Article:

Tse, Ying Kei orcid.org/0000-0001-6174-0326, Loh, Han Lin, Ding, Juling et al. (1 more author) (2018) An investigation of social media data during a product recall scandal. Enterprise Information Systems.

https://doi.org/10.1080/17517575.2018.1455110

An investigation of social media data during a product recall scandal

Ying Kei Tse[1], Hanlin Loh [2], Juling Ding[3], Minhao Zhang[4]

1. The York Management School, the University of York, York, UK

Email:mike.tse@york.ac.uk

2. The York Management School, the University of York, York, UK

Email: hanlin.loh@gmail.com

3. School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China

Email: julingding@jxufe.edu.cn

4. Newcastle University Business School, Newcastle University, Newcastle, UK

Email: minhao.zhang@york.ac.uk

Abstract

As social media has become an important part of modern daily life, users often share product opinions online and these tend to spike when large companies undergo crises. This paper investigates customer online responses to a large company crisis by uncovering hidden insights in social media comments and presents a framework for handling social media data and crisis management. Analysis of textual Facebook data from users responding to the 2013 horsemeat scandal is presented. In this study, we used a novel comprehensive data analysis framework alongside a text-mining framework to objectively classify and understand customer perceptions during this horsemeat scandal. This framework provides an effective approach for investigating customer perception during a company crisis and measures the effectiveness of crisis management practices which the company has adopted. Our analyses show that social media can provide important insights into customer behaviour during crisis

communications.

Keywords: Social Media, Horsemeat Scandal, Crisis Management, Sentiment Analysis

1. Introduction

The horsemeat scandal in the British Isles began in January 2013, when the Food Safety Authority of Ireland disclosed that prominent supermarkets such as Tesco, Iceland, Aldi and Lidl had undeclared horse DNA in their frozen beef burgers that were currently on sale. Later, reports of horsemeat being sold as beef were found in many different European countries that related to products such as burgers and ready meals. This led to many products being recalled by the seller or manufacturer.

The impact of this horsemeat scandal affected not only customers and supermarket retailers, but the entire food industry. This incident led to a severe loss of customer trust in the food industry. Nevertheless, companies may benefit to have good communications with customers post-crisis to restore customer trust, and this should still be the priority for companies. Therefore, understanding how customers feel is an essential step in developing an effective product recall and crisis communication strategy for professionals (Tse et al., 2016). In order to establish an effective risk management strategy in product recall crisis and win back customer trust, company managers need to investigate how their risk communication works among customers and monitor customer perception during the outbreak of a crisis. One important tool that can be used to facilitate the formulation of a successful risk strategy in such situations is customer responses via social media platforms.

In this study, we aim to investigate how companies communicate to their customers during product recall crises by scrutinising customer responses on social media towards the communication. In recent years, companies have paid much attention to managing relations with customers using social media (Zhang et al. 2017; Jin, Liu, and Austin 2014). The data generated by social media and its potential in social media analysis have attracted a growing interest within academia, in part spurred by recent excitement over the concept of Big Data (Tse et al., 2016; Basili et al., 2017). This is at the top of the agenda for many business leaders who have sought to establish ways in which social media platforms, such as Facebook and Twitter, can be used to improve the competitiveness of their organisations (Kaplan and Haenlein 2010). Academics and practitioners have recently made use of social media data to manage operations and to direct decision-making processes.

This article proposes a new research framework for empirical analysis of social media data to facilitate understanding and exploration of customer sentiments and responses during the 2013 horsemeat crisis. A secondary aim is to examine the following research questions: (i) What are customer sentiments towards enterprising companies that are caught in the midst of a food recall crisis? (ii) How do customers respond to social media messages from such companies during a product recall crisis? (iii) How can text-mining help companies make better crisis management decisions in the future?

In this study, we investigate Facebook data relating to Tesco during the 2013 horsemeat scandal. Tesco was chosen because Tesco is the largest supermarket chain in the UK. Moreover, Tesco was in the news headlines during the outbreak of the horsemeat scandal for selling frozen *Everyday Value Beef Burgers* that contained horsemeat. Tesco suffered a lot during the scandal and was seeking a better method of interaction with customers to win back trust and lost market share. The reaction of customers towards Tesco during this crisis can provide important insights for crisis managers in affected companies when faced with similar crises in the future (Senadheera, Warren, and Leitch 2017; Tse et al. 2016). Facebook was chosen as the social media platform in this analysis because Tesco used Facebook as one of its main platforms to communicate information to the customers during the scandal. The company took crisis management actions (e.g. offering public apologies via Facebook posts) and updated latest product information on the Facebook platform. Thus, customer responses to these relevant posts are a rich data source for researchers to use to understand public perception during product recall crisis and for measuring the effectiveness of the firm's remedial actions.

2. Literature Review

In this section, relevant research in social media, crisis management and the applications of Big Data and Text Mining are reviewed. Additionally, this section reviews two essential techniques used to analyse social media data – Machine Learning (ML) for text classification and Sentiment Analysis (SA).

2.1 Social Media

Malita (2011, p.748) defines social media as "*the tools that facilitate the socialisation of content and encourage collaboration, interaction, and communication through discussion, feedback, voting, comments, and sharing of information from all interested parties*". Nearly all online platforms include social components such as a comment field for users to share and exchange information. In the current study, social media is broadly defined as a set of digital tools and applications that facilitates interactive communication and content exchange among and between the public and organisations (Jin, Liu, and Austin 2014). In recent years, social media has progressively become more and more essential for companies to use to directly communicate information to a large audience (Lee, Hutton, and Shu 2015).

Currently, Facebook, Twitter and LinkedIn are the social media sites most frequently used by businesses. All these platforms are regarded as Social Networking Sites (SNS), a common form of social media. Due to social media's strengths of instant response and direct communication, practitioners and professionals have extensively studied customer relationship management and explored customer attitudes using social media (Jin and Phua 2014; Ashley and Tuten 2015; Baird and Parasnis 2011). The opinions expressed by customers on their social networks play a significant role in influencing public opinions and behaviour (Mostafa 2013; Lee, Hutton, and Shu 2015).

2.1.1 Social Media and Big Data

Big Data commonly refers to very large datasets and is often used in the fields of popular media, business and computer science. Katal, Wazid, and Goudar (2013, p.404) define Big Data as "*a large amount of data which requires new technologies and architectures so that it becomes possible to extract value from it by capturing and analysis process*". Each individual social media user directly generates data that can be obtained from SNS providers or gathered manually with little to no effort compared with traditional data collection methods (Tufekci 2014). Big Data has revealed interesting insights to social phenomena, such as mood oscillation in millions of people across the world (Golder and Macy 2011).

Recent research indicates that the number of people using social media applications, particularly Facebook and Twitter, is constantly growing. And an increasing number of large organisations have integrated or are exploring the option of integrating social media into their business strategies to gain value in areas such as customer traffic, customer loyalty and retention, sales and revenues, customer satisfaction and brand awareness and reputation (Montalvo 2016; Trainor et al. 2014; Herzig et al. 2016). For instance, the hospitality chain Starwood Hotels and Resort has been actively using social media to seek feedback from guests, address customer concerns and complaints, give advice to potential customers and simply to keep in touch with customers (Duan et al. 2015). A study conducted by Pearson and

Wegener (2013) from Bain and Company, a renowned management consulting firm, discovered that large organisations significantly outperform their competitors if cutting-edge data analytics were adopted for Big Data usage. They also found that these companies were able to make good decisions five times faster than their competitors and were twice as likely to perform better financially within their markets because of the insights gained from data analytics.

2.1.2 Social Media and Crisis Management

Social media is slowly taking the place of traditional media and it has become very popular to report events on social platforms. Because of this, the public tends to turn to social media during crises to find out the latest news. One study showed that, during company crises in different part of the world, citizens, media and organisations use social media such as Twitter, Facebook and YouTube extensively to express their feelings and to share opinions and information about happening crises (Terpstra and Vries 2012). In addition, local eyewitnesses use social media to provide first-hand information immediately after the outbreak of a crisis (Bruns and Stieglitz 2013). Public participation on social media platforms has become a new normal in crisis management.

