Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 8, Issue, 03, pp. 28352-28356, March, 2016

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

ASPECT AND EMOTION CLASSIFICATION OF RESTAURANT AND LAPTOP REVIEWS USING SVM

[1,*]Kirange, D. K. and [2]Dr. Ratnadeep R. Deshmukh

[1]Department of Computer and IT, J T Mahajan College of Engineering, Faizpur, Tal, Yawal, Jalgaon, India
[2]Department of Computer Science and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabaad, India

ARTICLE INFO

Article History:
Received 24th December, 2015
Received in revised form
27th January, 2016
Accepted 29th February, 2016
Published online 31st March, 2016

Key words:

Aspect,
Sentiment Analysis,
SVM,
KNN,
Naïve Bayes,
Neural Network.

ABSTRACT

The "Sentiment Analysis" task focuses on the recognition and classification of emotions (positive, negative, conflict, neutral) in reviews for the aspect. In this paper we propose the system for recognizing and analyzing the sentiments using SVM for the restaurant and laptop review dataset. Here we also apply the proposed system of text classification for identifying the aspect in the review sentences. We compare the performance of the system with well-known KNN, Naïve Bayes and Neural Network classifiers.

Citation: Kirange, D. K. and Dr. Ratnadeep R. Deshmukh, 2016. "Aspect and emotion classification of restaurant and laptop reviews using SVM", International Journal of Current Research, 8, (03), 28352-28356.

INTRODUCTION

Sentiment analysis (or opinion mining) is defined as the task of finding the opinions of authors about specific entities. The decision-making process of people is affected by the opinions formed by thought leaders and ordinary people. When a person wants to buy a product online he or she will typically start by searching for reviews and opinions written by other people on the various offerings. Sentiment analysis is one of the hottest research areas in computer science. A sentiment analysis model is used to analyze a text string and classify it with one of the labels that you provide; for example, you could analyze a tweet to determine whether it is positive or negative, or analyze an email to determine whether it is happy, frustrated, or sad. Aspect-based sentiment analysis is the research problem that focuses on the recognition of all sentiment expressions within a given document and the aspects to which they refer. It is common to classify sentences into two principal classes with regard to subjectivity: objective sentences that contain factual information and subjective sentences that contain explicit opinions, beliefs, and views about specific entities.

Corresponding author: Kirange, D. K.
Department of Computer and IT, J T Mahajan College of Engineering, Faizpur, Tal, Yawal, Jalgaon, India

Here, I mostly focus on analyzing subjective sentences. As an example, here is a review about a hotel in Manhattan. "The king suite was spacious, clean, and well appointed. The reception staff, bellmen, and housekeeping were very helpful. Requests for extras from the maid were always provided. The heating and air conditioning functioned well; this was good as the weather was variable. The sofa bed was the best I've ever experienced. The king size bed was very comfortable. The building and rooms are very well soundproofed. The neighborhood is the best for shopping, restaurants, and access to subway. Only "complaint" has to do with high-speed Internet access. It's only available on floors 8–12." Overall the review is very positive about the hotel. It refers to many different aspects of the hotel including: heating, air conditioning, staff courtesy, and bed, neighborhood, and Internet access. In this paper we propose sentiment analysis systems which are able to provide a sentiment score for the whole review as well as analyze the sentiment of each individual aspect of the hotel. In this paper we propose a sentiment analysis task of Sem Eval which deals with the restaurant and laptops review dataset. Particularly we have considered the four subtasks

- Identification of sentiment term which is usually present in the sentence,

- recognition of sentiment category and
- Identification of aspect term which is present in the sentence,
- Recognition of aspect category.

The restaurant reviews datasets provided by SemEval 2014, SemEval 2015 and SemEval 2016 are used for classification of data using SVM, Naïve Bayes, KNN and Neural Network. Here we have classified the reviews data in various aspects such as Ambience, Food, Price, Service, Drinks, General, Location, and Misc. The experimental evaluations predict the usefulness of SVM in text classification as compared to other classification techniques. The text classification system is applied for emotion classification of restaurants and laptops reviews data provided by SemEval 2014, SemEval 2015 and SemEval 2016. The dictionary of WordNet Affect is used for emotional keywords. Here we have classified the reviews data in various emotions such as Anger, Disgust, Fear, Joy, Sadness, Surprise, Positive and Negative. The experimental evaluations predict the usefulness of SVM in text classification as compared to Naïve Bayes, Neural Network and KNN.

