CHAPTER 1

INTRODUCTION

1.1. Background

Technology advancement has brought us upon a new dawn of Internet, which is Internet of things era. Internet of things is an evolvement from internet, that enables aggregation of objects that inter-connect each other, gaining intelligence, communicate, information sharing to other objects and people (as user) anywhere, anytime, anyplace, anyone, and anything (Bhuvaneswari and Porkodi 2014). Where internet plays the role as the media of usage in internet of things, it is also having an important role of people's daily life nowadays. The data of internet user (shown in Figure 1.1) has grown significantly over the years describes how big is internet to people's life. Interconnecting function of internet also connects people to people through a media that is called as social media. Social media usage is the most-used internet branch of service in Indonesia.

The use of social media an undisputable fact that is becoming the major social lifestyle among social interaction of people, the development of its features and function attracts the interest of people as one of the primary needs of them. Things that are developed inside the interaction of social media also attracts many engineers around the globe to find such a fruitful insight of knowledge to be used as a research. For instance, in social media, people can just post what are they thinking about of someone or products through a written post that is formed from texts. The amount of data that is generated on each day has reached approximately 2.5 quintillion bytes (Marr 2018). Another thing that is also becomes one of the habits in this era is online reviews references. People share their experience and thoughts through social media to give a detailed experience towards product/services that is also crucial for a company/business organization to understand

and getting into those customer's voice in form of complaints or also opinions. Known as product reviews, it can be a good feedback for the company, or it can also be a bad omen for them (Bailey 2005). This phenomenon pushes company to evolve and adapt to gain competitive advantage in these market dynamics. The exploration of competitive advantage meaning is to search the distinct values from all the available competitors, yet it still needs some factors to maintain it sustainability, one of the factor is Market Dynamics, meaning to get the idea of the rate of change of customer requirements (Mcgee 2015). These adaptations are aimed to establish themselves to this phenomenon to gain the most-probable competitive advantages for them.

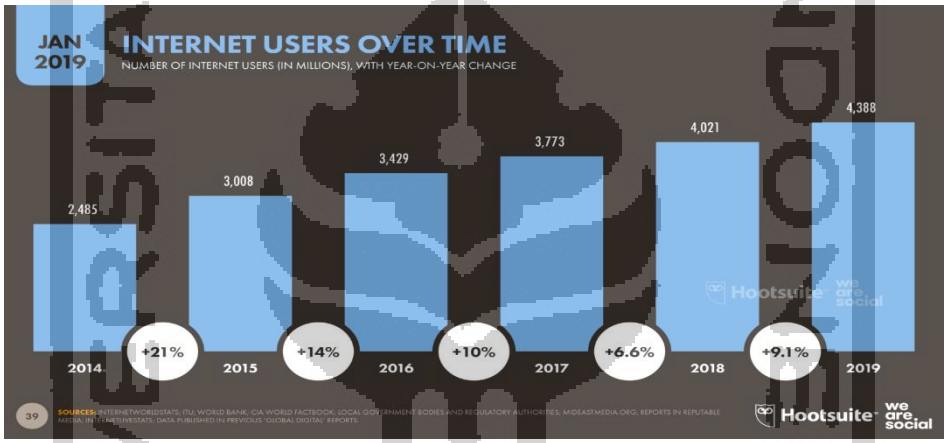

Figure 1. 1. Internet Usage over the years

Source: https://datareportal.com/reports/digital-2019-global-digital-overview

One of the highly reviewed all over the social media platforms is Smartphone. Since smartphone is a product that everyone poses in their pocket, it become a hot topic amongst reviewer and review seeker. One of the notable smartphone manufacturers is Samsung, which has the market share worldwide in 2018 18,7 percent, and Samsung has shipped more than 292 million smartphones all over the world (taken from https://www.statista.com/statistics/271496/global-market-share-held-by-smartphone-vendors-since-4th-quarter-2009/). One of notable products from Samsung is Samsung Galaxy S Series. Recently, it has developed S Series smartphones like S8, S9, and S10 that have significance pattern of sales in terms of these products' sales. According to research firm Canalys, Galaxy S9 sales for the first four weeks after being launched closely matched to the Galaxy S8. But the Galaxy S7 sold one million more units in the first month after

being launched, than the Galaxy S9. And according to (https://en.yna.co.kr/view/AEN20190702008800320) The first quarter sales of Samsung Electronics Co.'s Galaxy S10 series smartphones were 12 percent higher than sales of the previous model. According to industry tracker Counterpoint Research, 16 million units of the Galaxy S10 were sold between March and May, which was more than sales of the Galaxy S9 series in the same period last year. This shows a market slow sales for Samsung Galaxy S9 Series. This phenomenon shows the opportunity for a product evaluation and improvement. The current condition of data availability on the internet, a research on deploying machine learning and combined with Natural Language Processing might be beneficial for this case. It can deploy an exact answer towards the problem in a faster way and precise supporting data for a decision making to react to the market's response to a product released to the market and fill up the void that is found from the market response. Somehow, a fast way to react and adapt to the market demand is an important part of manufacturers readiness towards the era. If these conditions are not handled properly it might give a bad omen to the manufacturers.

