

Applying MCDM Techniques for Ranking Products Based on Online Customer Feedback

J. Santhana Preethi*, A.M. Abirami**, A. Askarunisa***, G. Sathya Priya****, E. Sankaragomathy*****

Abstract

Text analytics is to distil out structured information from unstructured or semi-structured text. User feedback analysis or sentiment analysis on products enables to highlight the best and worst of features and recommend the product to new buyers. The model extracts the positive and negative comments and identifies the emotions in the piece of text or n-way analysis and classification like very-positive, positive, neutral, negative or very-negative. Natural Language Processing (NLP) tools play vital role in classifying the sentiment polarity of sentences while data analytics has the role in recommendation of the product. In this paper, we propose a recommender system model to rank the products based on the feedback given by the users. Features, the topics of interest, are identified from the set of review text. Sentiments are detected from each review and thus senti-score is calculated for each feature of the product. We use the Analytic Hierarchy Process and the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), which are Multi-Criteria Decision Making techniques to rank a set of products. This method provides a logical framework to determine the benefits of each product based on the features and thus the products are ranked.

Keywords: Sentiment Analysis, User Feedback Analysis, Multi-Criteria Decision Making, Technique for Order Preference by Similarity to Ideal Solution, Analytic Hierarchy Process

Introduction

Sentiment analytics applications analyse feeds from social media data such as blogs, twitter, Facebook, etc. and help business get deep insights into how customers perceive their brand and products and what their competitors are doing to excite the market. This is achieved by extracting words from the messages posted in the social media that refer to positive or negative sentiments about a particular topic of interest. Information Extraction (IE) from these texts becomes essential in order to make automatic decisions, to improve business intelligence, etc.

Sentiment analysis is one of the applications of IE. Sentiment analysis is defined as the automatic extraction of subjective content from digital text and predicting the subjectivity as positive or negative. Sentiment analysis or opinion mining is the computational study of people's opinions, sentiments, evaluations, attitudes and emotions expressed in written text according to Liu (2012).

Analytic hierarchy process is a MCDM method that can be applied for decision making problems. Problems are decomposed into a hierarchy of criteria and alternatives. The information is then synthesised to determine relative rankings of alternatives. Both qualitative and quantitative criteria can be compared using informed judgments to derive weights and priorities. AHP technique can be used

* Department of Information Technology Thiagarajar College of Engineering, Madurai, Tamil Nadu, India.
Email: santhanapreethi28@gmail.com

** Department of Information Technology, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India.
Email: abiramiam77@gmail.com

*** Department of Computer Science and Engineering, Vickram College of Engineering, Enathi, Tamil Nadu, India.
Email: nishanazer@yahoo.com

**** Department of Information Technology, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India.
Email: sathyapriyagurusamy@gmail.com

***** Department of Information Technology, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India.
Email: sankaragomathy.it@gmail.com

in determining the best product by doing the comparative study on the features of products.

Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is a well-known technique that can be applied for decision making problems. This technique can be used to solve problems with 'm' criteria and 'n' alternatives. The distance of each alternative from the ideal solution and the non-ideal solution or the negative ideal solution is considered. The objective is to choose the alternative that is the closest to the ideal solution and farthest from the non-ideal solution.

User feedback analysis has become increasingly an active research topic as they can be seen as a main control mechanism for decision making. Measure of success relies on analyzing the emotions represented in the text. The level of complexity increases as the opinion holders uses different linguistics and images to represent their views. This paper proposes a model to apply AHP and TOPSIS techniques to rank the products based on the feedback given by the users in the web forums.

This paper is organised as follows: second section describes the related work, third section explains the proposed methodology, fourth section presents the implementation details that have been undertaken and the fifth section concludes the paper.

Related Work

Hwang and Yoon (1981) found nearly 17 methods of MCDM, and they are analysed for their efficiency in

decision making. Basic notations for these MCDM methods, their computing procedure are also explained elaborately. Azar (2000) calculated the performance of various imaging techniques in detecting the breast cancer by various MCDM methods like SAW, WPM and TOPSIS.

