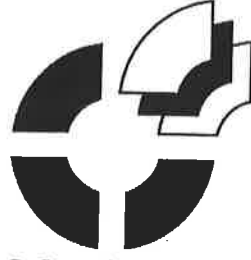


SENTIMENT ANALYSIS ON ONLINE PRODUCT
REVIEWS USING MACHINE LEARNING TECHNIQUES



सिद्धिमूलं प्रबन्धनम्
भा. प्र. सं. इन्दौर
IIM INDORE

A Doctoral Dissertation Submitted in Partial Fulfillment of the Requirements for the
Fellow Programme in Management

Indian Institute of Management Indore

By
Anuj Sharma

March 2013

Thesis Advisory Committee:

S. Dey
Prof. Shubhamoy Dey
(Chair)

Prabin K. Panigrahi
Prof. Prabin K. Panigrahi
(Member)

Ashish Sadh
Prof. Ashish Sadh
(Member)

Abstract

The growing popularity of Internet and Web 2.0 social media have led to the advent of many e-commerce websites, discussion forums and Weblogs. These in turn have facilitated the sharing of user provided feedback about post-purchase experience; specifically related to various products and services. This user-generated content comprises of opinions, appraisals, recommendations and evaluations associated with virtually anything people care about in products or services. The opinions and post purchase experiences shared by users through online product reviews constitute a major part of online word-of-mouth (WOM) communication. Online WOM is valuable to potential consumers for making product choices and purchase decisions. At the same time, by analyzing these reviews, business organizations can gain insights into what people are discussing about their products and services. Thus, the analysis of customer sentiments from 'freely available' online reviews can be a potentially cost effective and time efficient solution for eliciting consumer preferences. Also, deeper exploration of user opinions and feedback may lead to the discovery of interesting pattern of product usage (e.g., brand experience), product weaknesses and product-feature related opinion.

As more and more user-generated reviews are created and aggregated, a strong demand for automatic approaches capable of extracting overall as well as specific opinion from these unstructured texts has emerged. Sentiment analysis, often referred to as opinion mining, is a recent area of active research. It deals with the computational treatment of opinion and extraction of subjectivity knowledge from online user-generated content. Thus, sentiment analysis is the task of retrieving aggregated and fine-grained opinions related to an object or its attributes as expressed by users in the form of free text. However, there are many problems and challenges associated with extracting meaningful opinions articulated in unstructured user-generated texts, like product reviews.

Based on the broad objective of mining sentiments and opinions from online reviews, four conjoint studies were conducted. As the main step, a comprehensive study of automatic extraction of overall and fine-grained opinions from online reviews is presented. This study focuses on sentiment based product review classification to

discover product-sentiment. As an output, it identifies reviewed product(s) as recommended/not-recommended along with extraction of discussed features, feature-level opinion mining, and opinion summarization and visualization.

Sentiment based classification of text documents is a more challenging task compared to traditional topic based classification. Discrimination based on opinions, feelings, and attitudes is inherently more complex than classification based on topics due to the semantic relationships of the natural language involved. Further, extraction of the features or attributes about which opinion has been expressed is one of the major challenges of opinion mining. Feature-level opinion mining aims at identifying the relevant opinions associated with specific features or attributes of a product or service from a set of reviews. However, identifying and determining the relevance of features and the accuracy of the expressed opinion continue to pose challenges for this task.

This research work addresses some of the critical issues related to sentiment based classification of online reviews. Document-level sentiment analysis using supervised machine learning techniques faces many challenges like feature selection, dimensionality reduction, sentiment based visualization and domain dependency of sentiments. Text sentiment classification requires deep analysis and understanding of textual features and natural language semantics. Therefore, a part of this work has been devoted to the empirical comparative study of the applicability of feature selection techniques to sentiment analysis of text documents. This study also compares the performance of different machine learning classifiers on a benchmark dataset for document-level sentiment analysis and explores the synergy between feature selection techniques and various machine learning based classifiers.

Another contribution of this research to existing literature is the formulation of novel sentiment classification models using back-propagation artificial neural network (BPANN) and self-organizing maps (SOM). Domain independent sentiment classification models exploit sentiment lexicons in an attempt to classify online reviews from diverse domains. We have investigated the problems associated with domain dependency through sentiment analysis on documents from two different domains. By using large sentiment lexicons along with appropriate handling of negation, this study has shown that encouraging results are obtainable for domain independent sentiment analysis. Further, we have also demonstrated the efficacy of

supervised and unsupervised self-organizing map based approaches for sentiment based classification and visualization of opinion in text documents. Finally, the study establishes how the proposed sentiment analysis framework can be successfully employed for deriving marketing intelligence from online product reviews.

Sentiment analysis may be as simple as overall sentiment based categorization of text documents; but could as well be more complex and advanced procedures to extract opinion at different granularity levels. All the document-level and feature-level sentiment analysis approaches described in this study have been tested on a publicly available benchmark dataset and a real-life dataset created by us. The proposed methods have been found to yield significantly better accuracy in dealing with online subjective text compared to those previously reported. Thus, we have devised an effective way of domain-independent opinion summarization from online customer reviews using our unified framework for opinion retrieval, classification and summarization at various granularity levels.

