A Personalized Recommender System using Machine Learning based Sentiment Analysis over Social Data

Meghana Ashok, Swathi Rajanna and Pradnyesh Vineet Joshi
Department of Information Technology
National Institute of Technology Karnataka
Surathkal, Mangalore, India
Email: {m6gnaa,swathirbombay,pradnyeshj94}@gmail.com

Sowmya Kamath S
Department of Information Technology
National Institute of Technology Karnataka
Surathkal, Mangalore, India
Email: sowmyakamath@nitk.ac.in

***Abstract*—Social Media platforms are already an indispensable part of our daily lives. With its constant growth, it has contributed to superfluous, heterogeneous data which can be overwhelming due to its volume and velocity, thus limiting the availability of relevant and required information when a particular query is to be served. Hence, a need for personalized, fine-grained user preference-oriented framework for resolving this problem and also, to enhance user experience is increasingly felt. In this paper, we propose a such a social framework, which extracts user's reviews, comments of restaurants and points of interest such as events and locations, to personalize and rank suggestions based on user preferences. Machine Learning and Sentiment Analysis based techniques are used for further optimizing search query results. This provides the user with quicker and more relevant data, thus avoiding irrelevant data and providing much needed personalization.**

Keywords*—*Personalized Search Engine, Machine Learning, Sentiment Analysis, Natural Language Processing.

I. Introduction

Over the past five years, a marked development and uprising of social media applications over the Web. Social media has provided a platform, which combined with computer-mediated tools, allows users to create and also trade pictures, information, opinions et cetera amongst many online communities and networks. Social Media is an Internet-based application group known to build idea and technology based foundations of Web thereby enabling the creation and exchange of data related to a user. Also, social media is dependent on mobile and web based technologies, producing highly interactive platforms for communities to collaboratively create, share, discuss and modify user content.

Social media technologies manifest into different forms like, blogs, business networks, forums, products/services review, social bookmarking, micro-blogs, social networks, video and photo sharing, social gaming and virtual world. Social media analysis has proven to be very useful in the field of business, education, sports, news etc. The main advantages include - flexibility (where its adaptability makes content management more flexible), globally available (where it enables people to connect even when they are miles away), measurability (where the social media stats can be easily measure) and its simplicity of use. But this abundance of data may contain irrelevant data not required for a particular user. Thus, serving a particular query may result in irrelevant results and also can be time consuming due to the extended process of searching for the right results and segregating all available information. Thus, there is a need for a personalized fine-grained preference system to solve this problem thereby optimizing user experience.

In this context, *fine-grained user preference* refers to the smaller parts of the larger components, meaning, collecting the likes and dislikes of the user using his comments, reviews etc. This enables the elimination of irrelevant results in the overall big data, thus providing better suited results based on sentiment analysis. Sentiment Analysis (SA) uses NLP (natural language processing), textual and computational linguistics analysis for identifying and extracting information related to subjects from sources. In general, SA aims to understand and clearly determine a user's standpoint with respect to certain aspects or the gross polarity of documents.

In social media, heterogeneous data is available across diverse social networking websites in abundance, thus giving rise to the need of segregating the important ones from the miscellaneous ones. Many algorithms have been implemented to perform sentiment analysis on the given set of data. For example, consider a user's comments,*"I love the burger in that restaurant but not the salad"*. This would imply that the user likes one item while he dislikes another item at a particular restaurant. Machine learning incorporates the study of algorithms that self-learn and make forecasts on such multi-polarity, large-scale data. Such algorithms work by developing models from sample inputs to make data-driven decisions, rather than following pre-defined instructions.

For this paper, we attempt to compare the different approaches to sentiment analysis. The rest of the paper is organized into four sections. In section II, we present a discussion on the existing techniques used in Sentiment Analysis. Section III outlines the details of the proposed approaches. Section IV presents the corresponding experimental results, followed by section V, which outlines conclusion and possible directions for future work.

II. Related Work

Several researchers have tried to tackle the problem of personalized social media search through varied methods. Medhat et al [11] provided an elaborated overview of the work existing in the field of Sentiment Analysis of social media. Enhancements in various recently proposed algorithms and various Sentiment Analysis (SA) applications were investigated and

978-1-4673-7918-2/16/$31.00 © 2016 IEEE

presented briefly in this survey work. Various approaches such as machine learning based (probabilistic classifiers - Naive Bayes, Maximum Entropy and Linear, Support Vector machines (SVM), neural nets) and non-machine learning based techniques (statistical approaches) were looked into.

