Product Sentiment Trend Prediction

Vatsal Gala[1], Varad Deshpande[1], Ibtihal Ferwana[2] and Mariofanna Milanova[3]

[1] Mumbai University, Mumbai 400086, India,
[2]Prince Sultan University, Riyahd, Saudi Arabia
[3]University of Arkansas at Little Rock, USA

vatsalpgala@gmail.com, clustervarad@gmail.com, ibtihalferwana@gmail.com, mgmilanova@ualr.edu,

Abstract. The prospects of spectrum sentiment analysis are great and is a field that has been given very little research focus. We develop a system that can recognize human recognizable emotions and quantify them, the system can then predict the trend in the spectrum sentiments provided a chronological data. This paper discusses a lexicon-based approach for spectrum sentiment analysis. It further describes a quantification method to factor in the effects of time in trend prediction and a novel idea of using consecutive calculated values for current trend value calculation. The system is designed for e-commerce data but has flexibility to be used for other fields too. The system uses a simple neural network with image and text features as input and the trend values as output. This system can then be used to predict sentiment trend for newer or existing products. The system shows great prospects for multi-modal sentiment analysis of sentiments on spectrum range and can be advanced by using more complex approach.

Keywords: Trend Prediction, Multi-modal Approach, Spectrum Sentiment Analysis, Text Mining

1 Introduction

The internet boom has been advantages to many sectors. One of them being commerce. E-commerce is one of the sector that has greatly prospered with the advent of internet (online shopping). With online shopping becoming the trend, the focus of sellers has shifted from profit margins and stock management to customer satisfaction and understanding the customer sentiment. E-commerce websites have contributed greatly to this shift of focus. With features such as descriptions of products and giving a space for the customer to give reviews for the products they are buying the websites have assisted a smoother exchange of information between the customer and the seller. This has profited both sides greatly. However now that the websites such as Amazon or E-bay have been growing at greater pace, the data generated through exchange of information between the customer and the seller has increased exponentially. Which in turn has given way for more deeper analysis of this exchange of information.

adfa, p. 1, 2011.

Before moving on to the study it is important to understand the exact process we are trying to study and enhance. The process being predicting the trend in sentiments for customer for a particular product. For a customer the first contact with the product on any e-commerce website is the image of the product. After which it starts to read its description and later browse through the reviews to check for any anomalies in the product reviews. Thus, we could say that the image and the product reviews act as key factors in influencing a customer sentiment towards a product. We give less importance to product description as it has mostly technical descriptions of the product and has little impact to the final customer sentiment[1]. We have now to process 2 kinds of data namely image and text. However merely analyzing text data cannot allow us to predict a trend. Trend can be predicted by bringing time as one of the factors for our analysis. The term trend in the context of our study means to find out how the sentiment of product changed overtime mainly from its date of launch till current time and using these data have estimation of how it will perform in the near future.

Images contain lots of information, however while considering the images to be the first impression for a product it becomes necessary to not analyze the detailed features of the image but stick to the features that give off a subtle information to the customer. Such an analysis partly falls into the domain of phycology and thus must be viewed so. Certain colors and certain levels of brightness act as pleasing feature for certain genre of products while sometimes a highly contrasted image and opposing colors can instigate the customer. Thus, the concept of what kind of image is necessary for influencing the customer becomes a topic of debate. However, it is quite clear that for a particular category the product images must have a similar appearance. Keeping this in mind we won't have to worry about the difference in appearance of different products as long as we operate on similar kind of products.

Text found on e-commerce websites in forms of reviews generally have a standard structure with minimal variations. Most of the reviews follow normal grammatical rules and don't have much spelling errors or use of slangs. However, it is quite important to understand that not all the reviews can be genuine and certain reviews can have a degree of deceit to it. There are methods to identify such reviews and segregate them from the dataset. The smaller the review the lesser importance it can have in the customer sentiment but if a review is longer than a few paragraphs, its authenticity must be questioned. Thus, though text plays a key role in sentiment analysis, it must be taken care that deceitful text can introduce errors in our prediction system.

