



Wong, T.C., Chan, H. K. and Lacka, E. (2017) An ANN-based approach of interpreting user-generated comments from social media. *Applied Soft Computing*, 52, pp. 1169-1180. (doi:[10.1016/j.asoc.2016.09.011](https://doi.org/10.1016/j.asoc.2016.09.011))

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/157154/>

Deposited on: 12 February 2018

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

An ANN-based approach of interpreting user-generated comments from social media

Name of Author 1: T. C. Wong (corresponding author)

Department: Department of Design, Manufacture and Engineering Management

University/Institution: University of Strathclyde

Town/City: Glasgow, Scotland

Country: UK

Email: andywtc@graduate.hku.hk

Name of Author 2: Hing Kai Chan

Department: Nottingham University Business School China

University/Institution: University of Nottingham

Town/City: Ningbo

Country: China

Email: Hingkai.Chan@nottingham.edu.cn

Name of Author 3: Ewelina Lacka

Department: Strathclyde Business School

University/Institution: University of Strathclyde

Town/City: Glasgow, Scotland

Country: UK

Email: ewelina.lacka@strath.ac.uk

An ANN-based approach of interpreting user-generated comments from social media

Abstract

The IT advancement facilitates growth of social media networks, which allow consumers to exchange information online. As a result, a vast amount of user-generated data is freely available via Internet. These data, in the raw format, are qualitative, unstructured and highly subjective thus they do not generate any direct value for the business. Given this potentially useful database it is beneficial to unlock knowledge it contains. This however is a challenge, which this study aims to address. This paper proposes an ANN-based approach to analyse user-generated comments from social media. The first mechanism of the approach is to map comments against predefined product attributes. The second mechanism is to generate input-output models which are used to statistically address the significant relationship between attributes and comment length. The last mechanism employs Artificial Neural Networks to formulate such a relationship, and determine the constitution of rich comments. The application of proposed approach is demonstrated with a case study, which reveals the effectiveness of the proposed approach for assessing product performance. Recommendations are provided and direction for future studies in social media data mining is marked.

Keywords: neural network application; statistical methods; product performance; social media.

1. Introduction

No one can deny that computers, computer-based information systems, smart devices, wireless technologies, etc. have greatly transformed human behaviour and the way in which we interact with each other as well as with products and brands. This transformation facilitates the growth of social media networks, which generate a number of business opportunities [1, 2]. For example, investing in social media technology can lead to better customer relationship management [3], and social media marketing can influence customer purchasing intention [4]. Social media platforms also allow users to post and exchange comments online, which results in profusion of potentially valuable data via the Internet. Since these data are readily available and free-to-access, many business intelligence

applications have targeted them for data mining purposes [5] in hopes of generating some benefits. Moreover, social media platforms can exist in different industries such as tourism management [6], marketing [1, 7], management [8], insurance and financial sectors [9] as well as education [10]. Therefore, the advantages deriving from appropriate analysis and interpretation of social media data are infinite.

Mining social media data however is easier said than done. Due to the recent advancement in data collection technologies [5], so called big data research has flourished. Big data are characterised by their volume (size of the database), velocity (speed of change of database), and variety (different sources or format of data) [11]. Being a specific type of big data, social media data are often qualitative, unstructured and subjective, all of which make the mining process difficult, and sometimes out of focus [12]. This paper aims to propose an ANN-based approach to social media data analysis in the most structured manner possible. The main contribution of this study will not only advance academic knowledge providing direction for future research but it will also generate direct implications to practitioners who are keen to better serve their consumers while improving business operations.

The rest of this paper is organised as follows. Section 2 reviews relevant studies on social media data and data mining processes with well-known mining algorithms. In this section, the characteristics of social media data are also discussed. Section 3 presents the mechanisms of the proposed approach for interpreting social media data in a form of user-generated comments. The main outcome of the proposed approach is to address what constitute a “rich comment” and how the value of such a comment can be assessed. It is noted that we do not intend to propose a novel data-mining algorithm but a more structured way of interpreting user-generated comments. Next, Section 4 employs a case to demonstrate the operations of the proposed approach. Finally, Section 5 concludes this paper.

2. Literature Review

In this section, an overview of social media data will be provided. Some existing data mining methods to social media data will then be discussed and some existing knowledge gaps will be highlighted. Lastly, the algorithms that can help fulfil the gaps will be reviewed.

2.1 Social Media Data

Social media data are mostly user-generated and take form of user-posted comments and their exchanges [1, 7, 13, 14]. These data are often highly unstructured [11]; as they are mostly user-posted comments, qualitative and extremely diversified and thus cannot be examined directly without pre-processing [7]. Moreover, many comments are incorporated with specific “texts” such as tags [13], which makes them unmanageable. Apart from being qualitative and unstructured, social media data are also highly subjective [12]. As the real value of the comments is hidden, the need to introduce mining algorithms cannot be overlooked [11]. These algorithms need to employ techniques that convert them into manageable quantitative dataset (details will be discussed), which later can be processed by statistical approaches to generate more meaningful and objective interpretations.

The three characteristics of big data (volume, velocity and variety) may further hinder the extraction of valuable information from user-generated comments. Therefore, there is a need to develop an approach, which can be used to interpret the data for knowledge extraction in a more structured way [1, 11, 12]. This paper proposes an ANN-based approach to achieve this objective in order to overcome the challenges of social media data research.

2.2 Data Mining

The objective of data mining is to extract value from the dataset to support decision-making [7, 15]. They can be achieved through the use of machine learning algorithms [16] and such algorithms should support automated search and analysis of the data [14, 15] as well as improve the objectivity of the data [15].

Statistical assessment of social media data using various quantitative metrics can be found in previous studies [11]. For example, [17] analysed the relationship between several companies' Twitter metrics (namely, number of followers, number of followings, and number of tweets) and their stock market performance. The data were first clustered by the K-means algorithm, followed by pairwise correlation analysis. Similar metrics from an online discussion forum was extracted in [18] which included the length of message (both in words and characters), occurrences of some characters, etc. Probabilistic clustering methods were then applied to categorise the data, followed by a regression analysis to explore the relationship between the metrics and the firms' market performance. A similar study using econometric modelling can be found in [19]. Above studies share two common characteristics from social media data mining point of view: (i) learning algorithm was not adopted; and/or (ii) analysis of qualitative comments was missing.

Similar observations can be also made in studies concerning other industries. For instance, [7] observed the trend of some pizza companies based on the usage data from their social media sites. The metrics used are number of fans/followers, number of postings, frequency of posts, and so on. They aimed to analyse the market competitiveness of different companies. But, only high level statistics were summarised. Nevertheless, they did apply content analysis to identify the main themes based on the qualitative comments. This study is similar to many existing papers that did not quantify social media data using statistical tools. Hence, they missed the opportunity to generalise the findings to the broader segment of the population and as such they were unable to predict future trends.

More recently, [20] made use of a number of Facebook user profile (e.g. gender, check-in app, etc.) as the predictors to anticipate the increases in Facebook usage rate. To fully utilise social media data, predictive algorithms are often required in order to foresee future events [5, 15, 21]. This is also the reason why data mining methods are commonly coupled with learning algorithms. In this regard, Artificial Neural Network (ANN), which is a common machine learning method [16, 21], is also employed in our study. Details of ANN will be discussed in the next section.

2.3 Algorithms

In this section, correlation analysis that helps identify the input-output relationship from the social media data will be discussed. Next, ANN, a well-known machine learning technique, for formulating such relationship will be reviewed together with details about how useful information can be extracted from its hidden structure.

2.3.1 Correlation Analysis

In order to facilitate ANN analysis, independent variables (inputs) and dependent variables (outputs) must be identified for model training and validation. This can be done by the analysis of correlation which has been successfully applied in various domains such as healthcare [22], marketing [23], organisational management [24, 25], etc. However, there is a limited application of its kind to social media data. In this study, correlation analysis and partial correlation analysis are used to uncover the input-output relationship. Details of these methods will be discussed in Section 3.2.

2.3.2 Artificial Neural Network (ANN)

After addressing the input-output relationship from the social media data, ANN can be used to formulate such a relationship mathematically. ANN is a non-linear regression model which has been widely used for regression and predictive purposes such as stock market forecasting [26], credit rating [27], and consumer analysis [28]. Its application to social media data however is somehow limited. Nevertheless, some studies applying ANN to user-generated online content analysis can be found. For example, [29] applied ANN to identify potential bloggers and their influential strength in the blogosphere based on three dimensions of blog characteristics. The results of their study can help advertisers promote their products or services in a more effective fashion. [30] employed ANN to construct a recommendation system of social applications based on the popularity and reputation of an application as well as preferences and social relationship of an application user. They suggested that the ANN-based method can outperform other benchmarking methods in recommending the top three social applications to the evaluation group on a Facebook test platform. [31] utilised an ANN-based method of analysing Twitter messages in order to determine consumer sentiment towards a brand.

