Public Sentiment Analysis on IOCL Petrol Pumps

Chirag Garg[#1], Parth Jindal[*2], Kushagra Singh[#3]

[#]*Computer Science Department, Netaji Subhas University of Technology*
Dwarka, Delhi

chirag.garg.ug20@nsut.ac.in

kushagra.singh.ug20@nsut.ac.in

parth.jindal.ug20@nsut.ac.in

***Abstract*— The public comments on the articles of the general populace on various Petrol Pumps to present their opinions. However, not all the comments will be relevant to our field of study. This project analyzes the public comments on Petrol Pumps to get clues to what IOCL considers worth promoting. It extracts the explicit aspects of the comments and detects the sentiment orientation on different aspects to predict the usability index and areas of improvement of a selected petrol pump. Opinion mining from customer reviews has become prevalent in recent years. Usually, sentences are classified independently even though they are integral to the review's argumentative backbone. Intuitively, sentences in a review build and elaborate upon each other; knowledge of the review structure and sentential context should thus inform the classification of each sentence.**

***Keywords*—– Opinion Analysis · Aspect Extraction · Sentiment Orientation · Text Classification· IOCL**

I. Introduction

This project explores the public's opinions on the various aspects of an independent and fully functioning Petrol Pump and tries to find the best-suited petrol pump in the vicinity. The hypothesis is that every disappointed user would take out time to punch out a negative review in equal proportions to that of a happy consumer typing out an equally lovely review across various regions without any bias.

Opinion analysis is an active research field, and thanks to many researchers' work, plenty of methods for opinion analysis have been proposed. The project's methodology is mainly generated from Sentiment Analysis and Opinion Mining (Bing Liu, 2012)[1]. It is an aspect-based analysis and tries to use the sentiment orientation of comments on different aspects to perform the binary classifications.

II. Related Work

There are many comments in the dataset that was used for analysis. All of the comments are cleaned up for the subsequent phases. Aspect extraction [3, 6, 7], opinion identification [8, 9], and aspect-based sentiment categorization [10, 11, 12] are some of the key subtasks of aspect-based sentiment analysis, which is a basic task in the dynamic analysis inquiry field [4, 5].

Several earlier studies [13, 14] attempted to handle these several subtasks simultaneously while concentrating most of their research efforts on a single subtask. The reason for this is that the remaining research tasks are still difficult.

Previous methods have relied on expensive hand-crafted classifiers with features based on n-grams, parts of speech, negation words, and sentiment lexica [16,17]. The only method that takes numerous sentences into account is the model developed by Zhang and Lan[15]. However, because it merely extracts elements from the previous and next sentences without any structuring, it is less expensive than this. An LSTM that analyses sentiment toward a target word based on its position [18] and a recursive neural network that needs parse trees [19] are two examples of neural network-based techniques. This model, in contrast, does not need feature engineering, positional data, or parser outputs, which are frequently unavailable for low-resource languages.

III. Problem Statement and Methodology

The main challenge in this work is to fetch the reviews of multiple users from our given database and find the emotions behind the same to determine the average level of satisfiability of the respective consumer.

The analysis regards the frequent nouns or noun phrases as aspects, which are all explicit aspects, and detects the sentiment orientation of comments on these aspects to build a sentiment orientation matrix.

The opinion of the public discussed in this project is a quadruple (*ei, aij, sijk, hk*)

where *ei* is the topic of the comment, *aij* is an aspect of *ei*, *sijk* is the sentiment orientation on aspect *aij* of the topic *ei*, *hk* is the opinion holder. In this paper, the topic *ei* is fixed at IOCL Petrol Pumps.

This work tries to solve the problem by performing classification using the sentiment orientation matrix. The whole procedure is implemented as the following five steps.

A. Text Cleaning

The dataset used for analysis is a bagful of comments. For the following steps, all the comments are cleaned based on sentence-level. All the sentences are tokenized, and words in the sentences are lemmatized to get unigrams by Part-of-Speech Tagging. However, at the same time, stop words are not abandoned.