Caterino, Iervolino, Manfredi and Cosenza (2008) found the different methods of MCDM and their applicability to some decision making problems are observed. The effectiveness of the various MCDM techniques is examined. Bhutia and Phipon (2012) identified some effective criteria which affect the process of supplier selection. Based on Analytic Hierarchy Process (AHP), they have calculated the weights for each criterion and used those weights to the TOPSIS method to rank suppliers.

Kang and Park (2012) proposed a framework for measuring customer satisfaction of service based on the user generated contents using sentiment analysis and MCDM approach. Velasquel and Hester (2013) analysed the merits and demerits of all the twelve MCDM methods and the correlation between a MCDM method and its advantages and disadvantages are drawn to find what type of method could be applied to a particular decision making problem.

Triantaphyllou and Mann (1995) observed and solved major issues that occur during the application of AHP (Analytical Hierarchy Process) in engineering applications. Arabameri (2014) combined the indexing system method with the AHP to assess locating fire stations.

Fig. 1: Sentiment Analysis Model for Online Reviews

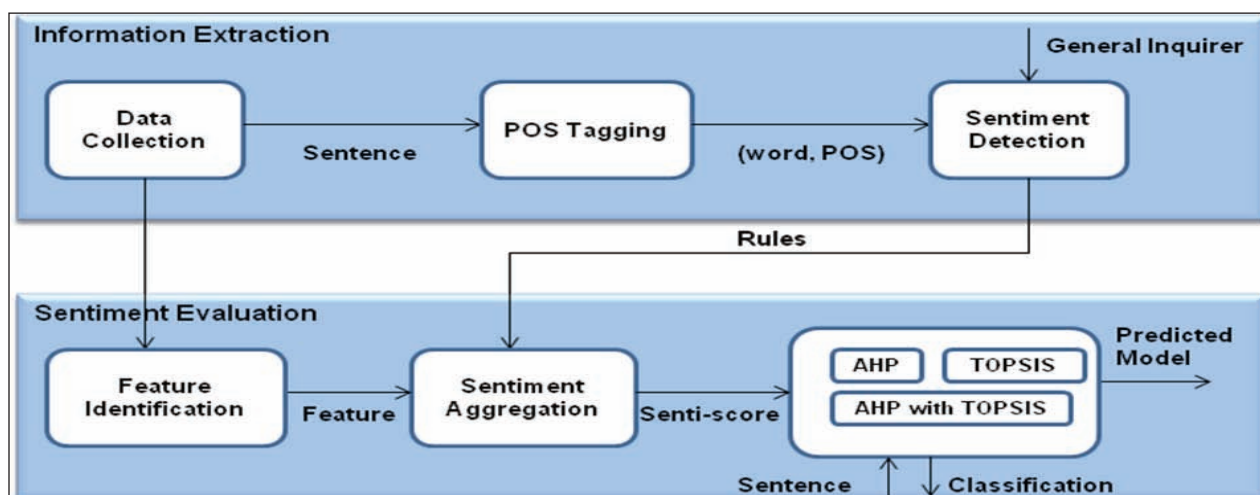
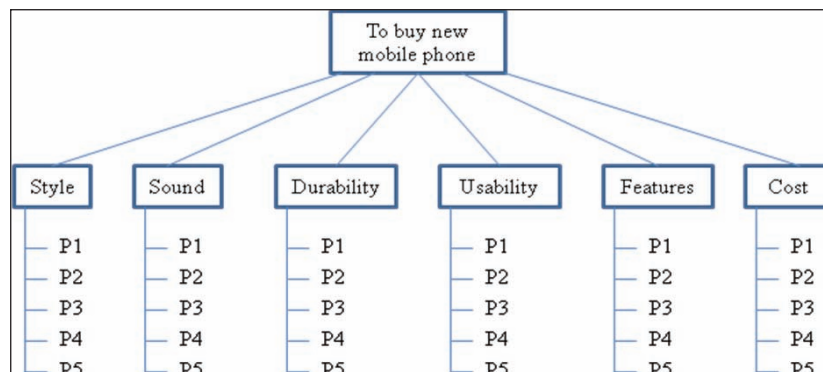


Fig. 2: Analytic Hierarchy Process

Methodology

This section describes the methodology proposed in this paper. Fig. 1 shows the proposed model and each step involved in the process are explained in the subsequent sections.