Social media is an important technology for crisis response, primarily because of the tools that enable open exchange of timely information through conversation and interaction(Lachlan et al. 2016; Senadheera, Warren, and Leitch 2017). Even in the relatively unexplored context of local government, the adoption and use of social media tools for crisis communication and social media's role in crisis management has been investigated by academics(Graham, Avery, and Park 2015). The public is growing increasingly reliant upon mobile and social media technologies during crises and other unanticipated events (Lachlan et al. 2016). Social media has created new venues that can be used to examine how individuals and organizations communicate during the crisis lifecycle. Coombs (2007) highlights that it is vital to have evidence-based guidelines in place to support the integration of social media into crisis management practices. The Social-mediated crisis communication model introduced by Jin, Liu, and Austin (2014) aims to assist practitioners in dealing with information gathered during crises. Furthermore, this model illustrates the benefit for such online communication to be strategically employed, as results show that suggest that public emotions such as anger and disgust tend to worsen when customers receive crisis information via third-party social media. Hence, in dealing with crisis situations, companies can likely benefit from having an active social media presence to demonstrate their involvement. Moreover, management of information propagation during crisis management is another important element that managers and policy makers would benefit from doing well. Shan and Lin (2017) developed an emergency information dissemination model and demonstrate their model's effectiveness by using emergency data in two earthquake cases from social media platforms.

Public crises associated with food-related issues have been drawing the attention of social media researchers. This is because organisations involved in communication and responding to public enquiries handle food crises regularly, such as the horsemeat scandal incident in 2013, and simultaneously engage in promotional campaigns (Rutsaert et al. 2013).

In this decade, social media has become the one of the most popular channels which is used for direct interaction between the companies and their customers. Company invests more resources to operate a dedicated Facebook page as their core social media marketing activity (Maecker, Barrot, and Becker 2016). Facebook is not only a channel to contact the customer as it now a new platform that can be used to take remedial actions when during company crises. For instance, Facebook can be used to interact with the worried customer, provide accurate risk information, and announce follow-up actions.

2.2 Text Mining

Text Mining (TM) is a data mining method that has been applied to many fields to identify useful information and insights from a large volume of textual data. One of the major applications of TM is to identify and analyse the sentiment expressions in textual data, particularly, sentimental analysis techniques are used extensively by researchers and managers to evaluate public polarity in social media data. Some examples of TM using social media data include: Incorporating social media competitive analysis to businesses (He, Zha, and Li 2013), studying customer attitudes towards well-known brands (Mostafa 2013), investigating public political opinions (Meduru et al. 2017), identifying the satisfaction levels of customers (Liau and Tan 2014), predicting stock market activity (Nguyen, Shirai, and Velcin 2015) and categorising customers' positive and negative reviews (Turney 2002; Yang, Chen, and Chang 2014).

2.2.1 Sentiment Analysis

Sentiment Analysis identifies the sentiment polarity (positive, negative or neutral) or sentiment strength of the opinions expressed on subject matter (Gamon et al. 2005; Mostafa 2013; Hogenboom et al. 2013) or assigns a general polarity to a set of textual data (Pang and Lee 2008). Sentiment analysis is now widely adopted to analyse unstructured text data to obtain insights relating to public perception, comments, sentiments, attitudes and emotions from written language (Medhat, Hassan, and Korashy 2014). Lexicon-based approach and machine learning approach are the most popular sentiment analysis techniques to extract sentiment orientation from text messages.

The lexicon-based approach utilises dictionaries with words that are predefined in their semantic orientation (polarity and strength) to analyse the customer opinions towards the core subject matter (Taboada et al. 2011). Wilson, Wiebe, and Hwa (2006) introduced context, negations, superlatives and idioms to the process through the polarity predicting method. This is achieved by using the grammatical framework, usually in conjunction with a lexicon.

For the machine learning approach, polarities of a set of manually annotated texts are used to identify text features related to the positive, negative and neutral categories to train a predictive model. Nevertheless, there are few algorithms built to detect both sentiment strength and polarity (Bravo-Marquez, Mendoza, and Poblete 2013) and they operate on the assumption that individuals are able to distinguish between different levels of emotion in text documents.

2.2.2 Machine Learning (Text Classification)

The Machine Learning (ML) technique is a general basic process that automatically builds a classifier model by learning from a set of pre-classified documents which contain the characteristics (i.e. features) of the classes. This is a superior approach within the research community in dealing with large volumes of data. ML technique are further classified into supervised and unsupervised methods (Kotsiantis 2007). In a situation where labels are known, then the learning is called supervised; if labels are not given, the learning is called unsupervised. Typically, ML approaches use training datasets labelled with its classifications to construct a predictive model.

As for automated text classification in particular, it is one of the most widely accepted learning models in data mining (Ahmed 2004). It is regarded as a necessity in handling and processing documents of large volume and aims at developing a predictive model to predict customers' attitudes through categorising datasets into pre-classified categories based on pre-set criteria (Nassirtoussi et al. 2014). Essentially, text classification organises texts by seeking to classify a document under a predefined category.

3. Framework

Figure 1 shows an analysis framework diagram for our Facebook message analysis procedure. Firstly, the raw data was crawled from Tesco's Facebook messages which consists of information such as name, comments, time of action and Facebook likes. The raw data are collected and transferred to a spreadsheet and subsequently processed by an extraction programme coded in R-Programme (Danneman and Heimann 2014) in order to change the raw data into a structured format. Secondly, we screened for non-customer opinions and Tesco customer service responses , and these were removed from the structured format data. A frequency table of keywords was generated simultaneously. Thirdly, the co-occurrences of high frequency keywords were evaluated, and the main discussion themes were identified by clustered groups, then a Multidimensional Scaling (MDS) diagram was created (Tse et al., 2016). Fourthly, to investigate the large volume of opinions made by customers towards Tesco during the horsemeat scandal, a machine learning technique was adopted to classify customers' comments into the relevant discussion theme. We prepared a training dataset in which 1000 samples were randomly selected and manually annotated based on the identified discussion themes. For developing a prediction model, the accuracy of machine classification was optimised to be above 70% and the prediction model was trained and the machine annotated message dataset was formed. Finally, cross-analyses were performed which include lexicon-based analysis, gender analysis, time series analysis and sentiment analysis.

[Figure 1 near here]

4. Data Analysis

4.1 Data Collection

Data was collected from Facebook social media platform. Compared with traditional data (such as survey data and interview data), social media data has a better data integrity and it contains the 'real signal' of a user expression.

[Figure 2a near here]

[Figure 2b near here]

Two of Tesco's Facebook message posts (shown in Figure 2a and Figure 2b) were analysed. The datasets consist of 5,865 and 1,197 comments respectively from Facebook users. The nature of these two messages are different. Specifically, message 1 is an apology and message 2 is about the updates and their follow-up actions. Customers' comments referring to message 1 between 16 and 22 January 2013 and those on message 2 from 30 January to 5 February 2013 were used to analyse customer attitudes and responses.

4.2 Text Data Pre-processing

A data processing phase was performed to convert unstructured raw textual data into a format that is recognisable by ML algorithms (Pang, Lee, and Vaithyanathan 2002). This process uses RapidMiner analytics software. First, textual data was broken down into discretising words, an action known as tokenisation (Verma, Renu, and Gaur 2014), so that RapidMiner was then able to set apart important keywords from a large volume of words and use them for prediction. The dataset was then separated into two sets, where the training set was used to help build a model and the test set was used to help predict classifications.

4.3 Classifying the Comments

4.3.1 Multidimensional scaling (MDS)

[Figure 3 near here]

Following the data pre-processing, we then investigated co-occurrence of keywords in the Facebook comments on the Tesco apology statement using the keywords frequency table. The MDS approach is a useful tool for identifying potential themes for focus in a discussion group (Tse et al. 2016). In Figure 3, the bubble represents important dataset keywords and lines between keywords represent the strength of their relationships to each other. The size of the bubble represents the frequency of the keywords. Additionally, higher co-occurrence of keywords is represented by closer bubbles.

We then used the MDS graph to identify different clustered discussion themes and their strength of relationship. After combining the Facebook output themes with the theme identified in our previous twitter research (Tse et al. 2016), we now propose four main discussion groups: (i) accepting apology, (ii) blame Tesco, (iii) Don't care about the issue and (iv) show customer behaviour. Table 1 below shows the four main discussion groups and descriptions of customer perceptions.

