SPRINGER BRIEFS IN APPLIED SCIENCES AND TECHNOLOGY · COMPUTATIONAL INTELLIGENCE

Surbhi Bhatia
Poonam Chaudhary
Nilanjan Dey

Opinion Mining in Information Retrieval

Springer

SpringerBriefs in Applied Sciences and Technology

Computational Intelligence

Series Editor

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

SpringerBriefs in Computational Intelligence are a series of slim high-quality publications encompassing the entire spectrum of Computational Intelligence. Featuring compact volumes of 50 to 125 pages (approximately 20,000–45,000 words), Briefs are shorter than a conventional book but longer than a journal article. Thus Briefs serve as timely, concise tools for students, researchers, and professionals.

More information about this subseries at http://www.springer.com/series/10618

Surbhi Bhatia · Poonam Chaudhary ·
Nilanjan Dey

Opinion Mining in Information Retrieval

Springer

Surbhi Bhatia
Department of Information Systems
King Faisal University
Al Hasa, Saudi Arabia

Nilanjan Dey
Department of Information Technology
Techno India College of Technology
Kolkata, India

Poonam Chaudhary
Department of Computer Science
and Engineering and Information
Technology
The NorthCap University
Gurugram, India

ISSN 2191-530X ISSN 2191-5318 (electronic)
SpringerBriefs in Applied Sciences and Technology
ISSN 2625-3704 ISSN 2625-3712 (electronic)
SpringerBriefs in Computational Intelligence
ISBN 978-981-15-5042-3 ISBN 978-981-15-5043-0 (eBook)
https://doi.org/10.1007/978-981-15-5043-0

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The core part of human actions that play a vital role on the conduct and behavior is the 'opinions.' Opinions can be considered as a feeling, review, sentiment, or assessment of an object, product, or entity. The way of expressing opinions on certain products that people purchase and the services that they receive in the various industries has been transformed considerably because of ubiquitous web-bing. So, people have been inclined to engage themselves more in online shopping. Review sites and social networking sites fascinate people to post feedbacks and reviews online on blogs, Internet forums, review portals, and on many more platforms. These opinions play a very important role for customers and product manufacturers as they tend to give better knowledge of buying and selling by setting positive and negative comments on products and other information which can improve their decision-making policies.

Mining of such opinions has focused the researchers to pay a keen intention in developing such a system which not only collects useful and relevant reviews online in a ranked manner but also produces an effective summary of such reviews collected on different products according to their respective domains. However, there is little evidence that researchers have approached this issue in opinion mining with the intent of developing a system. Existing opinion mining systems lacked several important features such as aspect detection, summaries generated on the basis of aspects, and classification of opinions using emerging learning algorithms. Existing research in sentiment analysis tends to focus on finding out how to classify the opinions using traditional machine learning algorithms and produce a collaborative summary in their respective domains. In spite of an increase in the field of opinion mining and its research, presenting all the areas such as aspect identification, opinion classification, and opinion summarization together by developing a coherent structure as an information retrieval system is still lacking.

This book intends to design a more comprehensive way of building a system to mine opinions and present an integrative framework with hybrid techniques of learning. Consequently, the aim of this book is to discuss the overall novel architecture of developing an opinion system that will address the remaining challenges and provide an overview of how to mine opinions. The people's opinions which are

extracted from the writings on the Web largely depend on the above-mentioned factors, and the summary formed should be structured and concise in the aggregated form covering the important aspects of the product.

Opinion mining is an integral part of many commercial applications and research projects today, in areas ranging from data mining, information retrieval, and machine learning to finding your friends on social networks. In this book, we want to show you how easy it can be to build information retrieval system and how to best go about it. The applications of opinion mining are endless and, with the amount of data available today, mostly limited by your imagination.

I dedicate my work to my beloved husband Mr. Sahil and my daughter Inaaya for encouraging me to pursue my research work. They have truly been there for me to give me all the support that I needed. My special thanks to my friends and colleagues for helping me in writing this research oriented book.

