

# Development of Decision Making and Analysis on Customer Reviews using Sentiment Dictionary for Human-Robot Interaction

Gurunath H Naragund<sup>1</sup>, Santhosh Kumar K L<sup>2</sup>, Jharna Majumdar<sup>3</sup>

M.Tech Scholar, Computer Science, NitteMeenakshi Institute of Technology, Bangalore, India<sup>1</sup>

Assistant Professor, Department of CSE (PG), NitteMeenakshi Institute of Technology, Bangalore, India<sup>2</sup>

Dean R&D, Professor & Head Department of CSE (PG), NitteMeenakshi Institute of Technology, Bangalore, India<sup>3</sup>

**Abstract:** Online Shopping is becoming more and more popular day by day because - most of the people will buy electronic products through e-commerce websites such as Amazon, Flipkart, Snapdeal etc. Now a day, the web provides valuable resources for obtaining consumer reviews which is useful for business analyst to analyse the data for marketing strategies. The online products receive hundreds or thousands of consumer reviews that is increasing rapidly. The reviews or comments expressed on the websites, forums, and blogs will vary the quality of expressiveness. If we consider prominent product, the various reviews expressed on such products may be in large. It is very hard for a dormant buyer to read them to make final decision whether to buy the product or not. It can be hard for the seller to keep track on particular product and to manage consumer opinions which receives enormous reviews. The Decision Making and Analysis (DMA) on customer reviews using sentiment analysis are to make firm decision whether to purchase the product or not. The decision making will purely depend on consumer reviews or sentiments expressed towards particular product. The work rely on automatic identification of consumer reviews, extracting opinions and features, data cleaning and building various dictionary. The work also concentrates on sentiment analysis which includes identification of positive and negative sentiment orientations. Further, validating the review results based on various analysis techniques. Finally, we embed this software into robot so that it can provide step by step voice instruction of running DMA tool and at the end it gives desired result to the end user about the selected product.

**Keywords:** Decision Making and Analysis, Opinion Mining, Sentiment Analysis Dictionary, Human Robot Interaction.

## I. INTRODUCTION

Online shopping or e – shopping, allows customer to purchase the products or services directly from seller or manufactures over the internet using a web browser. Online shopping websites getting more popular day by day, because of increase in the number of internet users. Now a days internet is available everywhere and anybody can access internet through desktop, laptop, smartphone and tablet. The online shopping websites are available in every device. The iphone, android and windows users can use their favourite shopping site through apps. So, there is increase in demand for online shopping.

The customer will buy the product only if the product has received good number of positive opinions. Even the manufacturer or seller will improve their product based on good and bad review received. So, the customer review analysis will be beneficial for both customer and manufacturer or seller. But, customers have to read all the reviews before buying the product. If the product received thousands of reviews then it is very difficult to read and make final decision on that product and also it is more difficult for manufacturer or seller to find out strengths and weakness of their product if it received thousands of reviews. The proposed decision making tool will tell whether the given product is good or bad depends upon the customer reviews in web.

The standard review structure commonly used while posting opinions on the web are as follows.

Type1: Only Pros and Cons

The buyer or customer will write pros and cons distinctly.

Example:www.cnet.com

Type2: Pros, Cons and Summary

The buyer will describe Pros and Cons with the summary.

Example:www.epinions.com

Type3: Free Structure

The buyer can describe freely in the web.

Example:www.amazon.com

The proposed work is mainly focuses on type 2 i.e. pros, cons and summary. Initially, the extraction and identification of pros and cons is done from product review sites. The pros and cons consisting of features or opinions that must be extracted and separated using efficient algorithms. The work also concentrates on building positive and negative dictionary from various sites and opinions expressed on other websites, which helps in comparing and analyzing the orientation of opinion. The opinion orientation identification methodology is carried out using sentiment analysis dictionary and SentiWordNet Lexical Analyzer.

The Decision making parameters and the product quality will be validated. Finally, we embed this DMA tools into robot for the Human-Robot Interaction.