The understanding of humans regarding emotions is equivocal however its quite clear that humans measure emotions on a scale and consider emotions to be of multiple types. Emotions range from anger to fear and from joy to sadness. Though we have slightly altered perception of how emotion is represented, we mainly use the same

[1] We assume that the sentiment of customer towards a product is an indirect perception of the seller. Thus, the product description will certainly be same for various sellers, but the reviews and the images used are different from seller to seller.

means of communicating the emotion i.e. words in a language. More often than not we use similar set of words to express similar emotions. The words “I ‘m furious.” Can hardly mean anything but expression of anger. Likewise, a similar lexicon of words can be used to express similar emotions. We can greatly benefit from this aspect of language to develop a system that can recognize sentiments on a spectrum rather than the existing methods of recognizing bi-polar sentiments (positive and negative). Also, a system that can track multiple types of emotions can benefit the sellers more than a bi-polar one. Thus, our study will greatly focus on achieving spectrum sentiment analysis and analyzing a trend in it.

2 Literature Review

A lot of work has been done in the field of Sentiment analysis and its supporting subject such as NLP[2], text processing, semantic lexicon generation, image processing etc. Much of this research lays a foundation for our study and helps us in moving in the right direction. Much of such work that was inspirational for the study is discussed in this section.

2.1 Image

[1] shows how images can act as supplementary for sentiment analysis. The use of Flickr image dataset [2] for training a CNN[3] along with transfer learning proves to be quite effective. An approach of using domain specific dataset for feature extraction and machine training is rousing. A low-level feature extraction approach is synonymous to our approach as described in [3] with the difference of using a CNN to identify the features. We will plainly use the raw features available directly from the image. The reason for using raw features is to keep the system in check and not over-fit it with a particular feature. Features change form image to image, but the general genre shows a resemblance. A good reason to choose this is inspired by [4]. The paper describes how basic features of an image such as saturation and brightness can be used to derive Pleasure, Arousal and Dominance values. These values correspond to one of the 2 widely acclaimed models i.e. PAD model of emotion [5]. The model described in [4, p. 395] provides a spectrum value that our study aimed to achieve, thus acting as a stepping stone for our study.

2.2 Text

Much of the research has already been done in the field of sentiment analysis using text. However much of what is found focuses on the polar nature of words apportioning them into positive or negative. A semantic similarity approach in [6] using multiple sources

[2] NLP: Natural Language Processing
[3] CNN: Convolution Neural Network

has been motivating for using semantic similarity approach for spectrum emotion recognition.

2.3 Multi-modal Approach

Recently a lot of focus has shifted from performing sentiment analysis on a single type of data to having a multi-modal approach. A multi-modal approach ensures that the sentiment resonates with actual human sentiments as we tend to express emotions in different forms. Thus, an integration of text mining or opinion mining along with image analysis has proved to be of paramount importance for sentiment analysis. [7] describes a method for individually analyzing image and text for sentiments and then using a similarity-based classifier classifying the data for sentiments. [8] have too had early attempts of multi-modal sentiment analysis approach. Another yet innovative process is described in [9] which uses CNN and DNN[4] for analyzing the key features from text and images. This approach yields superior results but still lacks the spectrum approach nevertheless having greater accuracy.

The majority focus of research in sentiment analysis has been in the direct or multi-modal analysis but lesser attention has been given to spectrum approach and involving the time factor. Our study tries to tackle this while keeping the domain of e-commerce into mind.

3 Methodology

The following flowchart explains the process the system follows.