Data Collection

Sentimental tweets or sentences are collected from multiple web sources for various mobile brands. Collection of tweets is automated by using the twitterR API in R tool. Comments and opinions on the different features of mobiles are also gathered from the website.

POS Tagging

In this step, the sentences in the dataset collection are tokenised using the POS tagger of Stanford. During this process, a part of speech such as noun, verb, adverb, adjective, conjunctions, negations, and the likes are assigned to every word in the sentences. It has been made sure that the conjunctions or negatives present in the sentences are tagged correctly using General Inquirer's word dictionary.

Feature Identification

Topics of interest are extracted based on the noun category identified from the POS tagging. The major features mentioned in the tweets and opinion sentences are identified and extracted irrespective of the product models. In the case of mobiles the features like sound, style, durability, usability, features and cost are the

most recurring words in the data collected. So they are identified as the prominent features and considered for sentiment analysis.

Sentiment Detection and Evaluation

Generally the words like adjective, adverb or verb expresses the sentiment in the sentence. Based on the category determined by the General Inquirer for the sentimental word, the senti-score is assigned for the word. Overall polarity of a sentence is then determined based on the number of positive and negative words.

Sentiment Aggregation

Sentiment aggregation is done using three MCDM techniques: (i) AHP method, (ii) TOPSIS method, and (iii) Combined AHP and TOPSIS. Considering a set of products and the identified features, the features are the criteria and the products are the alternatives.

Analytic Hierarchy Process

The weights for the features are determined by analysing the collected data. Term Frequency Document matrix (M_1) is built for the set of tweets/sentences. If a tweet talks about the feature say, battery, then the senti-score of the battery is assigned as explained earlier; all other features are assigned zero. The senti-scores are assigned to all the features of all the products using the tweets. Criteria matrix (M_2) is built for pair-wise comparisons for each feature. The relative importance of features is judged by comparing the coefficients obtained during multiple regression analysis. Weights of each feature and ranking of each product are thus evaluated using AHP technique.

Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

In TOPSIS method, the decision matrices are constructed for all the products as shown below.

$$\begin{matrix}
 & P_1 & P_2 & P_j & P_n \\
 F_1 & (X_{11} & X_{12} & \dots & X_{1j} & \dots & X_{1n}) \\
 F_2 & (X_{21} & X_{22} & \dots & X_{2j} & \dots & X_{2n}) \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 F_i & (X_{i1} & X_{i2} & \dots & X_{ij} & \dots & X_{in}) \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 F_m & (X_{m1} & X_{m2} & \dots & X_{mj} & \dots & X_{mn})
 \end{matrix}$$

where X_{ij} is the senti-score about the i^{th} feature of the j^{th} product, $i=1,2,\dots,m$ are the set of features and $j=1,2,\dots,n$ are the set of products

The normalised decision matrices are obtained using the formula

$$N_{ij} = X_{ij} / \sqrt{\sum_{i=1}^m X_{ij}^2}, j=1,2,\dots,m$$

The attribute weights W_j are also calculated by taking average of the decision makers' values. The normalised decision matrix is multiplied with the respective weights to obtain the weighted normalised decision matrix V .

$$V_{ij} = W_j * N_{ij}$$

The positive ideal solution V_j^* and negative ideal solution V_j^- are obtained as the maximum and minimum values respectively among the products for each feature under consideration. The separation measures from the positive ideal solution and the negative ideal solution are calculated for all the X_{ij} values.

The separation measure of the product from the positive ideal solution is given by

$$S_i^* = \sqrt{\sum_{j=1}^m (V_{ij} - V_j^*)^2} \text{ where } i=1,2,\dots,n$$

The separation measure of the product from the negative ideal solution is given by

$$S_i^- = \sqrt{\sum_{j=1}^m (V_{ij} - V_j^-)^2} \text{ where } i=1,2,\dots,n$$

The relative closeness to the ideal solution is calculated for all the products.

$$C_i^* = S_i^- / (S_i^* + S_i^-) \text{ where } i=1,2,\dots,n$$

The product which has the maximum value among the C_i^* values is the better choice among the products.

