ANALYSIS AND CLASSIFICATION OF SENTIMENTS USING DUAL SENTIMENT FILTRATION

Miss. Pooja J. Biradar[1],
Dr. K. V. Metre[2]

[1] M.E. Student, Department of Computer Engineering, Savitribai Phule University, Nasik, India
[2] Professor, Department of Computer Engineering, Savitribai Phule University, Nasik, India

ABSTRACT

Sentiment analysis and opinion mining are used to determine sentiment expressed in text reviews. In area of sentiment domain analysis, Bag of Words i.e. BoW is the statistical machine learning approach which suffers from some fundamental issues those occurred during the problem of polarity shifting. Therefore, DSA i.e. Dual Sentiment Analysis is proposed to overcome the problem in classification of sentiments. It consists of two phases such as, Dual training (DT) which determines original as well as reversed training reviews from learning sentiment classifiers and Dual Prediction (DP) sort/classifies the reviews by considering both negative and positive sides of single review. Dependencies of DSA are removed by implementing corpus-based methodology which construct pseudo antonym dictionary. It is used for reversion of review. Along with DSA proposed system contributes features for inconsistency classifier by syntactic construction which helps to improve system performance.

Keyword: - *Opinion mining, sentiment analysis, machine learning, Natural language processing*

---

1. INTRODUCTION

In text mining domain, sentiment analysis and opinion mining is the special domain to identify subjective attitude expressed in text. Sentiment classification is the process of classification of sentiments into positive or negative terms. For classification of sentiments, traditional topic-based text classification in which BoW model is used. It represents review texts as independent vector. Furthermore, to provide training on sentiment classifier statistical machine learning algorithms such as, naïve bayes, maximum entropy classifiers are employed. With efficient attributes BoW model has some limitations such as, it spoils the order of words, breaks its structures and drop some syntactic information. Several researches have been conducted for improvement of BoW but due deficiencies in BoW show very slight effect with accuracy improvement. Polarity shifting is the linguistic phenomenon which reverses the sentiment polarity of text. Text negation is an important type of polarity shift. It is applied by adding negative word to positive text. For example, “I like rose.” It converted into negative polarity by adding “don’t” i.e. “I don’t like rose.” In BoW, two sentiments opposite texts are considered to be very much similar hence machine learning algorithm fails under the circumstances of polarity shift. There are several approaches have been proposed to address the polarity shift problem, however complex linguistic knowledge is required for them.

According to research analysis, DSA i.e. Dual Sentiment Analysis is proved as an efficient solution for the problem of polarity shifting. Two class labels such as, positive and negative are used for sentiment classification. To create reversed review, data expansion technique is used. In this original review and reversed review are build in one-to-one correspondence. For training of statistical classifier DT and DP algorithms have proposed. To maximize the probable combinations of original and reversed training dataset DT classifier is utilized. In DT, two sides of one review are considered for prediction.

DSA framework is further extend for classification of polarities into 3-class i.e. neutral-positive-negative sentiment classification. A corpus based antonym dictionary is constructed to reduced dependencies in DSA model.

The dictionary is language-independent and domain adaptive. Using this dictionary DSA model is applied into a wide range of applications.

As the primary contribution is to make dual sentiment analysis of sentences to overcome polarity shift problem in NLP, to fulfill this approach there is aim to develop a corpus based pseudo antonym dictionary. With this approach the dependency of DSA on external antonym dictionary is reduced. As secondary contribution it is to show that the use of syntactic constructions as features for the inconsistency classifier can improve performance. In this more complex polarity shift patterns are going to be considered.

2. RELATED WORK

Reviews are trained and tested to identify features from scoring methods. Feature selection process HTML tags and the documents separate into sentences. Parser is used to parse single-word tokens. Negating phrases such as “not” or “never” are identified in language modification. The numbers of issues are identified such as, inconsistency in ratings, comparisons and uncertainty; scant data alter distribution etc [1]

A technique based on customer feedback data from the survey of web which is noisy and fragmentary. SVM filter is used for classification of sentiments. Basically, it is used in text based classification. The standard supervised machine learning task consists of training and classification of sequential SVM by M. Gamon[2] et al, M. Gamon [2].