Their proposed method would be useful in classifying and estimating important consumer sentiment on Twitter corpus. However, while making use of the generalisation ability of ANN, all the above mentioned studies did not fully explore the explanatory capability of ANN. Hence, limited knowledge was extracted from the data.

A typical ANN consists of an input layer, a hidden layer, and an output layer. Units in input and hidden layers as well as those in hidden and output layers are inter-connected forming a network. The strengths of the inter-connections are defined as the connection weights (also known as network parameters). Through proper training, ANN can learn from the sample data by adjusting the connection weights when mapping the inputs onto the outputs. Hence, a well-trained ANN is able to quantitatively describe the connectivity between inputs and outputs. In fact, ANN is well-known for its mathematical utility in acquiring knowledge about the input-output relationship without the need of presuming the data distribution. But, at the same time, it has been criticised as a “black-box” system from which it is hard to extract explanatory information about inter-relationships between network variables [32]. To address this criticism of ANN, different ANN-based methods have been applied to generate explicit interpretation of causal connectivity among variables of the network. Details of these methods will be discussed in Section 3.3.

3. Research Method

In this study, the main aim is to identify the rationale of posting comments, which can be reflected by their length. It is assumed that the longer the comment length, the more complex sentences and ideas users can express. Hence, the more information the comment contains [33, 34], i.e. information richness. Also, a positive correlation is noted between comment length and its perceived usefulness [35]. To achieve the main aim, an ANN-based approach is proposed as shown in Figure 1. It first characterises the comments with a list of attributes via content analysis. This is followed by statistical analysis to uncover the significant relationship between attributes and comment length. A training model based on ANN is then employed to measure such a relationship. As a result, the final outcome can be used to indicate which key attributes make up rich comments, and assess the richness of new

comments. Mechanisms of the approach are discussed in this section. In Section 4, a case study is presented to disclose more details of the approach and how it can be operated in practice.

“Figure 1”

3.1 Content Analysis

The first mechanism is to map each user-generated comment against a list of predefined product attributes. As discussed in Section 2, content analysis is a useful method to identify codes or themes in a qualitative database [2, 11, 12]. In this study, our objective is to determine what constitutes a “rich” comment and to predict the richness of future comments. It is believed that the richness of the comment can be assessed with respect to its length which is an objective measure. Hence, we aim to examine what attributes may affect the length, and how the length of new comments may be related to the attributes. To do so, a set of attributes is defined and mapped against the comments via content analysis. The application of the content analysis will be discussed in Section 4.1. It is noted that more than one attribute can be assigned to a comment. Consequently, the output of this step is the mapping between comments and attributes, which serves as the input for correlation analysis.

3.2 Correlation Analysis

Next, correlation analysis and partial correlation analysis are employed as the second mechanism to statistically identify the input-output relationship between attributes and comment length. The explanation of this step of the analysis will be presented in the case study; see section 4.2 and 4.3, respectively. Correlation analysis is used to determine the significance and strength of linear relationship between variables. Similar to correlation analysis, partial correlation analysis can be used to justify the same relationship in consideration of the third variable. In other words, we attempt to address the “direct” association between variables without the effect of the third variable. The relevant procedures are detailed in [25, 36]. The outputs of this step are the input-output models depicting the causal relationship between key attributes and comment length. These models are regarded as inputs for ANN analysis.

3.3 ANN Analysis

As the third mechanism, feed-forward backpropagation ANNs with Levenberg-Marquardt training function are used to formulate the relationship of the input-output models for constructing the connectivity models. For benchmarking purpose, MLR is also applied. After training and validation, the relative influence of each key attribute (input) towards the comment length (output) and that among key attributes can then be measured such that one can tell what constitute a rich comment and how the value of a comment can be justified. Since there is no universal method of identifying important inputs in all cases [37, 38], five different ANN-based methods (three connection weight methods and two model response methods) and five MLR-based methods are applied and compared in examining the social media data. The application of ANN analysis will be discussed in Section 4.4.

3.3.1 Training and Validation

In brief, each of the ANNs is trained and evaluated via a 10-fold cross validation, and 90% of the data are used for training whereas the remaining 10% are for validation. Using ANN, the number of neurons in the hidden layer (P) can be determined by the method of [39] in which $P = \text{number of inputs } (N)+1$ was suggested. A trial-and-error method is then used to confirm the value of P [38]. In general, it can be expected that the performance of ANN is improved as P increases. This is because a large P may enhance the ability of ANN to learn and memorise the data, but it may lack ability to generalise, i.e. over-fitting. However, the ANN may not be able to learn from the data if P is too small [38].

After model training and validation, ANN can be used to compute the relative influence among attributes and comment length. Using ANN, there are two alternative ways to measure such influence: connection weight method and model response method. The former can compute the relative importance among attributes by the internal connection weights of a well-trained ANN model while the latter can compute the same from the comparison between the model response and the actual response by changing inputs of a well-trained ANN model.

Connection weight methods:

There are three well-known connection weight methods: Garson's method [40] as shown in Equation (1), Yoon's method [41] as defined in Equation (2), and Tsaur's method [42] as defined in Equation (3), where RI_{ik} denotes the relative importance of input unit i towards output unit k , W_{ij} denotes the connection weight between input unit i and hidden unit j , W_{jk} denotes that between hidden unit j and output unit k , N is the total number of input units, and P is the total number of hidden units. All these methods utilise the model parameters of an ANN (i.e. the connection weights between input units, hidden units, and output unit) to determine the relative importance of each input in predicting the single output.

$$RI_{ik} = \frac{\sum_{j=1}^P \frac{|W_{ij}| |W_{jk}|}{\sum_{i=1}^N |W_{ij}|}}{\sum_{i=1}^N \sum_{j=1}^P \frac{|W_{ij}| |W_{jk}|}{\sum_{i=1}^N |W_{ij}|}} \quad (1)$$

$$RI_{ik} = \frac{\sum_{j=1}^P W_{ij} W_{jk}}{\sum_{i=1}^N \left| \sum_{j=1}^P W_{ij} W_{jk} \right|} \quad (2)$$

$$RI_{ik} = \frac{\sum_{j=1}^P (W_{ij} + W_{jk})}{\sum_{i=1}^N \sum_{j=1}^P (W_{ij} + W_{jk})} \quad (3)$$

Model response methods:

Alternatively, there are two well-known model response methods: change of mean square error (COM) method and sensitivity analysis (SA) method. The COM method is used to statistically rank the influence of the inputs [37]. In brief, it measures the change in the mean square error (MSE) of a prediction made by an ANN model after an input is removed. The MSE with input n removed is

determined by Equation (4) where A_i and E_i are the observed values and predicted values of i -th testing dataset respectively, and D is the total number of tested dataset. To be specific, the model with $N-1$ inputs is re-trained each time after input n is removed, and the absolute difference between the MSE of the model without input n (MSE_n) relative to that of the full model with all N inputs (MSE_{all}) is then computed. The input whose deletion induces the largest change in MSE is ranked as the most influential factor, since its removal from the full model causes the most variation in the model responses. Thus, the RI of attribute n can be measured based on its proportion of the change induced relative to the total change in MSE induced by all attributes, as defined by Equation (5).

$$MSE_n = \frac{1}{D} \sum_{i=1}^D (A_i - E_i)^2 \quad (4)$$

$$RI_n = \frac{|MSE_n - MSE_{all}|}{\sum_{i=1}^N |MSE_i - MSE_{all}|} \quad (5)$$

Another method; SA, examines the model responses by changing the inputs [43]. In brief, each model is trained by fixing all inputs at their average values except input n , which varies through its entire range. If input n is important, it should produce a high variance (V_n) measured by Equation (6), where y_{ni} is the model output when input n is set at its i -th level and other inputs are held at their average values, and \bar{y}_n is the average model output over L_n levels of input n . Thus, the RI of input n can be computed by Equation (7). It is noted that the number of levels of all attributes is two (binary) except the comment length. As comment length is a continuous variable, $L = 5$ is suggested [44]. To obtain a more robust estimation of the input influence, the RI of each attribute is averaged over 10-fold training in 100 trials.

$$V_n = \frac{1}{L_n - 1} \sum_{i=1}^{L_n} (y_{ni} - \bar{y}_n)^2 \quad (6)$$

$$RI_n = \frac{V_n}{\sum_{i=1}^N V_i} \quad (7)$$

3.3.2 Benchmarking with MLR

The general form of a linear model is presented as Equation (8) where Y is the output, β is a matrix of all unknown coefficients, X is a matrix of all inputs, and ε is an D -by-1 vector of random disturbances, where D is the number of observations. Using D training datasets, the working mechanism of MLR is to minimise the sum of the squared residuals by computing a closed-form expression for estimating the unknown parameter β .

$$Y = \beta X + \varepsilon \quad \text{where } \beta = [a \ b \ c \dots]^T \quad (8)$$

Using linear modelling, it is feasible to compute the relative influence of each variable towards the dependent variable directly from the regression equation [45]. In this study, five different methods are employed: R^2 contribution averaged over orderings among regressors by [46], squared standardised coefficient (β^2), the product of the standardised coefficient and the correlation by [47], R^2 decomposition by [48], and another R^2 decomposition by [49]. For ease of presentation, these methods are denoted by LMG, Beta2, Pratt, Gen, and ZS respectively.