[4] DNN: Deep Neural Network

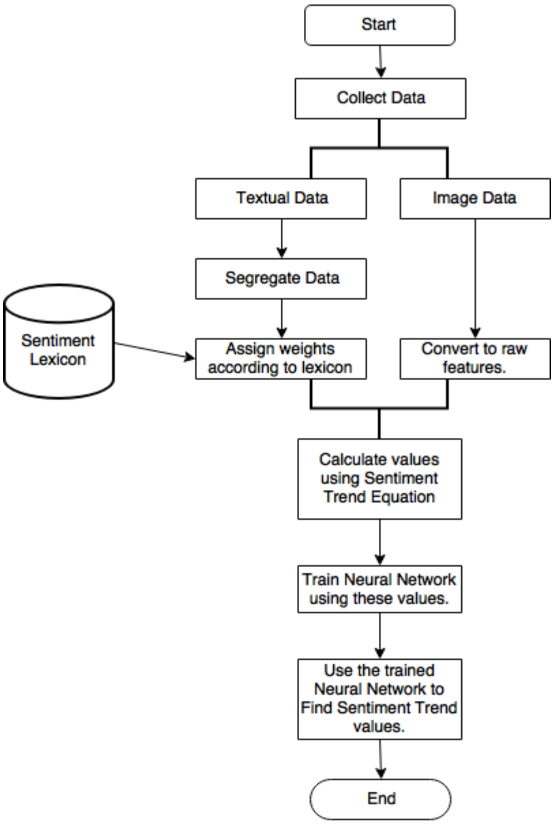

Fig. 1. The flowchart describing the process followed by the system for Trend Sentiment Prediction.

3.1 Data

For accurate systems it is quite obligatory to have genuine data to work with. Since our field of study rests with the domain of e-commerce we require a dataset that is authentic in nature and quite descriptive. [10,11] are a great source for such a data. [10] This dataset includes reviews (ratings, text, helpfulness votes), product metadata (descriptions, category information, price, brand, and image features), and links (also viewed/also bought graphs). Such data also contains the review timmings to aid involving of time factor into sentiment trend analysis. This data is required to be processed in ways that make extracting features from it simpler. Unnecceary fields are removed and only the text (review content, review title), review timing, product id and the image (image link) are kept.

3.2 Image Processing

The raw features to be extracted from the images included: color features such as hue, saturation, brightness and structural features such as contrast, correlation and entropy. Hue, saturation and Brightness represent the general visual aspects of the image while the other three represents the structural aspects and play subtle role in determining the nature of the image. The basic pre-processing of the image includes resizing all of the image to a size of 160px X 160px. Images that have resolution below this size are discarded. Then using the individual pixel data calculate the most frequent hue, brightness level and saturation values. These values represent[5] the most probable values that the customer is to glance at. The rest three structural values of image viz. contrast, correlation and entropy were found out using the Mahotas python library [11]. It allows us to calculate the concerned values easily. All of these 6 values are stored separately with their product id. This concludes the basic pre-processing of image data.

3.3 Text Processing

Text processing has more steps as compared to image. Any sentence in any language has words that are important and words that are supplementary (stop-words). Firstly, the reviews and review titles were segregated by their length. If a review was too small say few words it was discarded, also reviews that had multiple paragraphs were discarded to reduce any error by introducing exaggerated or diminished valued text. Basic pre-processing of the text includes cleaning of these stop words. NLTK python library [12] is an excellent tool that was utilized for doing so. Later the words left were categorized by their POS[6]. Each set was associated with a product id. All of this data is stored along with their product id.

3.4 Trend Methodology

The data from both the image and text are combined into a single data file and the combination is done based on product id. The time data is stored for each review. The dataset contains multiple entries of image features with different reviews features. The words or parts of words extracted from the reviews are then utilized for calculating sentiment values. A process of identifying semantic similarity and assigning weights to the words is used. For this a simple lexicon is maintained for different sentiments classified based on POS. These act as core words for the semantic similarities. While checking for words against the lexicon a check for similar words is checked too. If the match is based on similar words a lesser weight is assigned to the word or phrase. A direct match yield higher weightage. Besides assigning weights based on individual words a

---

[5] A better representation of the hue, saturation and brightness could be taken by taking median or mean values of the image, however such a measure will not accurately represent the image. However, better methods to represent the image can be used at the expense of more complex computation.

