

# Content Analysis Concerning Online Shopping in UAE: Evaluation of Impact Score from News

Jitendra Singh Rathore\*, Rajita Srivastava\*\*

## Abstract

The news in newspaper and magazine plays an imperative role to apprise public. People get influenced by the contents (either news or articles), and for businesses, this subjects to spontaneous decisions. It also corroborates greater defy for businesses owing to mammoth amount of data aggregated. The research conducts content analysis of “online shopping news” from two prominent newspapers (Gulf News and The National) and other protuberant sources (including Khaleej Times and Arabian Business) in UAE, summing to 49 news published during 2016-2017. The aim of collecting these documents is to find their impact (in terms of polarity using Plutchik emotional model) on public. The tests consist of impact identification phase, which associates news with impact aggregation and scoring.

**Keywords:** Content Analysis, Gulf News, Impact Score, Plutchik Emotional Model, The National

## Introduction

The study is grounded on two vital observations, wherein, on one side the competition in online shopping is getting tougher in UAE with Amazon acquired Souq (Turner & Wang, 2017), and brick-and-mortar retail company, Emaar, procured two online retail companies, Namshi and JadoPado (Reuters, 2017; Scott, n.d.). Secondly, these acquisitions are happening in a place where online shopping accounted for just 1% of total shopping value as compared to international average of 15-20% elsewhere (Euromonitor, 2017). Both these annotations are contrary, as companies see potential in the market and customers instead of trusting of local online shopping companies, do cross-border shopping. Almost 46% of UAE online

shoppers purchased online from websites outside of UAE (Jayakumar, 2016). This disintegration draws a major attention — does the news regarding local online shopping in UAE newspapers and magazine have any impact on consumer buying behaviour?

Newspapers and magazines do play an imperative role in informing public about the happenings; predominantly, in those areas in which audiences do not possess unswerving knowledge or experience (Happer & Philo, 2013). There is also an argument that newspapers and articles not only publicize opinion, but also generate stimulus for these opinions (Livingstone, 1994). Researchers focus on news articles and are mainly concerned about the polarity of these articles, whether they convey a positive or negative attitude towards the subject of discussion (Haider-Markel, Allen, & Johansen, 2006). These polarities generate shoppers’ mindset, referred to as “specific cognitive orientation”, resulting in different modes of emotions and information processing (Wind & Mahajan, 2001). The large linguistic miscellany in these documents makes these newsworthy, but challenging the cradle of information (Domann & Lommatzsch, 2017). Newspaper and magazine articles can be precarious, as customer can endorse or criticize a product through it. This information is of prodigious significance for a business because the existence of business depends on it (Lommatzsch, Bütow, Ploch, & Albayrak, 2017). As the reader of these news is emotional, the overall impact analysis can be evaluated using the eight different types of polarities (Plutchik, 2001).

In this paper we present a supervised methodology to compute impact analysis for news articles. Using text mining, the texts can be categorized inevitably with their polarities; thereby, overall impact score of each news can be assessed independently. In the course of this paper,

\* Assistant Professor, Faculty of Management Studies, WISDOM, Banasthali University, Rajasthan, India.

\*\* Research Scholar, Faculty of Management Studies, Rajasthan, WISDOM, Banasthali University, Rajasthan, India.  
Email: rajitariktesh@gmail.com

we describe the development of “News Analysis Model” that takes advantage of a machine learning approach to accomplish this objective. We first collect the news from three types of sources, and then conduct the evaluation of emotions using Plutchick Emotional Model. Furthermore, we compare our findings; the results are presented on a front-end. Finally, we draw an inference in which we also recommend points for additional research and augmentation of our current approach.

The complete paper is divided into four sections. Section 2 particularizes the proposed News Analysis model, which comprises of four steps. Section 3 expresses the results analysis for 49 news, contributed from three types of sources, exploring the comprehensive set of outcomes.

Section 4 accomplishes the paper with inferences and conclusion.