Related Work

In (Mesut Kaya, 2012) sentiment classification techniques are incorporated into the domain of political news from columns in different Turkish news sites. Authors have compared four supervised machine learning algorithms of Naïve Bayes, Maximum Entropy, SVM and the character based N-Gram Language Model for sentiment classification of Turkish political columns. We Also the problem of sentiment classification in the political news domain is discussed. It is observed from empirical findings that the Maximum Entropy and N-Gram Language Model outperformed the SVM and Naïve Bayes. Using different features, all the approaches reached accuracies of 65% to 77%. A generative probabilistic topic model is proposed in (Noriaki Kawamae, 2012) that detects both an aspect and corresponding sentiment, simultaneously, from review articles. Unlike existing sentiment analysis models, which generally consider rating prediction to be a side task, the proposal (Noriaki Kawamae, 2012), the hierarchical approach to sentiment analysis, identifies both an item and its rating by dividing topics, traditionally treated as one entity, into aspect and sentiment topics. Since the model is aware of both objective and subjective information, it can discover finegrained tightly coherent topics, and describe the generative process of each article in a unified manner. A Portuguese dictionary focused in a specific field of study was built in (Renata Lopes Rosa, 2013), in which tenses and negative words are treated in a different way to measure the polarity, the strength of positive or negative sentiment, in short texts extracted from Twitter. For the Portuguese Dictionary performance validation, the results are compared with the SentiStrength tool and are evaluated by three Specialists in the field of study; each one analyzed 2000 texts captured from Twitter. Comparing the efficiency of the SentiMeter-Br and the SentiStrength against the Specialists' opinion, a Pearson correlation factor of 0.89 and 0.75 was reached, respectively, proving that the metric used in the Sentimeter-Br is better than the one used in the Senti Strength. The polarity of the short texts were also tested through machine learning, with correctly classified instances of 71.79% by Sequential Minimal Optimization algorithm and F-Measure of 0.87 for positive and 0.91 for negative phrases. In (Jian Chen *et al.,* 2013) a method to handel sentiment Analysis is proposed for Cantonese opinion mining. In (Jian Chen *et al.,* 2013) authors use Hidden markov model (HMM) to conduct word segmentation, and then building opinion orientation dictionary for Cantonese. Also a method for feature orientation summarization is proposed. It is the first work of sentiment analysis for Cantonese opinion mining. The study in (Wenjing Duan *et al.,* 2013) aims to look beyond the quantitative summary to provide a more comprehensive view of online user-generated content. Authors obtain a unique and extensive dataset of online user reviews for hotels across various review sites and over a long time periods. Authors in (Jian Chen, 2013) use the sentiment analysis technique to decompose user reviews into five dimensions to measure hotel service quality. Those dimensions are then incorporated into econometrics models to examine their effect in shaping users' overall evaluation and content generating behavior.

Task Definition

Sentiment analysis (also known as opinion mining) refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials. A basic task in sentiment analysis is classifying the *polarity* of a given text at the document, sentence, or feature/aspect level whether the expressed opinion in a document, a sentence or an entity feature/aspect is positive, negative, or neutral. Sentiment analysis is increasingly viewed as a vital task both from an academic and a commercial standpoint. The majority of current approaches, however attempt to detect the overall polarity of a sentence, paragraph or text span, regardless of the entities mentioned (e.g. restaurants, laptops) and their aspects (e.g food, price, service). In contrast this task focuses on sentiment analysis where the goal is to identify aspect of the given target entity and sentiment expressed in each review. In particular, the paper focuses on four tasks of SemEval Conference.