[Table 1 near here]

4.3.2 Machine Learning

Machine Learning is a general process that automatically builds a classifier model which we then used to classify high volume customer responses to Tesco's Facebook announcements (in figure 2a and 2b). A Naive Bayes classifier was adopted as a simple probabilistic classifier based on the application of Bayes' (statistical) theorem, with strong (naive) independence assumptions. This step includes a document vector, created through a term frequency–inverse document frequency method (TF-IDF) that transforms the corpus of Facebook comments into a matrix of individual words that is then worked upon by the standard ML algorithms for classification. The four classifications applied to determine the responses of Facebook users to Tesco's messages were "accept", "blame", "behaviour" and "don't care" (Table 1).

4.3.2.1 Training Set. A training set refers to a set of data that has pre-classified targets and predictor variables (Elkan 2013; Kumar et al. 2013). A training set is used to train a model by tuning the parameters using a validation set to test the performance of the classifier on an unseen test set.

In this study, we allocated approximately 20% of the dataset to the training set via the following sampling method. The population of each dataset was divided into sections of *k* numbers, where the first *n* samples were taken from each section. For example, the first 100 samples were taken from a section of 500 samples in the population from dataset 1. This was done to avoid bias and to have a larger representative population sample. After forming the training set, the comments were manually labelled for supervised learning and evaluation purposes (see Table 1 for said classifications). Identifying key features is important when building a high-performing ML model. Particularly, filtering attributes using feature-selection methods allows attributes that have a strong correlation with the predicted or dependent variable to be included (Elkan 2013). The techniques ranked attributes based on the amount of information obtained from the example set and selected those that met or exceeded a chosen threshold or simply select ed the top-*k* attributes. The feature-selection methods used in this study were the Weight by Information and Weight by Support Vector Machine (SVM) operators.

The results of feature-selection show that familiar terms expected to correlate with users' response to the messages were observed. Terms such as "accepted", "apology" and "trust" scored highly. Also, there were some unexpected predictors, such as the term "quality", which appeared to be relevant to the classification of user response to the horsemeat scandal. The results from the two operators were compared and the feature-selection method that yielded the highest weighting was then chosen for predictive modelling in the next step of the analysis. Table 2 and Table 3 summarise the top 5 results obtained from Information-based and SVM-based features selection separately.

[Table 2 near here]

[Table 3 near here]

4.3.2.2 The Prediction Model. The prediction model is shown in Figure 4. It works by gathering the document vector and attribute weights, then the Select-by-Weights operator combines these for selection of top-k attributes. A series of trial-and-error experimentation is then conducted using several different ML algorithms and the validation operator used for modelling. The two variables that affected the accuracy of the modelling are the number of top-*k* values selected and the type of ML algorithm used. These variables are adjusted systematically for best results (Akthar and Hahne 2012). In addition, the overall accuracy of the classification prediction is over 75%.

[Figure 4 near here]

4.3.2.3 Classification analysis. Differences between messages 1 and 2 are highlighted in Table 4, which shows that message 1 received more responses in the "accept", "blame" and "don't care" categories, whereas message 2 received more responses in the "behaviour" category. The "don't care" category had the most responses for both messages.

[Table 4 near here]

4.4 Sentiment Analysis (SA)

This study so far has managed to automatically classify the Facebook comments into predefined categories; however, this methodology is not sufficient to obtain any insight into the emotional tone underlying these comments. We then used SA to obtain an understanding of the attitudes, opinions and emotions expressed by the Facebook users. As Taboada et al. (2011) remarked, SA supports researchers in determining the opinions of customers by extracting subjectivity and polarity from textual data. The lexicon-based approach was chosen to establish the polarities of the comments using a predefined dictionary or corpus of subjective words. The dictionary typically incorporates a wordlist and a matching Semantic Orientation (SO) value, where SO is characterised as a quantitative measure of subjectivity and a viewpoint in a text (Taboada et al. 2011). The lexicon of Hu and Liu (2004) was used to conduct the analysis, because this dictionary has been used successfully in a number of research projects (Mostafa 2013; Pang and Lee 2008). This lexicon consists of 4,783 negative words and 2,006 positive words, and they are words annotated with known orientation scores.

The mean sentiment score (total sentiment scores/total no. of comments) calculated using all comments within 7 days of message appearing online was -0.12 for message 1 and 0.06 for message 2. This result indicates that during the horsemeat crisis, the public's attitude was worse towards Tesco's first message than it was towards their second.

For both messages, the categories "blame" and "behaviour" remain negative (as shown in Table 5). However, sentiments for the categories "accept" and "don't care" are reversed, where the sentiment for "accept" becomes positive and for the class of "don't care" becomes negative in message 2.

[Table 5 near here]

4.5 Gender Analysis

Gender analysis is an important tool used by companies for segmentation. This segmentation is demonstrated by products that are targeted for sale separately for males and females from brands in many product groups. Consequently, deciding whether and how males and females vary in their attitudes towards product harm crises is an essential matter for affected companies to explore. This section will cover two types of gender analysis: gender sentiment analysis and gender classification analysis.

4.5.1 Gender Sentiment Analysis

The mean sentiment scores of Facebook users in response to message 1 reflects a more negative response from females (-0.19) than males (-0.06) (Table 6). On the contrary, female responses to message 2 are more positive (0.18) than males (-0.10). Additionally, it was observed across the two messages that females display a more extreme range of sentiment value, from -0.19 to 0.18.

[Table 6 near here]

4.5.2 Gender Classification Analysis

Gender classification analysis revealed that men are less inclined to react or consider taking an active approach in expressing themselves compared to women, who deliberate actions. Accordingly, data marking the responses to the statements made by Tesco on social media makes clear the gender differences.

[Table 7 near here]

4.6 Time Series Analysis

Another way to analyse the data is to break down entire datasets into separate dates to investigate the changes in comment frequencies and customer opinions with regards to the development of the scandal. In this section, the time range is narrowed to seven days for both datasets: dataset 1 from 16/01/2013 to 22/01/2013 and dataset 2 from 30/01/2013 to 5/02/2013.

[Figure 5a near here]

[Figure 5b near here]

Figure 5a and Figure 5b show that the changes in comment frequency across the 7 days are similar for both datasets. The number of comments was highest on the days Tesco made the two public statements, which were 16 January and 30 January. The frequencies then sharply decreased and activity eased offafter the first two days. In addition, a ratio diagram (Figure 6) shows the ratio of the volume of comments by customers before & after Day 7. The ratio of comments before and after Day 7 for message 1 and message 2 are 46 : 1 and 17 : 1, respectively (Before Day 7:After Day7).

[Figure 6 near here]

The mean sentiment of the comments in dataset 1 was negative, while dataset 2 was positive throughout the seven-day observation period (Figure 7). The worst and best sentiments of the two datasets appeared on the last few days of the duration where dataset 1 experienced two dips of mean sentiment on 17 Jan and 21 Jan and dataset 2 had two increases of average sentiment on 31 Jan and 4 Feb.

[Figure 7 near here]

4.7 Cross Analysis

[Figure 8 near here]

In the first few hours following Tesco's message 1, females were more willing to express their willingness to accept the apology message and also show off that they do not eat meat than males. For those who did not care about the event and even those messages in the "blame" category, there was no gender difference. For the "behaviour" category, females tended to be more active in showing their opinions about taking some measurement about the crisis. In general, the longer the crisis went on, the less the gender gap was.

5. Discussion

5.1 Strategic Use of Emotions in Crisis Communications

This study examined the reactions of individuals who responded to Tesco's Facebook messages and found that there was a more people who in the "don't care" category compared each of the other three categories in both messages 1 and 2. Responses were generally neutral or humorous in tone both in relation to message 1 and 2. For instance, an example of a neutral response is "*Well, at least we know what horse tastes like now lol*" and a joke commonly made by the public was "*Your burgers are part of a stable diet haha*". These findings are coherent with Tse et al.'s (2016) Tesco horsemeat scandal research, where we showed that humour was likely to be used by British customers to communicate their attitudes towards the scandal on social media.