Keywords: Machine Learning, Feature Selection, Sentiment Analysis, Classification, Opinion Mining, Performance, Experimentation.

Table of Contents

Abstract	i
Acknowledgements	iv
List of Tables	viii
List of Figures	ix
List of Abbreviations	x
1. Introduction	1
1.1. Sentiment and Opinion	2
1.2. Sentiment Analysis and Related Terminology	3
1.3. Motivation	5
1.4. Research Problems	7
1.5. Research Objective	10
1.6. Contributions of the Work	11
1.7. Business and Managerial Implications	11
2. Literature Review	13
2.1. Machine Learning Approach	14
2.1.1. Machine Learning for Document-level Sentiment Analysis	15
2.1.2. Machine Learning for Sentence-level Sentiment Analysis	21
2.1.3. Machine Learning for Feature-level Sentiment Analysis	24
2.2. Dictionary Approach	26
2.2.1. Dictionary Approaches for Document-level Sentiment Analysis	26
2.2.2. Dictionary Approaches for Sentence-level Sentiment Analysis	28
2.2.3. Dictionary Approaches for Feature-level Sentiment Analysis	29
2.3. Statistical Approach	31
2.3.1. Statistical Approaches for Document-level Sentiment Analysis	31
2.3.2. Statistical Approaches for Sentence-level Sentiment Analysis	32
2.3.3. Statistical Approaches for Feature-level Sentiment Analysis	33
2.4. Semantic Approach	34
2.4.1. Semantic Approaches for Document-level Sentiment Analysis	34

2.4.2.	Semantic Approaches for Sentence-level Sentiment Analysis	35
2.4.3.	Semantic Approaches for Feature-level Sentiment Analysis.....	36
2.5.	Applications of Sentiment Analysis and Opinion Mining	36
2.6.	Issues and Challenges of Sentiment Analysis Approaches	37
2.7.	Issues Addressed in this Research.....	39
3.	Research Methodology	40
3.1.	Opinionated Text Corpus	41
3.2.	Textual Pre-processing.....	42
3.3.	Sentiment Analysis.....	43
3.3.1.	Opinion Related Resource Generation.....	43
3.3.2.	Sentiment based Classification of Reviews	44
3.3.3.	Feature-level Opinion Mining.....	45
3.4.	Business Intelligence as Overall Consumer Sentiment.....	45
3.5.	Implications of the Results.....	46
4.	Feature Selection and Machine Learning Techniques for Sentiment Analysis ...	48
4.1.	Introduction	48
4.2.	Machine Learning Classifiers with Feature Selection.....	48
4.3.	Feature Selection Methods.....	49
4.3.1.	Document Frequency (DF)	50
4.3.2.	Information Gain (IG).....	50
4.3.3.	Gain Ratio (GR).....	51
4.3.4.	CHI Statistic (CHI)	52
4.3.5.	Relief-F Algorithm.....	52
4.4.	Machine Learning Techniques for Sentiment Classification	52
4.4.1.	Naïve Bayes (NB).....	52
4.4.2.	Support Vector Machine (SVM).....	53
4.4.3.	Maximum Entropy	54
4.4.4.	Decision Tree.....	54
4.4.5.	K-Nearest Neighbor Classifier.....	55
4.4.6.	Winnow Classifier	56
4.4.7.	Adaboost Classifier.....	56

4.5.	Dataset and Experimental Design	57
4.6.	Experimental Results.....	58
4.7.	Conclusions.....	62
5.	Artificial Neural Network Based Approach for Sentiment Analysis.....	64
5.1.	Introduction	64
5.2.	Need of BPANN for Sentiment Analysis.....	64
5.3.	Back-Propagation Artificial Neural Network (BPANN)	66
5.4.	Experimental Design and Datasets.....	69
5.4.1.	Corpora and Sentiment Lexicons.....	70
5.4.2.	Performance Evaluation.....	71
5.5.	Experimental Results.....	72
5.5.1.	Results on Movie Reviews Dataset.....	72
5.5.2.	Results on Hotel Reviews Dataset	74
5.6.	Discussion	76
5.7.	Conclusion.....	77
6.	Sentiment Analysis Using Self-Organizing Maps	79
6.1.	Introduction	79
6.2.	Need of SOMs for Sentiment Analysis.....	79
6.3.	Self-Organizing Maps (SOM).....	81
6.3.1.	Learning Algorithm for SOM	82
6.4.	Learning Vector Quantization (LVQ).....	84
6.4.1.	Learning Algorithm for LVQ.....	85
6.5.	Self Organizing Maps for Sentiment Visualization	86
6.6.	Experimental Design and Dataset	89
6.7.	Experimental Results and Discussion	91
6.8.	Conclusions	95
7.	Sentiment Analysis for Mining Marketing Intelligence	96
7.1.	Introduction	96
7.1.1.	Traditional Word-of-Mouth (WOM) Communication	97