Al Hasa, Saudi Arabia — Surbhi Bhatia
Gurugram, India — Poonam Chaudhary
Kolkata, India — Nilanjan Dey

About This Book

The core part of human actions that play a vital role on the conduct and behavior is the "Opinions." Opinions can be considered as a feeling, review, sentiment, or assessment of an object, product, or entity. The way of expressing opinions on certain products that people purchase and the services that they receive in the various industries has been transformed considerably because of ubiquitous webbing. So, people have been inclined to engage themselves more in online shopping. Review sites and social networking sites fascinate people to post feedbacks and reviews online on blogs, Internet forums, review portals, and on many more platforms. These opinions play a very important role for customers and product manufacturers as they tend to give better knowledge of buying and selling by setting positive and negative comments on products and other information which can improve their decision-making policies.

Mining of such opinions has focused the researchers to pay a keen intention in developing such a system which not only collects useful and relevant reviews online in a ranked manner but also produces an effective summary of such reviews collected on different products according to their respective domains. However, there is little evidence that researchers have approached this issue in opinion mining with the intent of developing a system. Existing opinion mining systems lacked several important features such as aspect detection, summaries generated on the basis of aspects, and classification of opinions using emerging learning algorithms. Existing research in sentiment analysis tends to focus on finding out how to classify the opinions using traditional machine learning algorithms and produce a collaborative summary in their respective domains. In spite of an increase in the field of opinion mining and its research, presenting all the areas such as aspect identification, opinion classification, and opinion summarization together by developing a coherent structure as an information retrieval system is still lacking.

This book intends to design a more comprehensive way of building a system to mine opinions and present an integrative framework with hybrid techniques of learning. Consequently, the aim of this book is to discuss the overall novel

architecture of developing an opinion system that will address the remaining challenges and provide an overview of how to mine opinions. The people's opinions which are extracted from the writings on the Web largely depend on the above-mentioned factors, and the summary formed should be structured and concise in the aggregated form covering the important aspects of the product. A complete dynamic information retrieval system is proposed giving summarized opinions according to the interest of the user based on the features. The fresh reviews are crawled using World Wide Web by proposing novel algorithms. The revisit frequency of the updated Web pages is computed, and the most recent reviews are extracted by opinion retriever. The pre-processing tasks are carried on the opinions for cleaning and removing irrelevant content. After the extraction is done, the following activities are carried out. Firstly, the features from the customer's opinions are identified by using dependency rules of natural language processing. Secondly, the orientation of each review based on the extracted feature is detected using deep learning algorithms. The proposed model of convolutional neural networks is used for binary opinion classification. Finally, abstractive and extractive techniques are used for summarizing the opinions by proposing novel algorithms. Comparative experiments on different datasets (standard and extracted) are conducted, and the accuracy is effectively measured using ROUGE tool.

The prime component that can enhance the quality of services is the opinion of the users. The two vital roles played by humans are sentiments and emotions which they share with their friends, relatives and associations. The ease of Internet access, economical computing devices leads the users to online shopping which has exponentially increased the data about the people's mood and opinion. This makes 'Opinion' as prime component that can enhance the quality of services. The new resources and interactive format of feedback system adopted by review sites and social media has opened huge heterogeneous data sources of user opinions. Researchers found that opinions of people are scattered on various online shopping sites and social media, so instead of traversing the different websites for the reviews with their polarity education of sentences, a complete IR system should be developed which directly gives the summarized opinions.

Opinion Mining in Information Retrieval makes available approaches and techniques for complete opinion oriented Information Retrieval systems and discusses in detail the trends of sentiment analysis emphasizing on "How online review and feedback reflects the opinion of users and lead to the major drift in the decision making process of the organization". The book focuses on providing different machine learning and deep learning approaches to solve the new challenges raised by opinion mining applications with the survey and comparison with traditional fact based studies. It includes the Opinion Score Mining System, Opinion Retrieval, Aspect Extraction. Finally, it gives the application of deep learning and machine learning approaches that develop the opinion oriented Information Retrieval systems along with the benchmark datasets, discussion on available sources, opinion summarization, and future work.

Opinion Mining in Information Retrieval is the advanced comprehensive survey of this challenging yet interesting field of Opinion Mining and will be of interest professionals, researchers and individuals have interest in this area.

Contents