Furthermore, we can see from Table 4 that message 2 has 5.38% more responses in the "behaviour" category than message 1 has. This could possibly be attributed to the different natures of the two crisis messages. The results suggest that the nature of message 2, an informative message, led to a more reactive emotional response from customers compared to message 1. Since message 1 is an apology message, its nature may not have called for an active customer response, illustrated by higher responses in the "accept", "blame" and "don't care" categories, which essentially reflect milder emotional responses that have a lower tendency of resulting in positive action. According to McKay-Nesbitt et al. (2011), an organisation's choice of message strategy influences individuals' emotional responses when reading the message as does how the individual perceives the organisation undergoing crisis. Our finding thus demonstrates that there are potential benefits for companies if they develop effective response strategies that take public emotional responses into account. To capitalise on strategic use of emotion, potential crisis responses from companies should include an emotional appeal when communicating with the public.

5.2 The Importance of the Nature of Crisis Messages

During the horsemeat scandal, Tesco showed initiative by recalling many product lines voluntarily without external intervention. This type of initiative whereby companies recall products voluntarily is a common approach that companies adopt as a crisis response strategy (Roshan, Warren, and Carr 2016). The potential value to the company in doing this rests on the confidence that their messages will have positive effects on the public by repairing reputational damages (Coombs and Holladay 2014), especially for a reputable company. However, given that Tesco is a reputable company, the mixed sentiment results of the current research are contrary to previous studies (Coombs and Holladay 2014) and to some extent challenge the assumptions concerning the positive effect of voluntary recall. To further understand this observation, individual messages were further analysed and compared with regard to the nature of the messages. The SA revealed that message 1 received a negative sentiment of -0.12 and message 2 a positive sentiment of 0.06. These results reflected that the public has a more positive response towards the informative message (message 2) than the apologising message (message 1), which received a negative response. This brings Bradford and Garrett's (1995) assumption into question; here they assumed that positive reactions and higher organisational reputation would result from a public message of apology where the company acknowledges responsibility (Bradford and Garrett 1995). This does not correspond with the current SA findings.

Furthermore, the integrated classification and SA results show that message 2 had a more positive sentiment of 0.32 for the "accept" category compared to negative sentiments for the other three classes. Responses included comments such as, "*I think it's very big of them to apologise - apology accepted Mr.Tesco*" and "*Appreciate the effort and honesty*" from message 2. It seems that providing information and updates on the crisis situation is something that appeals to customers. Classification results show that message 2 has the lowest number of individuals in the "blame" category, which suggests that this message may be a more appropriate message for engaging and changing customer perception by lessening customer perception of danger. This could indirectly affect customer perception of the company in a positive manner.

Therefore, the nature of a company's public message released on SNS could determine their success in crisis communication. In the horsemeat scandal, the sentiment scores for message 2 are above zero (0.06), which means that customer sentiment was high. Therefore, it is potentially beneficial for companies to include messages that provide information and updates to customers on social media like Facebook in crisis situations as part of their emergent customer communication strategy.

5.3 Gender-Specific Insights

Gender information for the Facebook users in this study was obtained by using coded programme and predicted the first name by using the US Census Bureau[1]. From the two sets

[1] Gender estimation programme - https://github.com/Bemmu/gender-from-name

of data combined, 6,030 user comments were successfully categorised by gender based on their usernames (i.e. representing 88.13% successful classification). While the rate of identification is relatively low, the aggregate amount of identified Facebook comments is still much larger compared to the identifiable accounts on Twitter, because Facebook users generally use their own names. Of the 6,030 gender categorised comments, 2,878 were made by men and 3,152 were made by women. This might imply that the issue of a supermarket selling horsemeat is more important to women than it is to men.

Furthermore, examination of gender data revealed that females had "extreme" responses towards the messages posted on Facebook and a greater tendency to blame Tesco and deliberate action as compared to men. Laufer and Gillespie (2004) discover that women feel more vulnerable to harm than men after reading about a product harm crisis. This is consistent with findings that women are more likely to view threatening incidents as more severe than men in a product harm crisis scenario. According to Harris and Miller (2000), this is attributed to biological and socialisation factors. This information has provided insights into female behaviour and response patterns and company managers have categorised these females that feel more vulnerable than men into a more specific customer group based on their responses and sentiments for customer retention purposes in marketing. Drawing such implications from these comments reflects that practitioners can formulate strategies based on statistical gender differences. This tool can potentially render strategies that are more effective in communicating to a specific group, such as gender.

5.4 Crisis Monitoring and Framework Proposal

In addition to analysing gender differences, we also investigated the changes in customer attitudes and the frequency of comments during the breaking of the horsemeat scandal. The chronological range of the data was narrowed to the first seven days from the day that the

Tesco messages were posted on Facebook. This was done because the highest volume of comments by customers was made during these first seven days. After day seven, the number of customers comments was small and their inclusion would not have added value to the analysis. The sudden drop in the frequency of comments was very similar for both Tesco messages as shown in Figure 5 and 6. It would be interesting to determine if this pattern was similar for all product harm crises. One reason for this may be that the crisis was more prominent in the press during the first week. A possible implication for Tesco and other companies that may face similar crises in the future is that they should prioritise and enhance social media monitoring and communications immediately following the company's online message(s).

Furthermore, the sentiment change of the customers for both messages is the exact opposite of each other as shown in Figure 7. For message 1, negative customer sentiments dips twice over the span of a week, with the decrease in sentiment value of the second dip tripling that of the first. The sentiment value trend of message 2 exhibits a striking resemblance to that of message 1, but within the positive range of sentiment. As such, the patterns of sentiment change for both public statements are similar, almost forming mirror images of each other. However, drastic change in sentiment towards the end of the period where there were low numbers of comments indicates that the sentiment results may not be very accurate. This is due to the problem of sample size limitations that affects result accuracy (Guess 2011).

This research gives a researcher or practitioner a more complete and objective method to analyse the text-based data. The combination of multiple approaches will be advantageous to managers because it allows them to not only identify how their customers feel but also how they react. This is also a gateway for managers to carry out a combination of integrated analyses, such as classification-sentiment analysis and classification-gender analysis, for greater insights into the data.

Table 8 summarises the key points in our discussion.

[Table 8 near here]

6. Conclusion

In the field of crisis management, the horsemeat scandal holds many valuable lessons for practitioners. This study has investigated customer responses during the scandal and has identified elements that are of key importance in understanding customer opinions and responses. This will allow businesses to act appropriately to either prevent or mitigate the impact of future crises on their business.

Firstly, the classifications analysis found that the nature of message 2 led to more emotional responses from customers as compared to message 1. Secondly, the results of the SA indicate that UK customers have mixed sentiment between both messages toward Tesco regarding the horsemeat scandal. The mixed sentiment results are contrary to previous studies, and, to a certain extent, challenge assumptions made previously about the positive effect of voluntary recall. Furthermore, the positive sentiment of message 2 shows that providing information and updates on the crisis is something that appeals to the customers and is an initiative they appreciate. Next, the gender analysis showed that the issue of a supermarket selling unlabelled horsemeat is more important to women than it is to men and females were observed to have “extreme” responses towards the messages and a greater tendency to blame Tesco and deliberate action, compared to men. Lastly, this study also investigated changes in frequency of comments and sentiment scores across time. This revealed a trend where activity by customers was highest during the first two days following the release of the public statement. The drastic changes in sentiment scores indicate that sentiment results may be unreliable and therefore it may be best to use SA in conjunction with ML to analyse text-based data.

It is also important to note that the two messages were analysed independently without considering the relation or potential impact of message 1 on message 2. There were several messages announced by Tesco on Facebook during the horsemeat scandal and this research only considered the first two messages. These insights and marketing strategies that lead directly from this study are valuable to crisis management social media research in addition to companies that are developing crisis management strategies.

Acknowledgement

This research is supported by BA/Leverhulme Small Research Grants SG141713 and the National Social Science Foundation of China (Grant No. 14CTQ031). We also thank the anonymous reviewers for insightful comments that helped us improve the quality of the paper.

Reference

Ahmed, S. R. 2004. "Applications of data mining in retail business." Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04).

Akthar, F., and C. Hahne. 2012. "RapidMiner5 Operator Reference." accessed 1st June 2017. Available at: https://rapidminer.com/wp-content/uploads/2013/10/RapidMiner_OperatorReference_en.pdf.

Ashley, C., and T. Tuten. 2015. "Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement." *Psychology & Marketing* 32 (1):15-27.