Yang et al [19] proposed a location based personalized recommendation system called SESAME, which incorporated hybrid user location preference model. This combines user preferences (mostly fine grained) obtained from check-ins and text-based tips and performs sentiment analysis on them. They employed social matrix factorization algorithm which takes into account both influence of the user on social media and also influence in location recommendation due to venue similarity. They proposed a hybrid preference model that marked better results in comparison to the state-of-the-art methods.

Chen and Tseng [8] used two multi-class SVM based classification approaches - One-Versus-All and Single Machine Multi-class SVM, to categorize digital cameras and MP3 reviews according to their quality. Their method was superior than the state-of the-arts methods in reference to performance in classifying the reviews based on quality. The futile use of the standard LS-SVM and PSVM in regression and multi-class classification applications directly was noted by Guang-Bin Huang [10] though the variants of the above system were able to handle such cases. Simplification of both LS-SVM and PSVM was presented by them, along with the unified learning framework of LS-SVM, PSVM, and other regularization algorithms known as Extreme Learning Machine (ELM).

Duric et al [9] proposed methods to select features that utilised *Content and Syntax model*. This enabled automatic learning of a set of features in a review document by distinguishing the entities that were reviewed from the subjective expressions that describe those entities in terms of polarities. With more focus on the subjective expressions without much focus on entities, abundant salient features can be selected for document-level sentiment analysis.

In order to predict the sentiments of online users from text papers, Xue Bai [7] proposed a heuristic model. This so called Markov Blanket model was search enhanced that could perceive the word dependencies and dispense a vocabulary that is enough for the purpose of sentiment extraction.Their performance indicated the ability to identify close-fitted set of predictive features at the same time yield comparable or better predictive results about the orientations of sentiments than many state-of-the-art feature selection algorithms and sentiment prediction methods.

Yessenov et al [20] built a system to perform sentiment analysis on Movie reviews. They utilized three supervised techniques of machine learning such as Naive Bayes, Decision Trees, Maximum-Entropy, and one unsupervised machine learning technique - K-Means clustering followed by a comparison of their accuracy. They graded each sentence of a review based on subjectivity and polarity. Uryupina et al [18] provided a corpus for sentiment analysis on YouTube which can be used as an input for our Sentiment Analysis and Text Classification.

Pang et al [15] examined the effect of various machine learning classifiers such as Naive Bayes, SVM and Maximum Entropy classifiers for traditionally classifying the documents by topic, as well as recognizing the overall sentiment i.e., to figure out if a particular review is positive or negative. The dataset used is the IMDb movie reviews database. Pavlopoulos [16] presented an Aspect Based Sentiment Analysis system that takes reviews and messages about any particular products type as input and attempts to deduce the main aspects of the entity and calculates the average sentiment of the texts corresponding to each aspect. His thesis proposed three main sub-tasks which included extracting, aggregating and estimating polarity of aspect terms. New evaluation measures are introduced for each sub-task.

III. Proposed Work

In this paper, we attempt to compare three different approaches to sentiment analysis as applied to social media data. The approached that were employed are discussed briefly here:

A. *Rule based Sentiment Analysis*

This is a statistical based technique where the machine marks polarity based on some words extracted from a sentence. It not only provides polarity of the sentence but also provides the polarity of the elements of the sentence. Figure 1 represents the architecture of the Rule Based Sentiment Analysis method.

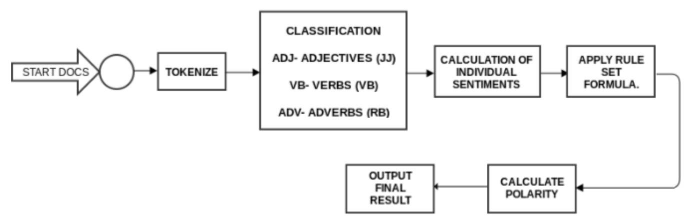

Fig. 1: Methodology of Rule Based Sentiment Analysis

Two approaches were considered as per this model. Both do not require any training data and follow the set of rules with changes only in the calculation of polarity stage.

