

## Mining “What they talk about” for a Private Healthcare Service Provider

**Angela Siew-Hoong Lee**

Sunway University, 5, Jalan Universiti, Bandar Sunway  
47500, Selangor, Malaysia

**Tong-Ming Lim**

Sunway University, 5, Jalan Universiti, Bandar Sunway  
47500, Selangor, Malaysia

**Mark P.C. Chia**

SAS Malaysia, Kuala Lumpur Sentral  
Malaysia.

**Sue-Lynn Ea**

Sunway University, 5, Jalan Universiti, Bandar Sunway  
47500, Selangor, Malaysia

**Mun-Yee Yap**

Sunway University, 5, Jalan Universiti, Bandar Sunway  
47500, Selangor, Malaysia

### Abstract

In every industry, customer feedback contains opinions with different types of sentiment on services and products provided by companies that they purchased from. Feedback from customers help business operators to understand their customers better in order to improve different aspects of their products and services. This research studies healthcare service consumers' perception of a private hospital on the quality of food, waiting time, services and customer expenses. This research intends to explore a set of customer's feedbacks from year 2013 to 2016 to investigate potential new findings and to create new value added improvements to the current process. Text mining technique were used to extract and discover hidden knowledge from the unstructured feedbacks. The techniques are text parsing, filter, topic and cluster as for sentiment analysis, term frequency and weight is used in conjunction with corpus of files of text to investigate the emotional elements of the feedbacks that can be classified into either positive, neutral or negative. The outcomes of this study have highlighted several new findings and supported hypothesis of this research.

**Keywords:** Text mining, sentiment analysis, patient's feedback, services, customer expenses, quality of food, waiting time

### INTRODUCTION

Customer feedback are vital for every business to continue improve their quality of service and products. Companies can leverage on these feedbacks to maintain their reputation in the industry. For instance, in trading, manufacturing and hospital sectors, product pricing is not the key factor that keeps customer satisfied; customer services are more important than price of products. Oja (2010) highlighted that customers feedback can help to improve services offer to customers. As healthcare business becomes more competitive, the growing desire of

consumers for better services has also grown rapidly. The desire for more effective organizational management has also become one of the key considerations of healthcare consumers as they choose which healthcare service provider to go to. With good customer services quality, people will be happy to continue to be patrons for the service provider. However, to be able to identify key opinions and sentiments of these customers who undertake their services is always a challenge. Hence, analysing these customers' feedback to identify the 'likes' and 'dislikes' experienced by these customers is a good practice for the management to understand their customers better.

For any operator in the healthcare industry, the management team can attempt to manually read and analyse all the feedbacks provided by the customers but it would be time-consuming and inconsistent due to the vast amount of unstructured text. In view of the tedious analysis, many management might choose to neglect these feedbacks that might be useful and constructive to the hospital operators. These healthcare service operators know that when their services are not good enough, their patients will have bad impression and they will share their bad experiences with others. This will cause revenue lost and reduce the number of healthcare customers that will visit them to take up their services due to the bad feedbacks on social media or business related sites (Barnard, 2002). In addition, healthcare customers will also express their opinions through surveys, forums, and feedback forms on their experiences on the hospital services. In many small and large hospitals, comments on the social media sites and surveys allow hospitals to effectively identify positive and constructive or negative feedbacks for the services betterment.

In this research, a private Malaysian healthcare service provider has chosen to participate in this research study. Three years of customer's feedback from this private healthcare service provided had been used to analyse using SAS Enterprise E-miner. The methodology of this research study will be discussed in further section.

### **Related Works**

A statistic shows that there are approximately 85% of businesses are using various form of text to store data such as surveys (Hotho, et. al 2005). For any industries that have a huge amount of data, analytics tools can be used to draw out many findings within a short period of time on the data. Text mining is defined as a technique that able to automatically extract all information from various sources (Gupta & Lehal, 2009). In text mining, it involves information retrieval, clustering, text analysis and visualization. Text mining helps businesses to discover unknown knowledge and analyze data into new findings. Besides that, text mining can benefit business by saving time and resources, which will provide a fast, timely and accurate feedback for their clients.

According to Lay, Lee, & Ho (2005), text mining is vital and can be used also in hotel industries to develop different competitive strategic planning ahead. Text mining is used in their research to help extract information about other competitor fare for room price and facilities. Other than that, text mining also helps them to mine out things such as customer attitudes and demographics to understand customer's behavior patterns on staying with the hotel.

Although text mining applications are more popular in hospitality industry, text mining can also be applied in the banking and finance industry. Sources of textual data for banks comes from news, online articles, banking contracts, social media, financial reports etc. Even though the usage of text mining applications in the banking business is limited, the benefits of text

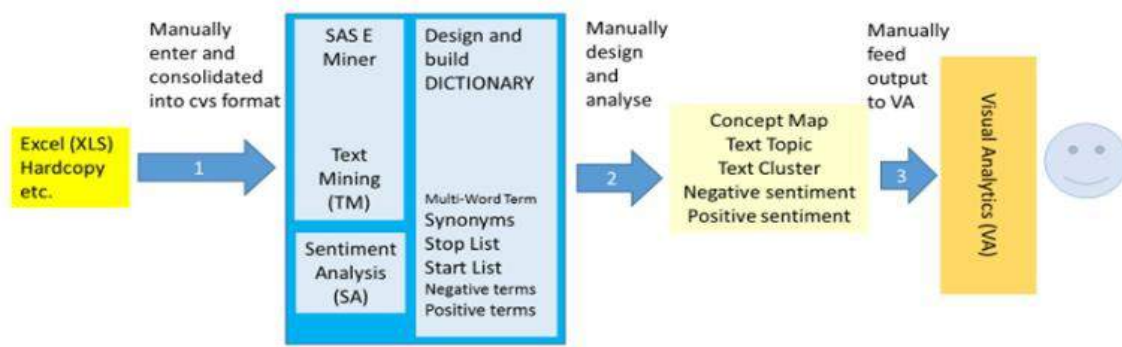
mining in helping the banks perform task more effectively can be seen (Bholat, et al., 2015). Other than that, Bank of Canada also utilize the benefits of text mining to extract significant information to determine the association of the information from Bank of Canada communications with interest rates and monetary policy. Additionally, text mining can also be set to extract information that are related with Fixed Announcement Dates(FADs) and Monetary Policy Reports to measures the movement of the market returns and instability. These methods are effective for banks to determine short-term rates therefore text mining tools is considered a valuable tool in the banking businesses (Hendry & Madely, 2010).

In today's real world, most of the organization and business would like to find out consumer opinions about certain products and services that provided by the industry. Not only business and organizations, individual customer also wish to know more about the opinion about the existing product from another regular user before buying it. Before the Web 2.0 is invented, individual need to ask opinions about the products through family and friends (Liu, 2012). Sentiment Analysis is defined as a type of information extraction and Natural Language Processing that will track the user's comment it is positive or negative feeling in the document text (Vinodhini & Chandrasekaran, 2012). Some methods such as opinion polls, social media platform and surveys are used to collect feedback on the business product and services. Later on, sentiment analysis is used to study the opinions and emotions that are expressed in the feedback from the respondents. Most of the businesses or domain in services, financial services, healthcare, product services, political and social events are using sentiment analysis applications. For instance, getting feedback from consumer or events can measure how success is the product, services or event is going on. Hence, business organization people can adjust and know how consumer feel about the particular product. Other than that, consumer opinions help business to increase business revenue and build a better product from the feedback given by consumer.

### **METHODOLOGY**

This section explains the research methodology for this research study. First of all, the target population for this research are patients from one of the local private hospital service provider. SAS Enterprise Guide will be used to extract and transform the raw data collected which are in an unstructured format to a structured format. Text preparation will be done using SAS Enterprise Miner to ensure the quality and reliability of the text data for Text Mining Analysis. In this research, secondary resources will be used. Data has been collected from the department of customer service whereby patients from different departments has given feedback on the hospital. A preliminary investigation has been conducted to understand more in depth about the current situation in hospital. Next, data cleaning will be conducted to avoid missing values and duplicate data. Therefore, SAS Enterprise Guide will be used to export raw data sets and convert into SAS data set. SAS Enterprise Miner will then be used to carry out the text mining and sentiment analysis. Text preparation will be applied using the functions multi word term, stop words and start list. This process is used to prevent duplication and removed unnecessary terms. After the text preparation step, the data will be easier to analyze and categorized those data into positive feedbacks and negative feedbacks. SAS Enterprise Guide are used to import the excel file into it and change it into SAS data set. To analyze the data, SAS Enterprise Miner will be used to analyze all the data of patient's feedbacks. All the words that are related to positive and negative feedbacks will be extracted out using the text mining technique. The results will be written in a report to explain the important information that extracted out from all the data. SAS Enterprise Guide will be used to make graphs so that it will be easier to understand.

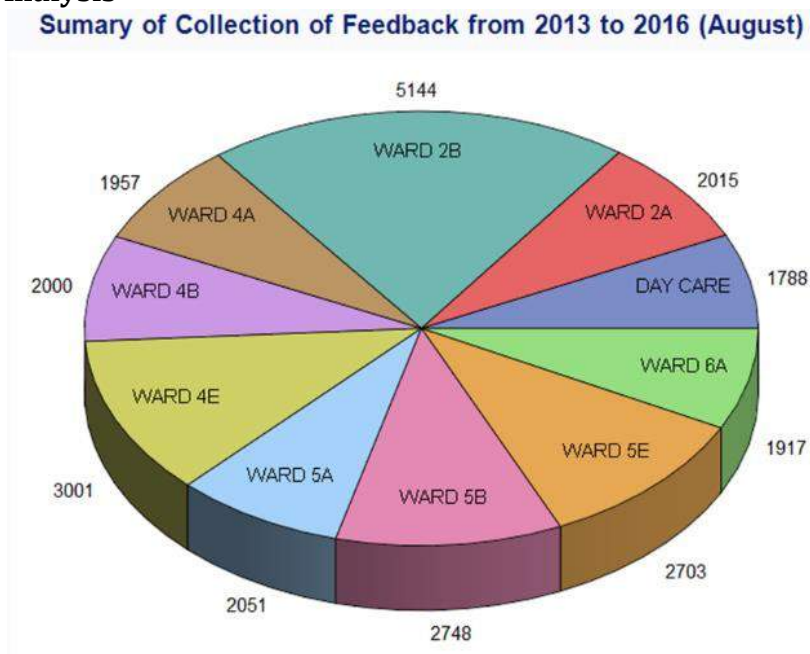
### TEXT AND SENTIMENT ANALYSIS MODEL



**Figure 1: Text and Sentiment Analysis Model**

Figure 1 shows the process flow of the different phases that were carried out in this research study. The Model starts from manually organizing the data given by the hospital. Data given will be in excel format which will be organize manually and consolidated into csv format before transferring the data into SAS Enterprise Guide. SAS Enterprise guide will be used to convert the text data into data sets before importing into SAS Enterprise Miner as a data source. SAS Enterprise Miner will be used to perform data preparation using multi-word term, stop list, start list and synonyms. Data preparation is essential to ensure the quality of the data and results. Once the data preparation is complete, text mining analysis will be conducted to generate the concept map, text topic and text cluster. Next, positive and negative terms will be manually identified and list down to carried out the sentiment analysis and the final results for the hospital will be visualize using SAS Visual Analytics.