Combined AHP and TOPSIS

In the combined method, we use the weights of the features which are obtained by performing pair-wise comparisons among the features. Using these weights, we apply TOPSIS technique and obtain the ranking of the products.

Results and Discussion

We have implemented the proposed model for online product reviews of mobiles to decide the best alternative of the product, considering all the six attributes specified and the five products under examination.

Method 1: Analytic Hierarchy Process

The weights of the features obtained after analysis of the collected dataset is shown in Table 1.

Table 1: Weights of Features

Feature	F1	F2	F3	F4	F5	F6
Weight	0.061	0.082	0.338	0.229	0.127	0.162

Table 2: Pair-wise Comparison Based on Criteria

	F1	F2	F3	F4	F5	F6
P1	0.31	0.32	0.32	0.31	0.27	0.26
P2	0.12	0.06	0.12	0.12	0.13	0.14
P3	0.21	0.22	0.22	0.22	0.22	0.24
P4	0.06	0.05	0.05	0.06	0.06	0.07
P5	0.30	0.34	0.29	0.30	0.33	0.29

The results of the pair-wise comparison of the products based on each of the considered features are listed in Table 2. The ranks obtained for the products by the AHP method are $P5 > P1 > P3 > P2 > P4$. So, the product P5 is ranked highest by this technique.

Method 2: TOPSIS

The weights of the features are calculated based on the judgements of the decision makers and the decision

matrix is constructed. Table 4 shows decision matrix obtained after normalisation.

Table 3: Decision Matrix

Criteria	P1	P2	P3	P4	P5
F1 (Sound)	4.09	3.55	4	3.25	3.125
F2 (Style)	4.25	3.86	4.07	3.61	2.938
F3(Durability)	3.83	3.31	3.74	2.795	2.5
F4(Usability)	4.06	3.62	4.15	3.22	3
F5(Features)	3.62	3.83	4.26	3.46	3.125
F6 (Cost)	3.1	3.86	4	2.71	3.438

Table 4: Normalised Decision Matrix

Criteria	P1	P2	P3	P4	P5
F1 (Sound)	0.504	0.438	0.493	0.401	0.385
F2 (Style)	0.503	0.457	0.482	0.427	0.348
F3(Durability)	0.522	0.451	0.510	0.381	0.341
F4(Usability)	0.499	0.444	0.510	0.395	0.368
F5(Features)	0.440	0.465	0.517	0.420	0.379
F6 (Cost)	0.401	0.499	0.517	0.350	0.445

Table 5: Weighted Normalised Decision Matrix

Criteria	P1	P2	P3	P4	P5
F1 (Sound)	2.02	1.75	1.97	1.6	1.54
F2 (Style)	1.91	1.74	1.83	1.63	1.32
F3(Durability)	1.57	1.36	1.53	1.14	1.02
F4(Usability)	1.6	1.42	1.63	1.27	1.18
F5(Features)	1.58	1.68	1.86	1.51	1.37
F6 (Cost)	1.53	1.9	1.97	1.33	1.69

Table 5 shows the matrix obtained after applying the weights on the normalised decision matrix and Table 6 shows the relative closeness to the ideal solution.

Table 6: Relative Closeness to Ideal Solution

Criteria	P1	P2	P3	P4	P5
s*i	0.35	0.74	0.65	0.83	1.2
s'i	1.14	0.7	1.07	0.74	0.28
s*i+s'i	1.49	1.44	1.72	1.57	1.48
s'i/(s*i+s'i)	0.765	0.486	0.622	0.471	0.189

P1 has the highest value of 0.765, hence P1 is the better choice among the five products considered. P1 is the closest to the ideal solution as well as farthest from the non-ideal solution.

Method 3: Combined AHP and TOPSIS

The weights obtained by AHP method shown in Table 1 are used in the combined method.

Table 7: Normalised Decision Matrix

Criteria	P1	P2	P3	P4	P5
F1 (Sound)	0.505	0.438	0.494	0.401	0.386
F2 (Style)	0.504	0.458	0.482	0.428	0.348
F3(Durability)	0.523	0.452	0.511	0.382	0.341
F4(Usability)	0.499	0.445	0.51	0.396	0.369
F5(Features)	0.44	0.466	0.518	0.421	0.38
F6 (Cost)	0.401	0.5	0.518	0.351	0.445

Table 7 shows the normalised decision matrix obtained for this method.