4. Case and Results

In order to demonstrate how the proposed approach operates, a case is employed. In this study, social media data from the official “Samsung Mobile” Facebook page were extracted. Facebook page is chosen because it allows creation of online brand or product communities and encourages consumers to freely express their feelings and sentiments regarding product and brand with less restriction. For example, unlike Twitter, Facebook does not limit the number of signs/characters used in the comments. The official Facebook page can facilitate the centralisation of feedback related to its product as well as comments posted by users from all around the world. The intrinsic features of

Facebook enable the users to express their opinions more accurately and hence the written comments may be more representative as compared to other social media platforms.

Since the products of Samsung can be considered as “hi-tech” type of products, product innovation is chosen as the high level theme for the mapping. For the sake of analysis, 86055 user-generated comments were downloaded in a two-month window. Only comments written in English and comments related to Samsung smart phone Galaxy S4 are studied. With above control, the database was reduced to 1674 comments for final analysis. Below we present the mechanisms of the approach and the results.

4.1 Characterisation of the Online Comments (the first mechanism)

As mentioned earlier, product innovation is the theme of the mapping. In this connection, a list of attributes regarding product innovation is identified for this mapping process via systematic review of the literature, which is useful and reliable source to informed audiences [50].

During the literature review, a number of high level determinants of new product performance are first identified. They are a combination of product, strategic, development process, organisational, and/or market environment factors [51]. The sub-factors are then extracted from top-tier publications from the ABS journal list. The list of factors, sub-factors and the corresponding references are listed in the appendix. After identifying the attributes of product innovation, the researchers then map the 1674 comments to these 37 attributes using a 3-step approach, i.e. initialisation-crosschecking-confirmation. There are two main reasons why a manual approach was adopted to perform the mapping over a computerised approach. Firstly, manual mapping has been deemed as equally effective as computerised mapping [52]. Secondly, during manual mapping, researchers could map comments with specific texts, tags and occasional spelling errors, which would not be handled by computerised approach. Assignments are agreed by all authors to improve reliability of the mapping. Take the following comments as an example: “I update my galaxy s4 to android 4.3 and now don't work fine, have wifi problems, and some times I can't unlock screen...”, “My Phone Galaxy S4 I9500

4.3 Turkey Update is Too bad! Decreased by 24% over 6 hours battery WiFi off, mobile data off, sync when closed”, and “Please fix the 4.3 update on the s4 (sprint). Keep getting kicked from wifi. Not cool”. All these comments are related to technology synergy and product quality, hence they can be mapped to attributes S1 and P1 (see appendix). This step essentially characterises all the comments. The relationship between attributes and comment length can then be identified with the second mechanism.

4.2 Results of Basic Statistics

Table 1 shows that the total number of comments we examined is 1669 (5 comments are excluded as no mapping can be done). On average, 2.93 attributes can be identified with a comment. The average length of comments is 210.89 characters including spacing. Also, users use an average of 69.29 characters to construct a comment which can be mapped with one attribute. Table 2 reports that the total number of attributes we examined is 30 (7 attributes are excluded as no mapping can be done). On average, 163 comments can be mapped with each attribute. Also, users use an average of 69.18 characters to address one attribute. The ratio of attributed comments to the total number of comments for each attribute is also examined. For example, attribute S1 can be found in 19.11% of the comments, i.e. 319 out of 1669 comments. All attributes can be then ranked by this percentage. As a result, M5, P4, P1, and O2 can be found in more than 40% of the comments while the remaining attributes are addressed in 0.06%-19.11% of the comments. Since the comment lengths of the samples are skewed, log-transformation is performed to the raw data for ensuring normality for later analysis.

“Table 1”

“Table 2”

The correlation coefficient (r) between attributes per comment and length of comments (log-transformed) is 0.49 ($p=0$). It means that a significant correlation with medium strength (r : 0.4 – 0.7) can be found. In other words, it informs the fact that users may need to use more characters to cover more attributes related to the product performance. Insignificant correlation is found between

comments per attribute and characters used per attribute ($r=0.05$, $p=0.7863$). It implies that the number of characters used to describe an attribute is not affected by the prevalence of that attribute. For example, M5 is one of the most common attributes found in the comments, but an average of 55.7 characters is used to address it. In contrast, an average of 72 characters is used to address O6 which is the least common attribute.

4.3 Statistical Correlation Analysis (the second mechanism)

As a complement to the correlation analysis, Table 3 reports the coefficients of partial correlation analysis (r), which are significant at 95% confidence level ($p<0.05$) without the effect of controlling the third variable. Thus, for example, weak but significant correlation ($r=0.3$, $p<0.05$) between S1 and D4 is found and this correlation may be affected by a third variable such as S3, D1, D3, etc. Partial correlation can be then used to confirm the connection between S1 and D4. If the connection is still found significant while controlling for the third variable, it must be sustained, otherwise, it is removed [22, 25, 36]. Accordingly, the linkages between S1 and S3, and S1 and D1 are removed due to the impact of the third variable. More information about partial correlation can be found in the study of [53].

“Table 3”

Figure 2 shows all the direct as well as indirect association between the key output (comment length) and inputs (important attributes). In brief, there are total of six attributes (D10, M6, O4, P1, P3, and P5) directly related to the comment length (LEN). Particularly, mutual influence is found between $O4 \leftrightarrow D10$, and $O4 \leftrightarrow P1$. Also, there are total of eleven attributes (D7, D9, M1, M3, D3, P6, D4, M5, O2, M7, and S1) indirectly connecting to LEN.

“Figure 2”

However, Figure 2 cannot show a thorough picture how LEN can be affected. In addition to the main effects of attributes, 2-factor interaction effects are also considered. ANOVA is performed to examine the impact of both main effects and interaction effects on LEN. The results of ANOVA are reported in Table 4. From Table 4, the main effects of P1, P3 and P5 are deemed as insignificant ($p > 0.05$) while only three interaction effects are found significant ($p < 0.05$), i.e. D10*P1, O4*P1, and P1*P5. Hence, Figure 2 can be revised as Figure 3 in which a more complete and informative description about the relationship between comment length and attributes can be obtained.

“Table 4”

“Figure 3”

To facilitate ANN analysis, the model (Figure 3) can be split into four different input-output models, i.e. Model 1-Model 4 corresponding to four different hubs: comment length (LEN), D10, O4, and M6, as shown in Figure 4. A hub is formed when a factor is connected with two other factors or more. By examining the input-output relationship, all factors (inputs) should be directed to the hubs (outputs). That's why all connections in the four models are pointing at the hubs. Even though P6 has more than one connection as shown in Figure 3, it can hardly form a hub as all of its connections are not pointing towards itself.

“Figure 4”

4.4 Analysis Results of ANN and MLR (the third mechanism)

Each of the input-output models is used to define a single ANN, hence, four different ANNs are developed. Since the output of Model 1 is a continuous variable and that of the other models is a binary one, two different measures are used for model validation: mean absolute percentage error (MAPE) for Model 1 and prediction correctness (PC) for other models. MAPE is defined by Equation (9) where A_i and E_i are the observed values and predicted values of i -th testing dataset respectively, and D is the total number of tested dataset. PC is defined by Equation (10) as the ratio of the total

number of correct predictions ($\sum\theta_i$) over the total number of tested dataset (D). After training and validating (Table 5), it is found that, in yielding the best performance, P must be 7 in Model 1 (N=6) where both MAPE and standard deviation (SD) are the minimum. In maximising PC, P must be 10, 4, and 3 for Model 2 (N=9), Model 3 (N=3), and Model 4 (N=2), respectively.

$$\text{MAPE} = \frac{1}{D} \sum_{i=1}^D \frac{|A_i - E_i|}{A_i} \times 100 \quad (9)$$

$$\text{PC} = \left(\sum_{i=1}^D \theta_i \right) \times \frac{100}{D} \quad (10)$$

where $\theta_i = \begin{cases} 0 & \text{if } A_i \neq E_i \\ 1 & \text{if } A_i = E_i \end{cases}$

“Table 5”

For benchmarking purpose, Table 6 shows the comparison results between ANN and MLR in modeling the connectivity in Model 1-Model 4 over 10-fold cross-validation. The table shows that ANN can outperform MLR at 95% significance level for Model 1 ($p=0$) and Model 2 ($p<0.05$), and it would confirm the non-linearity of the relationship between the variables [54, 55]. But, no significant difference between ANN and MLR is found for Model 3 ($p>0.1$) and Model 4 ($p>0.5$). A possible reason is that Model 3 and Model 4 are much simpler and smaller in size with only 2-3 inputs as compared to Model 1 and Model 2. In modeling small problems, simple regression methods may yield comparable performance as more sophisticated methods such as ANN [38]. Nevertheless, ANN is deemed to be reliable in capturing the connectivity between attributes and the comment length.