Bagheri, A., M. Saraee, and F. de Jong. 2013. "Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews." *Knowledge-Based Systems* 52:201-13.

Basili, R., D. Croce, & G. Castellucci. 2017. "Dynamic polarity lexicon acquisition for advanced Social Media analytics." *International Journal of Engineering Business Management,* 9:1-18.

Baird, C. H., and G. Parasnis. 2011. "From social media to social customer relationship management." *Strategy & Leadership* 39 (5):30-7.

Bradford, J. L., and D. E. Garrett. 1995. "The effectiveness of corporate communicative responses to." *Journal of Business Ethics* 14:75-892.

Bravo-Marquez, F., M. Mendoza, and B. Poblete. 2013. "Combining Strengths, Emotions and Polarities for Boosting Twitter Sentiment Analysis." WISDOM'13, Chicago, USA., August 11 2013.

Bruns, A., and S. Stieglitz. 2013. "Towards More Systematic Twitter Analysis: Metrics for Tweeting Activities." *International Journal of Social Research Methodology* 16 (2):91-108.

Coombs, W. T. 2007. Crisis Management and Communications. Institute for Public Relations.

Coombs, W. T., and S. J. Holladay. 2014. "How publics react to crisis communication efforts." *Journal of Communication Management* 18 (1):40-57. doi: 10.1108/jcom-03-2013-0015.

Danneman, N., and R. Heimann. 2014. *Social Media Mining with R*. BIRMINGHAM: Packt Publishing Ltd

Duan, W., Y. Yu, Q. Cao, and S. Levy. 2015. "Exploring the Impact of Social Media on Hotel Service Performance." *Cornell Hospitality Quarterly* 57 (3):282-96.

Elkan, C. 2013. "Predictive analytics and data mining." accessed 1st June 2017. Available at: http://cseweb.ucsd.edu/~elkan/255/dm.pdf.

Gamon, M., A. Aue, S. Corston-Oliver, and E. Ringger. 2005. "Mining Customer opinion from free tree." IDA 2005.

Golder, S. A., and M. W. Macy. 2011. "Diurnal and seasonal mood vary with work, sleep, and Daylength Across Diverse Cultures." *SCIENCE* 333 (30 SEPTEMBER):1878-82.

Graham, M. W., E. J. Avery, and S. Park. 2015. "The role of social media in local government crisis communications." *Public Relations Review* 41 (3):386-94.

Guess, A. R. 2011. "The Possibilities and Limitations of Sentiment Analysis." accessed 1st June 2017. Available at: http://www.dataversity.net/the-possibilities-and-limitations-of-sentiment-analysis/.

Harris, M. B., and K. C. Miller. 2000. "Gender and Perceptions of Danger." *Sex Roles* Vol. 43 (Nos. 11/12).

He, W., S. Zha, and L. Li. 2013. "Social media competitive analysis and text mining: A case study in the pizza industry." *International Journal of Information Management* 33 (3):464-72.

Herzig, J., G. Feigenblat, M. Shmueli-Scheuer, D. Konopnicki, and A. Rafaeli. 2016. "Predicting Customer Satisfaction in Customer Support Conversations in Social Media Using Affective Features." UMAP '16, Halifax, NS, Canada, July 13-17.

Hogenboom, A., D. Bal, F. Frasincar, and M. Bal. 2013. "Exploiting Emotions in Polarity Classfication of Text." *Journal of Web Engineering* 0 (0):1-19.

Hu, M., and B. Liu. 2004. "Mining and Summarizing Customer Reviews." In Proceedings of the 10th international conference on knowledge discovery and data mining (ACM SIGKDD 2004), Seattle, Washington, USA.

Jin, S.-A. A., and J. Phua. 2014. "Following Celebrities' Tweets About Brands: The Impact of Twitter-Based Electronic Word-of-Mouth on Consumers' Source Credibility Perception, Buying Intention, and Social Identification With Celebrities." *Journal of Advertising* 43 (2):181-95.

Jin, Y., B. F. Liu, and L. L. Austin. 2014. "Examining the Role of Social Media in Effective Crisis Management." *Communication Research* 41 (1):74-94.

Kaplan, A. M., and M. Haenlein. 2010. "Users of the world, unite! The challenges and opportunities of Social Media." *Business Horizons* 53 (1):59-68.

Katal, A., M. Wazid, and R. H. Goudar. 2013. "Big Data Issues, Challenges, Tools and Good Practices." Contemporary Computing (IC3), Sixth International Conference.

Kotsiantis, S. B. 2007. "Supervised Machine Learning A review of classification technique." *Informatica* 2007 (31):249-68.

Kumar, A., S. Vembu, A. K. Menon, and C. Elkan. 2013. "Beam search algorithms for multilabel learning." *Machine Learning* 92 (1):65-89.

Lachlan, K. A., P. R. Spence, X. Lin, K. Najarian, and M. Del Greco. 2016. "Social media and crisis management: CERC, search strategies, and Twitter content." *Computers in Human Behavior* 54:647-52.

Laufer, D., and K. Gillespie. 2004. "Differences in consumer attributions of blame between men and women: The role of perceived vulnerability and empathic concern." *Psychology and Marketing* 21 (2):141-57.

Lee, L. F., A. P. Hutton, and S. Shu. 2015. "The Role of Social Media in the Capital Market: Evidence from Consumer Product Recalls." *Journal of Accounting Research* 53 (2):367-404.

Liau, B. Y., and P. P. Tan. 2014. "Gaining customer knowledge in low cost airlines through text mining." *Industrial Management & Data Systems* 114 (9):1344-59.

Maecker, O., C. Barrot, and J. U. Becker. 2016. "The effect of social media interactions on customer relationship management." *Business Research* 9 (1):133-55.

Malita, L. 2011. "Social media time management tools and tips." *Procedia Computer Science* 3:747-53.

McKay-Nesbitt, J., R. V. Manchanda, M. C. Smith, and B. A. Huhmann. 2011. "Effects of age, need for cognition, and affective intensity on advertising effectiveness." *Journal of Business Research* 64 (1):12-7.

Medhat, W., A. Hassan, and H. Korashy. 2014. "Sentiment analysis algorithms and applications: A survey." *Ain Shams Engineering Journal* 5 (4):1093-113.

Meduru, M., A. Mahimkar, K. Subramanian, P. Y. Padiya, and P. N. Gunjgur. 2017. "Opinion Mining Using Twitter Feeds for Political Analysis." *International Journal of Computer* 25 (1):116-23.

Montalvo, R. E. 2016. "Social Media Management." *International Journal of Management & Information Systems* 20 (2).

Mostafa, M. M. 2013. "More than words: Social networks' text mining for consumer brand sentiments." *Expert Systems with Applications* 40 (10):4241-51.

Nassirtoussi, A. K., S. Aghabozorgi, T. Y. Wah, and D. C. L. Ngo. 2014. "Text mining for market prediction: A systematic review." *Expert Systems with Applications* 41 (16):7653-70.

Nguyen, T. H., K. Shirai, and J. Velcin. 2015. "Sentiment analysis on social media for stock movement prediction." *Expert Systems with Applications* 42 (24):9603-11.

Pang, B., and L. Lee. 2008. "Opinion Mining and Sentiment Analysis, Foundations and Trends." *Information Retrieval* Vol. 2 (No. 1–2): 1-135.

Pang, B., L. Lee, and S. Vaithyanathan. 2002. "Thumbs up Sentiment Classification using Machine Learning." Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Philadelphia, July 2002.

Pavlopoulos, I. 2014. "Aspect based Sentiment Analysis." PHD, Department of Informatics, Athens University of Economics and Business.

Roshan, M., M. Warren, and R. Carr. 2016. "Understanding the use of social media by organisations for crisis communication." *Computers in Human Behavior* 63:350-61.

Rutsaert, P., Á. Regan, Z. Pieniak, Á. McConnon, A. Moss, P. Wall, and W. Verbeke. 2013. "The use of social media in food risk and benefit communication." *Trends in Food Science & Technology* 30 (1):84-91.

Senadheera, V., M. Warren, and S. Leitch. 2017. "Social media as an information system: improving the technological agility." *Enterprise Information Systems* 11 (4):512-33.