### Empirical Analysis



**Figure 2: Top 10 Departments with highest number of feedback forms**

Figure 2 is a pie chart that shows the top 10 departments that has the highest number of feedback forms collected from 2013 till 2016 August. From the results above, Ward 2B have

the highest amount of feedback forms collected with a total of 5144 feedback forms. The second department with the highest collection of feedback forms is Ward 4E with a total of 3001 of feedback forms collected followed by Ward 5B with a total of 2748 feedback forms collected. Therefore, the hospital should look into Ward 2B and Ward 4E due to the high number of feedbacks received from these departments.

## FINDINGS AND RESULTS

### Text Mining Analysis

#### Concept Link

Four key focus are discussed in this section on quality of food, waiting time, services and bill price. In each part of the discussion, the concept links visually shows factors that are linked and contributed to the explanation of what causes the key focus or concept to occur.

#### a. Quality of Food

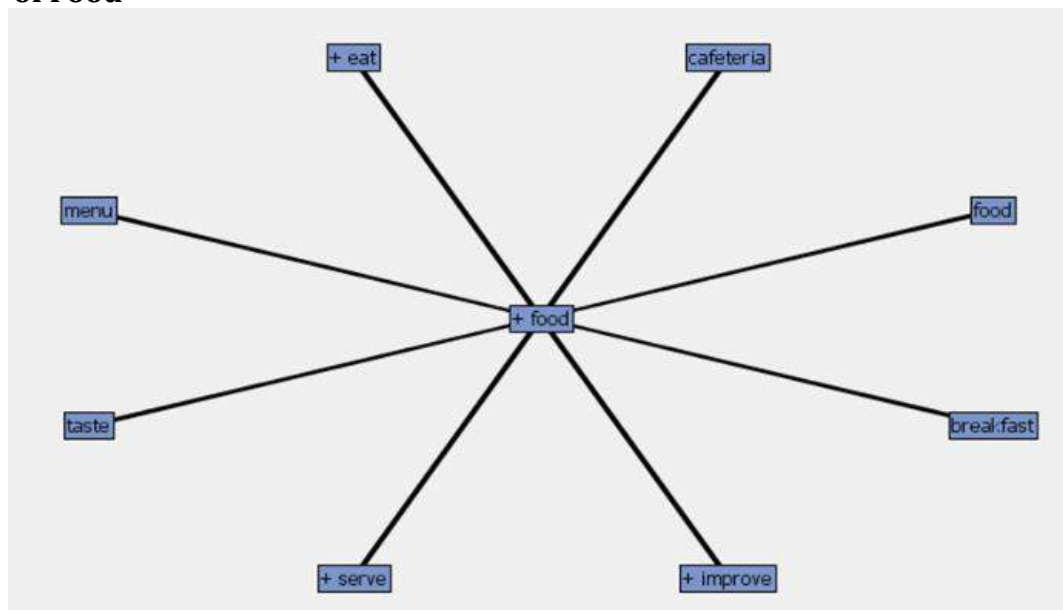
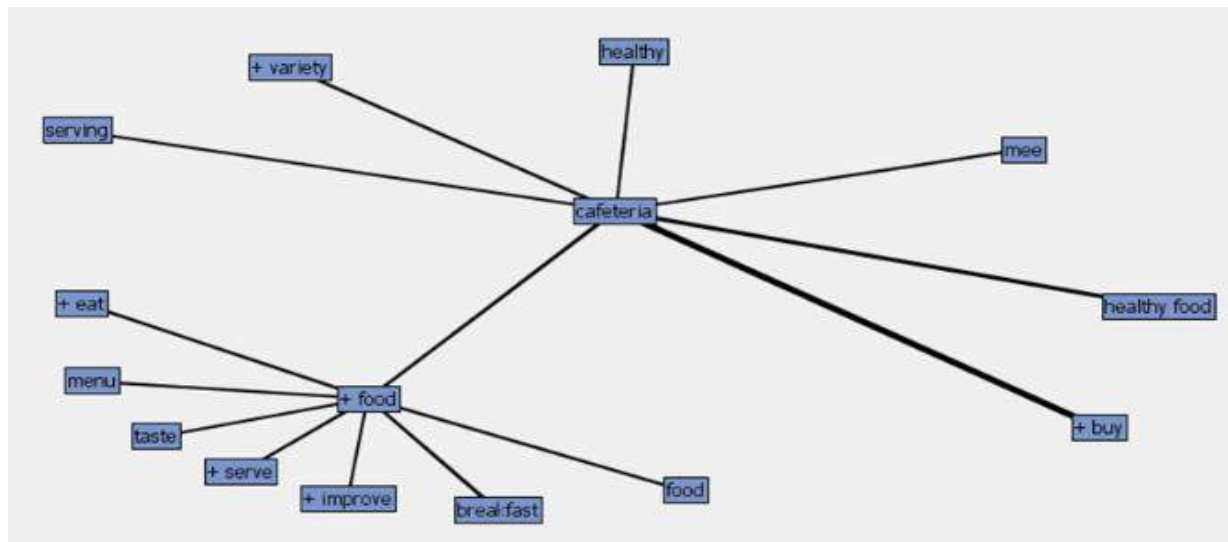


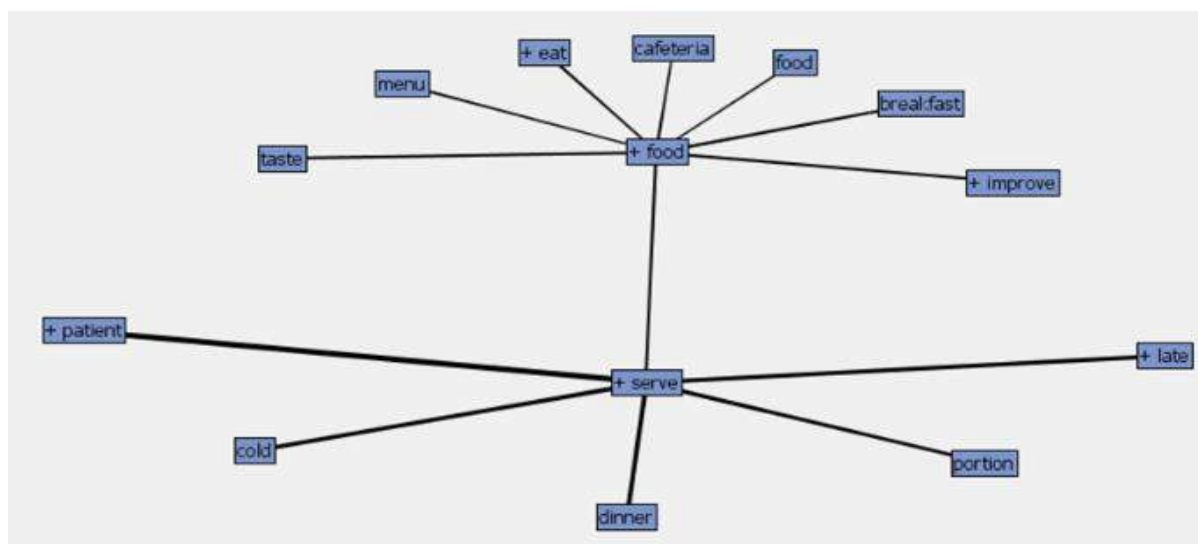
Figure 3: Concept Linking for “+food”

Figure 3 shows the concept linking for the term “+food” which occurred in 159 documents and the other terms that co-occur with the term “+food” consist of “+eat”, “cafeteria”, “food”, “breakfast”, “+improve”, “+serve”, “taste” and “menu”. This indicates that whenever patient’s talks about the hospital food they will relate with these terms. In addition, based on the concept map in Figure 3 it shows that the line thickness which connects the terms to “+food” have the same thickness and width which indicates the frequency of these terms to occur together with “+food” in the documents are almost the same.