In this today's world, the research area of Robotics and Artificial Intelligence has got great demand. It is very much required to make the intelligent robots which can perform in the real time and also interact with the Human (Human Robot Interaction – HRI). As an initiative, the Centre for Robotics Research at NMIT, Bangalore has developed Humanoid Robotic Head - ARIA(Advanced Robot for Interactive Application), which obey's the human instructions. This is a multi-disciplinary work consisting of undergraduate and postgraduate students from various disciplines such as Mechanical, Electrical, Electronic and Computer Science. The proposed DMA system has been integrated into the Robotic Head, where the robot will be more interactive and useful in the real time.

## II. METHODOLOGY

The architecture of the DMA system model is shown in figure 1. The DMA system works in two stages. First is the training phase, during which it will build sentiment analysis dictionary and stop word dictionary for comparison and orientation extraction. In the second phase, it will extract the pros and cons for review analysis which in turn help in decision support. The input to the system is URL of the web page or Saved HTML page containing customer reviews. The output gives the classified and summarized reviews along with decision support results.

Given the URL of product page as input, at first, the system will build all required dictionary database for sentiment analysis. Further, it will extract customer review from the given product page and extract the features and opinions using POS (Part of Speech) tagger. The extracted features and opinions are stored in separate database. The comparison is done between standard dictionary, opinion word score and review. Finally, it gives the appropriate result based on the decision making and analysis.

The various components included in the proposed system model are as follows:

### A. Website Loader:

In this module, the website is loaded if the given product's URL is valid otherwise it will give an error message.

*Example:* [http://www.epinions.com/reviews/elec\\_Camera-Digital\\_CanonPowerShott-Canon\\_PowerShot\\_S10/1944](http://www.epinions.com/reviews/elec_Camera-Digital_CanonPowerShott-Canon_PowerShot_S10/1944)

### Source Code Extraction:

In this module, The source code of webpage is extracted from the specified URL. The source code will be in the form of HTML tags.

### C. Pros and Cons Extraction:

The input to this module is the source code of webpage containing reviews and the output is the extracted pros and cons.

This works in two stages:

- i. Pros and Cons Extraction
- ii. Storing extracted Pros and Cons into a database.

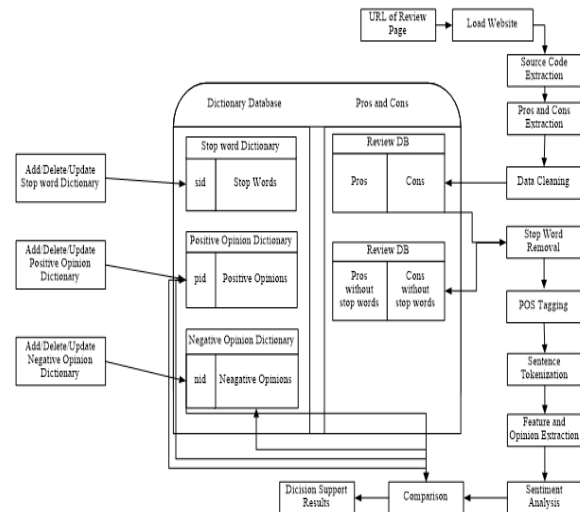


Figure 1. The architecture of the proposed DMA model

### D. Data Cleaning:

In this module, the pros and cons may or may not contain noise such as comma, full stop, special characters etc that should be removed.

#### Example:

*Before Data Cleaning:* I stay on my "phone" and it has to be able to just about anything.

*After Data Cleaning:* I stay on my phone and it has to be able to just about anything

### E. Stop word Removal:

The extracted pros and cons are input to this module. Here, the stop words such as to, for, and etc., are removed and stored into database.

#### Example:

#### Pros and Cons with stop words:

better image display compared to the ipad has the ability for HD movies

#### Pros and Cons without stop words:

betterimage display compared ipad ability HD movies

### F. POS Tagger:

The pros and cons without stop words are the input to this module. The POS tagger will tag the sentence in terms of noun(NN/NNS/NNP/NNPS), adjective(JJ/JJS/JJR), verb(VB/VBN), adverb(RB/RBR/RBS) etc.