Shan, S., and X. Lin. 2017. "Research on emergency dissemination models for social media based on information entropy." *Enterprise Information Systems* DOI: 10.1080/17517575.2017.1293300.

Taboada, M., J. Brooke, M. Tofiloski, K. Voll, and M. Stede. 2011. "Lexicon-Based Methods for Sentiment Analysis." *Computational Linguistics* 37 (2):267-307.

Terpstra, T., and A. d. Vries. 2012. "towards a Realtime Twitter Analysis during Crises for Operational Crisis Management." Proceedings of the 9th International ISCRAM Conference, Vancouver, Canada, April.

Trainor, K. J., J. Andzulis, A. Rapp, and R. Agnihotri. 2014. "Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM." *Journal of Business Research* 67 (6):1201-1208.

Tse, Y. K., M. Zhang, B. Doherty, P. Chappell, and P. Garnett. 2016. "Insight from the horsemeat scandal." *Industrial Management & Data Systems* 116 (6):1178-1200.

Tufekci, Z. 2014. "Big Questions for Social Media Big Data Representativeness, Validity and and Other Methodological Pitfalls." Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media.

Turney, P. D. 2002. "Thumbs up or thumbs down Semantic orientation applied to unsupervised." Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, 417-424.

Verma, T., Renu, and D. Gaur. 2014. "Tokenization and Filtering Process in RapidMiner." *International Journal of Applied Information Systems* 7 (2):16-8.

Wilson, T., J. Wiebe, and R. Hwa. 2006. "Recognizing strong and weak opinion clauses." *Computational Intelligence* 22 (2):73-99.

Yang, C.-S., C.-H. Chen, and P.-C. Chang. 2014. "Harnessing consumer reviews for marketing intelligence: a domain-adapted sentiment classification approach." *Information Systems and e-Business Management* 13 (3):403-19.

Zhang, M., L. Guo, M. Hu, and W. Liu. 2017. "Influence of customer engagement with company social networks on stickiness: Mediating effect of customer value creation." *International Journal of Information Management* 37 (3):229-40.

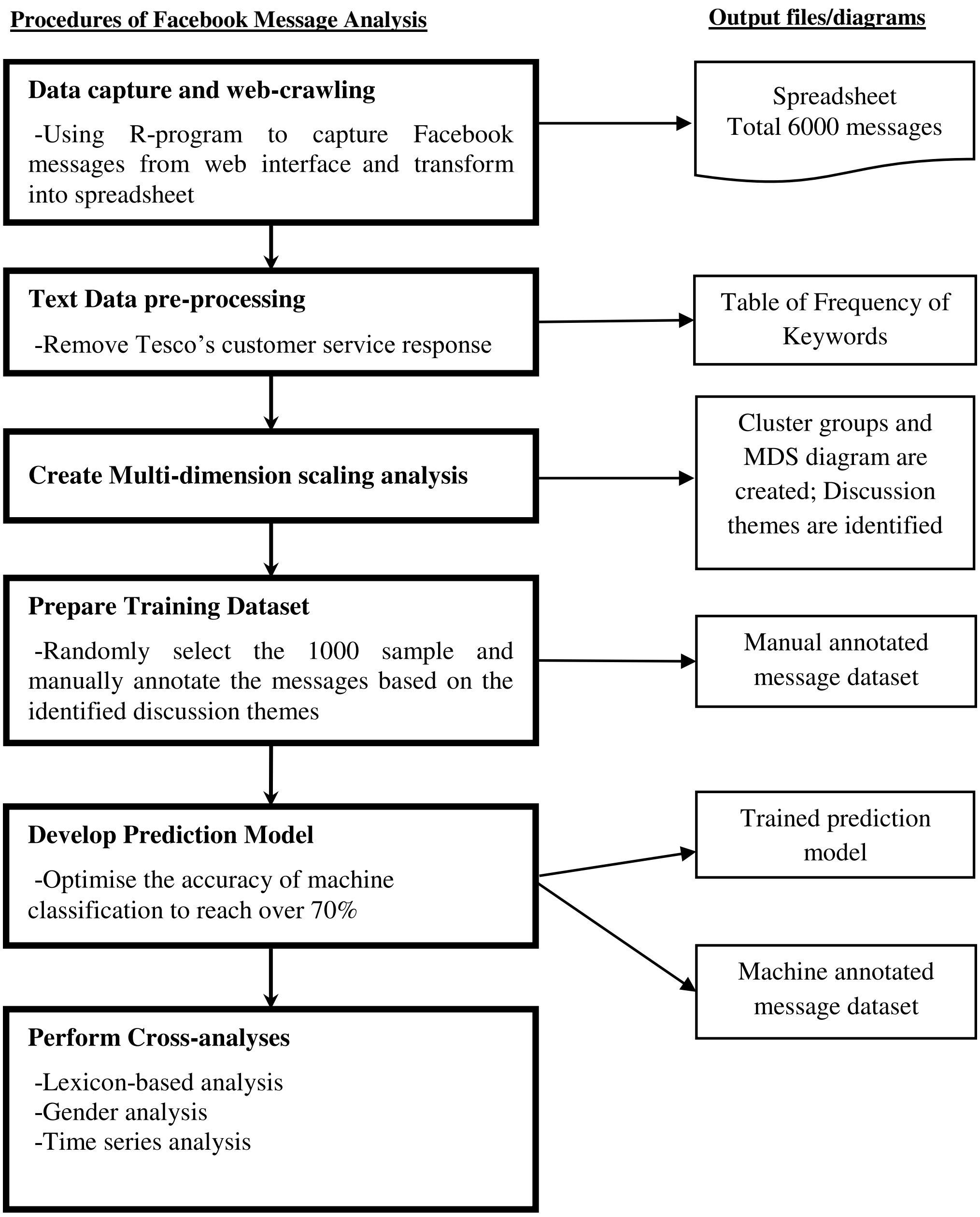

Figure 1. Diagram of Framework

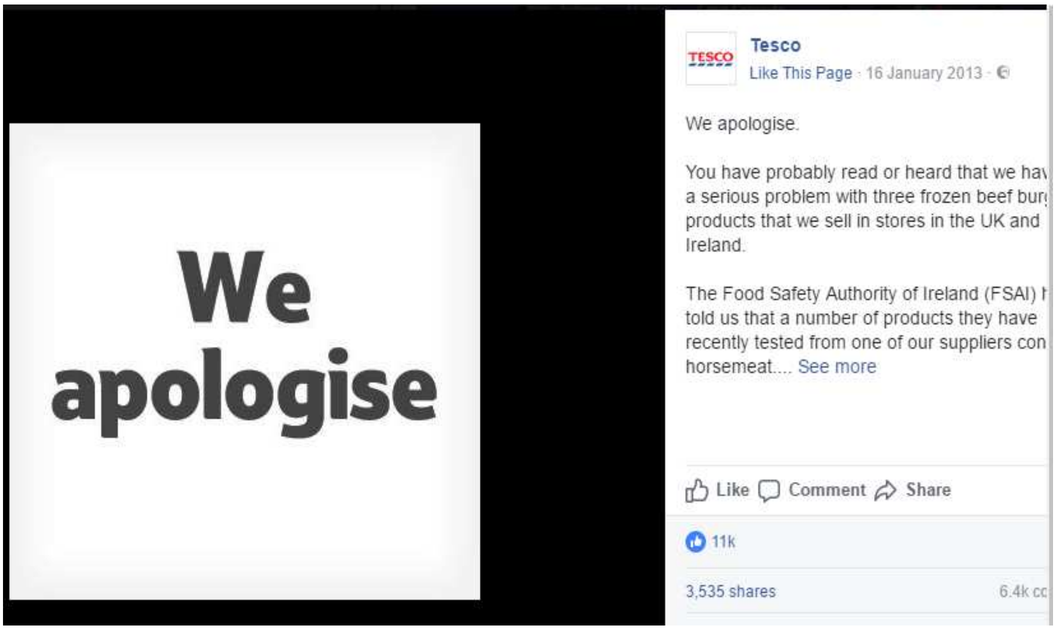

Figure 2a. Tesco's Facebook Message 1: 'We apologise'