### News Analysis Model

There are four steps in news analysis model, as mentioned in Figure 1. Step 1 is collection of news about pertinent topic of interest. For the study, we selected three types of sources (*Gulf News*, *The National*, and Others). Step 2 is preprocessing, which is done to remove common words and URLs from the texts. Step 3 does tokenization, which chops the document into individual words. Step 4 conducts the impact of these words using Plutchik Emotional Model.

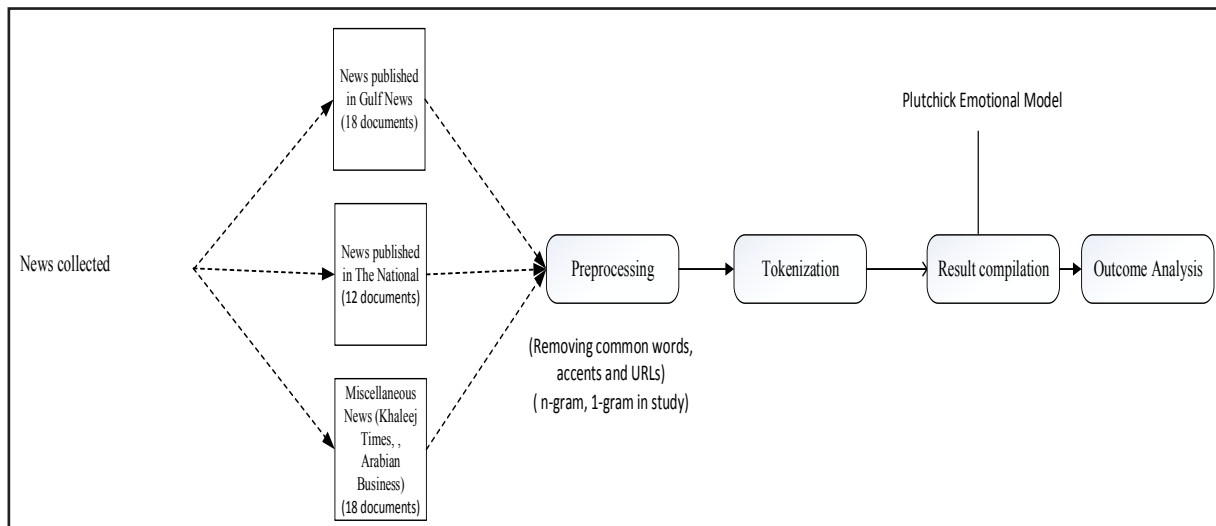


Fig. 1: News Analysis Model

#### Preprocessing

The documents collected from newspapers and magazines are in raw format. Preprocessing converts this raw format into more meaningful format for analysis. The steps for transformation are:

- Convert all text into lowercase.
- Remove all diacritics or accents in text, as héllö → hello.

- Documents also include links. These tags act as a single word and results in wrong identification of emotions. `<a href> i love online shopping </a>` gets converted into `→ i love online shopping`.
- Remove all URLs from documents.

The steps can be represented mathematically by equation (1) as:

$$\sum_{i=1}^n \prod_{j=1}^m d_{ij} = \sum_{i=1}^n \left[ \prod_{j=1}^m \begin{bmatrix} (d_{11}) \\ (d_{12}) \\ \vdots \\ (d_{nm}) \end{bmatrix} \right] = \sum_{i=1}^n \left[ \prod_{j=1}^m \begin{bmatrix} (Word_{11} Word_{12} \dots Word_{1j}) \\ (Word_{21} Word_{22} \dots Word_{2j}) \\ \vdots \\ (Word_{n1} Word_{n2} \dots Word_{nm}) \end{bmatrix} \right] \tag{1}$$

where,  $n$  is number of documents where each documents contain  $m$  words. Notice from equation (1) that each word of documents is not disjointed and embodied as sentences.

