

Chapter 22

Exploiting Chi Square Method for Sentiment Analysis of Product Reviews

Nilesh M Shelke

Priyadarshini Indira Gandhi College of Engineering, India

Shrinivas P Deshpande

DCPE, HVPM, India

ABSTRACT

Sentiment analysis is an extension of data mining which employs natural language processing and information extraction task to recognize people's opinion towards entities such as products, services, issues, organizations, individuals, events, topics, and their attributes. It gives the summarized opinion of a writer or speaker. It has received lot of attention due to increasing number of posts/tweets on social sites. The proposed system is meant to classify a given text of review into positive, negative, or the neutral category. Primary objective of this article is to provide a method of exploiting permutation and combination and chi values for sentiment analysis of product reviews. Publicly available freely dictionary SentiWordNet 3.0 has been used for review classification. The proposed system is domain independent and context aware. Another objective of the proposed system is to identify the feature specific intensity with which reviewer has expressed his opinion. Effectiveness of the proposed system has been verified through performance matrix and compared with other research work.

INTRODUCTION

As the publication of posts/opinion/thoughts/tweets on social sites/blogs/forums became easy, internet sites are flooding with product reviews. Reviews are useful to customers for taking decisions for purchasing, institutes/companies for branding, politicians for making image in the public, etc. This public generated review holds sentiments for various products/persons/movies etc. But their browsing is tedious as they are huge in number. Sentiment Analysis from text is multi-faceted research issue. Even though

DOI: 10.4018/978-1-7998-0951-7.ch022

Exploiting Chi Square Method for Sentiment Analysis of Product Reviews

there are tremendous challenges and many difficulties in sentiment analysis from text; its contribution and applications to industry is not ignorable. Because sentiment analysis bases its results on factors that are so inherently humane, it will become one the major drivers of many business decisions in future. Practical applications of sentiment analysis are huge. Sentiment analysis can be successfully applied to learn user preferences and interests from users' personal writings. These methods are often studied for consumer feedback analysis. Similarly, e-learning systems can benefit from affective tutoring approaches.

Sentiment Analysis is an extension of data mining which employs Natural Language Processing and Information Extraction task to recognize people's opinion towards entities such as products, services, issues, organizations, individuals, events, topics and their attributes. Sentiment Analysis analyze each text and detect which part contain opinionated word, which is being opinionated and who has written the opinion. It gives the summarized opinion of a writer or speaker. Sentient analysis can be done at word level, sentence level and document level. In a broader sense, sentiment analysis is focused to detect the attitude of a speaker or a writer with respect to some topic or about overall coverage of a document (Zhao et al., 2015). Sentiment can be defined as

Sentiment:= $\langle o_j, f_{jk}, so_{ijkl}, h_i \rangle$,

o_j is the target: what/whom the sentiment is expressed to.

f_{jk} is a feature of the object o_j ,

so_{ijk} is the sentiment value of the opinion of the opinion holder h_i on feature f_{jk} of object o_j ; so_{ijkl} is +ve, -ve, or neutral, or a more granular rating.

h_i is holder: who expresses the sentiment (Pawar, 2016).

Identification of sentiments or moods embedded in the (written) natural language is critical and a challenging issue. Researchers are attempting to apply different approaches and techniques and extract useful information from this data.

As such demand of sentiment analysis applications is increasing rapidly. Primary tasks in sentiment analysis is finding the "aspects" or "features", finding their corresponding sentiments, identifying the intensity of the sentiments and finally calculating the overall opinion of the reviewer. Consider some actual reviews from the dataset.

1. The bread is top *notch* as well.
2. In fact, this was not a Nicosia salad and was *barely eatable*.
3. I have to say they have one of the *fastest* delivery times in the city.
4. The food is always *fresh* and *hot* - ready to eat!
5. Did I mention that the coffee is *outstanding*?
6. Certainly not the *best* sushi in New York, however, it is always *fresh* food, and the place is *very clean, sterile*.

In the above example reviews, underlined terms are the product features and italic terms indicates the sentiments about those features termed as sentiment features.

16 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/exploiting-chi-square-method-for-sentiment-analysis-of-product-reviews/239948

Related Content

Understanding Quantum Computing Implications for Cybersecurity

Angel Justo Jones (2025). *Leveraging Large Language Models for Quantum-Aware Cybersecurity* (pp. 29-66).

www.irma-international.org/chapter/understanding-quantum-computing-implications-for-cybersecurity/366975

Author Profiling Using Texts in Social Networks

Iqra Ameerand Grigori Sidorov (2021). *Handbook of Research on Natural Language Processing and Smart Service Systems* (pp. 245-265).

www.irma-international.org/chapter/author-profiling-using-texts-in-social-networks/263105

NLP and the Representation of Data on the Semantic Web

Jose L. Martinez-Rodriguez, Ivan Lopez-Arevalo, Jaime I. Lopez-Veyna, Ana B. Rios-Alvaradoand Edwin Aldana-Bobadilla (2021). *Handbook of Research on Natural Language Processing and Smart Service Systems* (pp. 393-426).

www.irma-international.org/chapter/nlp-and-the-representation-of-data-on-the-semantic-web/263114

Exploring the Relationship Between Customer Services and Fintech Adoption Among Generation X: An Empirical Study in Odisha, India.

Dusmant Kumar Sahoo, B.C.M. Patnaik, Ipseeta Satpathy, Vishal Jainand Shiva Ram Patnaik (2025). *Intersecting Natural Language Processing and FinTech Innovations in Service Marketing* (pp. 203-216).

www.irma-international.org/chapter/exploring-the-relationship-between-customer-services-and-fintech-adoption-among-generation-x/377508

The Media-Sphere as Dream: Researching the Contextual Unconscious of Collectives

Stephen Brock Schafer (2020). *Natural Language Processing: Concepts, Methodologies, Tools, and Applications* (pp. 343-371).

www.irma-international.org/chapter/the-media-sphere-as-dream/239944