“Table 6”

Tables 7-10 show the relative importance of factors in Model 1-Model 4 computed by ten different methods as discussed in Section 3. Since ANN significantly outperforms MLR in modelling the

connectivity of Model 1 and Model 2, it can be assumed that such connectivity may be non-linear, hence, ANN-based methods of computing the input influence must be more accurate and reliable. In Model 1, the two most influential factors (M6 and P1*P5) in affecting the comment length can be addressed by all ANN-based methods except SA. Similarly, in Model 2, the three most influential factors (O4, P6, and LEN) can be identified by all ANN-based methods except SA.

“Table 7”

“Table 8”

Since there is no significant difference between ANN and MLR in modeling the connectivity of Model 3 and Model 4, all ten methods are compared. In Model 3, all methods can address the two most influential factors (P6 and LEN) except SA. Likewise, all methods can identify the most important factor (LEN) in Model 4 except SA. Based on the above observations, it can be concluded that SA method is found to be unreliable in measuring the input influence from our datasets. Same as the study of [36], the ranking results by the COM method are proved to be consistent with that of Garson’s and Yoon’s methods. Moreover, the COM method is able to generate ranking results comparable to expert judgments [24] and differentiate factors with real-world consent [22, 25, 54, 56]. Given its authenticity, the COM method is used to derive the input influence to the comment length.

“Table 9”

“Table 10”

Using the COM method, the relative importance of factors in Model 1-Model 4 can be summarised in Table 11. In Model 1, the two most influential attributes linking to the comment length (LEN) are “P1*P5” and “M6” with overall influence = 72%. In Model 2, the four most influential factors contributing to “D10” are “P6”, “O2”, “LEN” and “O4” with overall influence = 72%. Interestingly, LEN is the most influential factor in both Model 3 and Model 4 with overall influence = 83%.

“Table 11”

If only connections with more than 70% overall influence are examined, the connectivity model can be shown as Figure 5. From Figure 5, all significant direct as well as indirect influence between comment length and attributes can be highlighted. It is noted that, given any empirical data, the true importance of variables is usually unknown [56]. Therefore, this model can only provide explanatory information to support decision-making.

“Figure 5”

5. Discussions

From the connectivity model (Figure 5), it can be observed that if a user concerns about product quality and product technological performance (P1*P5), and/or the legal regulation (M6), he/she tends to write a long comment regarding the target product. These two main factors account for more than 70% of influence towards the comment length (LEN) while customer input (D10) and organisational support (O4) only account for the remaining 30% of influence. It implies that the user needs to apply more words to define P1*P5 and M6 in the comment. This is in line with the study of [33]. While P5 alone is objective (low- or hi-tech), it becomes highly subjective when users link it up with P1. As we know, quality, which is customer-oriented, is achieved only when customer needs can be met. In other words, hi-tech (or low-tech) products do not necessarily increase (or decrease) customer satisfaction. This can be justified by the findings of [57] that the technical performance of the product was found not significantly correlated with market share in high-tech industries. That's why users are more conscious about such products and generally need a longer comment to define such opinions (P1*P5). The explanation of P1*P5 interaction can also derive from consumer's growing sense of empowerment observed among users of social media platforms. [58] noticed that nowadays consumers desire to play a greater role in product development and co-creation which can be facilitated by the Internet-based platforms.

Also, M6 is rigid and users are required to specify explicitly which regulation or standard they are concerned about. Hence, a longer comment is also created. This can indicate consumers' growing role in the process of product development and co-creation. It is apparent that now consumers are not only concerned about product features but also legal aspects that may affect the product development and its introduction to the market [59]. In contrast, LEN is less sensitive to D10 and O4. In brief, if only D10, M6, O4 and P1*P5 are considered, longer comments are expected if the users concern more about M6 and/or P1*P5.

The aggregated model also indicates the “bi-directional” relationship between $LEN \leftrightarrow D10$, $LEN \leftrightarrow O4$, and $LEN \leftrightarrow M6$. Regarding $LEN \leftrightarrow M6$, it implies that longer comments are created if customers want to emphasise M6. On another hand, if a new comment is long, we are confident to say that it must contain information related to M6 (e.g. environmental issues). Regarding $LEN \leftrightarrow O4$, whereas O4 does not have huge impact on the LEN as compared to P1*P5 and M6, LEN does have substantial influence on O4. In other words, if the new comment is long, it must also contain information related to O4 (e.g. customer service, product recall, etc). This relationship reveals that consumers do not request to take part in the early stages of product development such as idealisation and product development, but they want to be involved in post-launch activities [60]. Given the weak impact of LEN on D10, the new comment may or may not contain information related to customer input (e.g. good sense of touching, fashionable appearance, etc).

There are four attributes which are indirectly linked to LEN: D10, communication (O2), O4, and product innovativeness (P6). Due to the mutual link between D10 and O4, they have direct as well as indirect impact on LEN through each other. Their impact however is deemed to be weak. Both O2 and P6 are indirectly linked to LEN through D10. It implies that most of the customer-specific needs are related to O2 (e.g. internal communication within the company) and P6 (e.g. creativeness of the product). Moreover, P6 is indirectly linked to LEN via O4. Another implication is that top management support is critical to enhance the creativity of the product, at least from the users' perspective.

Based on the above discussions, there are several key findings:

- If users are concerned about P1*P5 and/or M6, they tend to write long comments.
- If the new comments are long, they are most likely related to M6 and/or O4.
- Concerns about D10 are mostly associated with that of O2 and P6 in the comments.
- Concerns about O4 are mostly associated with that of P6 in the comments.

6. Conclusions

In order to overcome the challenges of social media data mining, this study provides a structured mechanism to extract values from the data. In the proposed approach, a simple yet comprehensive mapping between social media data and a number of pre-defined attributes is firstly conducted. Next, input-output models are developed to identify the relationship between attributes and comment length. Finally, neural networks are created, allowing such relationship to be formulated quantitatively. The final outcome of the procedure is a list of key attributes that are of consumer concern and which have to be addressed by the company to improve the product performance.

For illustration purpose, the proposed approach was used to analyse user-generated comments concerning Samsung Galaxy S4 model from official Samsung Mobile Facebook page within a 2-month window. An interesting finding was observed that consumers may convey less information through common attributes. In other words, less common attributes may help elicit useful information from the consumers, e.g. an important attribute, legal issues regarding market environment (M6), was found in less than 3% of the comments. Our analysis further revealed that consumers demonstrate their major concerns over the coupled impact of product quality and product technological performance (P1*P5) and M6 by writing long comments. The analysis also confirmed that the longer the comments, the more likely the information is about M6 and support which the consumer receives from the company (O4), where consumer input (D10) has marginal impact. Thus it is suggested that Samsung should give priority to quality and product technological performance as well as possible legal issues when developing its products.

In this study, the focal point is the information richness of the comments concerning the product performance; hence the authors did not evaluate the sentiment of the comments when characterising the comments. In other words, the comments are considered rich (characterised by the identified attributes) regardless of whether they are negative or not. Another limitation of this study is that only comment length is considered for the analysis. In the future, more characteristics of the comments, such as sentiment, attitude and personality of the users posting the comment, will be considered. Incorporation of such attributes into the investigation will allow multi-dimensional way of addressing information richness of comments. The proposed method can also be applied to different areas or topics by a clear definition of the mapping. More importantly, this study does provide a first attempt to pave the path for quantitative analysis in the realm of social media research.

Acknowledgement

This research was partially supported by the Ningbo Soft Science Programme, the Ningbo Science and Technology Bureau (reference number: 2016A10037).

References

1. He, W., H. Wu, G. Yan, V. Akula, and J. Shen, A novel social media competitive analytics framework with sentiment benchmarks. *Information & Management*, 2015. **52**(7): p. 801-812.
2. Chan, H.K., E. Lacka, R.W.Y. Yee, and M.K. Lim, The role of social media data in operations and production management. *International Journal of Production Research*, 2015: p. 1-10.
3. Trainor, K.J., J. Andzulis, A. Rapp, and R. Agnihotri, Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM. *Journal of Business Research*, 2014. **67**(6): p. 1201-1208.
4. Kim, A.J. and E. Ko, Do social media marketing activities enhance customer equity? An empirical study of luxury fashion brand. *Journal of Business Research*, 2012. **65**(10): p. 1480-1486.
5. Chaudhuri, S., U. Dayal, and V. Narasayya, An Overview of Business Intelligence Technology. *Communications of the Acm*, 2011. **54**(8): p. 88-98.
6. Zeng, B. and R. Gerritsen, What do we know about social media in tourism? A review. *Tourism Management Perspectives*, 2014. **10**: p. 27-36.
7. He, W., S.H. Zha, and L. Li, Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, 2013. **33**(3): p. 464-472.
8. Ballings, M. and D. Van den Poel, CRM in social media: Predicting increases in Facebook usage frequency. *European Journal of Operational Research*, 2015. **244**(1): p. 248-260.
9. Smith, K.A., R.J. Willis, and M. Brooks, An Analysis of Customer Retention and Insurance Claim Patterns Using Data Mining: A Case Study. *The Journal of the Operational Research Society*, 2000. **51**(5): p. 532-541.
10. Shen, C.-w. and C.-J. Kuo, Learning in massive open online courses: Evidence from social media mining. *Computers in Human Behavior*, 2015. **51**: p. 568-577.