#### Example:

#### Before POS tagging:

better image display compared ipad ability HD movies

#### After POS tagging:

better/RBR image/NN display/NN compared/VBN  
ipad/NN ability/NN HD/NN movies/NNS

### Sentence Tokenization:

In this module, the pros and cons without stop word sentences are input. The each sentences is split into tokens or words, which will be helpful for the next module.

Example: {better} {image} {display} {compared} {ipad} {ability} {HD} {movies}

**H. Feature and Opinion Extraction:**

The tagged sentence of pros and cons are input to this module. This will extract the feature and opinions and stored into database.

**I. Stop Word Dictionary:**

In this module, we will build the stop word dictionary which will be stored in the database for stop word removal function.

**J. Positive Opinion Dictionary:**

In this module, the positive opinion dictionary is built using positive opinion words list and stored in database for Sentiment Analysis.

**K. Negative Opinion Dictionary:**

In this module, the negative opinion dictionary is built using negative opinion words list and stored in database for Sentiment Analysis.

**L. Sentiment Analysis and Comparison:**

The output of feature and opinion extraction will be input to this module. This module will compare with standard opinion dictionary database and calculate the score of pros and cons. It will then compare with threshold value and gives the results whether to buy the product or not.

**III. INTEGRATION OF DMA SYSTEM TO HUMANOID ROBOTIC HEAD**

The integration of DMA system model in the Humanoid Robotic Head is shown in figure 2. The workflow of the integrated system is as follows:

- i. The proposed DMA, will take the input from the authorized (valid login Id and password should be given) user to review the product.
- ii. Once the input is given, the user can select the review from the online/offline mode.
- iii. The robot will search for the appropriate URL (offline/online) for the given product.
- iv. At the end the robot will make decision based on Sentimental Analysis and tells the appropriate opinion about the product.
- v. At each and every module, the robot will instruct the user to operate the API and gives the background process work through voice instruction. The robot will wait for the user to interact with the API for further process.
- vi. In every module, the appropriate commands are given to the robot control system to work in the real time and appropriate voice files are generated to instruct the users.

Table 1, gives the sample analysis result for digital camera. The positive and negative score of each product is tabulated and result is shown.

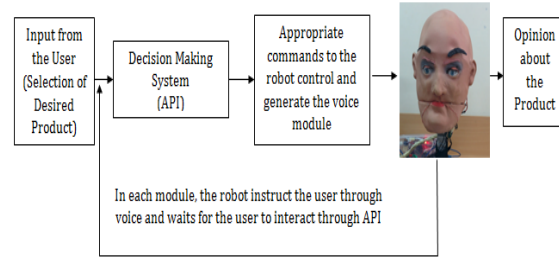


Figure 2. Integration of DMA system with Robotic Head

**IV. RESULTS AND ANALYSIS**

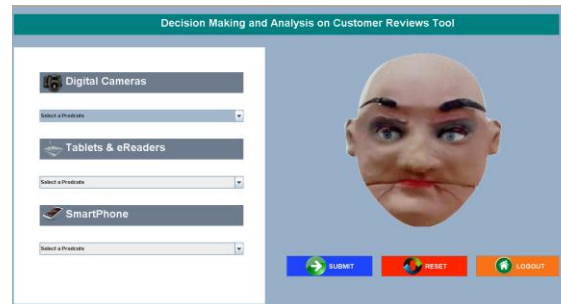


Figure 1. Selection Panel for different categories of Product

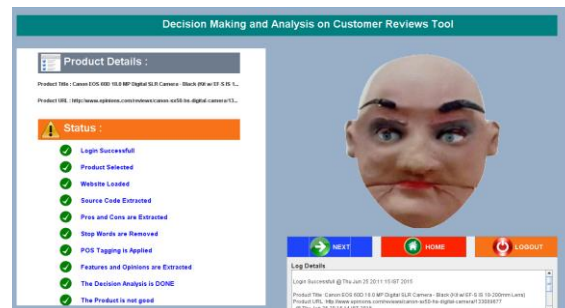


Figure 2. Updating of status and log details for each module

The decision analysis is done on the basis of positive and negative score of the respective product selected. The positive score must be high than the negative score of a product to be considered as the good quality of the product. The rule-based classification is considered to provide the appropriate decision based on the customer's review. The figure 3-5 gives the overview of the application interface designed for the DMA, which is given to the users.