Research in *Automatic Text Classification* seeks to construct models for designating category labels to new documents; this approach is based on a training set of documents that have been protective by experts of domain. Lot of studies of automatic text classification has focused on “topical classification”. Documents are classified as per different subjects. “Sentiment Classification”, is automatically classifying documents according to the overall sentiment expressed in them. There are two machine learning methods are introduced for classification of reviews into two phases such as, positive sentiment(recommended) and negative sentiment(not recommended) by J. Na, et al [3].

A. Kennedy et al[4], analysed three types of bearing shifters such as, negations, intensifiers and diminishers. A corpus based semantic orientation values of terms are evaluated with association scores having small group of +ve and –ve terms. Two classification mechanisms have been represented such as; first count is introduced for positive and negative terms in the review. Both are taken from GI i.e. General Inquirer.

The problem of document classification not by topic also specified [5]. Examining factors of the sentiment classification make the classification more challenging problem. They were concentrating on only refining between positive and negative sentiments. There are three standard algorithm proposed by them for experimental analysis, they are namely, naïve Bayes classification, ME classification and SVM etc by B. Pang and L. Lee et al. [5].

Sentiment and simultaneous topics from texts have used automated tools. A joint sentiment-topic (JST) model is introduced for the framework of probabilistic modeling. It based on latent Dirichlet allocation (LDA) and its main target is to detect topics from text simultaneously. The reverse series of sentiment and the process of topic generation are also implemented to obtain Reverse-JST. It is highly portable unlike the supervised approaches of sentiment classification. Author mainly focused on document-level sentiment classification for general domains in conjunction with topic detection and topic sentiment analysis, based on the proposed weakly supervised joint sentiment-topic (JST) model. Topics as well as topic sentiment identified by joint sentiment-topic (JST) model is more informative. In future work they have planned for incremental learning of JST by Lin et al[6].

In 2009, S. Li, Y. Chen et al [7], proposed machine learning algorithm to integrate polarity shifting information into the document level sentiment classification. In this primarily feature selection method is used to automatically generate training data for classification of binary classifier on polarity shifting detection of sentences. Polarity shifting means the contradiction of sentence is distinct from the contradiction demonstrated by the sum of content words in the given sentences. Author referred such type of polarity shifting as polarity shifting structure. In polarity shifting, lack of relevant training data creating a large database of polarity shifting sentences which seem as time-consuming task.

In 2009, S. Li and C. Huang [8] proposed two kinds of linguistic phenomena which are able to reverse the sentiment polarity. Their main focus is on sentiment shifting. It includes negation and contrast transition, hence they can identify type of shifting even fully reverses the sentiment polarity. BoW modeling approach is used for complete reverse sentiment polarity. Initially From training data T terms are extracted in the form of word ,phrases and n-grams. Let *N* be the number of terms. To calculate weights of the terms i.e. importance in the training data techniques like tf and idf are used. These technique provide the term statistics. The given text is then represented in the form of statistical vector as: $X(T) = \langle stat(t1), stat(t2), \ldots, stat(tN) \rangle$. After analysis it generates positive or negative values represents the positive or negative polarity. Following is the example for reversed polarity representation:

Original Statement from Training Dataset: *This is not a good movie and I hate it.*

Reverse Statement: *This is such a good movie and I do not hate it at all.*

As above 2 statements contains almost similar words. The Bag-Of-Word technique gives almost similar word collection. The polarity of training dataset statement is positive. Using bag of word technique the reverse statement also gives the same polarity by simply matching tf and idf using bag of word collection. Hence Bag-of-word technique is not at all suitable for reversing the polarity and sentiment classification.

In 2011, A. Abbasi, S. France, et al. [9], proposed FRN i.e. Feature Relation Network. It is a rule based multivariate text feature selection method. It is method intended to enhanced sentiment classification by enabling sets of heterogeneous n-gram features. There are several important phases are included in data mining which contains sentiment polarity and intensity assignment. In this study author proposed the use of higher set of n-gram feature spanning with multiple variable and fixed categories of n-gram. For opinion classification they have combined the extended feature set with the method of feature selection. The proposed technique, FRN includes semantic information that is inherited from lexical resources of n-gram features. Therefore, feature selection method shows improved classification performance than the existing feature selection methods.