- *Approach 1:* A review may contain more than one sentence and the polarity of all the individual sentences must be calculated. As a result, the reviews are broken down to elemental sentences, say S_i as shown in Equation (1). Each S_i is tokenized using parts of speech taggers and thus further divided into elemental words W_i. Every W_i is classified into a category of verb, adjective and adverb (nouns basically is more or less the entity the review is concerned with and hence is omitted). Each classified W_i is assigned a polarity from a WordNet corpus. The elemental word adverbs are checked for inverse property like *no, never*,etc. which help to invert the polarity sign of the sentence. The polarities range between [-1,1] and are added using the formula:

$$S_i = max/min(-1/1, \alpha \times adj_i \pm \beta \times adv_i) \quad (1)$$

$$F\_pol = \frac{\sum_{i=1}^{n} S_i}{n} \quad (2)$$

where,

$$\alpha > \beta, \alpha = (0,1]\ and\ \beta = (0,0.5) \tag{3}$$

where β is chosen depending on how much an adverb influences an adjective. The final polarity of the sentence is given by the sign of F_pol as shown in Equation (2) and the value is given as a percentage of its magnitude. In case of an absence of adjectives, verbs are taken into consideration.

- *Approach 2:* A review may contain more than one sentence and the polarity of all the individual sentences must be calculated. As a result, the reviews are broken down to elemental sentences say S_i as shown in Equation (4). Each S_i is tokenized using parts of speech taggers and thus further divided into elemental words W_i. Every W_i is classified into a category of verb, adjective and adverb (nouns basically is more or less the entity the review is concerned with and hence is omitted). Each classified W_i is assigned a polarity from a WordNet corpus. The elemental word adverbs are checked for inverse property like *no, never*,etc. which help to invert the polarity sign of the sentence. The polarities range between [-1,1] and are added using the formula:

$$S_i = max/min(-1/1, \alpha \times vb_i \pm \beta \times adj_i \pm \gamma \times adv_i) \tag{4}$$

$$F\_pol = \frac{\sum_{i=1}^{n} S_i}{n} \tag{5}$$

where,

$$\alpha + \beta + \gamma = 1\ and\ \alpha > \beta > \gamma \tag{6}$$

The final polarity of the sentence is given by the sign of F_pol as shown in Equation (5) and the value is given as a percentage of its magnitude.

B. Sentiment Analysis using Machine Learning Techniques

Sentiment analysis is conducted on the movie reviews dataset from Cornell[13], which contains thousand positive reviews, and thousand negative reviews. 70% of the data was used for training, and 30% was used as test data. The amount of positive and negative samples were equal in both training set as well as the test set.

A. Feature Extraction: For feature extraction, vector representation of words is used. The Word2Vec model is built using the text contained in all the reviews. From this, a vector representation for each word is obtained.

- *Approach 1 (Average Vector as feature vector):* For each review, the average of all the vectors corresponding to the words in the review is computed. This average vector is used as a feature vector. Thus, we represent each review with a vector of dimension 300.
- *Approach 2 (Bag-of-centroids as feature vector):* Word2Vec creates clusters of semantically related words, so we can create clusters of similar words. This grouping of vectors is known as vector quantization. To accomplish this, the centers of the word clusters are found using K-means clustering algorithm. Here, the parameter that needs to be set is the number of clusters, K. Using trial and error, we found that good performance is obtained when number of words per cluster is low, around 5 words per cluster. K-means gives us cluster/centroid assignment for each word. Thus, we can convert a review into "bag-of-centroids", which works just like "bag-of-words", except here, semantically related clusters are used instead of individual words. In short, for each review, we get a feature vector whose dimension is equal to the number of clusters.

B. Classification: For classification, Naive Bayes, SVM, random forest and maximum entropy classifiers are used, and their performance is compared in terms of accuracy.

- ***Naive Bayes classifier:*** It is a probabilistic model for classification. If there are two classes ω_1 and ω_2, the probability that a sample **x** belongs to a class ω_j, as given by *Bayes' formula*, is:

$$P(\omega_j|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\omega_j)\, P(\omega_j)}{p(\boldsymbol{x})} \tag{7}$$