B. Aspect Extraction

Under the reviews, people are discussing different sub-topics, which are called aspects of this work, such as staff behavior. This project focuses only on explicit aspects based on the frequency, which are nouns and noun phrases appearing in the comment body. In this case, all the noun phrases are assumed to be composed of two words, i.e., all the noun phrases are bi-grams.

Based on the uni-grams, bi-grams are built at the sentence level to find candidate noun phrases. This project assumes that the candidate noun phrases should be in the pattern of two nouns, or stop word plus noun or adjective plus noun. So, to efficiently get the bi-grams, bi-grams without nouns inside are abandoned. To obtain the valid noun phrases and rank them, the model calculates Pointwise Mutual Information (PMI) scores for all the bi-grams. As shown in equation 1, PMI values consider the correlation between the two words inside the noun phrase, avoiding the case where the noun phrases, in fact, are the aspects of the aspects. 50 bi-grams with the highest PMI scores are selected, but the meaningless ones among them are abandoned.

EQUATION I
PMI SCORE

$$PMI(a, b) = log(\frac{(P(a,b)}{(P(a)*P(b))})$$

Except for these selected noun phrases, all the other bi-grams are split into uni-grams again, and the corpus is recleaned by removing the stop words. The comments as documents are used to calculate the term frequency-inverse document frequency (TF-IDF) for all the nouns and candidate noun phrases. Twenty features with the highest TF-IDF (refer to equation 2) are selected as aspects.

C. Aspect Categorization

The aspects found from the last step are the explicit aspects of this analysis, but these aspects have other expressions because different people may have different describing habits. So the target of this step is to group aspect expressions into aspect categories.

Assuming that aspects of expressions belonging to the same category have the same context, it is necessary to compare the context of words and phrases. A Word2Vec model is trained to represent all the unigrams and bigrams as vectors; thus, it is possible to compare the context between them. These vectors can be seen as a description of the context of each element in the vocabulary. Skip-gram model is used to get the input for the word embedding representations. Through skip-gram, all the sentences in all the comments are organized into sequences. The length of the context window is 3, and each gram concerns the 2 grams surrounded as neighbors.

EQUATION 2
TF- IDF

$$TF -- IDF - e = \sum_{j=1}^{n} \frac{TF-IDFj,d}{e^{distance(i,j)}}$$

$$TF -- IDF_{j,d} = TF * IDF$$

$$TF = n_{id} / \sum k n_{kd}$$

$$IDF = log(\frac{|D|}{N_i})$$

Since all the words and phrases are in a numeric form, the cosine similarity (refer to equation 3)distance between them can be calculated to see the difference in their contexts. With the word embedding model, the model takes the top 10 words or phrases closest to it for each aspect category extracted in the last step, regarded as different aspect expressions. However, some expressions may not be nouns or noun phrases, so these are removed from the category.

EQUATION 3
COSINE SIMILARITY

$$cos(\theta) = \frac{A.B}{|A|.|B|}$$

D. Sentiment Orientation

This paper uses the sentiment orientation matrix to represent the public's opinions. Rows of sentiment orientation matrix represent comments, and columns represent aspects. The values of the matrix are among -1, 0, and 1, representing positive, neutral, and negative, respectively. Since there are no labels on polarity, an unsupervised learning method is adopted.

The sentences containing at least one aspect are called opinion sentences here. For each comment, if it does not contain any opinion sentence on one aspect, then 0 is assigned to it on that aspect, regarded as neutral. For each aspect, all the comments with opinion sentences on this aspect are collected, and these comments abandon the sentences which do not contain this aspect. Since there is no sentiment orientation label and comments show only two different sentiment orientations, positive and negative, K-Means is implemented to cluster the comments into two groups.

The comments from the Bag-of-Word model are used to represent the comments by the TF-IDF approach. The two comments which are closest to the centroid of each cluster are called center comments. The two center comments represent two sentiment polarities of the aspect. After checking the text of center comments, label them positive and negative manually. Then the comments in the same cluster are labeled after the center comment. In this way, -1 is assigned to comments labeled with negative on one aspect, and otherwise, one is assigned to comments labeled with positive.