11. Gandomi, A. and M. Haider, Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 2015. **35**(2): p. 137-144.
12. Chan, H.K., X. Wang, E. Lacka, and M. Zhang, A Mixed-Method Approach to Extracting the Value of Social Media Data. *Production and Operations Management*, 2016. **25**(3): p. 568-583.
13. Zeng, D., H. Chen, R. Lusch, and S.-H. Li, Social Media Analytics and Intelligence. *IEEE Intelligent Systems*, 2010. **25**(6): p. 13-16.
14. Dang, Y., Y.L. Zhang, P.J.H. Hu, S.A. Brown, Y.C. Ku, J.H. Wang, and H.C. Chen, An integrated framework for analyzing multilingual content in Web 2.0 social media. *Decision Support Systems*, 2014. **61**: p. 126-135.
15. Bose, R., Advanced analytics: opportunities and challenges. *Industrial Management & Data Systems*, 2009. **109**(1-2): p. 155-172.
16. Nassirtoussi, A.K., T.Y. Wah, S.R. Aghabozorgi, and D.N.C. Ling, Text mining for market prediction: A systematic review. *Expert Systems with Applications*, 2014.
17. Liu, L., J. Wu, P. Li, and Q. Li, A social-media-based approach to predicting stock comovement. *Expert Systems with Applications*, 2015. **42**(8): p. 3893-3901.
18. Jiang, S., H.C. Chen, J.F. Nunamaker, and D. Zimbra, Analyzing firm-specific social media and market: A stakeholder-based event analysis framework. *Decision Support Systems*, 2014. **67**: p. 30-39.
19. Yu, Y., W.J. Duan, and Q. Cao, The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems*, 2013. **55**(4): p. 919-926.
20. Ballings, M. and D. Van den Poel, CRM in Social Media: Predicting Increases in Facebook Usage Frequency. *European Journal of Operational Research*, 2015.
21. Moro, S., P. Cortez, and P. Rita, A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 2014. **62**: p. 22-31.
22. Xu, M., T.C. Wong, and K.S. Chin, Modeling daily patient arrivals at Emergency Department and quantifying the relative importance of contributing variables using artificial neural network. *Decision Support Systems*, 2013. **54**(3): p. 1488-1498.
23. Bose, I. and X. Chen, Quantitative models for direct marketing: A review from systems perspective. *European Journal of Operational Research*, 2009. **195**(1): p. 1-16.
24. Wong, T.C., S.Y. Wong, and K.S. Chin, A neural network-based approach of quantifying relative importance among various determinants toward organizational innovation. *Expert Systems with Applications*, 2011. **38**(10): p. 13064-13072.
25. Wong, T.C., S.C. Ngan, F.T.S. Chan, and A.Y.L. Chong, A two-stage analysis of the influences of employee alignment on effecting business-IT alignment. *Decision Support Systems*, 2012. **53**(3): p. 490-498.
26. Cao, Q., M.E. Parry, and K.B. Leggio, The three-factor model and artificial neural networks: predicting stock price movement in China. *Annals of Operations Research*, 2011. **185**(1): p. 25-44.
27. Yu, L., S.Y. Wang, and K.K. Lai, Credit risk assessment with a multistage neural network ensemble learning approach. *Expert Systems with Applications*, 2008. **34**(2): p. 1434-1444.
28. Lam, H.Y., G.T.S. Ho, C.H. Wu, and K.L. Choy, Customer relationship mining system for effective strategies formulation. *Industrial Management & Data Systems*, 2014. **114**(5): p. 711-733.
29. Li, Y.M., C.Y. Lai, and C.W. Chen, Discovering influencers for marketing in the blogosphere. *Information Sciences*, 2011. **181**(23): p. 5143-5157.
30. Li, Y.M., H.W. Hsiao, and Y.L. Lee, Recommending social network applications via social filtering mechanisms. *Information Sciences*, 2013. **239**: p. 18-30.
31. Ghiassi, M., J. Skinner, and D. Zimbra, Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with Applications*, 2013. **40**(16): p. 6266-6282.
32. Olden, J.D. and D.A. Jackson, Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, 2002. **154**(1-2): p. 135-150.
33. Mudambi, S.M. and D. Schuff, What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.Com. *Mis Quarterly*, 2010. **34**(1): p. 185-200.

34. Halpern, D. and J. Gibbs, Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Computers in Human Behavior*, 2013. **29**(3): p. 1159-1168.
35. Liu, Z.W. and S. Park, What makes a useful online review? Implication for travel product websites. *Tourism Management*, 2015. **47**: p. 140-151.
36. Wong, T.C., K.M.Y. Law, H.K. Yau, and S.C. Ngan, Analyzing supply chain operation models with the PC-algorithm and the neural network. *Expert Systems with Applications*, 2011. **38**(6): p. 7526-7534.
37. Sung, A.H., Ranking importance of input parameters of neural networks. *Expert Systems with Applications*, 1998. **15**(3-4): p. 405-411.
38. Wong, T.C. and A.H.S. Chan, A Study of the Impact of Different Direction-of-Motion Stereotypes on Response Time and Response Accuracy Using Neural Network. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans*, 2012. **42**(5): p. 1077-1087.
39. Huang, Z., H.C. Chen, C.J. Hsu, W.H. Chen, and S.S. Wu, Credit rating analysis with support vector machines and neural networks: a market comparative study. *Decision Support Systems*, 2004. **37**(4): p. 543-558.
40. Garson, G.D., Interpreting neural-network connection weights. *AI Expert*, 1991. **6**: p. 47-51.
41. Yoon, Y., T. Guimaraes, and G. Swales, Integrating Artificial Neural Networks with Rule-Based Expert-Systems. *Decision Support Systems*, 1994. **11**(5): p. 497-507.
42. Tsaor, S.H., Y.C. Chiu, and C.H. Huang, Determinants of guest loyalty to international tourist hotels - a neural network approach. *Tourism Management*, 2002. **23**(4): p. 397-405.
43. Kewley, R.H., M.J. Embrechts, and C. Breneman, Data strip mining for the virtual design of pharmaceuticals with neural networks. *Ieee Transactions on Neural Networks*, 2000. **11**(3): p. 668-679.
44. Cortez, P., M. Portelinha, S. Rodrigues, V. Cadavez, and A. Teixeira, Lamb meat quality assessment by support vector machines. *Neural Processing Letters*, 2006. **24**(1): p. 41-51.
45. Howes, P. and N. Crook, Using input parameter influences to support the decisions of feedforward neural networks. *Neurocomputing*, 1999. **24**(1-3): p. 191-206.
46. Lindeman, R.H., P.F. Merenda, and R.Z. Gold, Introduction to bivariate and multivariate analysis. 1980, Glenview, Ill.: Scott, Foresman. 444 p.
47. Pratt, J.W. Dividing the indivisible: using simple symmetry to partition variance explained. in *Proceedings of Second Tampere Conference in Statistics*. 1987. University of Tampere, Finland.
48. Genizi, A., Decomposition of R(2) in Multiple-Regression with Correlated Regressors. *Statistica Sinica*, 1993. **3**(2): p. 407-420.
49. Zuber, V. and K. Strimmer, High-Dimensional Regression and Variable Selection Using CAR Scores. *Statistical Applications in Genetics and Molecular Biology*, 2011. **10**(1): p. 1-27.
50. Tranfield, D., D. Denyer, and P. Smart, Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, 2003. **14**(3): p. 207-222.
51. Cho, J. and J. Lee, Development of a new technology product evaluation model for assessing commercialization opportunities using Delphi method and fuzzy AHP approach. *Expert Systems with Applications*, 2013. **40**(13): p. 5314-5330.
52. Morris, R., Computerized Content-Analysis in Management Research - a Demonstration of Advantages and Limitations. *Journal of Management*, 1994. **20**(4): p. 903-931.
53. Kalisch, M. and P. Buhlmann, Estimating high-dimensional directed acyclic graphs with the PC-algorithm. *Journal of Machine Learning Research*, 2007. **8**: p. 613-636.
54. Gevrey, M., L. Dimopoulos, and S. Lek, Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling*, 2003. **160**(3): p. 249-264.
55. Lek, S., M. Delacoste, P. Baran, I. Dimopoulos, J. Lauga, and S. Aulagnier, Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling*, 1996. **90**(1): p. 39-52.
56. Olden, J.D., M.K. Joy, and R.G. Death, An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling*, 2004. **178**(3-4): p. 389-397.