**Figure 4: Concept Linking for “+food” expand to “cafeteria”**

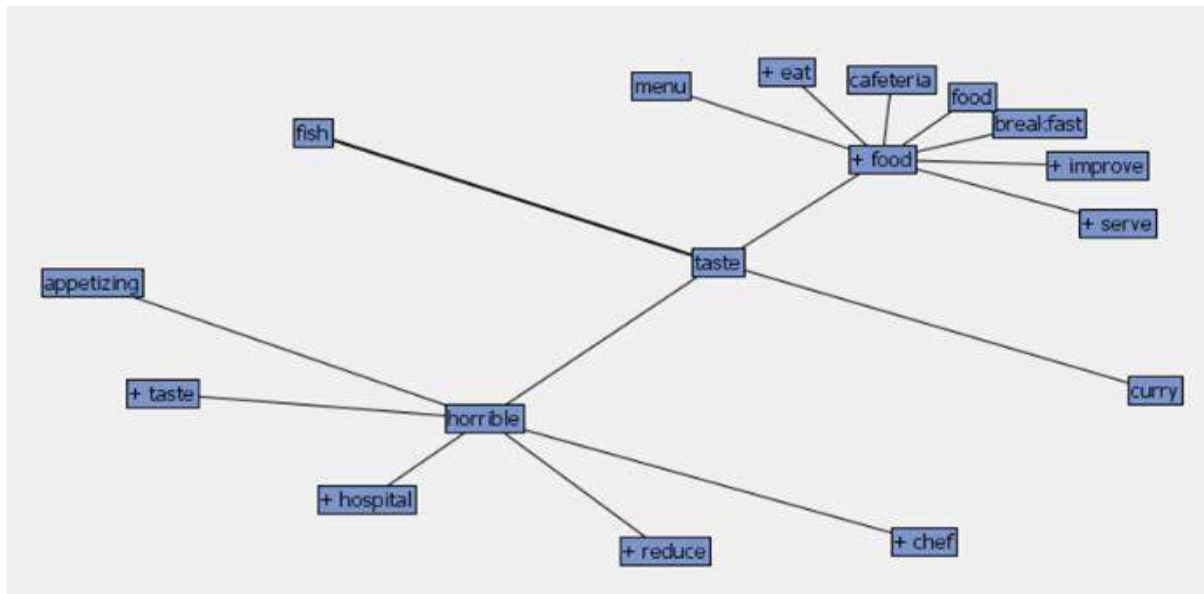
Figure 4 shows the concept linking for “+food” which was expanded to second level on the term “cafeteria”. The term “cafeteria” appeared in 27 documents and 17 of the documents contains both the term “+food” and “cafeteria”. The terms that are associated with the term “cafeteria” includes “serving” “+variety”, “healthy”, “mee”, “healthy food” and “+buy”. The term with the strongest association with the term “cafeteria” is “+buy” as the line connecting to it is the thickest compared to the other terms. The reason the term “buy” is strongly associate with the term “cafeteria” is because the hospital cafeteria is one of the frequent place patients or patient’s friends and family members always go to buy food. Other than that, it seems that patients are requesting for healthier and variety of food in the cafeteria such as more choices for vegetarian food for those who are vegetarian.



**Figure 5: Concept Linking for “+food” expand to “serve”**

Figure 5 shows the concept linking for “+food” which was expanded to the second level for the term “serve”. The term “serve” appeared in 99 documents and which 40 out of the 99 documents contains both the term “+food” and “+serve”. The terms that are associated with the term “+serve” includes

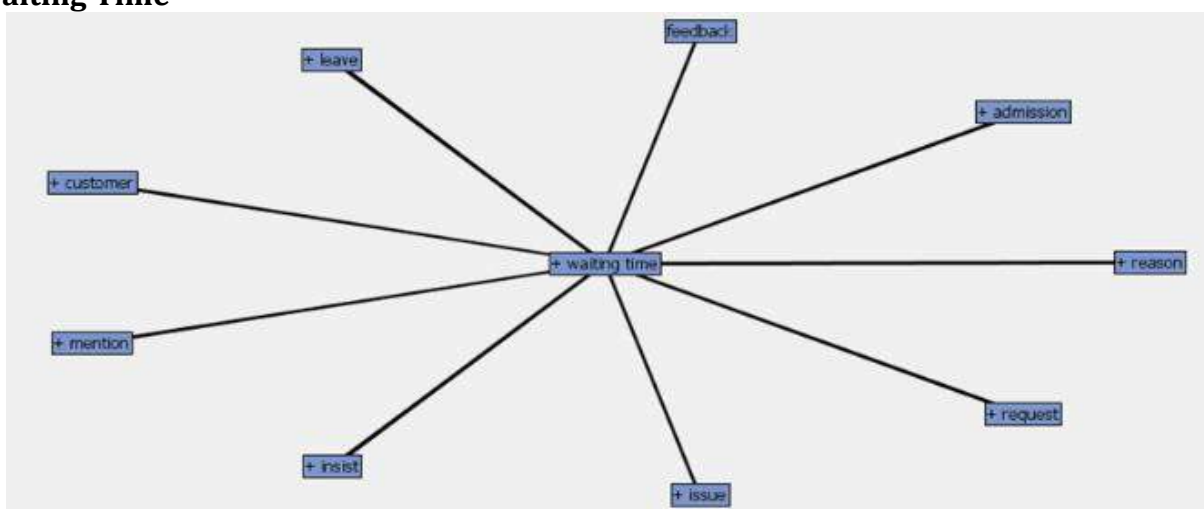
“+patient”, “cold”, “dinner”, “portion” and “late”. The term with the strongest association with the term “+serve” is the term “+patient” as the line connecting these two term is the thickest. This is because the food prepared are for hospital patients and their satisfaction on the quality of the food serve by the hospital is important. Besides that, patient’s also give feedbacks that food is serve late and cold which are not to the patient’s liking. In addition, the term “+serve” is also links with “portion” indicating the food portion size is small and insufficient for the patients.



**Figure 6: Concept Linking for “+food” expand to “taste”**

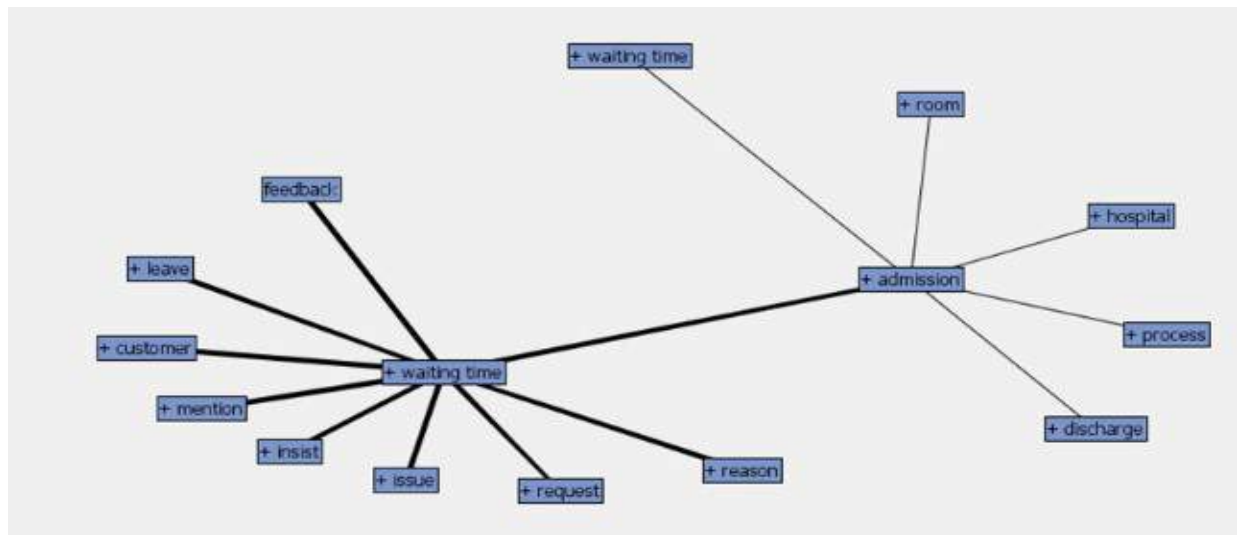
Figure 6 shows the concept linking for “+food” which was expanded to the second level for the term “taste”. The term “taste’ are associated with the terms “fish”, “curry” and “horrible”. The term “horrible” was expanded to the third level to explore what are the terms that patients used when relating with the term “+food”, “taste” and “horrible”. By expanding the term “horrible” to the third level it relates terms such as “appetizing”, “+taste”, “+hospital”, “+reduce” and “chef”. This indicates that the hospital should prepare foods that are more appetizing and tasty to meet different patient’s preference on the food taste.

## b. Waiting Time



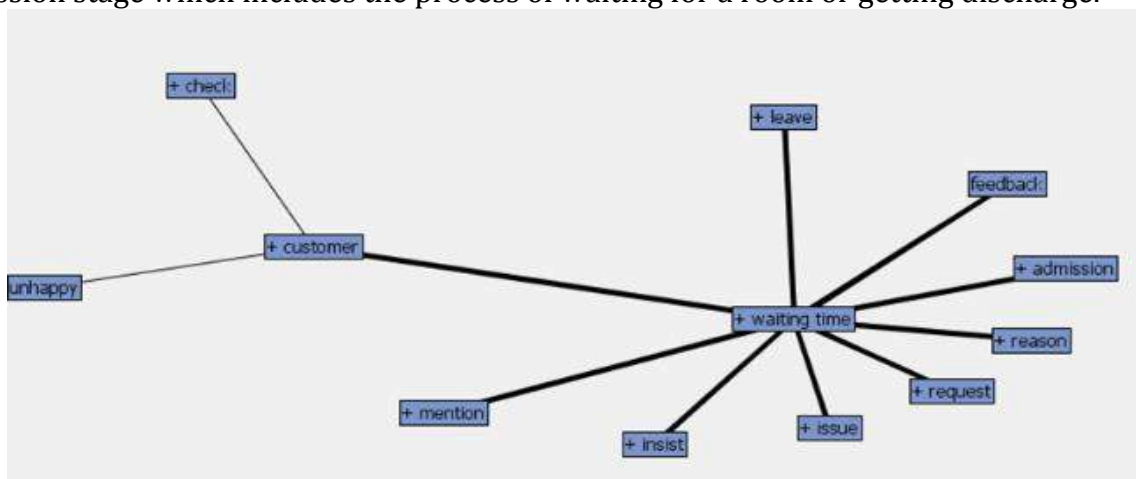
**Figure 7: Concept Linking for “+waiting time”**

Figure 7 shows the concept linking for the term “+waiting time” which occurred in 477 documents and the other terms that are associate with the term “+waiting time” consist of “feedback”, “+admission”, “+reason”, “+request”, “+issue”, “+insist”, “+mention”, “+customer” and “+leave”. This are the terms that are often related to each other whenever patients talk about waiting time in the feedback forms. By taking a closer look in it shows that the terms that are strongly associate with “+waiting time” includes “+customer”, “+admission” “reason” and “+request”. This is because these terms have thicker lines connecting to “+waiting time” compared to other terms which also indicates that these terms have higher frequency to occur together in the documents.