EQUATION 4
COMPARISON PARAMETERS

$$precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1 = \frac{2*precision*recall}{precision + recall}$$

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN}$$

$$Specificity = \frac{TP}{TP+FP}$$

E. Text Classification

The final step is to build the relationship between the sentiment orientation matrix and the editor’s selection by classification. This step uses Random Forest and Logistic Regression as tools.

IV. EXPERIMENT RESULTS

This column is a textual representation of the output of our model.

A. Data Description

The data analyzed in the project is taken directly from IOCL to figure out the worst-performing states from the bunch and the aspects that need attention and improvement. The features of data used for analysis are the followings:

- State - It tells the national state of the respective petrol pump.

-Rating - Customer Rating Level on a scale of 5.

-Review_description- Summary of the review as per user.

-Review_data - Detailed timeline of events leading to queries, suggestions, or praises.

B. Results and evaluation

After implementing all the procedures introduced above, we successfully get three aspect categories and their corresponding aspect expressions, as listed in Table 1. Figure 1 shows the sentiment orientation distribution of data on different aspects and the distribution of target variables.

The comments with no opinion sentences are abandoned, while a total of 26,717 comments are used for classification tasks. A few of them are not selected, while 25,907 are selected for further evaluation. Due to unbiased labels, this paper uses Adaptive Synthetic Sampling (ADASYN) to get a more balanced dataset to prevent overfitting and bad performance on True Positive. After adjustment, the data is transformed into 25,907 not selected and 26,092 selected.

TABLE I

Aspect category and its aspect expression

Aspect category	Aspect Expression
staff	staff be, all staff, bad, behavior, customer, behave, area
station	Pump, road, place, petrol, fuel station, big, station with
card	Part, cash, debit card, credit card, upi, accept, credit, payment, card payment, card accept
place	road, visit, location, this place, busy, all time
staff	staff be, all staff, bad, behavior, customer, behave, area
all facility	Available, facility, toilet, water, drinking water, washroom, clean, free air,air
work	Machine, fill, guy, put, person, cheat, ask, wait
quiality	Quantity, and quantity, good quality, accurate, petrol, product, petroleum, pure
oil	Corporation, outlet, operate, engine, company, petroleum, leg, product, city
diesel	gas,petrol diesel, provide, premium,good,leg,pure,station with
fill	put,tank,bike,refill, person, puncture, wait,guy,u,car

Figure 2 describes the output of each model keyed using Table 2. It is a graphical representation of the confusion Matrix of each model. Figure 3 describes the Receiver Operating Characteristic of each model to give a better look at the model.

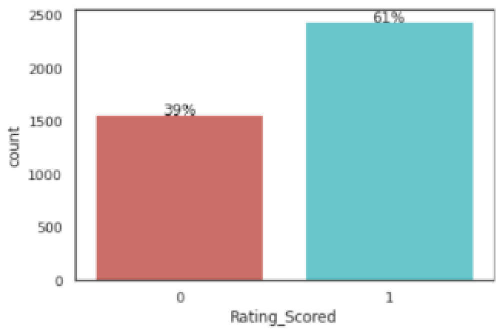
TABLE 2

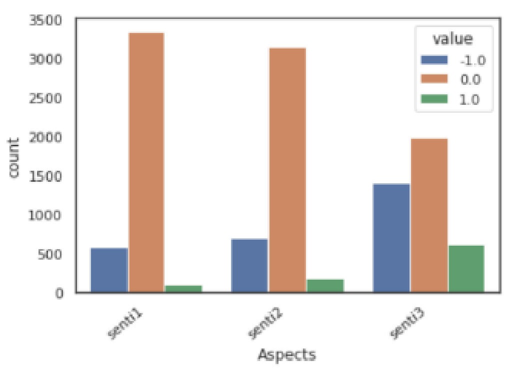
Comparison of results

No.	Description	CV Score	Test Score
Model A	Random Forest(Num)	0.75	0.747
Model B	Logistic Regression(Num)	0.562	0.561
Model C	Random Forest(Categorical)	0.676	0.672
Model D	Logistic Regression(Categorical)	0.592	0.60

FIGURE I

DISTRIBUTION OF INPUT AND TARGET VARIABLES





The classification is implemented by Random Forest and Logistic Regression on numeric input and categorical input (values of sentiment orientation transferred to dummy and 0 as baseline) respectively. The evaluation strategy to evaluate the performance of the model is 10-fold Cross Validation. They classification report of four classifier models is presented in Table 3.