In 2016, Rui Xia, Feng Xu, et al [10], proposed DSA approach to addressed polarity shift problem in semantic analysis. The DSA is Dual Sentiment Analysis model which can be better solution for sentiment classification. DSA can be classified into two stages like, DT i.e. dual training and DP i.e. dual prediction. DT is used to determine original as well as reversed training reviews from learning sentiment classifiers whereas, DP classifies the reviews by considering both sides i.e. negative and positive of single review. They were aiming to remove dependency of DSA by building a corpus-based strategy for constructing pseudo antonym based dictionary.

3. SYSTEM ARCHITECTURE

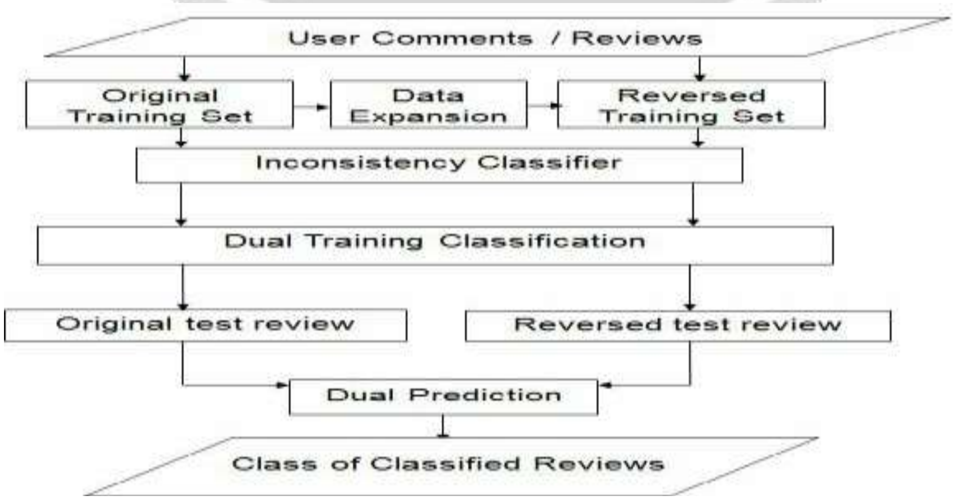

Fig.1 System architecture diagram

This is the sentiment analysis system, in which the user opinions/reviews text regarding any things is taken as input. By using the POS tagging and by DSA technique reviews are categorized in sections as positive, negative or neutral. Hence the system generates a classified analysis report. The actual modules used in the system are as discussed in the system architecture.

1. **XML processing:**
   At the beginning of system execution, XML file is uploaded into system. XML file contains multiple opening and closing tags in it. While processing XML file, system verify the structure of it i.e. opening tag is appropriately closed at specific place or not.
   In case of any mismatch tags system ignore that XML and generates alert message for end user.
   For valid XML file system carries following processing steps.

2. **Comment analysis:**
   In this phase, comments or reviews from given input XML file get extracted. Following are some techniques used to analyze comments,

   2.1 Statement extraction: Primary comments (statements) get extracted in this phase. This task identifies all the aspect entities which are present in a sentence. It determines all the aspects entities in advance and finds them in the reviews.

   2.2 Word extraction: Words are associated with different kind of sentiments, like negative, positive or neutral get extracted.

   2.3 Lemmatization: It is the process of doing things properly with the use of a vocabulary and morphological analysis of words, normally aiming to remove inflectional endings only and to return the base or dictionary form of a word, which is known as the lemma. For example:

   Caresses → caress

   Ponies' → poni

   Cats → cat

   2.4 POS tagging: It gives part-of-speech of each word. Stanford POS-Tagger is used. From the observation of POS output we come to know that words in noun phrases are separately tagged as noun.

   2.5 Relational Tree construction: It is used primarily in computational linguistics. Dependency is a one-to-one correspondence.

   To find out dependencies between words, sentences are sort to dependency parsing using dependency parser. Using parser system can identify relations.