57. Mallick, D.N. and R.G. Schroeder, An Integrated Framework for Measuring Product Development Performance in High Technology Industries. *Production and Operations Management*, 2009. **14**(2): p. 142-158.
58. Hoyer, W.D., R. Chandy, M. Dorotic, M. Krafft, and S.S. Singh, Consumer Cocreation in New Product Development. *Journal of Service Research*, 2010. **13**(3): p. 283-296.
59. Chiesa, V. and F. Frattini, Commercializing Technological Innovation: Learning from Failures in High-Tech Markets. *Journal of Product Innovation Management*, 2011. **28**(4): p. 437-454.
60. Crawford, C.M. and C.A. Di Benedetto, *New products management*. 7th ed ed. Irwin/McGraw-Hill series in finance, insurance and real estate. 2003, Boston: McGraw-Hill. xx, 567 p.

Appendix

Table A1. Key factors of new product performance

Categories	Label	Factors	Description	Sources
Strategic	S1	Technological synergy	Congruency between the existing technological skills of the firm and the technological skills needed to execute a new product initiative successfully	Montoya-Weiss and Calantone 1994; Henard and Szymanski 2001; Pun et al. 2010;
	S2	Company resources	Focused commitment of personnel and R&D resources to a new product initiative	Montoya-Weiss and Calantone 1994; Henard and Szymanski 2001; Krishnan and Ulrich 2001;
	S3	Business strategy	This factor indicates the strategic impetus for the product development (for example, defensive, reactive, proactive, imitative).	Montoya-Weiss and Calantone 1994; Hultink et al. 1997;Im and Workman 2004;
	S4	Marketing synergy	Congruency between the existing marketing skills of the firm and the marketing skills needed to execute a new product initiative successfully	Montoya-Weiss and Calantone 1994; Cooper and Kleinschmidt 1995; Hultink et al. 1997; Henard and Szymanski 2001; Krishnan and Ulrich 2001; Pun et al. 2010;
	S5	Innovation strategy	A plan made by an organisation to encourage advancements in technology or service by investing in research and development activities.	Hultink et al. 1997;
Development process	D1	Technical competitiveness	Proficiency of a firm's use of technology in a new product initiative	Montoya-Weiss and Calantone 1994; Henard and Szymanski 2001; Cho and Lee 2013; Cankurtaran et al. 2013;
	D2	Marketing activities	Proficiency with which a firm conducts its marketing activities	Montoya-Weiss and Calantone 1994; Cooper 1994; Cooper and Kleinschmidt 1995; Henard and Szymanski 2001; Cankurtaran et al. 2013;
	D3	Protocol	Protocol refers to the firm's knowledge and understanding of specific marketing and technical aspects prior to product development	Montoya-Weiss and Calantone 1994; Cankurtaran et al. 2013;
	D4	Speed to market	Speed in the concept-to-introduction time line (i.e., time to market)	Cooper and Kleinschmidt 1994; Padmanabhan 1997; Hendricks and Singhal 1997; Gruner and Homburg 2000; Henard and Szymanski 2001; Krishnan and Ulrich 2001;Chen et al. 2005; Mallick and Schroeder 2005;
	D5	Financial/business analysis	The proficiency of ongoing financial and business analysis during development, prior to commercialisation and full-scale launch.	Montoya-Weiss and Calantone 1994; Cooper 1994; Carrillo 2005;
	D6	Cost	Development cost including measures of production, R&D, or marketing cost overruns or expenditures.	Cooper 1994; Carrillo 2005; Chen et al. 2005; Mallick and Schroeder 2005;Pun and Chin 2005;

	D7	Design and testing	Product design, and performance testing and validation.	Krishnan and Ulrich 2001; Pun and Chin 2005; Cankurtaran et al. 2013;
	D8	Process development and improvement	Employment of formalised product development procedures	Pun and Chin 2005; Pun et al. 2010;
	D9	Well-defined plan/roadmap	Well-defined plan-roadmap to develop new product	Cooper and Kleinschmidt 1995; Carrillo 2005; Pun and Chin 2005; Cho and Lee 2013;
	D10	Customer input	Incorporation of customer specifications into a new product initiative	Henard and Szymanski 2001; Ernst 2002; Pun et al. 2010; Cankurtaran et al. 2013;
	D11	Product Launch	Proficiency with which a firm launches the product	Hendricks and Singhal 1997; Hultink et al. 1997; Padmanabhan et al. 1997; Gruner and Homburg 2000; Henard and Szymanski 2001; Krishnan and Ulrich 2001;
	D12	Process concurrency	Synchronisation of activities of multiple processes and it requires good communication among processes.	Chen et al. 2005; Cankurtaran et al. 2013;
Market environment	M1	Market potential	Anticipated growth in customers/customer demand in the marketplace	Montoya-Weiss and Calantone 1994; Hultink et al. 1997; McGrath 1997; Boer 1998; Henard and Szymanski 2001;
	M2	Market competition	Degree, intensity, or level of competitive response to a new product introduction	Montoya-Weiss and Calantone 1994; Hultink et al. 1997; Slater and Narver 1998; Henard and Szymanski 2001; Cankurtaran et al. 2013;
	M3	Market turbulence	The factor refers to the rate of change in the composition of customers' needs and their preferences	Montoya-Weiss and Calantone 1994; Henard and Szymanski 2001; Carrillo 2005; Chen et al. 2005; Pun et al. 2010;
	M4	Entry barriers	The factor refers to obstacles that make it difficult to enter a given market.	Slater and Narver 1998; Cho and Lee 2013; Cankurtaran et al. 2013;
	M5	Customer needs	Expectations and requirements from customers when purchasing the product	Mishra et al. 1996; Henard and Szymanski 2001; Pun and Chin 2005; Cho and Lee 2013;
	M6	Legal regulation	The factor refers to regulations that could affect the product development i.e. environmental issues	Cho and Lee 2013;
	M7	Technological turbulence	The factor refers to the rate of change associated with technology used to develop new products in an industry.	Chen et al. 2005; Cankurtaran et al. 2013;
Organisational	O1	Internal and external relations	This factor refers to the coordination and cooperation within the firm and between firms;	Montoya-Weiss and Calantone 1994; Henard and Szymanski 2001; Carrillo 2005; Pun and Chin 2005; Cankurtaran et al. 2013;
	O2	Communication	Level of communication among the team and cross-departments in a new product initiative	Henard and Szymanski 2001; Ernst 2002; Pun and Chin 2005;
	O3	Experience and competencies	It refers to organisation's experience & alignment with core competencies.	Cankurtaran et al. 2013;
	O4	Organisational support	Degree of senior management support for a new product initiative	Montoya-Weiss and Calantone 1994; Souder and Song 1998; Henard and Szymanski 2001; Ernst 2002; Bastic 2004; Cankurtaran et al. 2013;
	O5	Organisational integration	Degree of multiple-department participation in a new product initiative	Ernst 2002; Chen et al. 2005; Cankurtaran et al. 2013;
	O6	Organisational structure	It includes measures of organisational climate, size, centralisation, reward structure, and job design.	Montoya-Weiss and Calantone 1994; Pun and Chin 2005; Cankurtaran et al. 2013;
Product	P1	Quality	Product ability to perform its primary function	Gruner and Homburg 2000; Pun and Chin 2005;
	P2	Product	Superiority and/or differentiation	Montoya-Weiss and Calantone 1994; Hultink et

		advantage	over competitive offerings	al. 1997;Henard and Szymanski 2001; Pun and Chin 2005;
	P3	Product price	Perceived price-performance congruency (i.e., value)	Hultink et al. 1997; Henard and Szymanski 2001;
	P4	Product meets customer needs	Extent to which product is perceived as satisfying desires/needs of the customer	Gruner and Homburg 2000; Henard and Szymanski 2001;
	P5	Product technological performance	Perceived technological sophistication (i.e., high-tech, low-tech) of the product	Gruner and Homburg 2000; Henard and Szymanski 2001; Mallick and Schroeder 2005;
	P6	Product innovativeness	Perceived newness/originality/uniqueness/radicalness of the product	Hultink et al. 1997; Henard and Szymanski 2001;

References:

- Anderson, E. G., & Joglekar, N. R. (2005). A hierarchical product development planning framework. *Production and Operations Management*, 14(3), 344-361.
- Bastic, M. (2004). Success factors in transition countries. *European Journal of Innovation Management*, 7(1), 65–79.
- Boer, F. P. (1998). Traps, pitfalls and snares in the valuation of technology. *Research Technology Management*, 41, 45-54.
- Cankurtaran, P., Langerak, F., & Griffin, A. (2013). Consequences of New Product Development Speed: A Meta-Analysis. *Journal of Product Innovation Management*. 30(3), 465-486
- Carrillo, J. E. (2005). Industry clockspeed and the pace of new product development. *Production and Operations Management*, 14(2), 125-141.
- Chen, J., Reilly, R. R., and Lynn, G. S. (2005). The impacts of speed-to market on new product success: The moderating effects of uncertainty. *IEEE Transactions on Engineering Management* 52 (2): 199–212.
- Cho, J., and Lee, J. (2013) Development of a new technology product evaluation model for assessing commercialization opportunities using Delphi method and fuzzy AHP approach, *Expert System with Applications*, 40(13), 5314-5330
- Cooper, R. G. (1994). New products: The factors that drive success. *International Marketing Review*, 11(1), 60–76.
- Cooper, R. G. and Kleinschmidt, E. J., 1994, Determination of timeliness in product development. *Journal of Product Innovation Management*, 10, 112-125.
- Cooper, R. G., & Kleinschmidt, E. J. (1995). Benchmarking the firm's critical success factors in new product development. *Journal of product innovation management*, 12(5), 374-391.
- Ernst, H. (2002). Success factors of new product development: a review of the empirical literature. *International Journal of Management Reviews*, 4(1), 1-40.
- Gruner, K. E., & Homburg, C. (2000). Does customer interaction enhance new product success?. *Journal of business research*, 49(1), 1-14.
- Henard, D. H., & Szymanski, D. M. (2001). Why some new products are more successful than others. *Journal of marketing Research*, 362-375.
- Hendricks, K.B., and V.R. Singhal. 1997. Delays in new product introductions and the market value of the firm: The consequences of being late to the market. *Management Science*. 43(April) 422–436.
- Hultink, E. J., Griffin, A., Hart, S., & Robben, H. S. (1997). Industrial new product launch strategies and product development performance. *Journal of Product Innovation Management*, 14(4), 243-257.
- Im, S., and Workman, J. P. Jr., (2004). Market orientation, creativity, and new product performance in high-technology firms. *Journal of Marketing*, 68(2), 114–132.
- Krishnan, V., and Ulrich, K. T. (2001). Product development decisions: A review of the literature. *Management science*, 47(1), 1-21.
- Mallick, D. N., & Schroeder, R. G. (2005). An integrated framework for measuring product development performance in high technology industries. *Production and Operations Management*, 14(2), 142-158.
- McGrath, R.G., (1997). A Real Options Logic for Initiating Technology Positioning Investments, *Academy of Management Review*, 22(4), 974-996
- Mishra, S., Dongwook, K. & Dae, H. L. (1996) Factors affecting new product success: cross-country comparisons, *Journal of Product Innovation Management*, 13(6), 530–550.
- Montoya-Weiss, M.M., and Calanone, R., (1994) Determinants of new product performance: a review and meta-analysis, *Journal of Product Innovation Management*, 11, 397-417
- Padmanabhan, V., S. Rajiv, K. Srinivasan. 1997. New products, upgrades, and new releases, a rationale for sequential product introduction. *Journal of Marketing Research*. 34(November) 456–472.
- Pun, K. F., & Chin, K. S. (2005). Online assessment of new product development performance: an approach. *Total Quality Management and Business Excellence*, 16(2), 157-169
- Pun, K. F., Chin, K. S., & Yiu, M. Y. R. (2010). An AHP approach to assess new product development performance: An exploratory study. *International Journal of Management Science and Engineering Management*, 5(3), 210-218.
- Slater, S. F., & Narver, J. C. (1998). Research notes and communications customer-led and market-oriented: let's not confuse the two. *Strategic Management Journal*, 19(10), 1001-1006.
- Souder, W. E., & Song, X. M. (1998). Analyses of US and Japanese management processes associated with new product success and failure in high and low familiarity markets. *Journal of Product Innovation Management*, 15(3), 208–223.