  where $p(\mathbf{x}|\omega_j)$ is the probability distribution function of **x** within a class ω_j, and $P(\omega_j)$ is the prior probability of class ω_j. We used the Gaussian model for probability distribution.
- ***Support Vector Machines:*** Support vector machines are binary linear classifiers. However, nonlinear classification is also possible, by mapping the inputs to higher dimensional feature spaces. These mappings are known as kernel functions. The amount of fitting on the training set can be increased by using higher order kernel functions. The performance of SVM classifier depends on the choice of kernels as well as the soft margin parameter C. The higher is the value of C, the lower the amount of regularization, more over-fitting may occur.
- ***Random forest classifier:*** Random forests function by building multiple decision trees and yielding a class which is basically, the mode of all classes. Decision trees that grow deep often try to learn irregular patterns, in other words, over-fit the data. Random forests is a way to average multiple deep decision trees, which is done by training different subsets of the training set at a time, with the aim of increasing the variance. The algorithm used to train in case of random forests, uses two techniques:
  1) *Bagging/ Bootstrap aggregation:* Random samples of training data are taken and trees are fit to these samples. The prediction for an unseen sample is determined by taking a majority vote of all the decision trees. It decreases the variance of the model, without increasing the bias.
  2) *Feature bagging:* A subset of features is selected at each sampling. This is because if some of the features are very strong predictors, they will get selected in many decision trees, causing them to get correlated.

- ***Maximum entropy classifier:*** The MaxEnt is a generalization of the principle of logistic regression for multinomial cases. The probability mass function is given by :

$$p\left(y|\boldsymbol{x}\right)=\frac{exp\left(\boldsymbol{\theta}_{\boldsymbol{k}}^{T}\boldsymbol{x}\right)}{\sum_{c=1}^{C}exp\left(\boldsymbol{\theta}_{\boldsymbol{c}}^{T}\boldsymbol{x}\right)};\left(y=k\right) \tag{8}$$

Here, we have one vector of parameters and one hypothesis by class. Each hypothesis is a softmax function:

$$h_{\theta}\left(\boldsymbol{x}\right)_{k}=\frac{exp\left(\boldsymbol{\theta}_{\boldsymbol{k}}^{T}\boldsymbol{x}\right)}{\sum_{c=1}^{C}exp\left(\boldsymbol{\theta}_{\boldsymbol{c}}^{T}\boldsymbol{x}\right)} \tag{9}$$

In a similar way to the logistic regression, a maximum log likelihood is used to estimate θ. The log likelihood is given by:

$$l\left(\boldsymbol{\theta}\right)=\sum_{n=1}^{N}log\left(\frac{exp\left(\boldsymbol{\theta}_{\boldsymbol{k}}^{T}\boldsymbol{x}\right)}{\sum_{c=1}^{C}exp\left(\boldsymbol{\theta}_{\boldsymbol{c}}^{T}\boldsymbol{x}\right)}\right) \tag{10}$$

which implies,

$$l\left(\boldsymbol{\theta}\right)=\sum_{n=1}^{N}\left(\theta_{y_n}^{T}\boldsymbol{x}_{n}-log\sum_{c=1}^{C}exp\left(\theta_{y}^{T}\boldsymbol{x}\right)\right) \tag{11}$$

Gradient descent optimization is used to find the values of θ that maximize the likelihood. Thus, MaxEnt model is a generalization of the logistic regression to C classes.

C. Aspect Based Sentiment Analysis (ABSA)

Aspect Based Sentiment Analysis was first introduced as a SemEval task in 2014 (SE-ABSA14)[2] along with datasets of various product reviews in English. It was basically built for e-commerce websites; the datasets were represented by selected aspect terms such as a single word such as *pizza* or a set of words such as *hard disk*, and their respective polarities for laptop and restaurant reviews, as well as coarser aspect categories such as *FOOD* and their polarity determined only for the restaurants domain. The task was repeated in SemEval 2015 (SE-ABSA15)[3] aimed to accommodate more in-depth research in which all the recognized entities of the expressed comments or reviews (aspects, opinion target expressions and sentiment polarities) meet a set of rules or specifications and are linked to each other within tuples.

The Aspect based method is based on the concept of structuring data and dividing each sentence into a set of features and the opinion expressed about that particular feature. These features are called Aspects and are represented as *E#A* or *Entity#Attributes*. These Aspects provide a feature based review for products or places. For example, if the review is something like *"Great Ambiance, mediocre food"*. We consider ambiance and food, as the too entities within the review and the corresponding opinion are extracted for the two entities. Figure 2 represents the methodology of the Aspect Based Sentiment Analysis. The ABSA method is generally divided into four subtasks as specified in SemEVAL 2014[2] :

1) *Aspect term extraction:* The term that can be used to represent the sentence can be extracted as aspects and to return the distinct aspect terms. Aspect terms are values to be extracted from each review. Suppose, we find a food related set of words used in a review, we consider this as the aspect term.
2) *Aspect term polarity:* To identify the polarity of the the extracted aspect and categorize as positive, negative, neutral and conflict. We check the polarity of the of the attribute and its value words, thus getting a general opinion regarding that particular aspect term.
3) *Aspect category detection:* A group of words (for example, cost, service, ambiance, miscellaneous, food), that help categorize the needed aspects, are decided and the aspect terms are checked and appended to the corresponding aspect list. In the SIEL paper, they have chosen a certain number of static entities and attributes, for the dataset category as specified.
4) *Aspect category polarity:* Identify the aggregate overall polarity of each aspect category.