## Tokenization

Tokenization is a process of breaking documents into words (as stated in equation (1)). These words are smaller

$$\sum_{i=1}^n \prod_{j=1}^m \begin{bmatrix} (d_{11}) \\ (d_{12}) \\ \vdots \\ (d_{nm}) \end{bmatrix} = \sum_{i=1}^n \prod_{j=1}^m \begin{bmatrix} (Word_{11})(Word_{12}) \dots \dots (Word_{1j}) \\ (Word_{21})(Word_{22}) \dots \dots (Word_{2j}) \\ \vdots \\ (Word_{n1})(Word_{n1}) \dots \dots (Word_{nm}) \end{bmatrix} \quad (2)$$

Notice, that each word from collected documents is now separated into expressive tokens. Based on equation (2), the number of tokens for the collected documents is mentioned in Table 1.

**Table 1: Tokens Collected from Documents**

Document Types	Total number of documents	Number of words	Number of Tokens
Gulf News	18	10093	5146
The National	13	7243	3705
Miscellaneous	18	7731	4166

## Impact Analysis

The impact analysis is conducted using Plutchik model (Plutchik, 2001). Plutchik model divides the words into polarities of eight categories – Joy, Surprise, Trust, Anticipation, Anger, Disgust, Fear, and Sadness. The first four categories give positive impact score and later four give negative impact score, as mentioned in equations (3)

$$\sum_{i=1}^n \prod_{j=1}^m \begin{bmatrix} (IS_1) \\ (IS_2) \\ \vdots \\ (IS_j) \end{bmatrix} = \sum_{i=1}^n \prod_{j=1}^m \begin{bmatrix} (Word_{11}) \rightarrow \langle s1_j, s2_j \rangle (Word_{12}) \rightarrow \langle s1_j, s2_j \rangle \dots \dots (Word_{1j}) \rightarrow \langle s1_j, s2_j \rangle \\ (Word_{21}) \rightarrow \langle s1_j, s2_j \rangle (Word_{22}) \rightarrow \langle s1_j, s2_j \rangle \dots \dots (Word_{2j}) \rightarrow \langle s1_j, s2_j \rangle \\ \vdots \\ (Word_{n1}) \rightarrow \langle s1_j, s2_j \rangle (Word_{n1}) \rightarrow \langle s1_j, s2_j \rangle \dots \dots (Word_{nm}) \rightarrow \langle s1_j, s2_j \rangle \end{bmatrix} \quad (6)$$

## Research Analysis

The research analysis includes the independent content analysis for three types (18, 13, 18) of documents collected

components called tokens. To do so, broad rules of splitting each word in the documents by:

- Punctuations are also considered as tokens separately from words. These tokens are represented by either whitespace or full stop.
- Emoticons are also used as tokens. The impact of emoticons is relatively substantial in categorizing impact analysis.

Equation (1) can now be defined as in equation (2) as:

and (4), respectively.

$$\sum_{j=1}^m s1_j = \sum_{j=1}^m \left[ \prod_{j=1}^m Joy_j + \prod_{j=1}^m Surprise_j + \prod_{j=1}^m Trust_j + \prod_{j=1}^m Anticipation_j \right] \quad (3)$$

$$\sum_{j=1}^m s2_j = \sum_{j=1}^m \left[ \prod_{j=1}^m Anger_j + \prod_{j=1}^m Disgust_j + \prod_{j=1}^m Fear_j + \prod_{j=1}^m Sadness_j \right] \quad (4)$$

Based on equations (3) and (4), the overall impact score for each document is calculated as mentioned in equation (5) below.

$$IS_j = \left[ \sum_{j=1}^m s1_j - \sum_{j=1}^m s2_j \right] \quad (5)$$

Grounded from equation (5), if, the document is quantified to have positive impact, else if, the document provides negative impact, neutral otherwise. Overall, the impact score is revealed in equation (6).

through varied sources. Figures 2, 3, and 4 depict the content analysis of 49 news published concerning online shopping in UAE during 2016–2017. The selection is purely on news published (without revisions or public

feedback), as we intend to observe the direct impact of news and evaluate impact score as mentioned in equation (6).