Subtask 1: Sentiment Term Extraction

Given a set of sentences with pre-defined entities (e.g. laptops), identify the sentiment terms present in the sentence and return a list containing all the distinct sentiment terms. For example, "I *liked* the battery and the display, but not the sound,

Subtask 2: Sentiment Term Polarity

For a given set of sentiment terms within a sentence determine whether the polarity of each sentiment term is positive, negative, neutral or conflict. For example, I loved their Products

Subtask 3: Aspect Term Extraction

Given a set of sentences with pre-defined entities (e.g. laptops), identify the aspect terms present in the sentence and return a list containing all the distinct aspect terms. For example, "I liked the *battery* and the *display*, but not the *sound*,

Subtask 4: Aspect Term Classification

For a given set of aspect terms within a sentence classify the aspect in to its appropriate category.

Aspect Classification

The restaurant reviews datasets provided by SemEval 2014, SemEval 2015 and SemEval 2016 are used for classification of data using SVM, Naïve Bayes, KNN and Neural Network. Here we have classified the reviews data in various aspects such as Ambience, Food, Price, Service, Drinks, General, Location, and Misc. The experimental evaluations predicts the usefulness of SVM in text classification as compared to other classification techniques.

A. Dataset

Restaurant reviews

This dataset consists of over 3K English sentences from the restaurant reviews. The dataset included annotations for coarse aspect categories and overall sentence polarities; annotations for aspect terms occurring in the sentences, aspect term polarities, and aspect category-specific polarities.

B. Evaluations and Results

We have identified the various aspect categories as food, service, price, ambience, location, positive, negative etc. The well known SVM classifier is used for classification of a review sentence in appropriate category. The classification results for Semeval 2014 Restaurant Reviews Dataset are compared with the KNN classifier as shown in the table 2. The classification results for Semeval 2015 Restaurant Reviews Dataset are compared with the KNN, Naïve Bayes, and Neural Network classifier as shown in the Table 3.

Table 1. Datasets used for aspect category classification

Name of the Dataset	Number of Training Samples	Number of Testing Samples
Semeval 2014 Restaurant Reviews Dataset	3000	801
Semeval 2015 Restaurant Reviews Dataset	1846	949
Semeval 2016 Restaurant Reviews Dataset	74	74

Table 2. Classification Accuracy for Semeval 2014 dataset for aspect Term Classification

Aspect Category	Accuracy Using KNN	Accuracy Using SVM
Ambience	55.5556	91.6256
Food	87.0229	81.2808
Price	44.4444	94.5813
Service	61.2903	92.6108
Misc	60.9756	100.0000
Average	61.8578	92.0197

Table 3. Classification Accuracy for Semeval 2015 dataset for aspect Term Classification

Aspect Category	Accuracy Using SVM	Accuracy Using KNN	Accuracy Using Naïve Bayes	Accuracy Using Neural Network
Ambience	93.53	43.678	76.381	95.846
Drinks	100.0	26.562	76.000	72.897
Food	100.0	41.89	84.548	92.156
Location	50.00	0	76.000	10.526
Positive	94.94	23.333	92.202	83.720
Negative	91.42	10.344	100.00	73.170
General	100.0	9.3750	100.00	75.675
Service	55.07	0	76.000	94.871

Table 4. Classification Accuracy for Semeval 2016 dataset for aspect Term Classification

Aspect Category	Accuracy Using SVM	Accuracy Using KNN	Accuracy Using Naïve Bayes	Accuracy Using Neural Network
Ambience	100	100.0	100.00	100.00
Drinks	55.55	100.0	100.00	100.00
Food	100.0	100.0	73.584	100.00
Location	100.0	100.0	33.33	100.00
Positive	100.0	100.0	80.000	100.00
Negative	100.0	100.0	100.00	100.00
General	100.0	100.0	33.333	100.00
Service	100.0	100.0	33.33	100.00

Table 5. Datasets used for Emotion Classification

Name of the Dataset	Number of Training Samples	Number of Testing Samples
Semeval 2014 Restaurant Reviews Dataset	3000	801
Semeval 2015 Restaurant Reviews Dataset	1846	949
Semeval 2016 Restaurant Reviews Dataset	74	74
Semeval 2015 Laptops Reviews Dataset	2315	1065
Semeval 2016 Laptops Reviews Dataset	78	78

The classification results for Semeval 2016 Restaurant Reviews Dataset are compared with the KNN, Naïve Bayes, and Neural Network classifier as shown in the Table 4.