**Figure 8: Concept Linking for “+waiting time” expand to “Admission”**

Figure 8 shows the concept linking for “+waiting time” which was expanded to the second level for the term “+admission”. The term “+admission” occurred in 269 documents and out of these documents, 79 of the documents contain both the term “+waiting time” and “+admission”. The terms that are associated with the term “+admission” includes “+waiting time”, “+room”, “+hospital”, “+process” and “+discharge”. The term “+room” occurs in 316 documents and 44 of the documents contains the terms “waiting time” and “room”. Besides that, the term “+discharge” occurs in 144 documents and 31 of the documents contains both the term “+waiting time” and “+discharge”. These shows that waiting time often occurs during patient’s admission stage which includes the process of waiting for a room or getting discharge.



**Figure 9: Concept Linking for “+waiting time” expand to “Customer”**



Figure 9 shows the concept linking for “+waiting time” which was expanded to the second level for the term “+customer”. The term “+customer” occurred in 244 documents and 38 of these documents contains both the term “+waiting time” and “+customer”. The terms that are link with “+customer” are “+check” and “unhappy”. The term “unhappy” occurs in 80 documents with 25 of these documents contains the terms “waiting time” and “unhappy”. This indicates that customer are often unhappy if the waiting time at the hospital is long.

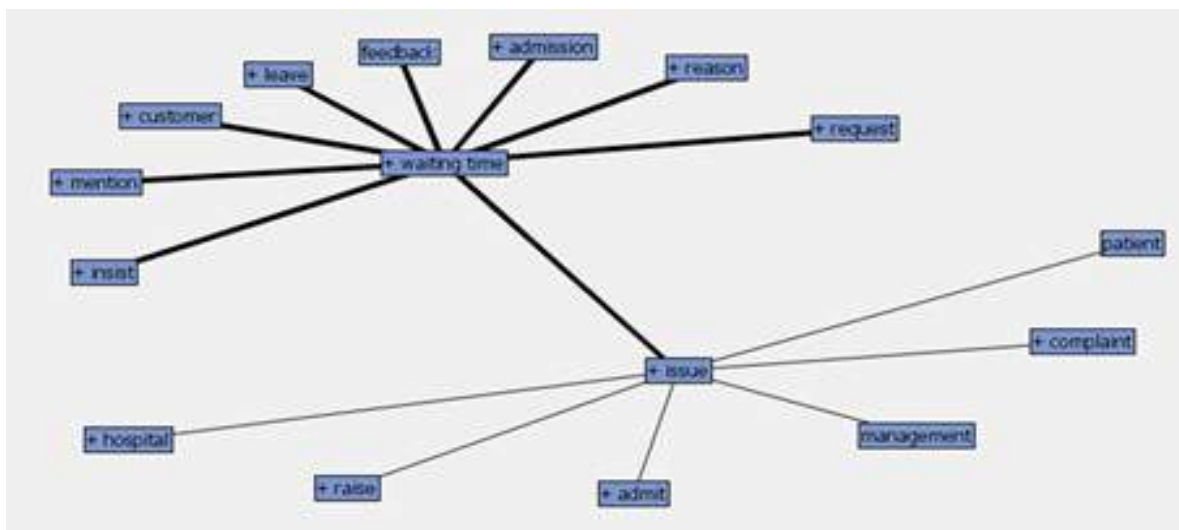


Figure 10: Concept Linking for “+waiting time” expand to “Issue”

Figure 10 shows the concept linking for “+waiting time” which was expanded to the second level for the term “+issue” to see what are the issues that relates with “+waiting time”. The term “+issue” occurred in 108 documents and 35 of the documents contains the terms “+waiting time” and “+issue”. The terms that are associated with the term “+issue” includes “+hospital”, “+raise”, “+admit”, “management”, “+complaint” and “patient”. This shows that waiting time issues often links with admissions and hospital management. Therefore, patients often expressed their hope for management to resolve this issue.

**c. Service**

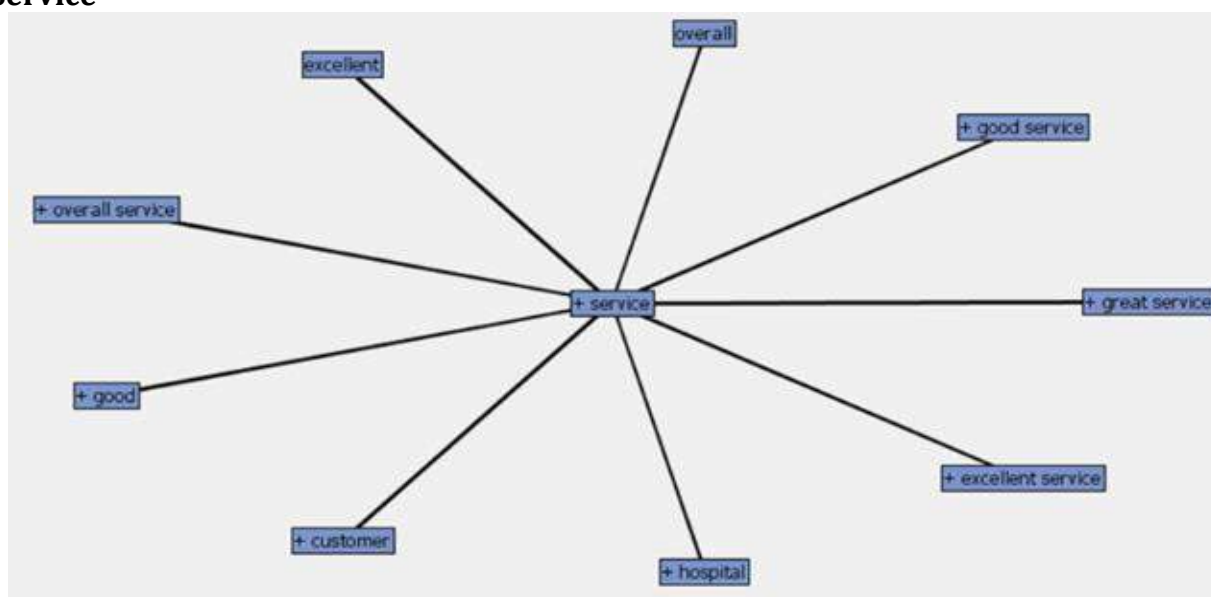
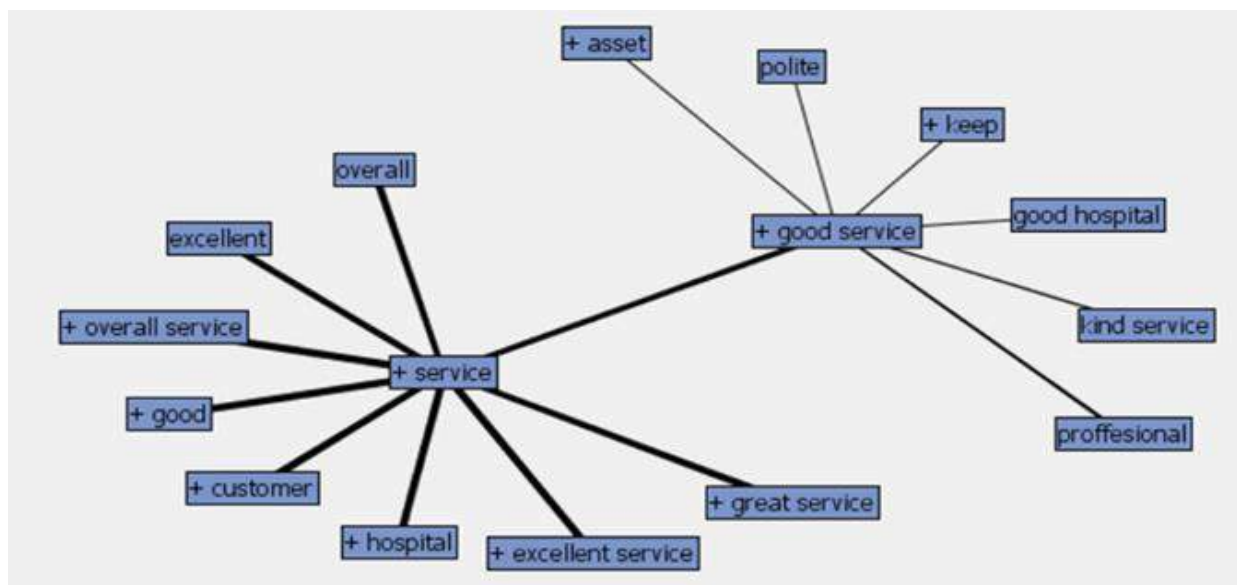


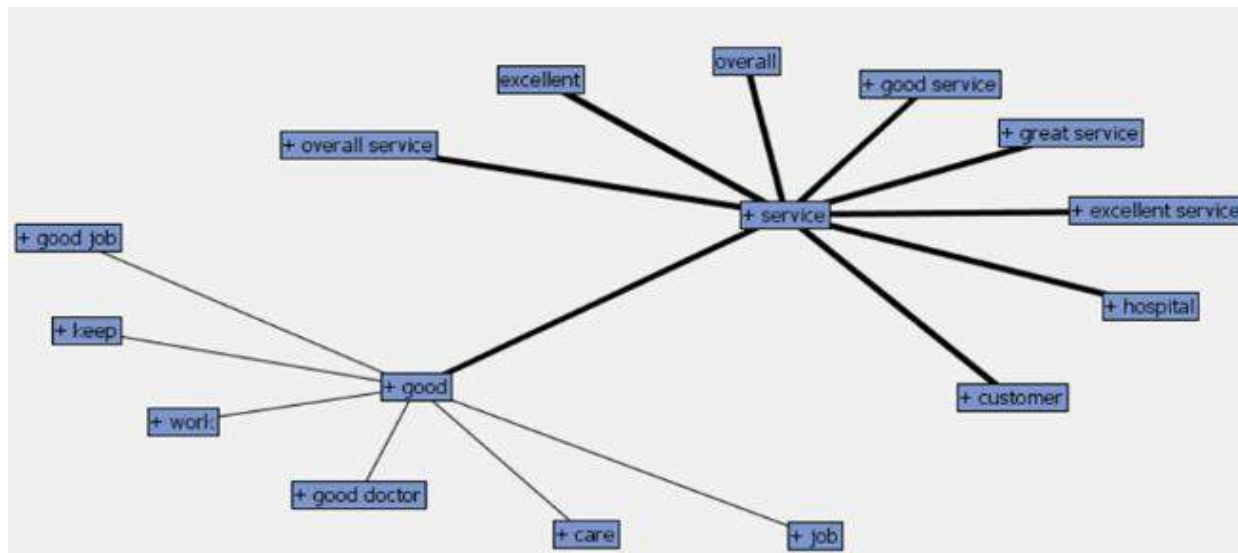
Figure 11: Concept Linking for term “+service”

Figure 11 shows a result of concept linking for term “+service” for the hospital. This concept linking reveals about the term “+service” is highly correlated with other terms such as “excellent”, “+overall service”, “+great service”, “+excellent service” etc. This is because all the lines that link with the term “+service” have all the slightly same darkness and thickness of the line connecting with other terms. Through this concept linking, it is easier to identify what kind of terms which is link to the term “+service”. For instance, this concept linking results shows a few of positive and good term for the hospital according to the hospital services towards the customers.



