covers the various approaches used in opinion summarization like Visualization, Abstractive, Aspect based, Query-focused, Real Time, Update Summarization, and highlight other Opinion Summarization approaches such as Contrastive, Concept-based, Community Detection, Domain Specific, Bilingual, Social Bookmarking, and Social Media Sampling. It covers the different datasets used in opinion summarization and future work suggested in each technique. Finally, it provides different ways for evaluating opinion summarization.

Keywords: Natural Language Processing, Sentiment Analysis, Opinion Summarization, Social Media, Opinion Mining, Tweet Summarization.

1. Introduction

Sentiment Analysis is a broad area that includes opinion mining, sentiment classification, and opinion summarization; Opinion Summarization is the process of automatically summarizing many opinions that are related to the same topic [1]. In sentiment analysis, Opinion Summarization involves many preprocessing steps such as tokenization, part of speech, stemming; making it different from traditional summarization [2]. It is one of the most valued and powerful NLP technologies [3]. In Social Media, it is about how to locate the most relevant posts with opinions to a given topic [4]. It will permit understanding hidden events and sentiments on different incidents [5]. Sentiment Summarization is also distinct from the factual data summarization, as sentences that were viewed as instructive from the factual point of view may not contain sentiment at all, making them useless from the sentiment perspective [6].

What makes this survey study important is that lately, there is an increased research interest in Opinion Summarization since it has turned into a pattern among individuals to give their sentiments on different features of products in blogs, review posts, and social networks [7]. As an example, on Amazon, some popular items could get thousands of reviews, making it hard for candidate clients to experience every one of the audits to settle on a choice to buy [8]. This large volume of data puts us in need for an automatic opinion summarization system and causes extraordinary challenges on the summarization system [9]. It would be useful for clients and manufacturers if the user reviews could be automatically processed and presented in a summarized form [10].

Opinion Summarization could be easily integrated into real-life applications, which will save users' time and effort [3]. For example, through Twitter opinions, politicians can review their public image and companies can check their customers' feedbacks [9]. It also plays a significant role in the semantic analysis of Social Media and Social Media Analytics [11].

There are two main approaches to generating textual summaries [12]: 1) **Extractive Summarization** in which the summary is consisting entirely of content extracted from the input [11] and 2) **Abstractive Summarization** in which the summary contains some content not present in the source for example as paraphrased material [11].

Developing opinion-tracking systems is commercially valuable [13]. As an example, Summly¹ (a mobile app that sorts news by topics, and uses an analytical tool to condense text) was sold to Yahoo for a reported \$30 million US dollars² making its author Nick D'Aloisio (born November 1, 1995) one of the youngest self-made millionaires ever.

¹ <http://summly.com/index.html>

² <http://www.stuff.co.nz/technology/digital-living/8474218/Teens-multi-million-dollar-Yahoo-payday>

Download English Version:

<https://daneshyari.com/en/article/6855729>

Download Persian Version:

<https://daneshyari.com/article/6855729>

[Daneshyari.com](https://daneshyari.com)