Emotion Classification: The task focuses on identification of sentiment term from the sentence, then accordingly identifying polarity for the laptop review sentence. We have used the dictionary based approach for recognizing the sentiment terms present in the sentence. To analyze the sentiment of the sentence we have used SVM classifier. The sentiments are classified as positive, negative, conflict, neutral etc.

Restaurants and Laptops Reviews Datasets for Emotion Classification

Restaurant Reviews

This dataset consists of over 3K English sentences from the restaurant reviews. The dataset included annotations for coarse aspect categories and overall sentence polarities; annotations for aspect terms occurring in the sentences, aspect term polarities, and aspect category-specific polarities.

Table 6. Emotion Classification for SemEval 2014 Restaurant Reviews Dataset

Aspect Category	Accuracy Using KNN	Accuracy Using SVM
Positive	82.6923	70.9360
Negative	43.5897	80.7882
Neutral	66.6667	90.6404
Conflict	87.0968	99.5074

Table 7. Emotion Classification for SemEval 2015 Restaurant Reviews Dataset

Aspect Category	Accuracy Using SVM	Accuracy Using KNN	Accuracy Using Naïve Bayes	Accuracy Using Neural Network
Anger	99.81	69.867	89.594	98.727
Disgust	71.42	0	76.470	78.571
Fear	98.09	51.219	75.555	91.588
Joy	77.77	18.181	92.857	93.333
Sadness	62.50	6.6667	25.000	22.222
Surprise	80.76	0	100.00	89.361
Positive	98.85	26.363	69.697	97.321
Negative	28.00	16.824	19.236	5.5556

Table 8. Emotion Classification for SemEval 2016 Restaurant Reviews Dataset

Aspect Category	Accuracy Using SVM	Accuracy Using KNN	Accuracy Using Naïve Bayes	Accuracy Using Neural Network
Anger	99.81	69.867	89.594	98.727
Disgust	71.42	0	76.470	78.571
Fear	98.09	51.219	75.555	91.588
Joy	77.77	18.181	92.857	93.333
Sadness	62.50	6.6667	25.000	22.222
Surprise	80.76	0	100.00	89.361
Positive	98.85	26.363	69.697	97.321
Negative	28.00	16.824	19.236	5.5556

Table 9. Emotion Classification for SemEval 2015 Laptops Reviews Dataset

Aspect Category	Accuracy Using SVM	Accuracy Using KNN	Accuracy Using Naïve Bayes	Accuracy Using Neural Network
Anger	99.70	73.765	99.064	100.00
Disgust	28.57	0	23.854	50.000
Fear	92.85	25.531	45.578	76.744
Joy	100.0	19.403	98.666	90.476
Sadness	28.35	20.000	36.000	73.076
Surprise	97.05	10.294	64.000	80.000
Positive	100.0	18.120	80.794	99.242
Negative	5.263	0	3.2259	2.5689

Table 10. Emotion Classification for SemEval 2016 Laptops Reviews Dataset

Aspect Category	Accuracy Using SVM	Accuracy Using KNN	Accuracy Using Naïve Bayes	Accuracy Using Neural Network
Anger	96.29	100.00	100.00	100.00
Disgust	30.00	100.00	100.00	100.00
Fear	100.0	100.00	50.000	100.00
Joy	100.0	100.00	100.00	100.00
Sadness	100.0	100.0	87.500	100.00
Surprise	71.42	100.00	0	100.00
Positive	100.0	100.00	67.567	100.00
Negative	100.0	100.00	0	100.00

Laptop reviews

This dataset consists of over 3K English sentences extracted from customer reviews of laptops. Experienced human annotators tagged the aspect terms of the sentences and their polarities. Part of this dataset will be reserved as test data.
Table 5 the details about the dataset used.