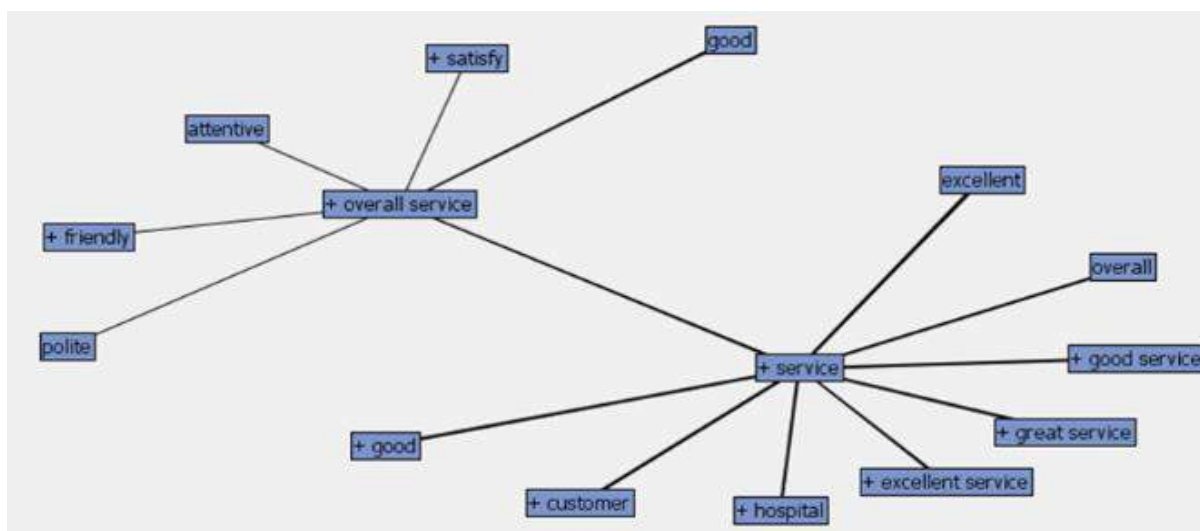
**Figure 12: Concept linking for term “+Service” expand to “+good service”**

For Figure 12, it shows a concept linking results for the term “+service” which is expanded to “+good service”. When the term “+good service” is expanded, it shows another second level of which reveals the other terms which are related with “+good service”. The results from this concept linking map shows there are some subjects which are related to each other. Some other terms were shown when the second level of “+good service” is expanded. Those terms include “polite”, “good hospital”, “kind service”, “professional”, etc. For instance, many of the hospital customers mention that the service is professional, have a kind and polite service. Hence, it leads a reputation that the hospital is a good hospital to other people.



**Figure 13: Concept linking for term “Service” expand to “+good”**

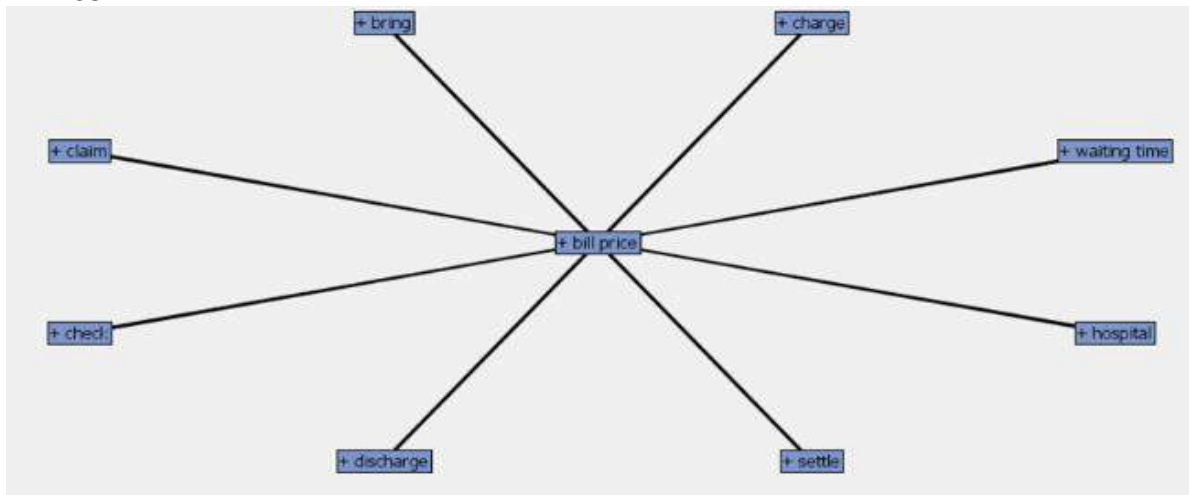
An illustration of concept linking map for term “Service” which is expanded to the term “+good” is shown above in Figure 13. The term “+good” can be expanded into the second level which show more terms that are interrelated with one another. The term “+good” is associated with term “+good job”, “+good doctor”, “+care”, and others as shown in Figure 13. This result prove that the hospital received a lot of feedback about having good doctor working in the hospital and constantly giving care to customers that need help. Nevertheless, the concept linking map also shows that all the doctors are doing good job in order to satisfy patient’s need.



**Figure 14: Concept linking for term “Service” expand to “+overall service”**

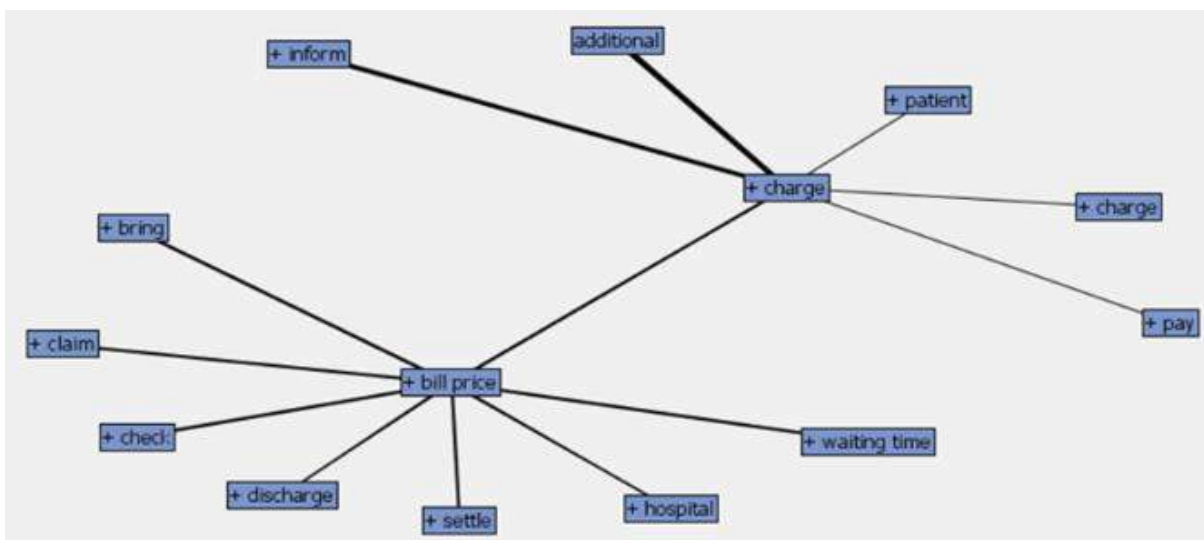
Figure 14 shows a concept linking for term “Service” expand to “+overall service”. When it is expanded to second level for the term “+overall service”, some terms can be seen such as “good”, “polite”, “+friendly”, “attentive” and “+satisfy”. The most highly associated term with “+overall service” is term “good” because the line is thicker than the other terms. From this result, it is believed that The hospital overall service is good to the customers. Not only that, the staff that work in the hospital are polite, attentive and friendly which then satisfy the customers of the hospital.