Among the popular websites for e-commerce service and product reviews references is Amazon.com. Over the years Amazon has become the popular e-commerce platform. One of their customer relation managements is to provide a platform for a product reviews that until now, various reviews and long story reviews about various products in Amazon are available. This enhances the chance of Amazon to get in touch to the customers as well to provide customers a relationship with the producer's engineers that generates profitable text data about product reviews. (Al Imran 2014). These data then can be taken using the techniques of web-scrapping and selected a certain product to be analysed on terms of the available text data about the product's reviews. Text data itself is an example of better sources in terms of the existence of a data that can be easily taken or scrapped while it is in form of unstructured information, which is also simple and can be generated in most of scenarios and available around us in social media. (Allahyari, et al. 2017). While text data currently trigger several projects on algorithm usage in social media, such that text mining for opinion and knowledge extraction on social media.

Text mining has gained a significant number of interests in recent years. Text mining is a process of patterns extractions to gain some insights from text data resources. It is a branch from information retrieval, data mining, machine learning, statistics and computational linguistics. Some text mining techniques such as, summarization, classification and clustering, etc, can be applied to extract knowledge, dealing with natural language. Some application fields of text mining are, industrial research, web development, customer relationship management, product suggestion analysis etc (Talib, et al. 2016). One of the interesting application fields of text mining is in product analysis and can be referred to product design as it is beneficial to gain customers trust in a business. In the early process of product design, one of the most important process is accosting and understanding the customer's voice. These customer's need or voice is used to determine and categorize the feature of upcoming products. A notable method of identifying such words in customer reviews and transform it to the product's features design characteristics. A method called Kansei Engineering, which conceptualized a way of an interaction between customers and product designer through the media of words of feedback (Kansei Words) in terms of what the customers felt about the experience of using a certain product. (Schutte, et al. 2004).

The interactive feedback concept in finding Kansei words in a product can be utilized with the use of text mining. It can be done easier in analysing such a huge amount of data and extracting opinion from the user experience that can help both sides in getting into their goals of the most probable beneficial outcomes (Mali, et al. 2016). Since opinion mining deals with a lot of unstructured data, a machine learning algorithm is applied to get the beneficial result of this analysis (Ireland and Liu 2018). One of algorithms that can be applied is Latent Dirichlet Allocation (LDA), which is one of the techniques of probabilistic topic modelling. It is categorized in unsupervised technique to extract thematic information (topics) of set of documents. The idea is to set documents as random mixtures of latent topics, where each topic is a probability distribution over words (Allahyari, et al. 2017). Thus, it makes the feature-based opinion mining possible, in applying such topic modelling techniques to gain further insight of knowledge of products' feature.

This research applies the method of text mining in scope of Kansei engineering concept by using the approach of feature-based opinion mining from a product review. The text data are scrapped from amazon customer review page, containing the reviews about Samsung Galaxy S9. An additional analysis on other smartphone reviews is used to filter the result of the Kansei words as a benchmarking process in the product improvement guideline. The idea of comparing other smartphone reviews is to get a better insight of what people perceives from other products features in this case acting as a direct competitor of Samsung S9 smartphone. This research applies a machine learning algorithm to find the most reviewed features of this phone and finding the related Kansei words from these features by looking up the most probable correlated adjectives words to the reviewed features using unsupervised machine learning algorithm named Latent Dirichlet Allocation (LDA). This algorithm implemented as a topic modelling method to get the features that are reviewed in amazon. A conceptual design approach by combining the concept of Kansei engineering and topic modelling is expected to find feedbacks on the reviews in a faster way.

1.2. Problem Formulation

Based on problem in background, the formulations of the problem to be addressed in this study are:

1. How to implement LDA topic modelling and Kansei engineering on Samsung S9 customer reviews on Amazon website?
2. How to develop a product improvement guideline from the result of LDA topic modelling implementation?

1.3. Research Objectives

The objectives of this research are:

1. To implement LDA topic modelling on Kansei Engineering for Samsung S9 smartphone product,
2. To develop a product improvement guideline from the LDA topic modelling implementation result.

1.4. Benefits of Research

The expected benefits of this research to some other involved parties are:

1. For author

   This study is aimed to fulfil the obligation of author as undergraduate student to accomplish and obtain Industrial Egineering Degree at the International Program, Department of Industrial Engineering, Universitas Islam Indonesia.

2. For manufacturers

   The result of this research can be used as a benchmarking data for further development on the product produced by the manufacturers.

3. For other parties

   The result of this research can be developed in the future since this research has its own limitation and this research can be beneficial as an additional reference for further development in terms of mining implementation and product designing.

1.5. Limitations

The author has set some limitations to avoid widespread discussion and understanding about this concept of research, which are:

1. The data used in this research, is the customer review on a product namely Samsung Galaxy S9 smartphone, taken from amazon customer reviews website,
2. The other smartphone brand reviews dataset is taken from Kaggle,
3. The data contain customer reviews limited to review in English only that will be analysed in this research,