3. **Search antonyms using Wordnet:** WordNet is a lexical database which groups English words into sets of synonyms called synsets, provides short, general definitions, and records the various semantic relations between these synonym sets. Using the antonym thesaurus it is possible to obtain the words and their opposites. It is simple and straightforward.

4. **Word classification:** In proposed system, three types of strategies follow for word classification such as, text reversion, domain specific word extraction and inconsistency identifier.
   For the task of sentiment analysis, understanding the polarity of words and classify them into positive or negative categories is important.

4.1 Text reversion: In the process of text reversion, if there exist negation then we first detect scope of negation. All sentiment words except negation are reversed to their antonyms using wordnet.

4.2 Domain specific text reversion: With a selected part of training reviews for data expansion, we might get better classification results. In this process, domain specific antonym words are extracted. For example, in Book domain, word "beautifully" can reversed using "reversed word "weak". In general, "beautifully" is revised by "ugly" but in case of domain specific strategy it is defined as, "weak"

4.3 Inconsistency text reversion: In this process, probabilities of each word are calculated. For example, "This book has cheap prize"

In this example, "cheap" is used to represent the prize of book. Therefore, the word "cheap" is reversed with antonym specified in "Book" domain.

But in case of following example,

"It is very cheap", here the domain is not specified. Then the word "cheap" may revised by antonym "expensive" as per text revised reversion. However, "cheap" and "expensive" both considered as negation in text reversion process. To overcome this contention, inconsistency classifier is implemented. The strategy of inconsistency classifier is explained below section.

1. **Dual Training:**
   In the training stage, all of the original training samples are reversed to their opposites. It is referred as original training set and reversed training set respectively. In data expansion technique, there is a one-to-one correspondence between the original and reversed reviews. The classifier is trained by maximizing a combination of the likelihoods of the original and reversed training samples. This process is called dual training.

2. **Dual Prediction:**
3. After uploading the review dataset revised sample set is created. Let x is the original sample and y be the class label whereas, x is the reversed sample and y be the revised class label. Document D can be represented as,
   $D = \{(x_i, y_i)\}^N_{i=1}$
   $\overline{D} = \{(\overline{x}_i, \overline{y}_i)\}^N_{i=1}$
   The aim is not to predict the class of x. but instead, x is used to assist the prediction of x. This process is called dual prediction. In DP, predictions are made by considering two sides of one review: When we want to measure how positive a test.

- Review x is, we not only consider how positive the original test review is, but also consider how negative the reversed test review is; Conversely, when we measure how negative a test
- Review x is, we consider the probability of x being negative as well as the probability of x being positive.
  1. **Logistic regression:** It uses the logistic function to predict the probability of a feature vector x belonging to the positive class as :
     $J(W) = \Sigma^n_{i=1}\, y_i \log[h(x_i)(1-h(N_i))] + (1 + y_i) \log[(1-h(x_i))hx_i]$
     W represents the weight of features
     Log likelihood of statement represents the polarity of a statement

  2. **Inconsistency Evaluation**:
     If polarity count is not completely reversed by the sentence context in which it occurs then these statements are considered as inconsistent statements.
     For these insistent statements polarity statement is identified by supervised training dataset.
     $h(x) = \dfrac{1}{1 + e_{x}^{wT}}$

3. **Probability Evaluation:**
   Probability is evaluated based on polarity classification,
   $P(+|x, X) = (1- \alpha).Pd(+|x) + \alpha.P(-|X)$,

   $P(-|x, X) = (1- \alpha).Pd(-|x) + \alpha.P(+|X)$,

   $P(*|x, X) = (1- \alpha).Pd(*|x) + \alpha.P(*|X)$
   Where,{+,-,*} denote the class labels of positive, negative and neutral, respectively.
   α is a tradeoff parameter(0 <=a <=1). The weight of p(.| x) is increased by choosing a larger α.
   Pd(y | x, x) denote the dual prediction of x based on an already-trained DT model.