7.1.2.	Online or Electronic Word-of-Mouth (Online WOM or eWOM).....	98
7.2.	Importance of Online Reviews for Marketing Intelligence	99
7.3.	Studies Related to eWOM and Online Product Reviews.....	100
7.4.	Sentiment based Classification of Online Reviews.....	106
7.5.	Feature-level Opinion Mining.....	107
7.5.1.	Extracting Features and Opinion Words.....	108
7.6.	Dataset and Performance Evaluation Metrics	110
7.7.	Experimental Results and Discussion	111
7.8.	Managerial Implications.....	118
7.9.	Conclusions	120
8.	Conclusions, Limitations and Outlook for Future Research.....	121
8.1.	Conclusions.....	121
8.2.	Major Contributions of the Thesis	122
8.3.	Limitations of the Current Research	123
8.4.	Outlook for Future Research.....	125
	References.....	127

List of Tables

Description	Page No.
Table 2.1.1: Machine Learning for Document-level Sentiment Analysis	19
Table 2.1.2: Machine Learning for Document-level Sentiment Analysis (2008-2010)	20
Table 2.1.3: Machine Learning for Document-level Sentiment Analysis (2011-2012)	21
Table 2.2: Machine Learning for Sentence-level Sentiment Analysis	22
Table 2.3: Machine Learning for Feature-level Sentiment Analysis	25
Table 2.4: Dictionary Approaches for Document-level Sentiment Analysis	27
Table 2.5: Dictionary Approaches for Sentence-level Sentiment Analysis	28
Table 2.6: Dictionary Approaches for Feature-level Sentiment Analysis	30
Table 2.7: Statistical Approaches for Document-level Sentiment Analysis	32
Table 2.8: Statistical Approaches for Sentence-level Sentiment Analysis	33
Table 2.9: Statistical Approaches for Feature-level Sentiment Analysis	33
Table 2.10: Semantic Approaches for Document-level Sentiment Analysis	35
Table 2.11: Semantic Approaches for Sentence-level Sentiment Analysis	35
Table 2.12: Semantic Approaches for Feature-level Sentiment Analysis	36
Table 4.1: Best Accuracy (in Percentage) for Different Classifiers	58
Table 5.1: The Confusion Matrix	71
Table 7.1: Studies Related to eWOM and Online Product Reviews (1)	104
Table 7.2: Studies Related to eWOM and Online Product Reviews (2)	105
Table 7.3: Feature-Opinion Tuple Extraction Rules	108
Table 7.4: Top 100 Frequent Features Extracted from 2000 Hotel Reviews	114
Table 7.5: Feature-Opinion Tuple for Top 5 Features	118

List of Figures

Description	Page No.
Figure 3.1: The Framework for Sentiment Analysis on Online Reviews	41
Figure 4.1: Performance of Machine Learning Classifiers with Document Frequency	60
Figure 4.2: Performance of Machine Learning Classifiers with Information Gain	60
Figure 4.3: Performance of Machine Learning Classifiers with Gain Ratio	60
Figure 4.4: Performance of Machine Learning Classifiers with CHI Statistic	61
Figure 4.5: Performance of Machine Learning Classifiers with Relief-F	61
Figure 5.1: Typical Three Layered BPANN for Sentiment Classification	67
Figure 5.2: Accuracy of BPANN based Sentiment Analysis on Movies Reviews	73
Figure 5.3: Comparison on Precision for Movies Reviews	73
Figure 5.4: Comparison on Recall for Movies Reviews	73
Figure 5.5: Accuracy of BPANN based Sentiment Analysis on Hotel Reviews	75
Figure 5.6: Comparison on Precision for Hotel Reviews	75
Figure 5.7: Comparison on Recall for Hotel Reviews	75
Figure 6.1: ESOM of Document Vectors with 100 Features	92
Figure 6.2: ESOM of Document Vectors with 700 Features	92
Figure 6.3: Comparison on Accuracy of SOM and LVQ Algorithms	93
Figure 6.4: Comparison on Precision of SOM and LVQ Algorithms	93
Figure 6.5: Comparison on Recall of SOM and LVQ Algorithms	94
Figure 6.6: Comparison on <i>F1</i> Score of SOM and LVQ Algorithms	94
Figure 7.1: Sentiment Classification Performance for Negative Online Reviews	112
Figure 7.2: Sentiment Classification Performance for Positive Online Reviews	112
Figure 7.3: Representing Feature Buzz with Top 30 Features in Online Hotel Reviews	114
Figure 7.4: Top 30 Most Frequent Negative Sentiment Words in Hotel Reviews	115
Figure 7.5: Top 30 Most Frequent Positive Sentiment Words in Hotel Reviews	116
Figure 7.6: Overall Positive and Negative Sentiment Words Used in Feature-Opinion Tuple	116