[6] POS: Parts of Speech. Included verbs, noun, adjective and adverb

n-gram of length 2 and 3 are used to have greater accuracy. Based on n-gram length, weights are assigned to semantic score found out between the n-gram and the available lexicon. At the end of the sentence all of these weighted values are summed up to calculate the final sentiment value for the particular sentence. This weighted sentence is then normalized based on length of important words in the sentence (Sentence without stopwords).

A heuristic approach is used to decide upon the weights for n-grams and similar words. All of the new sentiment values now found are stored alongside the image features. The next step involves factoring in the time constraint. The data now is grouped according to the product id and then chronologically. The data is then processed for individual product for factoring in time values.

For factoring in time, it is important to understand how time interacts with user sentiment. The 3 factors about time that are brought into the equation while factoring in time are:

1. Intensity: Intensity concerns with how often the value for a particular emotion for a product changes for a sentiment. A greater difference will indicate a greater change in the trend of customer sentiment. Also, consecutive increasing values will indicate an upsurge in the trend.
2. Direction: Direction involves the change in the direction of the sentiment. It involves the difference in sentiments from previous and current value.
3. Previous Value: For associating a trend it is apparent that the previous value should have certain impact on the next value to have a continuous trend.

The following equation explains the above factors:

$$y_n = f\left(t_{diff}, s_n, s_{n-1}, y_{n-1}\right) = g\left(t_{diff}\right) \times h(s_n, s_{n-1}) \times y_{n-1} + Sn \qquad (1)$$

Where, y_n is the value for a trend to be calculated, t_{diff} is difference in time between 2 consecutive elements, s_n and s_{n-1} are current and previous sentiments respectively and y_{n-1} is the previous value of trend. Each of the individual functions are as follows:

$$g\left(t_{diff}\right) = \frac{\tanh(t_{diff})}{t_{unit}} \qquad (2)$$

Here, t_{unit} is the idle time unit decided upon for measurement of time. The idle time unit is variable based on the frequency of the data available. Standard value can be assumed to be the minimum difference between any two consecutive values.

$$h(s_n, s_{n-1}) = \left| \frac{s_n - s_{n-1}}{r} \right|^r \quad (3)$$

Here, r is the range of possible values of sentiment.

The y_n values are individually calculated for all sentiments and the products are stored. These values i.e. hue, saturation, brightness, contrast, co-relation, entropy and the sentiment values (y_n) thus calculated will be used to train a neural network for recognizing the trend pattern for given values.

3.5 Training Neural Network

A neural network with 6 + n (n is the number of sentiments) input vectors is trained by taking the next successive y_n value to be the target value for $n - 1^{th}$ input values. Levenberg-Marquardt algorithm is used for training purpose and a simple 3-layer structure is used. Thus, a trained network will be able to generate the next value in trend for a particular product given current values for the product. A dataset of about 1900 reviews were used spanning more than 70 products from ‘pet supplies’ category. A separate dataset was maintained for testing purpose.

4 Observations:

(**Error! Reference source not found.**) shows the result for one of the product that was tested on the neural network after training the network on dataset. The dataset showed great results as the values for the reviews corresponded with an increase in the popularity for the product. There is an increase in positive sentiment and a decrease in the anger sentiment as the product reviews span from the day 247 (the date for the first review) to the day 705 (the date for the last review).