4. ALGORITHMS USED

Input: Customer's Reviews dataset D, Antonym Dictionary, Wordnet Dictionary, pseudo antonym Domain Dictionary

Output: Review Classification as D+, D-, D^N
Processing:

1. Obtain comments / reviews provided by the user for processing
2. Conversion of Review
   i. Apply sentiment word analysis to the input with respect to antonym Domain Dictionary
   ii. Searching for negative patterns using wordnet dictionary.
   iii. Search for positive sentiment patterns in the comments
   iv. Work with Inconsistent statement and positive pattern with NOT extension.
   v. If found just flip the polarity of the sentiment to negative.
3. Inconsistency Classifier
   For training inconsistency classifier proceeds for BOW i.e. Bag of words classifier

4. Dual Training
   Calculate logistic regression to predict probability of positive / negative class.

5. Dual Prediction
   Calculate posterior probabilities for each review in dataset and classify in D+, D- or D^N

5. MATHEMATICAL MODEL

Let 'S' be DSA system from online reviews such that

S = {I, F, O}

I: {I1, I2, I3, I4 }, input to the system	I1= User Login details I2= Review Dataset I3= wordnet dictionary I4= Inconsistent statement training set
F: {F1, F2, F3, F4, F5, F6, F7, F8, F9, F10, F11, F12, F13, F14 }, functions of system	F1= User login F2=Upload review dataset containing set of reviews {R1, R2 …Rn}

	F3= Statement parsing F4= Pos tagging F5= Sentiment word list extraction F6= Search antonyms F7= Generate revised comment F8= Review analysis F9 = Inconsistency analysis F10= Apply logistic Regression F11= Apply Log-Likelihood F12= Probability evaluation F13= Analyze type of comment F14= Generate comment analysis report\
O :{ O1, O2, O3, O4 O5}, Output of a system	O1=Positive words O2=Negative words O3 = Revised Review O4 = Category O5= Result of inconsistency evaluation

6. CURRENT IMPLEMENTATION STATUS

A. Experimental Setup:

A desktop based application for comment analysis is developed. The system in build java- Jdk1.7. Mysql 5.3 is used to store database. Core i3 machine with 8GB ram is used for development and testing. Wordnet 2.1 is required for antonym list generation.

B. Dataset Used:

For testing acl[11] dataset is used it contains English reviews for multiple domains like: Book, DVD, Electronics and Kitchen. These reviews are collected from Amazon.

7. RESULT TABLES AND DISCUSSION

Table 1: Performance Evaluation

Dataset	No. Of Comments	Positive Word List	Negative Word List	Execution Time
Book	100	453	156	899.2

DVD	100	325	219	937.93
Electronics	100	213	114	912.43
Kitchen and houseware	100	547	269	914.03

System's performance for 100 input comments in terms of number of comment extraction, Positive wordlist extraction, negative wordlist extraction and it's execution timing in sec.

For testing we have used three domains from dataset such as, Book, DVD and Electronics.

As per result observation, DVD domain required more execution timing.

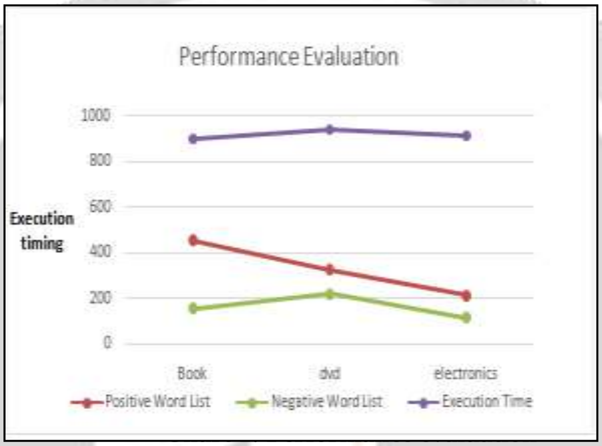

Chart 1 consists of graphical form of system performance evaluation in which X-axis consists of various domains and Y-axis consists of execution timing in second.

Table 2: Precision

Domains	**Review Analysis**	**Domain Specific Analysis**	**Inconsistency Handling**
Book	0.705	0.842	0.9
DVD	0.875	0.9	0.904
Electronics	0.809	0.85	0.944
Kitchen & House wares	0.761	0.904	0.95

In table 2, comparative precision value analysis between three system such as, Analysis review, Domain specific analysis and proposed system i.e. inconsistency handling is given.

According to analysis of table 2, it proved that proposed system outputs the improved as well as précised results for each domain than the existing system.