### Gulf News Observations

The three distinguishable factors for *Gulf News* are. The outcome, as mentioned in Figure 2 clearly indicates.

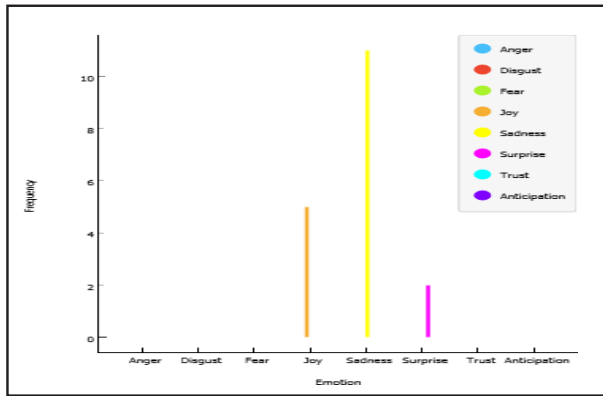


Fig. 2: Gulf News Impact Analysis

### Miscellaneous Observations

The miscellaneous collection of news includes collection from other important sources including *Khaleej Times* and *Arabian Business* and again three distinguishable factors are. The outcome of analysis clearly indicates.

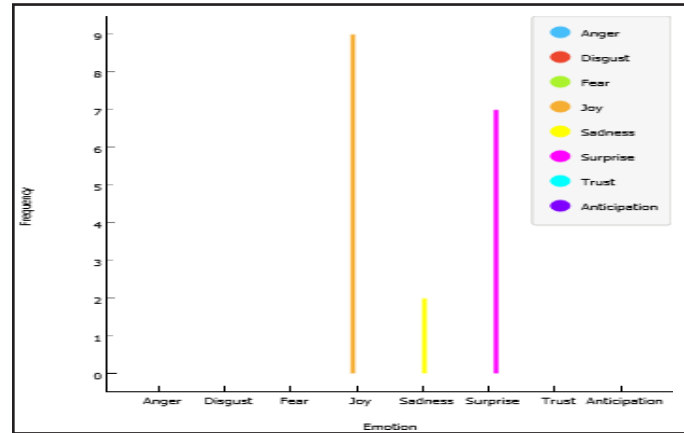


Fig. 4: Miscellaneous News Impact Analysis

### The National Observations

Again, the three distinguishable factors for *The National* are. The outcomes, as mentioned in Figure 3, illustrate.

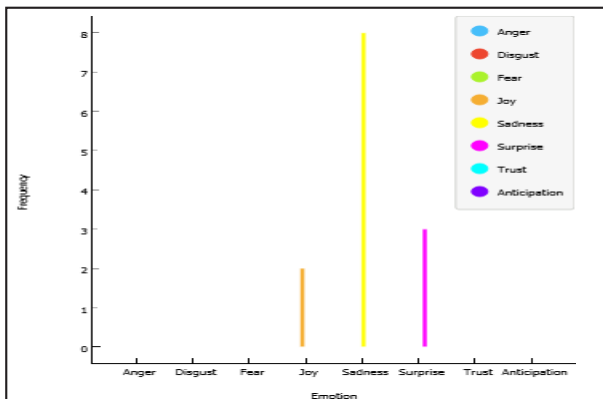


Fig. 3: The National Impact Analysis

### Inferences and Conclusions

Usage of published news in newspapers and magazines with the computational conduct of subjectivity of text becomes vital part of direct spurt on the curiosity of consumers. The results obtained through analysis of 49 such news are quite upsetting. Implementation of algorithm provides the results, as specified in Table 2. The three polarities in all three news analysis are, other polarities are set to 0 and thus ignored. Astoundingly, overall impact score of *Gulf News* and *The National* (two leading newspapers in UAE) are negative; whereas, impact of news from miscellaneous sources is exceedingly positive.