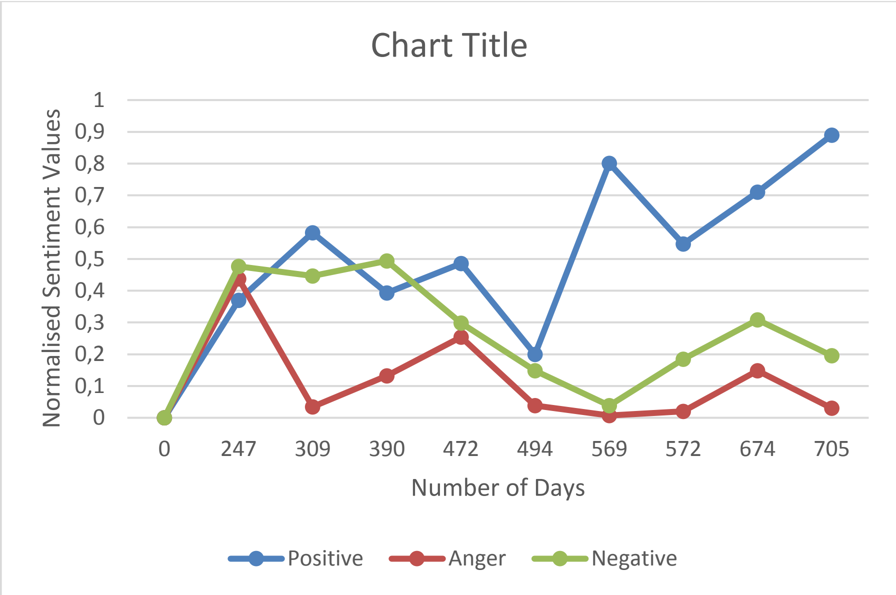

Fig. 2. Sentiment Trend graph for product labelled 'Natural Balance'.

5 Conclusion

Sentiment Analysis is a tricky subject to deal with and the data that is to be used for the analysis usually contains lots of anomalies. E-commerce websites have been generating lots of relevant data that can be used for a spectrum sentiment analysis. A good volume of research has been done in multi-modal approach for sentiment analysis which inspired us for developing a spectrum sentiment trend prediction system. Authentic data from e-commerce website can be vital for developing an accurate system. Images play an effective role in generating genre-specific sentiment. The text plays a major role in sentiment analysis. The pre-processing methods used are primitive but show remarkable results.

The system we developed shows good results in recognizing sentimental values from review text and images. Sentimental Analysis on a spectrum range can be achieved at simpler level using the above-mentioned system. The results matched the human perspective of the sentiments from the reviews fairly and quantifiable measure allowed us to have a more certain view on the sentiments. (1) has parameters r and t_{unit} that control the depth and detail of the trend graph. This approach is a novel one (as per the knowledge of the researchers) and can prove to be fundamental in getting more accurate results. Whole of the system requires minimum data i.e. text, image and the time for the data and yields trend in the sentiments with simple computations. Such a system can easily be scaled to a much larger database to generate greater accurate results.

5.1 Benefit

The field of spectrum sentiment analysis has been given very little research focus given to it. Our study brings attention to the simplicity of such a field and the possibilities of much advanced methods to give more accurate results than just simply having a bi-polar sentiment analysis. A lexicon-based sentiment analysis approach has a benefit of simplicity, easy manipulation and scalability. The trend calculation has taken time factor into account by bringing difference between time for 2 consecutive entities for sentiment analysis. Moreover, since the system is developed with keeping e-commerce into account the idea of keeping time into the equation while calculating trend enriches the concept of trend prediction. Another one of the novel idea of the (1) is keeping the previous sentiment value for the same entity into the equation. This is done with an understanding that the trend should follow a smooth curve and each consecutive sentiment if in positive respect should increment the current trend value but if negative should only cancel out the previous sentiment trend value in proportion to the accumulative of the previous sentiments. That is to say a single negative sentiment or a single positive sentiment unless excessively positive or negative should not be able to cancel out the accumulated positive or negative values of the trend. The excessively positive or negative values are taken care of during pre-processing of text by eliminating text segments that are too long or seem fake.