Figure 5. The final DMA result for the selected product

TABLE I: A SAMPLE RESULT OF DECISION MAKING ANALYSIS ON DIGITAL CAMERA PRODUCT

Sl No.	Product Title	Positive Score	Negative Score	Product Result
1.	Nikon D3000 10.2 MP Digital SLR Camera Black (Kit w/ AFS DX VR 1855mm Lens)	0.61	0.38	OK
2.	Canon EOS Rebel T3i / 600D 18.0 MP Digital SLR Camera Black (Kit w/ EFS IS 18135mm Lens)	0.80	0.20	GOOD
3.	Canon EOS 60D 18.0 MP Digital SLR Camera Black (Kit w/ EFS IS 18200mm Lens)	1.0	0.0	EXCEL LENT
4.	Canon EOS Rebel T5i / 700D 18.0 MP Digital SLR Camera Black (Body Only)	0.25	0.75	BAD
5.	Nikon D5100 16.2 MP Digital SLR Camera Black (Kit w/ AFS VR 1855mm Lens)	0.55	0.44	AVERA GE
6.	Sony Cybershot DSCTX1 10.2 MP Digital Camera Silver	0.5	0.5	TIE

### V. CONCLUSION AND FUTURE WORK

Decision Making and Analysis on Customer Reviews using Sentiment Analysis Dictionary (DMA) Tools developed for both customer and manufacturer or seller to check whether the selected product is a good or bad. This tool will make decision based on the reviews expressed in product webpage. This tool will reduce the time to read and understand all the comments or reviews expressed in review website or product webpage. The reviews or comments contains valuable information about product which the manufacturer selling. The purchased user know what are the defects, how the product works and experience about the product.

The different customer expresses different opinions about same the products on the webpage. By extracting these reviews, analyzing and comparing will give us valuable output which will help user for buying the product, the manufactures will also get to know what are the loop holes in the product. The manufacturer will improve their product quality by using this tool. The product is reliable when customer gives positive feedback when there is mixed opinions then there is conflict in decision making capability. The DMA tool was tested with product such as Digital Camera, Smartphone and Tablets. The output

produced by this tool is satisfactory and reliable. The review is fetched from epinions.com because they used structured review format. This will help the developer to extract the reviews easily.

In future, we extend fetching the resources from other popular shopping sites such as amazon.com, flipkart.com and ebay.com, where the reviews are in unstructured format. Another future work is creating web application of this version so that it is available on internet and worldwide. For creating web application requires huge storage device and maintenances. The developed tool is URL dependent means it will only take URL as input but if user gives product title then it won't recognize product title and hence we need to create search engine for product title and respective URL so that if user enters product title then it can fetch its respective URL in back end and process it further. The tool will work like intelligent software for Robotic applications, if we add various artificial intelligent techniques so that it automatically add the product which is recently entered in the market and analyze and give the result, which will enhances the Human-Robot Interaction.

### ACKNOWLEDGMENT

The authors express their sincere gratitude to **Prof. NRShetty**, Director, NitteMeenakshi Institute of Technology and **Dr. H C Nagaraj**, Principal, NitteMeenakshi Institute ofTechnology for providing encouragement and support. The authors extend their sincere thanks to Vision Group on Science and Technology (VGST), Government of Karnataka for providing the financial support to setup the infrastructure required to carry out the research.