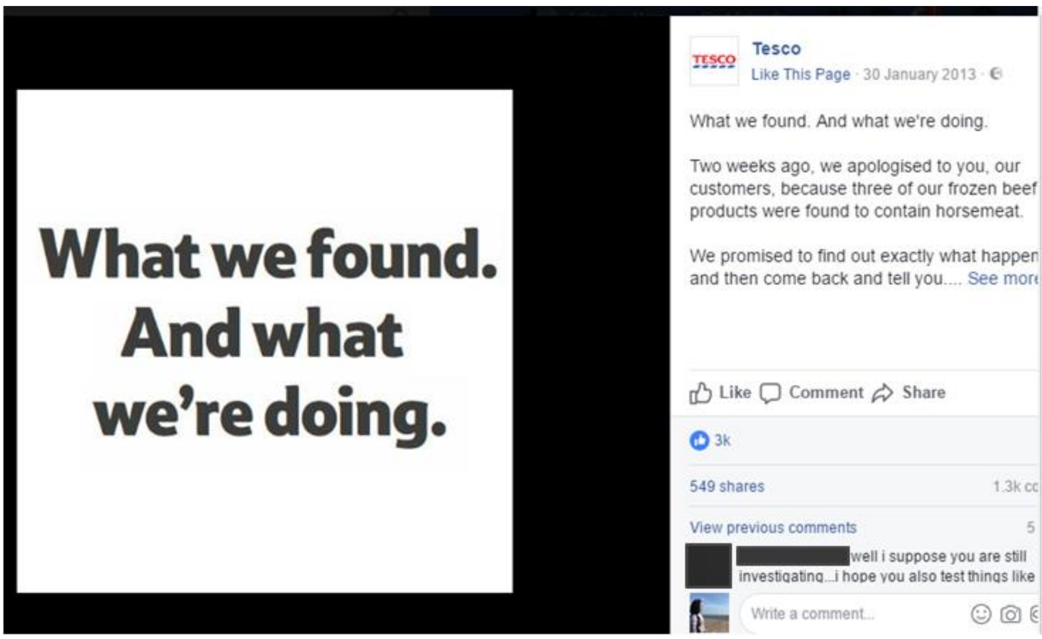

Figure 2b. Tesco’s Facebook Message 2 : ‘What we found. And what we’re doing’