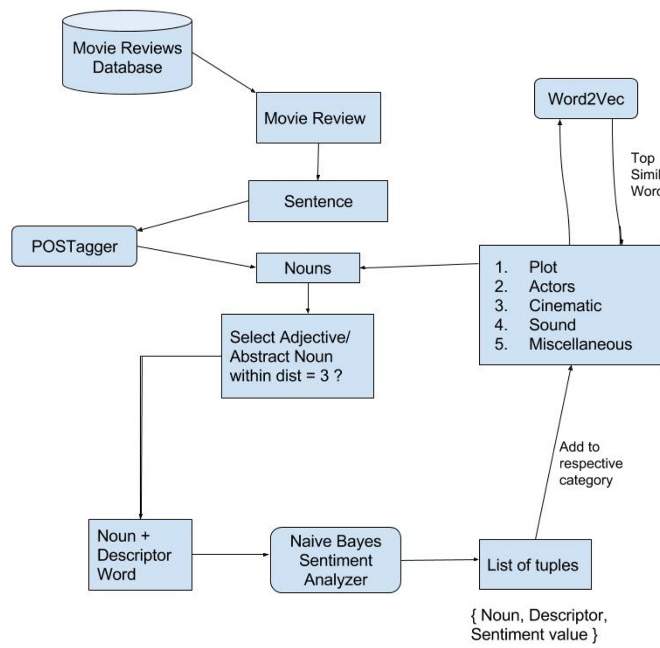

Fig. 2: Methodology of the Aspect Based Sentiment Analysis

IV. Experimental Setup and Results

For the Rule based sentiment analysis, the Python NLTK package was used for natural language processing. In-built tokenizers were used to convert the sentences to tokens. *Stanford PoS Tagger* was used in order to tag the tokens and thus helping in classifying them accordingly. Polarity was assigned to these token based on Sentiwordnet corpus[5] and AFINN corpus[12]. For rating emotions SentiStrength corpus[4] was used. The whole code which included 2 approaches were programmed in python. Datasets from Cornell[13] with respect to movie reviews were used as the test data to measure the accuracy. Also, another dataset from Cornell[14] containing a list of negative and positive sentences and Kaggle's testing dataset[1] were used to test the approaches to measure their efficiency.

For machine learning based techniques, the dataset used was the Cornell movie reviews dataset[13], containing a thousand positive reviews and thousand negative reviews. About seven hundred positive and seven hundred negative reviews were used as training set. The remaining three hundred of each positive and negative reviews were used as test set. Feature extraction using Word2Vec was done using gensim package in python. Random forest classifier was implemented using the sklearn package[17]. The number of estimators (decision trees) was set to be 100. Maximum entropy classifier was implemented using nltk package in python. Number of iterations was limited to 25, as the error remained almost constant after 25 iterations.

For Aspect based Sentiment Analysis, the entire system was built using the Cornell movie review corpus[13]. Five Aspects are singled out for movie reviews - actors, plot, cinematic, sound and miscellaneous. The Word2Vec function from gensim models was used om top most similar words to each of the aspects and added as dictionary values to their aspects respectively. Considering each review as a group of sentences and taking each sentence one at a time, The Stanford POS Tagger was used to label each word with its Part of Speech, and represented as a list of tuples. Drawing from the h&l algorithm[16], initially, the noun words from these sentences are separated out (this forms the entity part), and adjective present within three words before the noun (which is the attribute required for that entity) are extracted and added onto a list of lists that contains, the noun, the nearby adjective and its positive or negative polarity as required.