### REFERENCES

- [1] P.S Hiremath, Siddu P. Algur and Shivashankar S. "Web based Quality Assessment of Customer Reviews using Quartile Measure", International Journal of Recent Trends in Engineering, MAY 2009
- [2] Nitin Jindal and Bing Liu "Opinion Spam and Analysis", ACM Journal, February2008
- [3] XinfanMeng and HoufengWang "Mining User Reviews: from Specification to Summarization", ACL-IJCNLP 2009 Conference Short Papers, 2009
- [4] Mingqing Hu and Bing Liu "Mining and Summarizing Customer Reviews" ACM Journal, August 2004.
- [5] Su SuHtay and KhinThidar Lynn "Extracting Product Features and Opinion Words Using Pattern Knowledge in Customer Reviews", Hindawi Publishing Corporation The ScientificWorld Journal, September 2013
- [6] Desheng Dash Wu, LijuanZheng, and David L. Olson. "A Decision Support Approach for Online Stock Forum Sentiment Analysis" IEEE Transactions on Systems, Man and Cybernetics : Systems, Vol.44, No.8, August 2014
- [7] MsKrantiGhag and Dr. Ketan Shah, "Comparative Analysis of the Techniques for Sentiment Analysis", ICATE, 2013
- [8] KekeCai, Scott Spangler, Ying Chen and Li Zhang. "Leveraging Sentiment Analysis for Topic Detection", IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2008
- [9] A. Hamouda and M. Rohaim, "Reviews Classification Using SentiWordNet Lexicon," The Online Journal on Computer Science and Information Technology, Vol. 2, No. 1, 2011, pp. 120 - 123.
- [10] José M. Martín, Alvaro Ortigosa, Rosa M. Carro, "SentBuk: Sentiment analysis for e-learning environments", IEEE, Oct 2012, pp. 1 - 6.

- [11] Yin-Fu Huang and Heng Lin, "Web Product Ranking Using Opinion Mining", IEEE, August 2013
- [12] Viktor Hangya, Richárd Farkas, "Target-Oriented Opinion Mining from Tweets", 4th IEEE International Conference on Cognitive Infocommunications, December 2013
- [13] Wei Huang, Xin Chen, Haibo Wang, "Product Information Retrieval Based on Opinion Mining", IEEE, 2010
- [14] A. Esuli and F. Sebastiani, "SentiWordNet: A publicly available lexical resource for opinion mining", "Proceedings of Language Resources and Evaluation (LREC)", 2006
- [15] Weiping Wang, Yuanzhuang Zhou, "E-Business Websites Evaluation Based on Opinion Mining" International Conference on Electronic Commerce and Business Intelligence, 2009
- [16] Yiming Zhao, Kai Niu, Zhiqiang He, Jiaru Lin, and Xinyu Wang, Text Sentiment Analysis Algorithm Optimization & Platform Development in Social Network, IEEE, 2013

Hyderabad as the first Robotics Exhibit in Indian Museums. Dr. Majumdar has 3 patents with students of NMIT, Bangalore in last 4 years.

### BIOGRAPHIES



**Gurunath H Naragund** received B.E in Computer Science and Engineering from Visvesvaraya Technological University. He is currently Studying M.Tech Final Year in Computer Science & Engg. at Nitte Meenakshi Institute of Technology, Bangalore. His areas of interest are Web Mining, Data Mining, Sentiment Analysis and Opinion Mining.



**Santhosh Kumar K L** received B.E and M.Tech in Computer Science and Engineering from Visvesvaraya Technological University. He is currently working as Assistant Professor in the Department of Computer Science & Engg (PG), Nitte Meenakshi Institute of Technology, Bangalore. His areas of interest are Image Processing, Video Shot detection & Video Summarization, Activity Tracking and Facial Expression Recognition system.



**Dr. Jharna Majumdar** currently working as Dean R& D and Professor and Head of Computer Science and Engineering (PG) at the Nitte Meenakshi Institute of Technology, Bangalore. Prior to this Dr. Majumdar served Aeronautical Development Establishment, Defence Research and Development Organization (DRDO), Ministry of Defence, Govt. of India as Research Scientist and Head of Aerial Image Exploitation Division, Bangalore. Dr. Majumdar has 40 years of experience in R & D and Academics in the country and abroad. She has published large number of papers in National, International Journals and Conferences. Her Project with team of students from 7 engineering colleges and ISRO for building the first smallest satellite in India (a PICO satellite of weight less than 1 kg) had taken off to the orbit successfully by mid May 2010. Nitte Amateur Satellite Tracking Centre (NASTRAC) developed by a team of students from NMIT, Bangalore, under her guidance is the first Tracking Station of Small Satellites developed in the country. The research team of robotics under her guidance has developed a Robot with Innovative Vision System and installed at the Birla Science Centre,