Evaluations and Results

We have identified the various emotion categories as anger, disgust, fear, joy, sadness, surprise, positive, negative etc. The well-known SVM classifier is used for classification of a review sentence in appropriate category. The classification results for SemEval 2014 Restaurant Reviews Dataset are compared with the KNN classifier as shown in the table 6. The classification results for Semeval 2015 Restaurant Reviews Dataset are compared with the KNN, Naïve Bayes, and Neural Network classifier as shown in the table 7. The classification results for Semeval 2016 Restaurant Reviews Dataset are compared with the KNN, Naïve Bayes, and Neural Network classifier as shown in the table 8. The classification results for Semeval 2015 Laptops Reviews Dataset are compared with the KNN, Naïve Bayes, and Neural Network classifier as shown in the table 9. The classification results for Semeval 2016 Laptops Reviews Dataset are compared with the KNN, Naïve Bayes, and Neural Network classifier as shown in the table 10

Conclusion

This research paper investigates the emotion recognition in text by using various techniques. Experiments have been conducted using various classifiers such as SVM, Naïve Bayes and KNN and the results for emotion recognition from text varied. The main contributions of this research include aspect based emotion recognition from text using SVM. In the paper, we propose a method to handle sentiment analysis task of SemEval 2014, SemEval 2015 and SemEval 2016 workshop for Restaurant and Laptop review dataset. In our methods, we use Support Vector Machines (SVM) for recognizing sentiment categories and sentiment polarity as well as aspect term amd aspect category. The sentiment polarities are classified as anger, disgust, far, joy, sadness, surprise, positive, negative, neutral and conflict etc. The classification results are compared with KNN, Naïve Bayes and Neural Network classifiers. Due to its excellent performance, the proposed research work may be applicable to emotion recognition tasks in other Indian languages as well. More experiment will be carried out in this direction.

Acknowledgment

The authors would like to welcome the reviewers for their critical and constructive comments and suggestions. The authors would like to thank the Department of Computer Science of Dr. Babasaheb Ambedkar Marathwada University, Aurangabaad for providing the infrastructure to carry out the research.

REFERENCES

Kirange, D. K., Dr, R. R. Deshmukh, "Emotion Classification of news headlines using SVM" *Asian Journal of Computer Science and Information Technology,* 2 (5)

Deshmukh, R. R., Kirange D. K. "Classifying News Headlines for Providing User Centered E-Newspaper using SVM", *International Journal of Emerging Trends in Technology and Computer Science* (*IJETTCS*), June 2013

http://alt.qcri.org/semeval2014/task4/

http://www.WordNetAffect.net

Jian Chen, Yu Liu, Guangyi Zhang, Sentiment Analysis for Cantonese Opinion Mining, 2013 Fourth International Conference on Emerging Intelligent Data and Web Technologies.

Kirange D. K., Dr. Deshmukh R. R., "Aspect Based Sentiment Analysis: Semeval 2014- Task-4", Asian Journal of Computer Science and Information Technology (AJCSIT) Vol 4, No 8 (2014)

Mesut Kaya, Guven Fidan, Ismail H. Toroslu1J., Sentiment Analysis of Turkish Political News, 978-0-7695-4880-7/12, 2012 IEEE Computer Society

Noriaki Kawamae, Hierarchical Approach to Sentiment Analysis, 2012 IEEE Sixth International Conference on Semantic Computing.

Renata Lopes Rosa, Demostenes Zegarra Rodriguez, Graca Bressan, SentiMeter-Br: a Social Web Analysis Tool to Discover Consumers' Sentiment, 2013 IEEE 14th International Conference on Mobile Data Management.

Sciences, 1530-1605/12 2012 IEEE

Service Quality, 2013 46th Hawaii International Conference on System

Wenjing Duan, Qing Cao, Yang Yu, Stuart Levy, Mining Online User Generated Content: Using Sentiment Analysis Technique to Study Hotel

****