TABLE 3
CLASSIFICATION REPORT OF MODELS

Model	Label	Precision	Recall	F1 Score
A	0	0.70	0.85	0.77
	1	0.81	0.64	0.72
B	0	0.56	0.56	0.56
	1	0.56	0.56	0.56
C	0	0.75	0.51	0.61
	1	0.63	0.83	0.72
D	0	0.63	0.47	0.54
	1	0.58	0.73	0.65

It is found that random forest on numeric variables has the best accuracy performance and good results on precision and recall, while logistic regression's performance is not good, only a bit better than purely random classifiers. However, the classifiers with category input both have high true positives.

C. Concluding Remarks

From the result of random forest on numeric variable, Accuracy, Precision, Recall, and F1-Score (refer to equation 4)are not low for the classification task, all higher than 0.6, meaning that the sentiment orientation from the Public on different aspects has a correlation with the Binary rating scale. From the confusion matrix, it is found that most of the errors come from False Negative. However, the classifiers on categorical input have a good performance on False Negative. This phenomenon means that some sentiment polarities are wrong while the clustering performs well. However, this can prove that the aspects extracted are correct. The possible reason for the present situation may be because some other important aspects are not extracted. In summary, the experiment results manifestly demonstrate that shows their opinions, which means the relation found in the experiment, which connects the Public's opinions and Binary ratings, exists. To improve the result and find more clues on the relation, there is a lot to do for future work. Firstly, for sentiment orientation judgment, the number of labels for sentiment orientation is small, and thus it brings uncertainty to the analysis. On the one hand, to be more precise on the polarity, other information retrieval tools can be exploited. In the comments under Petrol Pump, not many sentiment words, but many Sarcasm and comparative words appear in the context. Other methods other than the Bag-of-Word model should be used to compare the emotion of the context. On the other hand, positive and negative are not sufficient to describe people's opinions on an aspect. Multi-classification can be implemented. Secondly, this project does not concern any implicit aspects. Part of the comments is abandoned because they do not contain aspects extracted. Topic models or sentiment words for detecting implicit aspects can be implemented to use more comments. It is necessary to explore this connection because it is the direct reason they wrote a comment. What is more, other classifiers or ensemble methods should be exploited for better classification models for this problem.

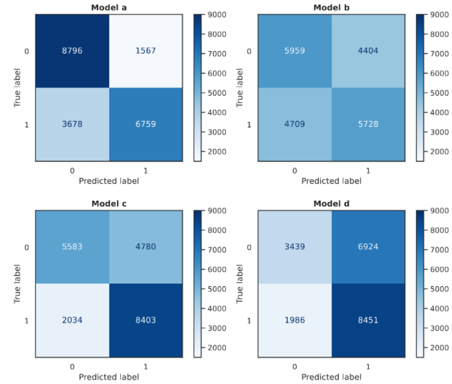
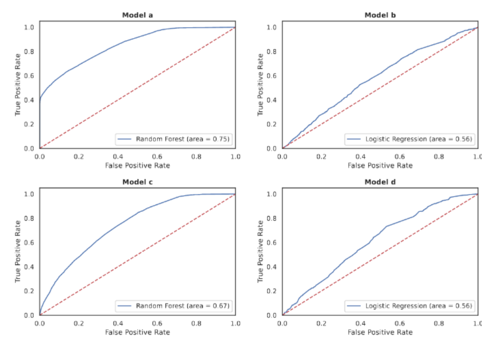


FIGURE 2

CONFUSION MATRIX

FIGURE 3

RECEIVER OPERATING CHARACTERISTIC



REFERENCES

[1] Hu, M., and Liu, B., 2004, August. Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 168-177).

[2] Liu, B., 2012. Sentiment analysis: A fascinating problem. In Sentiment Analysis and Opinion Mining (pp. 1-8). Springer, Cham..