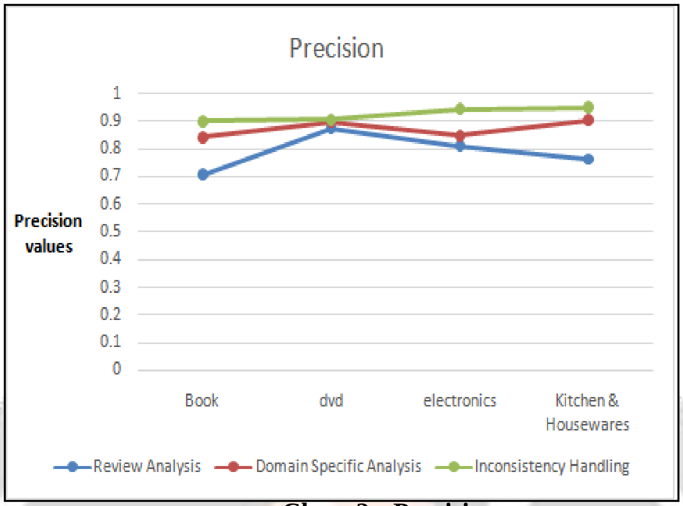

Chart 2:- Precision

Chart 2 represents precision in graph format such that X-axis consists on various domains and Y-axis consists of Precision values for each domain.

Table 3: Recall

Domains	Review Analysis	Domain Specific Analysis	Inconsistency Handling
Book	0.8	0.888	0.947
DVD	0.777	0.857	0.863
Electronics	0.739	0.894	0.894
Kitchen & House wares	0.842	0.904	0.904

Similar to precision strategy we also calculated recall values for three domain of dataset which is given in table 3, According to analysis of table 3, it recall for proposed system is more than other tow system's recall values.

Recall value is calculated by formula such that,

$$\frac{\textbf{Total number of true +ve + Total number of false -ve}}{\textbf{Total number of true +ve}}$$

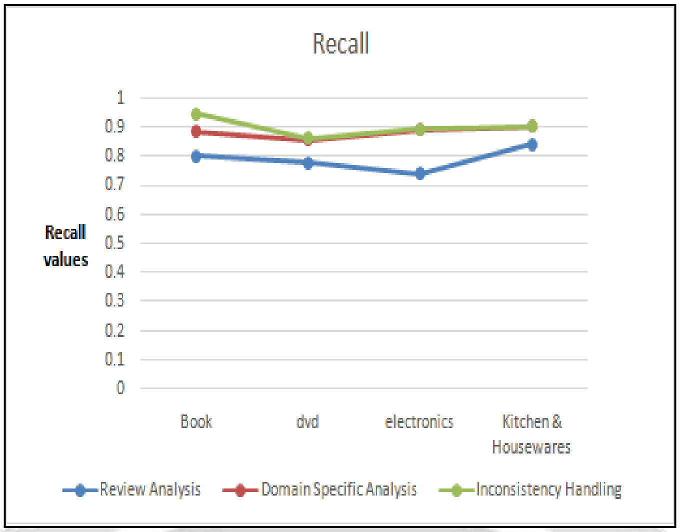

Chart 3:- Recall

Chart 3 represents recall of system in graphical format in which X-axis represents various domains and Y-axis represents recall values for each domain.

Table 4: F-measure

Domain	**Review Analysis**	**Domain Specific Analysis**	**Inconsistency Handling**
Book	0.749	0.864	0.922
DVD	0.824	0.877	0.883
Electronics	0.772	0.871	0.918
Kitchen & House wares	0.799	0.904	0.926

F-measure is calculated by formula given below,

$$\frac{\mathbf{2PR}}{\mathbf{(P+R)}}$$

Table 4 represents f-measure between existing and proposed system. It is seems that proposed system gives better f-measures than previous system due to more précised and improved values of its precision and recall.