The polarity was obtained using textBlob [6], that gave `textblob.sentiments`, a module which had 2 SA implementations - the *PatternAnalyzer* which was constructed on the pattern library and *NaiveBayesAnalyzer*, an NLTK classifier. The NaiveBayesAnalyzer was used to obtain the sentiment for each of the aspect terms. Once the entire file has been parsed and all such adjective-noun phrases are extracted, they are each classified into their respective aspects - based on the presence of the nouns in the aspect dictionary. The average sentiment for each of these aspects are then calculated and together the sentiment values for each feature are again added to get the total polarity of the document. This value is calculated for each file in the dataset and the percentage of correct result used to determine the efficiency of the method.

A. *Experimental Results and Analysis*

1) Rule based Sentiment Analysis (RBSA): The two approaches of rule based sentiment analysis do not require a training set and can directly be tested. The datasets Kaggle, Cornell-movie reviews[13] and Stanford-movie reviews were used for this evaluation. Figure 3 shows the results for the two approaches on various datasets. The efficiency as reported by these approaches are comparable to the Yang et al's study where an accuracy of 72% was reported, but the added advantage over the latter method is that no training datasets were required over here. The efficiency observed was more than that of Shih et al's study, which showed an accuracy of 69-72%.

2) Sentiment Analysis using Machine Learning Techniques: The number of training samples in our dataset is 1400.

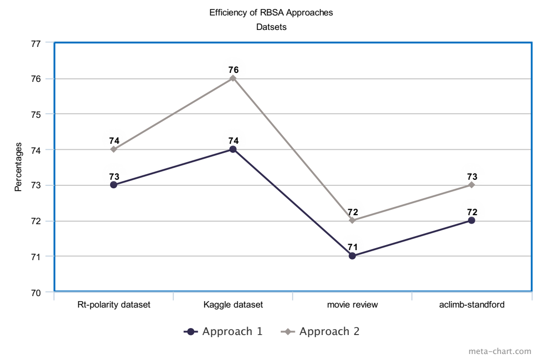

Fig. 3: RBSA Approaches on multiple datasets

Approach 1 and Approach 2 had 300 and 641 (which is one-fifth of the aggregate count of words in the dataset) number of features respectively. The performance of Naive Bayes classifier depends on the model chosen (in this case, Gaussian) and the prior probabilities. Since the amount of negative and positive samples as well as the model chosen is same in both the approaches, the accuracy is almost the same. With the number of features being comparable to the number of training samples in approaches used, we have a case of over-fitting. Therefore, SVM with linear kernel gives better results than the Gaussian kernel in the approaches used. The accuracy observed using approach 1 as well as approach 2 with different classifiers are showed in Table I and II respectively for five different Machine Learning methods.

TABLE I: Observed Classification Accuracy of ML techniques: Approach 1 *(Average Vector as feature vector)*

Classifier	True positives	False positives	True negatives	False negatives	Accuracy
Naive Bayes	180	95	205	120	64.17
SVM (Gaussian kernel)	288	254	46	12	55.67
SVM (Linear kernel)	195	109	191	105	64.33
MaxEnt	295	287	13	5	51.33
Random Forest	207	84	216	93	70.5

TABLE II: Observed Classification Accuracy of ML techniques: Approach 2 *(Bag-of-centroids as feature vector)*

Classifier	True positives	False positives	True negatives	False negatives	Accuracy
Naive Bayes	155	64	236	145	65.17
SVM (Gaussian kernel)	212	102	198	88	68.33
SVM (Linear kernel)	222	83	217	78	73.17
MaxEnt	225	62	238	75	77.17
Random Forest	243	237	63	57	51

Random forest classifier attempts to resolve the issue of over-fitting by taking random samples of data at a time. This classifier constructs multiple decision trees providing the class which is the mode of the classes of each decision tree. So, the performance achieved with this classifier is better than the maximum entropy classifier. As the number of features in approach 2 is higher compared to approach 1, it has more variance/over-fitting. As the amount of over-fitting is less in approach 1, it yields better performance in approach 1 with random forest classifier.

3) Aspect Based Sentiment Analysis: The result is obtained as tuples of the form : a noun (or *entity*), its adjective (or the *opinion word*) and the sentiment score for the same. For a given input text, the system extracts all the nouns, their respective opinion words and the sentiment score. These values are stored as a list of tuples. The entire list of tuples are saved and the values of sentiment categorized and sent to the respective entities, which are then added overall and averaged. The final values are the sentiment score for each category, with the last value being the overall sentiment of the review.The ABSA method achieved an accuracy of 92% when examined for 50 files of the Movie Reviews positive corpus and 48% for 25 files in the negative Movie Review corpus. Combined accuracy was around 70%.