#### d. Bill Price



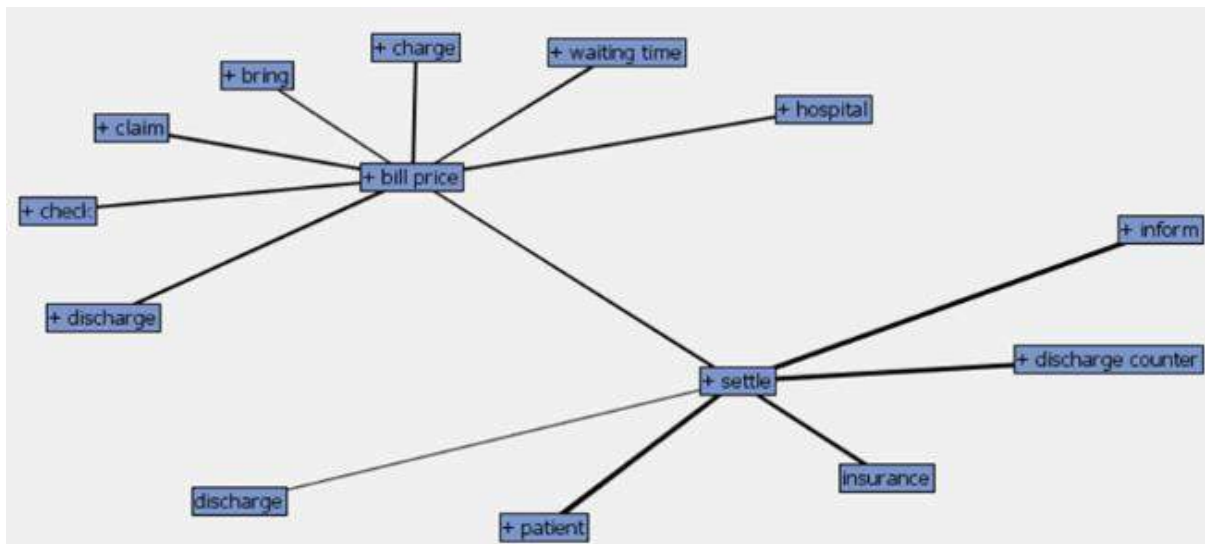
**Figure 15: Concept Linking for term “+bill price”**

Figure 15 shows a concept linking map for term “bill price”. There are some terms which is believed to provide meaningful results as most of the terms have dark lines connecting to term “bill price”. Term “bill price” have high association with the term “+settle”, “+charge”, “hospital” etc. as shown in Figure 15. From this concept linking map, it can identify what kind of feedback the hospital will receive from all the customers that have admitted and how their feedback link to medical fees that incur by the doctors.



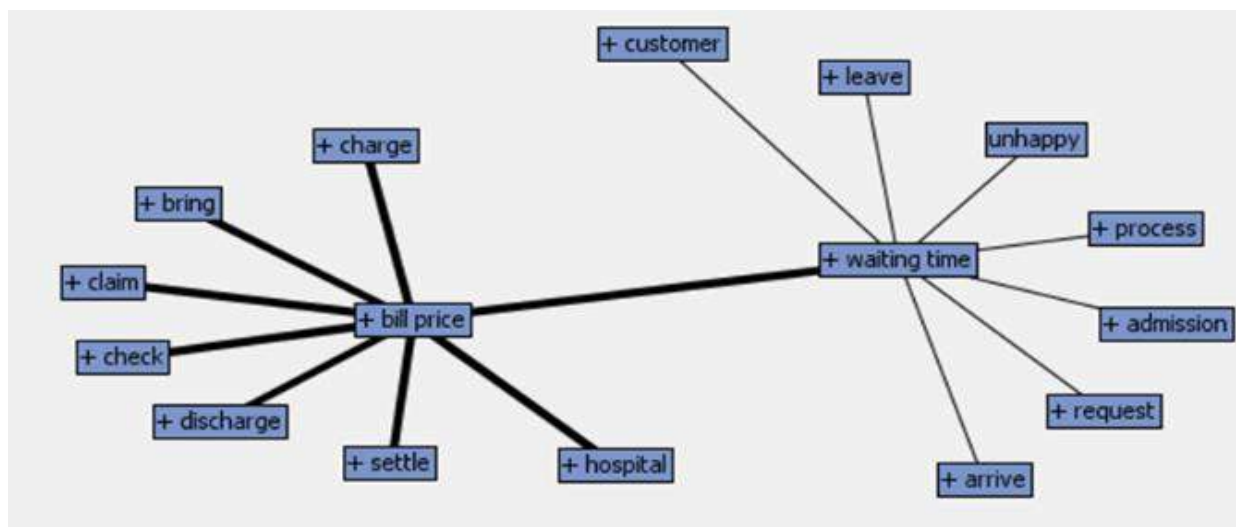
**Figure 16: Concept linking for term “+bill price” expand to “+charge”**

Figure 16 illustrate a concept link for term “+bill price” that expand to the term “+charge”. After expanding it into the second level, terms such as “+inform”, “additional”, “+patient”, “+charge”, and “+pay” appeared in the concept link. There are 2 black lines in the concept linking map which are the term “+inform” and “additional” which lead to high association with the term “+bill price”. This shows that hospital charge additional bill price to the customers that stayed in this hospital. Therefore, the customers of the hospital would like to inform the management that most of the customers have to pay additional charges that was shown in the results of concept linking map.



**Figure 17: Concept linking for term “+bill price” expand to “+settle”**

Figure 17 shows a concept linking for term “bill price” which is expanded to term “settle”. Then, term “settle” is expanded into the second level which consists of some terms such as “discharge”, “+patient”, “insurance”, “+discharge counter”, and also “+inform”. There is strong associated relationship with term “+patient”, “insurance”, “+discharge counter”, and “+inform” because these lines that connect with “+settle” have thick black lines. This concept link map results indicates that customers from the hospital have to settle insurance matters in the discharge counter.



**Figure 18: Concept linking for term “+bill price” expand to “+waiting time”**

Above figure 18 illustrate a result of concept linking map for term “+bill price” that expand to term “waiting time”. When this term “waiting time” is expanded into second level, some interrelated terms are connected with each other. For instance, terms like “+leave”, “unhappy”, “+admission”, “+process” etc. This results indicate that most of the feedback terms mentioned that customers mostly need to wait for the process to pay the bill and takes a very long time. Hence, this lead to unhappy customers that always need to wait to process the medical bill.

Topic ID	Topic	Topic Group
1	+good service,+service,+good,+keep,+excellent service	Service
2	+compliment,+staff,+care,+staff,+nurse	Staff
3	+patient,+admission,+waiting time,+waiting time,+hospital	Waiting Time
4	polite,+care,+care,efficient,helpful	Customer Service
5	excellent,+excellent service,+service,excellent job,excellent care	Excellent
6	+friendly,+staff,+keep,comfortable,+friendly staff	Friendly
7	courteous,polite,attentive,efficient,+friendly	Attitude
8	+doctor,+good doctor,+excellent doctor,excellent,+hospital	Doctor

**Table 1: Text Topic Results**

Text topic discover different topic groups by assigning the terms in each document to its topic of interest. By grouping the terms into different group of topics, results on area of interest can be found from the text documents.

Topic ID 1 group terms that relates with the topic group “service”. Terms that links with good service falls under this topic group. The topic included in this group includes “+good service”, “+service”, “+good”, “+keep” and “+excellent service”.

Topic ID 2 group terms that relates with the topic group “staff”. Terms that links with hospital staff such as nurse falls under this topic group. The topic included in this group includes “+compliment”, “+staff”, “+care”, and “+nurse”.

Topic ID 3 group terms that relates with the topic group “waiting time”. Terms that links with time and waiting time activities fall under this topic group. The topic included in this group includes “+patient”, “+admission”, “+waiting time” and “+hospital”.

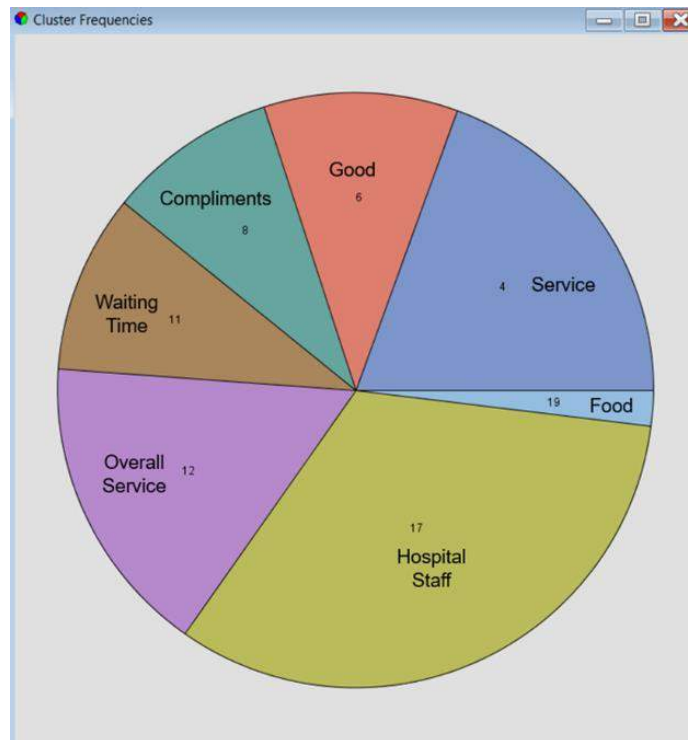
Topic ID 4 group terms that relates with the topic group “customer service”. Terms that relates on how customers thinks about the service received by the hospital falls under this topic group. The topic included in this group includes “polite”, “+care”, “efficient” and “helpful”.

Topic ID 5 group terms that relates with the topic group “Excellent”. Terms that are links with excellent in terms of service or treatment falls under this topic group. The topic included in this group includes “excellent”, “excellent service”, “+service”, “excellent job” and “excellent care”.

Topic ID 6 group terms that relates with the topic group “Friendly”. Terms that links with friendly attitude falls under this topic group. The topic included in this group includes “+friendly”, “+staff”, “+keep”, “comfortable” and “friendly staff”.

Topic ID 7 group terms that relates with the topic group “Attitude”. Terms that links with hospital staff attitude falls under this topic group. The topic included in this group includes “courteous”, “polite”, “attentive”, “efficient” and “+friendly”.

Topic ID 8 group terms that relates with the topic group “Doctors”. Terms that relates with patient’s views on the hospital doctors fall under this topic group. The topic included in this group includes “+doctor”, “good doctor”, “+excellent doctor” “excellent” and “+hospital”.



**Figure 19: Text Cluster Pie Chart - Hierarchical**

The result from the text cluster node using Hierarchical algorithm also shows that there are 7 different cluster group which is the same amount of cluster group using Expectation – maximization algorithm.