5.2 Drawbacks

The system shows good similarities with human perspective and has considerable accuracy. However, the system is still in its primitive stages and requires much development. It has some drawbacks as identified by the researchers. The system is weak against smaller bits of text and cannot extensively recognize slangs or sarcastic remarks. A lexicon-based approach requires the system to have a substantial sized lexicon for all the respective sentiments. This can be a problem as the language is an ever-changing concept and does not necessarily mean the same every time. The lexicon must be category specific cause, different words can express different emotions in different domains. Though the difference is little and many of the words used for expressing an emotion are domain-neutral, lack of a dynamic lexicon will overtime bring in more error to the system.

5.3 Future Prospects

Since the system has few drawbacks, it opens up a new opportunity for perfecting the system and adding more features to the system. Some of them being using a CNN or DNN on the images for recognizing the domain and simultaneously training the neural network on text, using methods for identifying sarcastic remarks or use of slangs and appropriately assigning weights to them to tackle smaller text problems, using facial analysis systems for understanding the customer sentiments on social media sites and developing a sub-system to create a dynamic lexicon that can filter out words periodically and learn new words for each sentiment and dynamically add them while using the system. We strongly believe that the mentioned system has great prospects for further development and in itself is novel is a novel work of research.

Acknowledgement: We would like to thank University of Arkansas at Little Rock, AR - USA and Vishwaniketan (iMEET), INDIA for providing the opportunity to collaborate and work on this research project.

Citations:

1. S. Jindal and S. Singh, "Image Sentiment Analysis using Deep Convolutional Neural Networks with Domain Specific Fine Tuning," in International Conference on Infromation Processing (ICIP), 2015.
2. S. jindal and S. Singh, 07 2015. [Online]. Available: http://dx.doi.org/10.6084/m9.figshare.1496534
3. H. Zhang, Z. Yang, M. Gonen, M. Koskela, J. Laaksonen, T. Honkela and E. Oja, "Affective Abstract Image Classification and Retrieval Using Multiple Kernel Learning," Springer-Verlag Berlin Heidelberg, vol. III, pp. 166-175, 2013.
4. P. Valdez and A. Mehrabian, "Effects of Color on Emotions," Journal of Experimental Phycology: General, vol. 123, no. 4, pp. 394-409, 1994.
5. "Wikipedia," [Online]. Available: https://en.wikipedia.org/wiki/PAD_emotional_state_model. [Accessed 07 2017].
6. Y. Li, Z. A. Bandar and D. McLean, "An Approach for Measuring Semantic Similarity between Words Using Multiple Information Sources," IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, vol. 15, no. 4, pp. 871-881, July/August 2003.
7. Y. Zhang, L. Shang and X. Zia, "Sentiment Analysis on Microblogging by Integrating Text and Image Features," Springer International Publishing Switzerland, pp. 52-63, 2015.
8. Maynard, Diana; Dupplaw, David; Hare, Jonathon; "Multimodal Sentiment Analysis of Social Media," in BCS SGAI Workshop on Social Media Analysis, 2013.
9. Y. Yu, H. Lin, J. Meng and Z. Zhao, "Visual and Textual Sentiment Analysis of a Microblog Using Deep Convolutional Neural Networks," MDPI, vol. 9, no. 41, 2016.
10. J. McAuley, "Amazon product data," [Online]. Available: http://jmcauley.ucsd.edu/data/amazon/. [Accessed 07 2017].
11. C. L.P., " Mahotas: Open source software for scriptable computer vision. Journal of Open Research Software 1(1):e3, DOI: http://dx.doi.org/10.5334/jors.ac," 2013.
12. E. Loper and S. Bird, "NLTK (Natural Language Toolkit)," ETMTNLP '02 Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics, vol. I, pp. 63-70, 2002.
13. R. He and J. McAuley, "Ups and Downs: Modeling the Visual Evolution of Fashion Trends with One-Class Collaborative Filtering," in International World Wide Web Conferences Steering Committee, Montréal Québec, 2016.
14. J. McAuley, C. Targett, Q. (. Shi and A. v. d. Hangel, "Image-based Recommendations on Styles and Substitutes," in International ACM SIGIR Conference on Research and Development in Information Retrieval, Santiago, Chile, 2015.