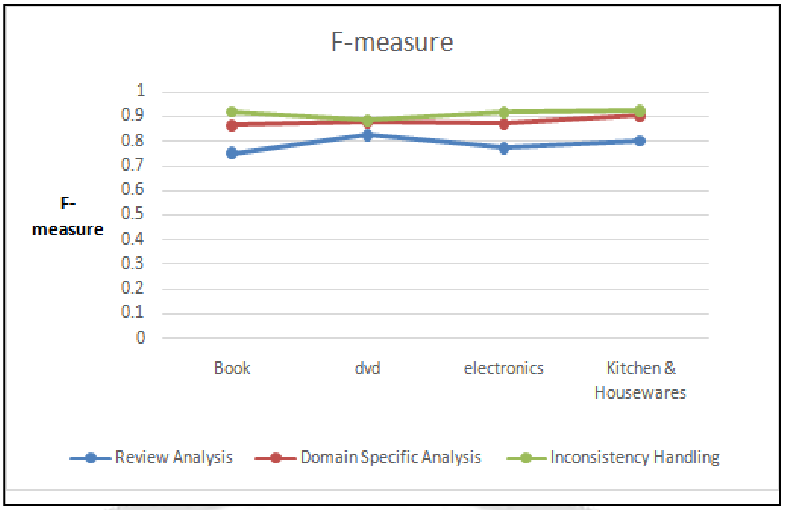

Chart:-3: F-measure

In above chart 3, X-axis represents dataset domains and Y-axis represents F-measure values.

8. CONCLUSIONS

A DSA algorithm can address the problem of polarity shifting in sentiment classification. There are two phases in DSA algorithm such as, Dual training and Dual prediction. DSA mainly works on generation of revised statement that has exactly sentiment-opposite polarity and provides pair of these two statements to the classifier as a training to generate prediction analysis report. DSA framework extends from polarity classification to three class classification as positive negative and neutral. DSA is extended to deal with inconsistent statement by using NLP term relationship analysis. Corpus based dictionary helps to increase precision and removes the direct dependency on external antonym dictionary. Proposed system generates more précised results as compared existing systems such as, review analysis and domain specific systems.

7. REFERENCES

[1] K. Dave, S. Lawrence and D. Pen-nock, “Mining the peanut gallery: Opinion extraction and semantic classification of product reviews,” in Proc. Int. World Wide Web Conf., 2003, pp. 519–528

[2] A.M. Gamon, “Sentiment classification on customer feedback data: Noisy data, large feature vectors, and the role of linguistic analysis,” in Proc. Int. Conf. Comput. Linguistics, 2004, pp. 841–847.

[3] J. Na, H. Sui, C. Khoo, S. Chan, and Y. Zhou, “Effectiveness of simple linguistic processing in automatic sentiment classification of product reviews,” in Proc. Conf. Int. Soc. Knowl. Org., 2004, pp. 49–54.

[4] A.Kennedy and D. Inkpen, “Sentiment classification of movie reviews using contextual valence shifters,” Comput. Intell., vol. 22, pp. 110–125, 2006.

[5] B.Pang and L. Lee, “Opinion mining and sentiment analysis,” Found. Trends Inf. Retrieval, vol. 2, no. 1/2, pp. 1–135, 2008.

[6] Lin and Y. He, “Joint sentiment/topic model for sentiment analysis,” in Proc. 18th ACM Conf. Inf. Knowl. Manage., 2009, pp. 375–384.

[7] S. Li, R. Xia, C. Zong and C. Huang, “A framework of feature selection methods for text categorization,” in Proc. Annu. Meeting Assoc. Comput. Linguistics, 2009, pp. 692–700.

[8] S. Li and C. Huang, “Sentiment classification considering negation and contrast transition,” in Proc. Pacific Asia Conf. Lang., Inf. Comput., 2009, pp. 307–316.

[9] Abbasi, S. France, Z. Zhang, and H. Chen, “Selecting attributes for sentiment classification using feature relation networks,” IEEE Trans. Knowl. Data Eng., vol. 23, no. 3, pp. 447–462, Mar. 2011.

[10] Rui Xia, Feng Xu, Chengqing Zong, Qianmu Li, Yong Qi, and Tao Li, “Dual Sentiment Analysis: Considering Two Sides of One Review”, IEEE Transactions on Knowledge and Data Engineering, vol. 27, no. 8, August 2015.

[11] http://www.cs.jhu.edu/~mdredze/datasets/sentiment/