Cluster ID	Descriptive Terms	Cluster group	Percentage
4	+friendly helpful excellent polite +service +staff courteous professional caring +'excellent service' +doctor +nice efficient attentive cheerful	Service	19%
6	+care +good care' +good job' +stay +job well +great special +appreciate +baby kindness 'special thanks' +well +good excellent	Good	11%
8	compliments caring polite kind +patience +smile gentle always responsible +appreciate smiling helpful courteous +professional +help	Compliments	9%
11	+ward nurses +staff discharge admitted +stay +care helpful discharged bill +admission wait hours especially time	Waiting Time	10%
12	+good +service +staff overall +'good service' services +doctor satisfied good happy staffs 'good services' 'overall good' doctors +friendly	Overall Service	16%
17	+good nurses +work +good work' +nurse +doctor +complaint 'keep up' +refer medical kindly informed claimed +admission +room	Hospital Staff	33%
19	+food served food breakfast meals cafeteria meal +serve menu chicken fish lunch rice improved +improve	Food	2%

**Table 2: Text Cluster Results**

Table 2 shows the 7-different cluster group using the Hierarchical algorithm for the clustering method. The cluster group consist of service, good, compliments, waiting time, overall service, hospital staff and food. These cluster group is similar using the Hierarchical -maximization clustering algorithm but differs in terms of the clusters percentage.

Cluster ID 4 represent the cluster group for "service". The terms that falls in this cluster includes "+friendly", "helpful", "excellent", "polite", "+service", "staff", "courteous", "professional" etc. These terms mostly describe the service received from the hospital such as nurses and doctors. This cluster group contributes 19% of the overall terms in the text documents.

Cluster ID 6 represent the cluster group for "good". The terms that falls in this cluster includes "care", "+good care", "+good job", "+stay", "job", "well", "excellent" etc. These terms are the words patients use when giving good feedback to the hospital. This cluster group contributes 11% if the overall terms in the text documents.

Cluster ID 8 represent the cluster group for "compliments". The terms that falls in this cluster includes "compliments", "caring", "polite", "kind", "+patience", "+smile" etc. These terms are words that patients use in the feedback to give compliments to the hospital staff. This cluster group contributes 9% of the overall terms in the text documents.

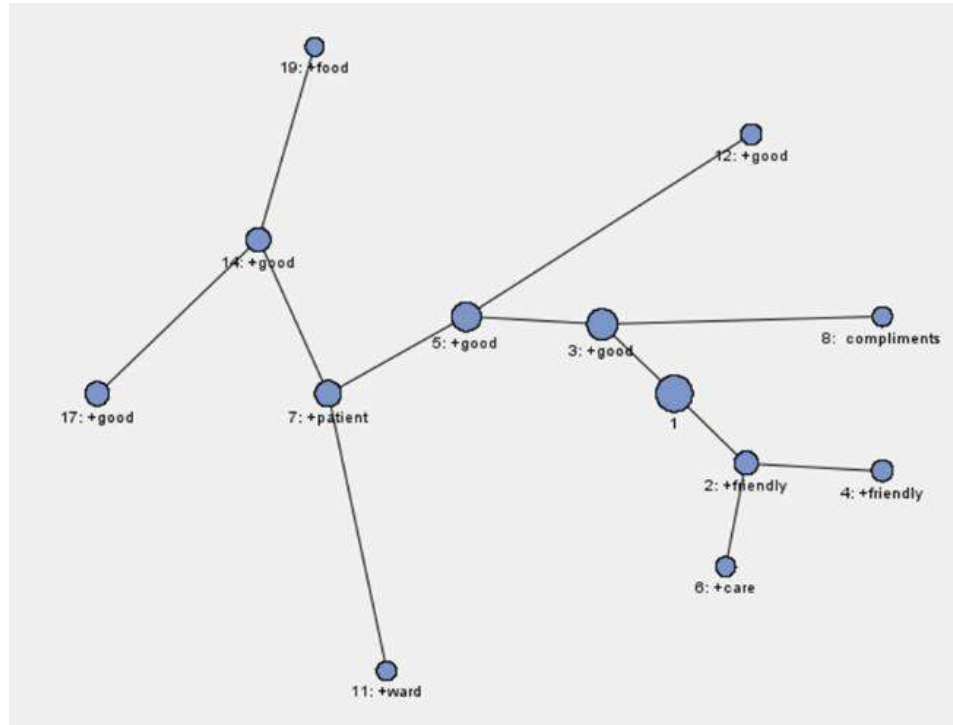
Cluster ID 11 represent the cluster group for "waiting time". The terms that fall in this cluster includes "+ward", "discharge", "admitted", "admission", "wait", "hours", "time" etc. These terms relate to the waiting time patients often faced by the hospital and activities that relates with waiting time such as discharge and admission process. This cluster group contributes 10% of the overall terms in the text documents.

Cluster ID 12 represent the cluster group for "overall service". The terms that fall in this cluster includes "good", "+service", "+staff", "overall", "+good service", "services", "overall good", "doctors" etc. These terms are related with the positive terms on overall service received from hospital staff. This cluster group contributes 16% of the overall terms in the text documents.

Cluster ID 17 represent the cluster group for "hospital staff". The terms that fall in this cluster includes "+good", "nurses", "+work", "+good work", "nurse", "doctor", "complaint" etc. These terms are categorise under the cluster group for hospital staff as the terms mostly relates with patient's feedback on the hospital staff in terms of compliments or complaints. This cluster group is the largest and contributes 33% of the overall terms in the text documents.

Cluster ID 19 represent the cluster group for "Food". The terms that falls in this cluster includes "+food" "served", "breakfast", "meals", "cafeteria", "menu", "chicken", "rice" etc. These terms are all relates with patient's views on the food serve in the hospital. This cluster group is the smallest and contributes only 2% of the overall terms in the text documents.





**Figure 20: Cluster Hierarchy**

Figure 20 shows a graphical representation for the Hierarchy Cluster. The size of the cluster is represented by the size of the circle. The clusters are link to each other to shows their association with another cluster group. Based on Figure 20, the cluster begins splitting from Cluster ID 1 therefore Cluster ID 1 is the parent cluster for Cluster ID 2 and Cluster ID 3. The cluster is then further split into smaller group of clusters and the size of the cluster gets smaller after each split. Therefore, figure 517 shows the links between the child cluster and parent cluster. For example, Cluster ID 14 is the parent cluster for Cluster ID 17 and Cluster ID 19 as these two cluster was split from Cluster ID 14 which splits topics on hospital staff good work and food.

**Sentiment Analysis Results**

Topic	Category	Term Cutoff	Document Cutoff	Number of Terms	# Docs
Negative Tone	User	0.001	0.001	355	1203
Positive Tone	User	0.001	0.001	355	7422

**Figure 21: Number of Documents for Positive and Negative Tone**

Figure 21 shows the term cut off, document cut off, number of terms and number of documents for negative tone and positive tone. The values for term cut off and document cut off is the same for both negative tone and positive tone topics. Term cut off is used to determine whether the term is eligible to fall under the topic group where else document cut off is used to determine whether the document is eligible to fall under the topic group. Topic weight greater than 0.001 indicates that terms and documents are eligible to fall under the topic for negative tone and positive tone.

Results shows that the number of terms for both positive tone and negative tone are the same amount which are 355 terms where else number of documents with positive tone is more compared to negative tone. The number of document for positive tone is 7422 where else the number of documents of negative tone is 1203. The number of documents with positive tone is

72% more than negative tone indicating that majority of the feedbacks received by this hospital are positive feedbacks.

## DISCUSSION

The goal of this research study is to identify factors affecting patient's perception on the hospital quality of food, waiting time, services and bill price using Text mining and Sentiment analysis. Besides that, the results from Sentiment Analysis is discussed because feedbacks with positive tone are consider the areas of strength for the hospital where else negative tone feedbacks are the areas where improvement can be made.

### Quality of Food

Hospital are often taking note on the quality of food serve to patients as the food serve affects patient's recovery and satisfaction. A research conducted by Dube, et al. in 1994, discovered that quality of food is the first determinant on patient's satisfaction during their stay in the hospital. The results on "Quality of Food" shows that there are many factors that links with this topic and many areas for improvement. Based on the concept linking results, there are a few factors that plays an important role on the quality of food serve. The factors that affects the hospital quality of food includes cafeteria, taste, food serving timeliness, food temperature, food variety and portion size. The results from the concept linking identified several issues on the quality of food that The hospital may take notice. The issues on Quality of Food identified are:

1. Food Serve is tasteless and not appetizing.
2. Not enough variety of food.
3. Unhealthy food serves in the Hospital cafeteria.
4. Food serve is late and cold.
5. Portion size of the food serve is small.

### ***Recommendation to Solve the issues on Quality of Food***

First of all, as patient's food should not include too much salt as they are bad for health, hospital chef can opt to using herbs to enhance the flavour of the food which also help to increase patient's appetite. Next, the issue on not enough variety of food can be solved by introducing new dishes into the menu. For issue on unhealthy food in the hospital cafeteria can be solved by determining what are the appropriate food to serve. Foods that are oily, salty and high in fat should not be sold in the cafeteria and also not serve to patient's. By providing healthy food options in the hospital cafeteria also allows hospital staffs and visitors to make healthier food options and maintain good health. Other than that, food serve late or cold is the responsibility of the kitchen staffs and serving staff. The kitchen staffs should always estimate sufficient time to prepare the meals and serving staff should serve the meals on-time. Lastly, different patients have different preference on the food portion size. Therefore, the hospital may solve this issue by giving patients the option to choose small or large food serving size. By doing so, the hospital can also avoid food wastage as some patients are small eater and prefer smaller serving size of food.

Although there are bad reviews on the quality of food, sentiment analysis result also shows that there are still some positive feedbacks on the quality of food by some patients. Based on the positive reviews on the quality of food, it shows that patients are satisfied with the quality of food whenever food is serve on time, tasty, healthy and good. These factors contribute to patient's treatment experience with the hospital. Therefore, the hospital should be consistent

on the quality of the hospital food and take into consideration of patient's reviews on the food quality.