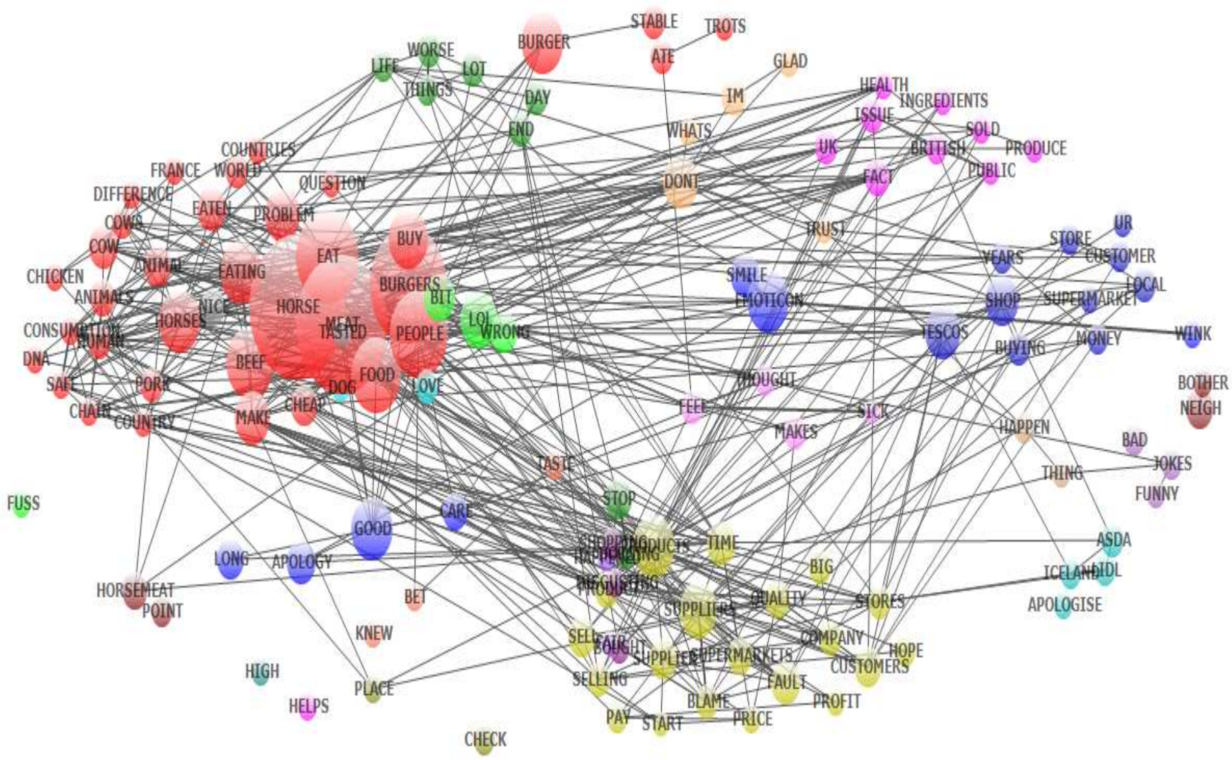

Figure 3. Map of the MDS

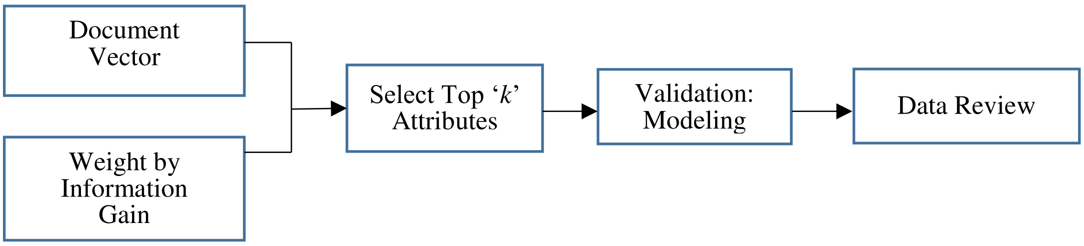

Figure 4. Process of Building a Predictive Model

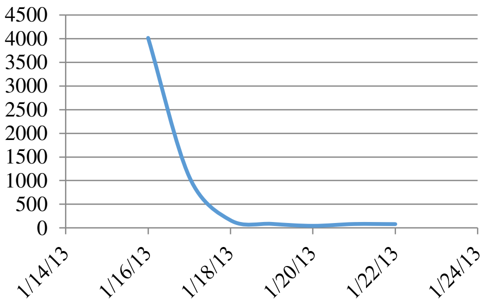

Figure 5a. Time series for Message 1

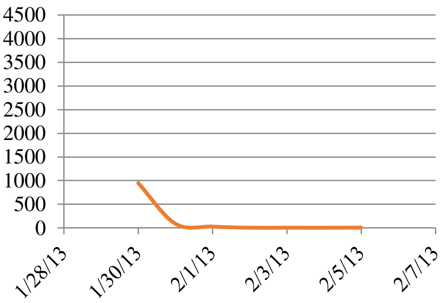

Figure 5b. Time series for Message 2

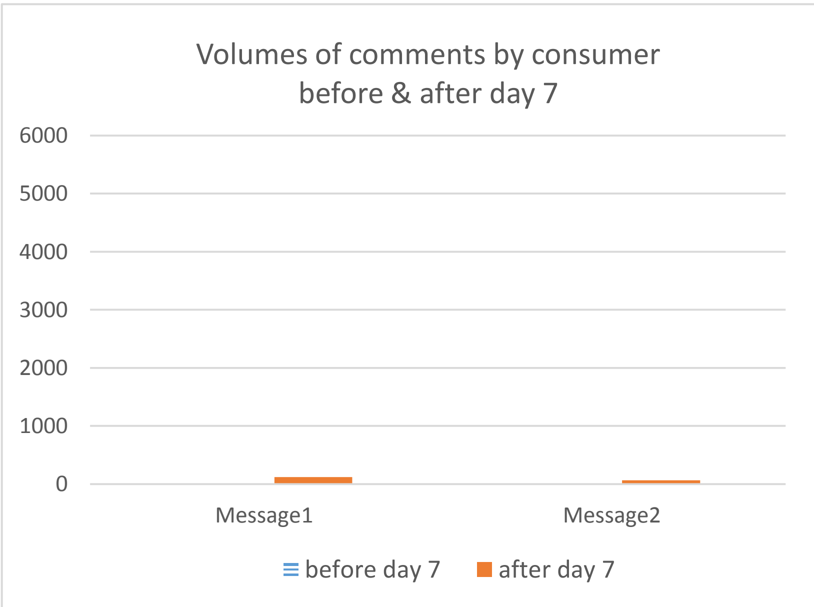

Figure 6. The ratio of comments before and after Day 7

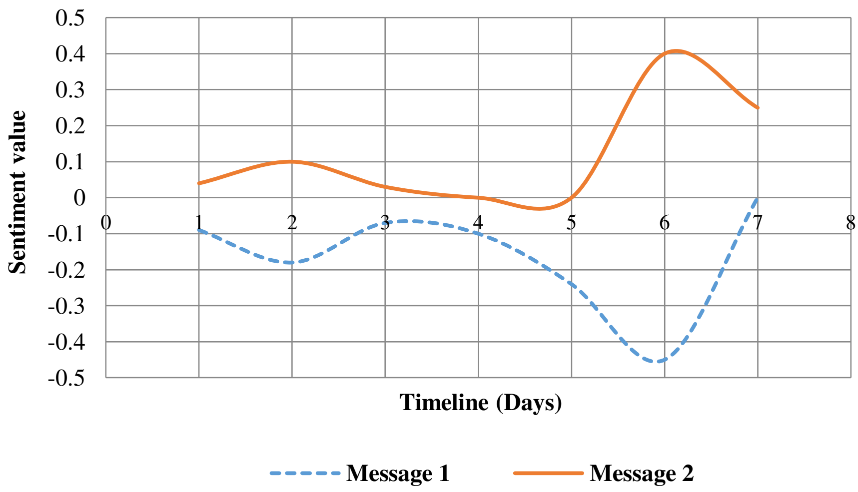

Figure 7. Overlaid Time Series Sentiment Change

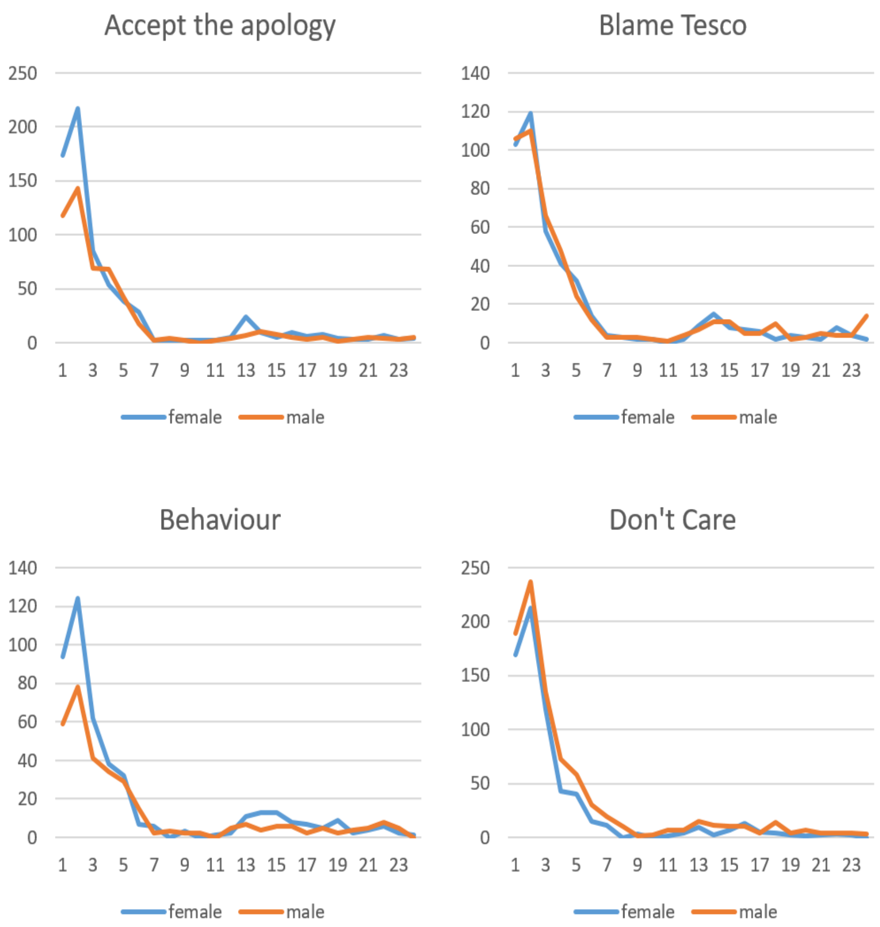

Figure 8. Cross Analysis Results

Classifications	**Perception details**
Blame (B)	Making Joke on Tesco
	Feeling sick
	Lost trust
Don’t care (DC)	Not a big deal
	Doesn’t matter to eat horse
Accept apology (A)	Not Tesco fault
	Accept apology
Behaviour (BH)	Don’t shop in Tesco
	Happy to be veggie
	Throw it
	Care about affected item

Table 1. Classifications of Samples

Attribute	Weight
1) SHOP	0.147
2) TESCO	0.053
3) APOLOGY	0.043
4) MEAT	0.041
5) ACCEPTED	0.038

Table 2. Top Five Results of Information-Based Feature Selection

Attribute	Weight
1) ACCEPTED	0.026
2) SHOP	0.026
3) LOL	0.026
4) HELPS	0.022
5) NEIGH	0.022

Table 3. Top Five Results of SVM-Based Feature Selection

	Accept	Blame	Behaviour	Don’t Care
Message 1	**27.06%**	**21.14%**	17.72%	**34.08%**
Message 2	25.30%	20.90%	**23.10%**	30.71%

Table 4. Overall Classification Results

Classifications	Message 1	Message 2

Accept	-0.27	0.32
Blame	-0.23	-0.01
Behaviour	-0.15	-0.06
Don't Care	0.07	-0.03

Table 5. Sentiment Classification Results

	Message 1	**Message 2**
Males	-0.06	-0.10
Females	-0.19	0.18
Overall	-0.12	0.06

Table 6. Gender Sentiment Analysis Results

	Accept		Blame		Behaviour		Don't Care	
	Male	Female	Male	Female	Male	Female	Male	Female
Message 1	11.95%	**15.46%**	**10.57%**	10.06%	7.37%	**10.21%**	**19.38%**	14.99%
Message 2	10.22%	**15.27%**	8.71%	**12.69%**	10.00%	**13.55%**	13.76%	**15.81%**

Table 7. Gender Classification Results of Message 1 and 2

	Message 1 (Apology)	**Message 2 (Informative)**
Classification Analysis	● Top response at 34.08% in class "don't care" ● Lowest response at 17.72% in class "behaviour" ● More responses in classes "accept", "blame" and "don't care" than in message 2	● Top response at 30.71% in class "don't care" ● Lowest response at 20.9% in class "blame" ● More responses in the class "behaviour" than in message 1

Sentiment Analysis	• Value is negative at -0.12497 • Among the four classes, only "don't care" has a positive sentiment at 0.0693; the rest are negative	• Value is positive at 0.05799 • Among the four classes, only class "accept" has a positive sentiment at 0.32; the rest are negative
Gender Analysis	• Both males and females have negative sentiment values at -0.0581 and -0.1884 respectively • More females classified as "accept" and "behaviour" than males • More males classed as "blame" and "don't care" than females, with a significantly higher number of males for class "don't care"	• Males have a negative sentiment value of -0.0974 • Females have a positive sentiment value of 0.1793 • More female than male responses across all four classes for message 2 • Notably higher percentage of females for classes "accept", "blame" and "behaviour"
Time Series Analysis	• Highest frequency on the first day • Exponential decrease in activity observed during the period • Highest activity on first two days • Customer's sentiment remained constantly negative throughout the period • Experienced two sharp dips on 17 Jan and 21 Jan	• Highest frequency on the first day • Exponential decrease in activity observed during the period • Highest activity on first two days • Customer's sentiment remained constantly positive throughout the period • Experienced two sharp increases on 31 Jan and 4 Feb

Table 8. Summary of Research Findings