Table 2: Overall Impact Score

News Source	Documents collected	Impact --> Joy	Impact --> Sadness	Impact --> Surprise	Impact --> +ve	Impact --> -ve	Overall Impact Score
Gulf News	18	27.8	61.1	11.1	38.9	61.1	-22.2
The National	13	15.38	61.54	23.08	38.46	61.54	-23.08
Misc.	18	50	11.11	38.89	88.89	11.11	77.78

As outcomes show, news published in newspaper may also be one of the reason that UAE has highest cross-border online shopping rate. The study indicates that foremost drive for UAE online shoppers to do cross-border online shopping is – 47% regarding secure payment, 44% because of free shipping, and 43% revealed proof of product genuineness (Jayakumar, 2016). The electronic commerce companies in UAE should focus that news published in newspaper should mention the positivity regarding the above three factors to boost confidence among consumers. Similarly, for magazines and other sources of information, these companies should lay stress on advertisements and product information.

The research will be further enhanced by evaluating opinions with UAE-based electronic commerce companies through social media analysis (especially on Facebook and Twitter). Combination of these two sources of information will provide a greater insight to companies and communication strategies can be deliberated and implemented accordingly.

## References

- Domann, J., & Lommatzsch, A. (2017). A Highly Available Real-time News Recommender based on Apache Spark. *CLEF' 17: Proceedings of the 8th International Conference of the (CLEF) Initiative*.
- Euromonitor. (2017, January 1). Retailing in the United Arab Emirates. Euromonitor. Retrieved from <http://www.euromonitor.com/retailing-in-the-united-arab-emirates/report>
- Haider-Markel, D. P., Allen, M. D., & Johansen, M. (2006). Understanding variations in media coverage of u.s. supreme court decisions comparing media outlets in their coverage of Lawrence v. Texas. *The Harvard International Journal of Press/Politics*.
- Happer, C., & Philo, G. (2013). The role of the media in the construction of public belief and social change. *Journal of Social and Political Psychology, 1*(1), 321-336. Retrieved from <https://doi.org/10.5964/jspp.v1i1.96>
- Jayakumar. (2016, December 7). Infographic: UAE Online Consumers Shopping Trends Revealed. Retrieved December 1, 2017, from <https://arabiangazette.com/uae-online-consumers-shopping-trends-revealed/>
- Livingstone, S. (1994). Watching talk: Gender and engagement in the viewing of audience discussion programmes. *Media, Culture and Society, 16*(3), 429-447.
- Lommatzsch, A., Bütow, F., Ploch, D., & Albayrak, S. (2017). Towards the automatic sentiment analysis of German news and forum documents. *Springer, 1-16*.
- Plutchik, R. (2001). The nature of emotions. *American Scientist, 89*, 344-350.
- Reuters. (2017, May 11). Fund led by Dubai billionaire Alabbar buys UAE website JadoPado. Reuters. Retrieved from <https://www.reuters.com/article/emirates-technology-jadopado/fund-led-by-dubai-billionaire-alabbar-buys-uae-website-jadopado-idUSL8N1ID1GK>
- Scott, A. (n.d.). Emaar Malls takes 51 per cent stake in online fashion retailer Namshi - The National. Retrieved December 1, 2017, from <https://www.thenational.ae/business/emaar-malls-takes-51-per-cent-stake-in-online-fashion-retailer-namshi-1.60869>
- Turner, G., & Wang, S. (2017, March 28). Amazon Buys Souq.com as Middle East Online Market Takes Off. *Bloomberg.Com*. Retrieved from <https://www.bloomberg.com/news/articles/2017-03-28/amazon-wins-battle-to-buy-middle-east-e-commerce-firm-souq-com>
- Wind, Y., & Mahajan, V. (2001). *Digital marketing: Global strategies from the world's leading experts* (1st ed.). Wiley.