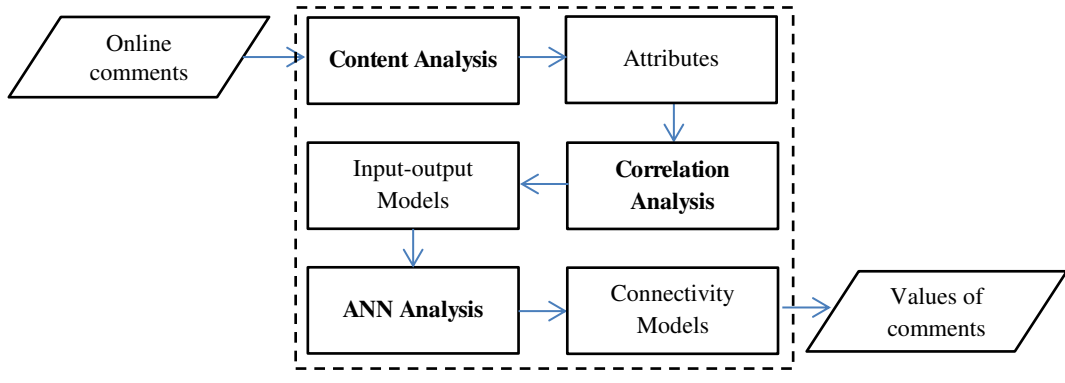


Figure 1: An overview of the proposed approach

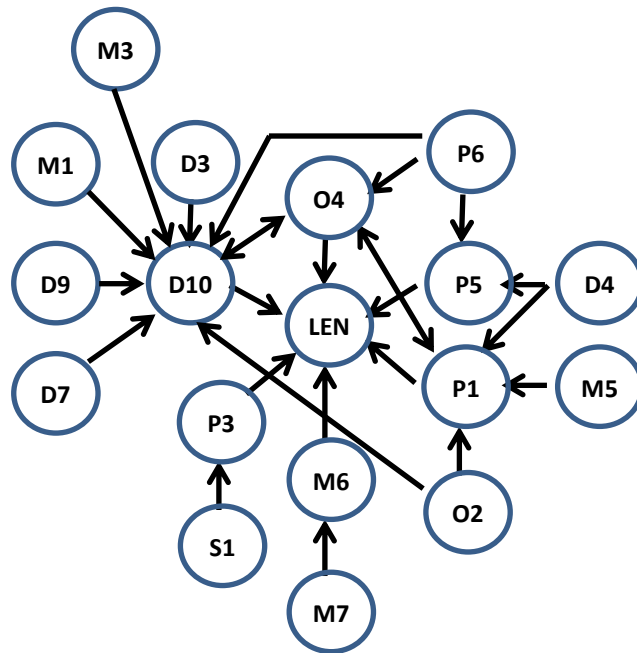


Figure 2: The overall model

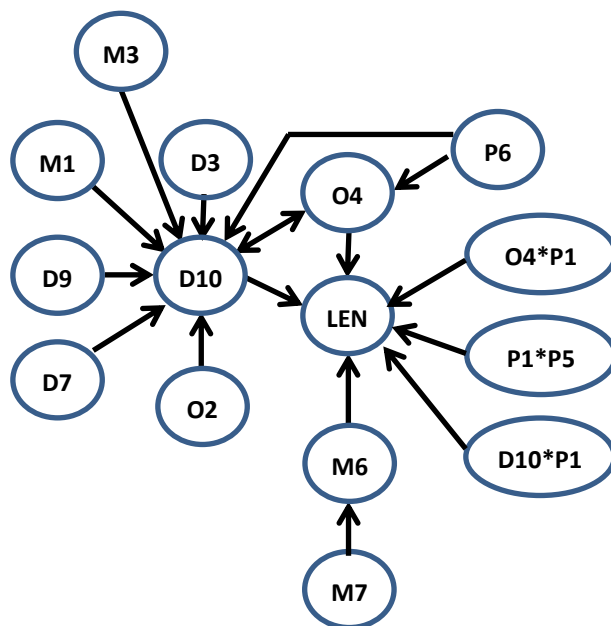


Figure 3: The overall model with interaction effects

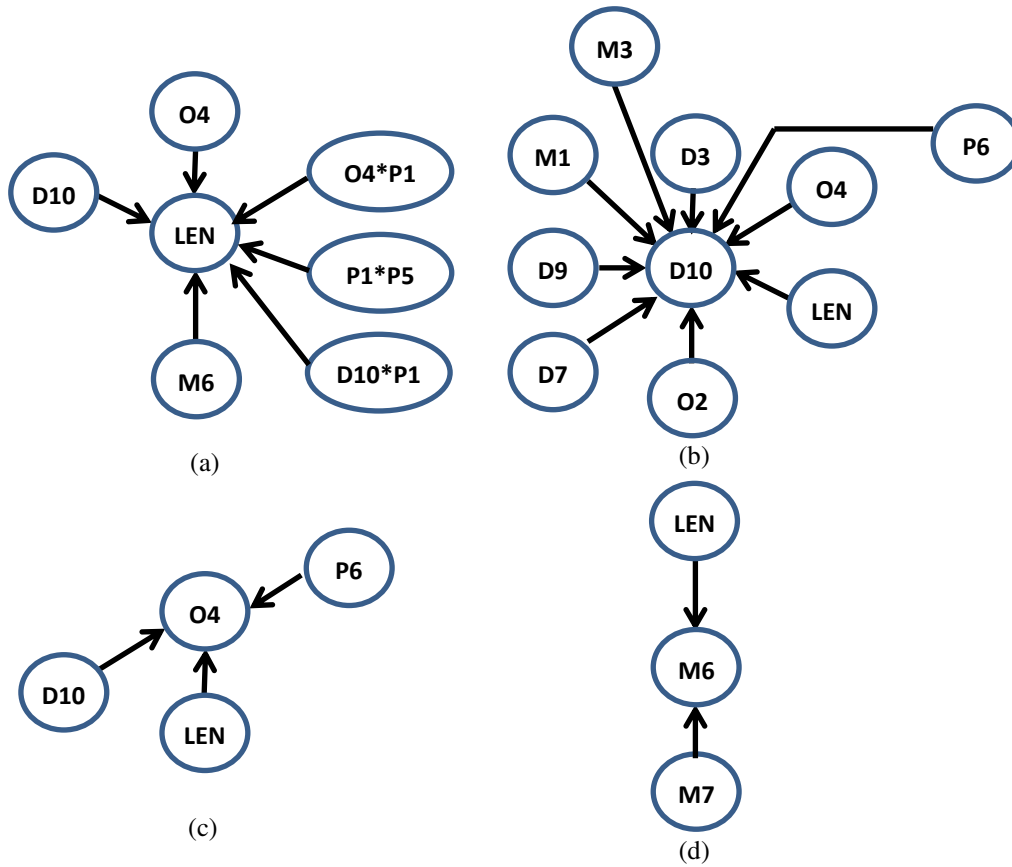


Figure 4: Input-output models. (a) Model 1; (b) Model 2; (c) Model 3; and (d) Model 4

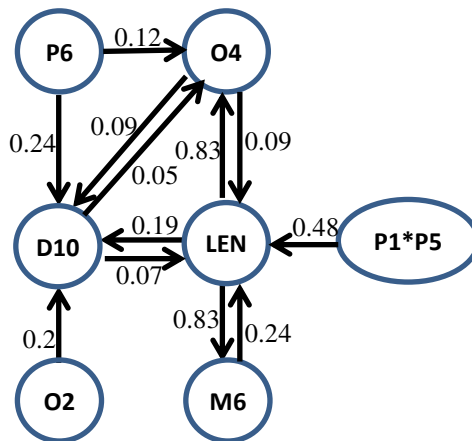


Figure 5: The connectivity model

Table 4: Results of ANOVA

Source	Sum Sq.	d.f.	Mean Sq.	F	p
D10	2.7713	1	2.7713	4.0047	0.045538
M6	3.6076	1	3.6076	5.2131	0.022544
O4	8.2435	1	8.2435	11.9121	0.000572
P1	1.8287	1	1.8287	2.6425	0.10423
P3	2.4036	1	2.4036	3.4733	0.062544
P5	2.1774	1	2.1774	3.1464	0.076277
D10*M6	0.073463	1	0.073463	0.10616	0.74461
D10*O4	0.19406	1	0.19406	0.28042	0.59649
D10*P1	9.25	1	9.25	13.3665	0.000264
D10*P3	0.096045	1	0.096045	0.13879	0.70954
D10*P5	0.49132	1	0.49132	0.70997	0.39958
M6*O4	0.3254	1	0.3254	0.47021	0.49299
M6*P1	2.4964	1	2.4964	3.6074	0.057699
M6*P3	0.93328	1	0.93328	1.3486	0.24569
M6*P5	0.44811	1	0.44811	0.64752	0.42112
O4*P1	6.126	1	6.126	8.8523	0.00297
O4*P3	0.26599	1	0.26599	0.38436	0.53536
O4*P5	0.05959	1	0.05959	0.08611	0.76922
P1*P3	2.4413	1	2.4413	3.5278	0.060525
P1*P5	13.191	1	13.191	19.0613	1.34E-05
P3*P5	0.036658	1	0.036658	0.052971	0.818
Error	1139.772	1647	0.69203		
Total	2158.731	1668			

Table 5: The performance of ANN with varying P

	P	2	3	4	5	6	7	8	9	10	11	12
Model 1	MAPE	8.8155	8.7324	8.7386	8.6528	8.6685	8.5901	8.6505	8.6911	8.6662	8.6551	8.6442
	SD	0.5255	0.5477	0.4477	0.4344	0.4470	0.3659	0.4477	0.4422	0.3959	0.4285	0.4226
Model 2	P	2	3	4	5	6	7	8	9	10	11	12
	PC	0.9090	0.9106	0.9110	0.9120	0.9122	0.9127	0.9131	0.9145	0.9169	0.9157	0.9163
Model 3	P	2	3	4	5	6	7	8	9	10	11	12
	PC	0.8970	0.9024	0.9060	0.9054	0.9018	0.9042	0.9024	0.9024	0.9018	0.9012	0.8976
Model 4	P	2	3	4	5	6	7	8	9	10	11	12
	PC	0.9717	0.9717	0.9657	0.9717	0.9711	0.9717	0.9717	0.9717	0.9717	0.9717	0.9717

Table 6: Comparisons between ANN and MLR

Model	Measure	ANN	MLR	t-test*
1	MAPE	8.59	10.36	p=0
	SD	0.37	0.76	
2	PC	91.69	90.84	p<0.05
3	PC	90.60	89.58	p>0.1
4	PC	97.17	97.17	p>0.5
* $\alpha=0.05$				

Table 7: Relative importance among factors of Model 1

Model 1 (LEN) [#]	ANN-based					MLR-based				
	Garson	Yoon	Tsaur	COM	SA	LMG	Beta2	Pratt	Gen	ZS
D10	0.09	0.04	0.13	0.07	0.11	0.06	0.11	0.06	0.07	0.07
M6	0.18	0.30	0.23	0.24	0.32	0.06	0.06	0.06	0.06	0.06
O4	0.18	0.18	0.12	0.09	0.38	0.26	0.42	0.30	0.27	0.27
D10*P1	0.16	0.09	0.15	0.06	0.10	0.27	0.12	0.22	0.24	0.26
O4*P1	0.17	0.14	0.11	0.06	0.07	0.17	0.22	0.22	0.19	0.18
P1*P5	0.22	0.25	0.26	0.48	0.02	0.18	0.07	0.14	0.17	0.16

[#]Output of the model

Table 8: Relative importance among factors of Model 2

Model 2 (D10) [#]	ANN-based					MLR-based				
	Garson	Yoon	Tsaur	COM	SA	LMG	Beta2	Pratt	Gen	ZS
D3	0.10	0.05	0.12	0.03	0.04	0.04	0.01	0.03	0.04	0.04
D7	0.10	0.06	0.19	0.07	0.35	0.09	0.07	0.10	0.09	0.10
D9	0.11	0.10	0.03	0.05	0.11	0.05	0.04	0.06	0.05	0.05
M1	0.05	0.00	0.02	0.06	0.13	0.11	0.10	0.11	0.11	0.11
M3	0.09	0.10	0.09	0.07	0.33	0.05	0.03	0.04	0.05	0.04
O2	0.10	0.12	0.05	0.20	0.01	0.16	0.18	0.16	0.16	0.16
O4	0.12	0.12	0.16	0.09	0.01	0.08	0.13	0.07	0.08	0.07
P6	0.19	0.27	0.21	0.24	0.01	0.21	0.14	0.20	0.21	0.21
LEN	0.15	0.19	0.14	0.19	0.01	0.21	0.30	0.23	0.21	0.22

#Output of the model

Table 9: Relative importance among factors of Model 3

Model 3 (O4) [#]	ANN-based					MLR-based				
	Garson	Yoon	Tsaur	COM	SA	LMG	Beta2	Pratt	Gen	ZS
D10	0.13	0.14	0.08	0.05	0.04	0.05	0.07	0.05	0.05	0.05
P6	0.22	0.26	0.09	0.12	0.03	0.08	0.08	0.07	0.08	0.07
LEN	0.65	0.60	0.83	0.83	0.93	0.87	0.85	0.88	0.87	0.88

#Output of the model

Table 10: Relative importance among factors of Model 4

Model 4 (M6) [#]	ANN-based					MLR-based				
	Garson	Yoon	Tsaur	COM	SA	LMG	Beta2	Pratt	Gen	ZS
M7	0.24	0.23	0.42	0.17	0.83	0.20	0.20	0.20	0.20	0.20
LEN	0.76	0.77	0.58	0.83	0.17	0.80	0.80	0.80	0.80	0.80

#Output of the model

Table 11: Relative importance computed by the COM method

Inputs	Model 1 (LEN) [#]	Inputs	Model 2 (D10) [#]	Inputs	Model 3 (O4) [#]	Inputs	Model 4 (M6) [#]
D10	0.07	D3	0.03	D10	0.05	M7	0.17
M6	0.24	D7	0.07	P6	0.12	LEN	0.83
O4	0.09	D9	0.05	LEN	0.83		
D10*P1	0.06	M1	0.06				
O4*P1	0.06	M3	0.07				
P1*P5	0.48	O2	0.20				
		O4	0.09				
		P6	0.24				
		LEN	0.19				
Total	1.00	Total	1.00	Total	1.00	Total	1.00

#Output of the model