4) Comparative Analysis of the Sentiment Analysis methods: All the above methods were executed on the Cornell movie Database[13] and the accuracy for each has been tabulated in Table III. As we can see, Maximum Entropy has the best accuracy value among all the Sentiment analysis techniques. Since our project requires entity based opinion of the user, The best option for the recommender would be to use ABSA integrated with Maximum Entropy Machine Learning Algorithm. Figure 4 shows the graphical representation of the results and thus we see, Machine Learning approaches clearly have higher accuracy.

TABLE III: Comparison of observed classification accuracy for all techniques

Sentiment Analysis Method	approaches	Accuracy %age %
Rule Based Sentiment Analysis	Approach 1	71
	Approach 2	73
Machine Learning	Naive Bayer's	65.17
	SVM	73.17
	Random Forest	70.5
	Maximum Entropy	77.17
Aspect Based Sentiment Analysis	ABSA	70

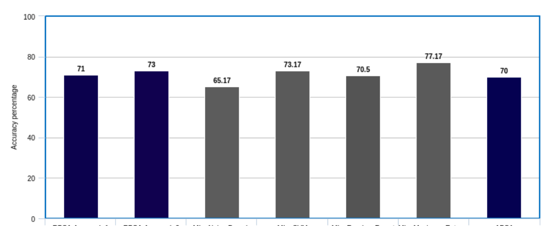

Fig. 4: Observed accuracy values during sentiment classification using the various approaches

V. Conclusions and Future Work

With the increase in the heterogeneous data, there is a need for fine-grained self aware system to eliminate the data not required. A personalized system which understands the users' needs is required to solve this problem. Thus through this project, a framework is implemented which will enable the user to choose from only those data which he likes and not waste time on the data irrelevant to his needs. This enables faster services, better results and also removal of irrelevant data. In future, a parallelized ranking algorithm will be incorporated, for improving result generation time. Also, both context based and preference based searching would be incorporated to achieve better accuracy. Based on these techniques, a recommendation system that would assist the user to exploit the knowledge contained in large volume of social data will be designed.

References

[1] Description - sentiment analysis on movie reviews | kaggle.

[2] International workshop on semantic evaluation semeval-2014 : Semantic evaluation exercises.

[3] International workshop on semantic evaluation semeval-2015 : Semantic evaluation exercises.

[4] SentiStrength - sentiment strength detection in short texts - sentiment analysis, opinion mining.

[5] SentiWordNet.

[6] TextBlob: Simplified text processing TextBlob 0.11.0 documentation.

[7] Xue Bai. Predicting consumer sentiments from online text. *Decision Support Systems*, 50(4):732–742, 2011.

[8] Tseng You-De Chin Chen Chien. Quality evaluation of product reviews using an information quality framework. *Decision Support Systems*, pages 50:755–68, 2011.

[9] Adnan Duric and Fei Song. Feature selection for sentiment analysis based on content and syntax models. *Decision Support Systems*, 53(4):704–711, 2012.

[10] Guang-Bin Huang, Hongming Zhou, Xiaojian Ding, and Rui Zhang. Extreme learning machine for regression and multiclass classification. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 42(2):513–529, 2012.

[11] Walaa Medhat, Ahmed Hassan, and Hoda Korashy. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4):1093–1113, 2014.

[12] F. Å. Nielsen. Afinn, mar 2011.

[13] Bo Pang and Lillian Lee. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the ACL*, 2004.

[14] Bo Pang and Lillian Lee. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In *Proceedings of the ACL*, 2005.

[15] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pages 79–86. Association for Computational Linguistics, 2002.

[16] Ioannis Pavlopoulos. Aspect based sentiment analysis. 2014.

[17] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[18] Olga Uryupina, Barbara Plank, Aliaksei Severyn, Agata Rotondi, and Alessandro Moschitti. Sentube: A corpus for sentiment analysis on youtube social media. *Cited by*, 4, 2014.

[19] Dingqi Yang, Daqing Zhang, Zhiyong Yu, and Zhu Wang. A sentiment-enhanced personalized location recommendation system. In *Proceedings of the 24th ACM Conference on Hypertext and Social Media*, pages 119–128. ACM, 2013.

[20] Kuat Yessenov and Saša Misailovic. Sentiment analysis of movie review comments. *Methodology*, pages 1–17, 2009.