[3] Zhiyuan Chen, Arjun Mukherjee, and Bing Liu. 2014. Aspect extraction with automated prior knowledge learning. In Proc. of the 52nd Annual Meeting of the Association for Computational Linguistics. ACL, Baltimore, Maryland, 347–358.

[4] Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. Removal-2014 task 4: Aspect based sentiment analysis. In Proc. of the 8th International Workshop on Semantic Evaluation. Dublin, Ireland, 27–35.

[5] Shuai Wang, Zhiyuan Chen, and Bing Liu. 2016. Mining aspect-specific opinion using a holistic lifelong topic model. In Proc. of the 25th International Conference on World Wide Web. International World Wide Web

Conferences Steering Committee, Montréal, Canada, 167–176.a

[6] Qiaozhu Mei, Xu Ling, Matthew Wondra, Hang Su, and ChengXiang Zhai. 2007. Topic sentiment mixture: modeling facets and opinions in weblogs. In Proc. of the 16th international conference on the World Wide Web. ACM, Banff, Alberta, Canada, 171–180.

[7] Soujanya Poria, Erik Cambria, and Alexander Gelbukh. 2016. Aspect extraction for opinion mining with a deep convolutional neural network. Knowledge-Based Systems 108, Supplement C (2016), 42–49.

[8] Ozan Irsoy and Claire Cardie. 2014. Opinion mining with deep recurrent neural networks. In Proc. of the 2014 Conference on Empirical Methods in Natural Language Processing. ACL, Doha, Qatar, 720–728.

[9] Yue Lu and Chengxiang Zhai. 2008. Opinion integration through semi-supervised topic modeling. In Proc. of the 17th International Conference on World Wide Web. ACM, Beijing, China, 121–130.

[10] Long Jiang, Mo Yu, Ming Zhou, Xiaohua Liu, and Tiejun Zhao. 2011. Target- dependent twitter sentiment classification. In Proc. of the 49th Annual Meeting of the Association for Computational Linguistics. ACL, Portland, Oregon, USA, 151–160.

[11] Svetlana Kiritchenko, Xiaodan Zhu, Colin Cherry, and Saif M. Mohammad. 2014. NRC-Canada-2014: Detecting aspects and sentiment in customer reviews. In Proc. of the 8th International Workshop on Semantic Evaluation. Dublin, Ireland, 437–442.

[12] Duyu Tang, Bing Qin, and Ting Liu. 2016. Aspect level sentiment classification with deep memory network. In Proc. of the 2016 Conference on Empirical Methods in Natural Language Processing. ACL, Austin, Texas, USA, 214–224.

[13] Samaneh Moghaddam and Martin Ester. 2013. The FLDA model for aspect-based opinion mining: addressing the cold start problem. In Proc. of the 22nd International Conference on WWW. ACM, Rio de Janeiro, Brazil, 909–918.

[14] Guang Qiu, Bing Liu, Jiajun Bu, and Chun Chen. 2011. Opinion word expansion and target extraction through double propagation. Computational Linguistics 37, 1 (2011), 9–27.

[15] Zhihua Zhang and Man Lan. 2015. ECNU: Extracting Effective Features from Multiple Sequential Sentences for Target-dependent Sentiment Analysis in Reviews. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)

[16] Maria Pontiki, Dimitrios Galanis, John Pavlopoulos, Haris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. SemEval-2014 Task 4: Aspect-Based Sentiment Analysis. Proceedings of the 8th International Workshop on Semantic Evaluation (Se- mEval 2014), pages 27–35.

[17] Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. 2015. SemEval-2015 Task 12: Aspect Based Sentiment Analysis. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), pages 486–495.

[18] Duyu Tang, Bing Qin, Xiaocheng Feng, and Ting Liu. 2015. Target-Dependent Sentiment Classification with sentiment Analysis. Proceedings of the 2015 Conference on Empirical Methods in NatLong Short Term Memory. arXiv preprint arXiv:1512.01100

[19] Thien Hai Nguyen and Kiyoaki Shirai. 2015. PhraseRNN: Phrase Recursive Neural Network for Aspect-based Natural Language Processing, (September):2509–2514.