Table 8: Weighted Normalised Decision Matrix

Criteria	P1	P2	P3	P4	P5
F1 (Sound)	0.031	0.027	0.03	0.025	0.024
F2 (Style)	0.041	0.037	0.039	0.035	0.028
F3(Durability)	0.177	0.153	0.173	0.129	0.115
F4(Usability)	0.115	0.102	0.117	0.091	0.085
F5(Features)	0.056	0.059	0.066	0.053	0.048
F6 (Cost)	0.065	0.081	0.084	0.057	0.072

Table 8 shows the matrix obtained after applying the weights on the normalised decision matrix and Table 9 shows the relative closeness to the ideal solution.

Table 9: Relative Closeness to Ideal Solution

Criteria	P1	P2	P3	P4	P5
s*i	0.017	0.042	0.02	0.039	0.079
s'i	0.072	0.044	0.078	0.026	0.01
s*i+s'i	0.089	0.086	0.098	0.065	0.089
s'i/(s*i+s'i)	0.806	0.507	0.796	0.406	0.113

P1 has the highest value of 0.806, hence P1 is the better choice among the five products considered.

Conclusions

The proposed model ranked the products using the user feedback by evaluating the sentiment orientation of the sentence. Using the various methods we identify the alternative which is the best choice compared to the other

alternatives. It enables for better decision making for the product buyers and sellers.

References

- Arabameri, A. (2014). Application of the Analytic Hierarchy Process (AHP) for locating fire stations: Case study Maku city. *Merit Research Journal of Art, Social Science and Humanities*, 2(1), 1-10.
- Azar, F. S. (2000). Multi-attribute decision-making: Use of three scoring methods to compare the performance of imaging techniques for breast cancer detection. University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-00-10. Retrieved from http://repository.upenn.edu/cis_reports/119/ (accessed June 29, 2015).
- Bhutia, P. W., & Phipon, R. (2012). Application of AHP and TOPSIS method for supplier selection problem. *IOSR Journal of Engineering*, 2(10), 43-50.
- Caterino, N., Iervolino, I., Manfredi, G., & Cosenza, E. (2008). *A comparative analysis of decision making methods for the seismic retrofit of RC buildings*. Paper presented at the 14th World Conference on Earthquake Engineering, Beijing, China.
- Hwang, C. L., & Yoon, K. P. (1981). *Multiple attribute decision making: Methods and applications*. New York, Springer-Verlag.
- Kang, D., & Park, Y. (2012). *Measuring customer satisfaction of service based on an analysis of the user generated contents: Sentiment analysis and aggregating function based MCDM approach*. Paper presented at the 6th IEEE International Conference on Management of Innovation and Technology, ICMIT, Bali, Indonesia.
- Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
- Senseforth Technologies. (2014). *Consumer electronics*. Retrieved from <http://senseforth.com> (accessed on November 15, 2014).
- The Harvard General Inquirer. (2015). Guide to general inquirer category listings: General inquirer basic spreadsheet. Retrieved from http://www.wjh.harvard.edu/~inquirer/spreadsheet_guide.htm (accessed on August 12, 2015).
- The R Project. (2015). *R version 3.2.2*. Retrieved from <http://www.r-project.org> (accessed July 19, 2015).
- The Stanford NLP Group. (2015). *Stanford log-linear part-of-speech tagger*. Retrieved from <http://nlp.stanford.edu/software/tagger.shtml> (accessed August 5, 2015).
- The twitterR API. (2015). *twitterR: R Based Twitter Client*. Retrieved from <http://cran.r-project.org/web/packages/twitterR> (accessed on July 19, 2015).
- Triantaphyllou, E., & Mann, S. H. (1995). Using the analytic hierarchy process for decision making in engineering applications: Some challenges. *International Journal of Industrial Engineering: Applications and Practice*, 2(1), 35-44.
- Velasquez, M. & Hester, P. T. (2013). An analysis of multi-criteria decision making methods. *International Journal of Operations Research*, 10(2), 56-66.