### **Waiting Time**

In a hospital, it is common to hear that patients have to wait for hours from registration up to discharge as there are many stages required to be complete. Patient's satisfaction towards the hospital drops as waiting time increases. In addition, hospitals are starting to measure patient's access to healthcare treatment based on waiting time statistic (Godden and Pollock, 2009). To provide a better care for patients, the hospital should understand to what extent is patient's satisfaction on the waiting time while receiving services and the factors influencing waiting time. The results on "Waiting Time" shows that there are many factors that links with waiting time and many areas for improvement. Based on the concept linking results, there are a few factors that affects patient's satisfaction that relates with waiting time. The factors that relates with patient's waiting time includes admission process, discharge counter, and payment. The results from the concept linking identified several issues on the long waiting time that The hospital may take notice. The issues on Long Waiting Time identified are:

1. Long waiting time for Admission Process.
2. Long waiting time for Discharge Process.
3. Long waiting time for payment and Insurance

### ***Recommendation to solve the issues on Long Waiting Time***

Some of the causes of long admission process is due to no empty room or beds which leads to long waiting time for patients. The hospital can allocate patient's that are very ill (example: dizzy, elderly etc.) to a temporary room to rest while waiting for the admission process. By doing so, the impact on patient dissatisfaction towards admission process can be lessen. Other than that, if the hospital is having shortage of staff at the admission counter, it is encourage to hire more staff members to help out with the admission process. Next, patients often complaints that admission process is long but discharge process is even longer. Discharge process are normally longer due to the long procedure that need to be taken before discharge. To prevent patient's from misunderstanding the hospital having bad management, the hospital should explain to the patients the process from admission to discharge so that the patients understand why discharge process are long. Besides explaining to patients on the process on discharge, hospital should also look at each discharge process to identify areas which can be accelerate to shorten the discharge waiting time. Lastly, long waiting time at the payment counter and insurance can be shorten by opening more payment counter and hiring more staff to help out.

Sentiment analysis results also shows that most of the feedbacks by patients on waiting time are negative reviews. However, there are only a few positive feedbacks on waiting time. This indicates that, waiting time is a current issue for hospital and an area that should be improve. Therefore, hospital should look into the issues on waiting time at admission process, discharge counter and payment counter by determining the cause of long waiting time at these processes in order to reduce patient's dissatisfaction on waiting time.

### **Services**

For this current study, one of the major category about patient's perceptions towards this hospital is the services. Services is very important to maintain the reputation of the hospital. A study shows that patient's satisfaction can give potential to a business to increase customer base and also business reputation. To succeed patient's satisfaction is vital to identify and to fulfilled customer's needs and able to satisfy customer. Any industries that rapidly understand and satisfy customer's requirements will be able to make more profit than industries who have

fail to identify and satisfy customers (Dominici & Guzzo 2010). If most of the patients complain that the hospital provides a bad service, other people would not choose the hospital as a best choice to go when they need to seek for medication. The results on the concept linking shows mostly good perception that The hospital provide to all customers. The factors that affect patient's perception includes:

1. Attentiveness of Hospital Staff
2. Great manners from hospital staff and professional doctors in the hospital

### ***Recommendations for better services***

Although concept linking shows the positive term for services, but sentiment analysis can identify the negative tone that were given by hospital customers in the feedback. For instance, certain feedbacks from customers are about the attitude of hospital staff is bad and services given by the staff is bad. To improve this, hospital staff should be given more training and meeting about how to provide great service to customers before working and serving hospital customers. This can remind all the hospital staff that services are very important to keep up the good reputation of the hospital. Since most of the feedbacks about this hospital is good, but there will also need to retain its best services to its customers. Therefore, Management can outline responsibilities and role to all hospital staff. Hospital staff are required to understand all its roles and responsibilities in terms of what kind of proper services to provide for customers. Operational standards can be developed in order to achieve the goal and objective by giving excellent services to its customers. Therefore, Management of this hospital can start developing the SMART acronym which are the specific, measureable, attainable, realistic, and timely. SMART acronym enable hospital staff to have a clear specific goal and achieved the goal measuring progress towards the achievement. It is also important to have commitment to provide better services in hospital staff behaviour. By having commitment in self-behaviour, individual will feel more responsibilities when providing services to the hospital customers.

### **Bill Price**

The part of pricing on customer satisfaction has been ignored by all industries when it comes to business. A study done by American Customer Satisfaction Index (ACI) had proved that customer satisfaction decreased dramatically as the room rates have increased extensively over the past few years (Mattila & Neill n.d.). Customer at every different business use price as a quality sign where the expectation will drive customer satisfaction. In other words, customer that pay above high price should expect the perceived of service quality is better than normal. With great performance and acceptable range from the industry, satisfaction from customer is developed. Also, the health care insurance has been increasing the cost and needs to be addressed. Many people are not having health care insurance due to the expensive cost and those who have insurance has claim that private hospital has increase in claim of cost in medical fees (McLouglin 2013.). Financial is the second major category that this hospital needs to be concerned about. In this study, financial will be defined as "Bill Price". When it comes to the bill price, customers will be aware because customers are paying for the services, medication and facilities in this hospital. From the concept linking results, there are a few perceptions that customer think about this hospital bill price. The results on the concept linking shows mostly bad perception on bill price which provide by this hospital. Sentiment analysis for "Bill Price" was studied to identify the positive and negative term in the comment. However, there were no positive term in the comment for "Bill Price". Therefore, it need to be addressed. The issues identified on Bill Price includes:

1. Waiting time for billing process
2. Settle Insurance Matter

### 3. Customers pay additional for extra charge

#### ***Recommendations to solve issues on "Bill Price"***

From the text mining analysis results, this hospital management can see what type of feedback with regards to bill pricing. For instance, management would like to look up to the billing process because patient's complaint about the process being very slow. Therefore, the management should encourage more staff to help out on billing process to avoid patients waiting too long. As for small amount of payment, the hospital can create 'Self payment machine' to pay small amount of bill such as paying for medicine from the pharmacy. Also, hospital management must take note on the insurance process and explain it to the hospital staff that is handling the insurance process. All hospital staff have to make sure to inform the patients about the insurance policy before any misunderstanding happens and lead to unhappiness when proceeding the bill. By solving this issue, hospital Management can consider having a specific counter that only handle insurance matters. This help to reduce miscommunication and misunderstanding about information when all counters are separated into different categories such as payment counter, insurance counter, customer service counter etc. This hospital management have to manage the bill price matter in order to decrease the negative feedback received from customers.

### **CONCLUSION**

Overall, Text Mining is a useful tool for understanding huge amount of text documents and discovering hidden knowledge. By using text mining and sentiment analysis to understand customer's perspective towards the services provided by the hospital, one can understand their customer better and make improvement based on the areas that makes customer unhappy. Other than that, the management can understand the huge amount of text data within a short time-frame and preventing from overlooking key issues compared to manual reading customer feedbacks.

### **References**

- Barnard, A. (2002). Feedback Seeking in Customer Service Relationships. [online] Available at: <http://citeseerx.ist.psu.edu/viewdoc/download>
- Barlow, J. & Moller, C. (1996) A Complaint is a Gift: Using Customer Feedback as a Strategic Tool. [ebook] San Francisco: Berrett-Koehler Publishers. Available from: <https://books.google.com.my/>
- Bholat, D., Hansen, S., Santos, P., & Bailey, C.S. (2015) Text Mining for Central Banks. Centre for Central Banking Studies.
- Blake, C. (2011) Text Mining. Annual Review of Information Science and Technology.
- Chakraborty, G., Pagolu, M. & Garla, S. (2013) Text Mining Analysis: Practical Methods, Examples, and Case Studies using SAS. Cary, NC: SAS Institute Inc.
- Chinsha, T.C. & Joseph, S. (2014) Aspect Based Opinion Mining from Restaurant Reviews. International Journal of Computer Applications.
- Dalmolen, S. (2010). Defining patterns in unstructured manifests in a volatile cross-domain environment. [online] Available from: <http://www.cs.rug.nl/~aiellom/tesi/dalmolen>.
- Dominici & Guzzo (2010), Customer Satisfaction in the Hotel Industry: A Case Study from Sicily, International Journal of Marketing Studies, 2(2).
- Fan, W., Wallace, L., Rich, S. & Zhang, Z. (2006) Tapping the Power of Text Mining. Communication of the Ach. [Online]. 49(9), 76-82. Available from: <http://dl.acm.org/citation.cfm?id=1151032>
- Godden, S. & Pollock, A.M. (2009) Waiting List and Waiting Time Statistics in Britain: A Critical Review. Public Health [Online] 123, 47- 51. Available from: [www.elsevierhealth.com/journals/pubh](http://www.elsevierhealth.com/journals/pubh).
- Gupta, V. and Lehal, G. (2009). A Survey of Text Mining Techniques and Applications. JOURNAL OF EMERGING TECHNOLOGIES IN WEB INTELLIGENCE, 1(1).

- Hendry, S., & Madeley, A. (2010) Text Mining and the Information Content of Bank of Canada Communications. Bank of Canada. [Online] Available from: <https://core.ac.uk/download/pdf/6699593.pdf>
- Hotho, A., Nurnberger, A. and Paab, G. (2005). A Brief Survey of Text Mining. [online] Available at: <http://www.kde.cs.uni-kassel.de/hotho/pub/2005/hotho05TextMining.pdf>
- Lau, K., Lee, K. and Ying, H. (2005). Text Mining for the Hotel Industry. 46(3).
- Lavrakas, P.J (2008). Encyclopedia of Survey Research Methods. 2nd Edition. California. Thousand Oaks, Calif.
- Liu, B., (2012). Sentiment Analysis and Opinion Mining. Sentiment Analysis and Opinion.
- Mattila & Neill n.d, Relationship between Hotel Room Pricing, Occupancy, And Guest Satisfaction: A Longitudinal Case of a Midscale Hotel in United States.
- McLoughlin. P (2013), Review of Measure to Reduce Costs in the Private Health Insurance Market.
- Oja, P. (2010). Significance of Customer Feedback An Analysis Of Customer Feedback Data In A University Hospital Laboratory. [online] Available at: <http://jultika.oulu.fi/files/isbn9789514262739.pdf>.
- Vinodhini, G., Professor, A. & Chandrasekaran, R., (2012). International Journal of Advanced Research in Computer Science and Software Engineering Sentiment Analysis and Opinion Mining: A Survey., 2(6).