Employee Sentiment Analysis Using Naive Bayes Classifier

Dessy Monica Ginting[1], Andry Alamsyah[2]

[1,2]Telkom University, Faculty of Economics and Business,
Bandung, Indonesia
dessy.ginting[at]gmail.com, andrya[at]telkomuniversity.ac.id

Abstract: *People today tend to use social media as their platform to share reviews or opinions, including about their work. Therefore, public opinion and sentiments are some of the most important factors for a company to take advantage of. It can be used as a source to be analyzed and extracted to valuable insights. When used appropriately, employee sentiment analysis can provide more effective tools for determining key factors such as job satisfaction than internal surveys or other conventional methods. By getting a clearer picture of employee sentiment, companies can identify areas where employees are dissatisfied and devise strategies to increase engagement and, in turn, improve employee productivity and retention.*

Keywords: Text Mining, Sentiment Analysis, Data Classification, Naive Bayes Classifier.

1. Introduction

In the current era of globalization where technological developments and information dissemination are increasing, every organization or company is required to do its best and be able to survive all challenges. To be able to maintain its existence, the company must have the right strategy to achieve its goal. The success of a company in achieving its goals cannot be separated from the existence of human resources that contribute to meet the needs of the company [1].

Employee retention, especially in the technology industry, has been a major concern for organizations today [2]. The employee turnover rate among Fortune 500 companies in the IT industry is the highest among all industries surveyed [2]. Some fairly reputable technology companies including Google and Amazon are at the top of the list [2]. One way for companies to retain their workforce is by prioritizing what matters most to their workforce, ensuring that employees are motivated and enthusiastic to work every day and better understand employee sentiments [3].

Employee sentiment analysis provides a more effective tool than internal surveys or other conventional methods [4]. Unfortunately, many organizations still use annual employee surveys to assess satisfaction, engagement, and culture in the workplace. Though comprehensive, this type of survey spends a lot of time on Human Resources and quickly becomes obsolete because as the Human Resources team will examine the data and turn the feedback into valuable insights, new problems will arise.

The penetration of global internet users around the world is amounted at 3.773 billion [5]. This means that active internet users in the world reach 50% from the 7.476 billion of total population [5]. The phenomenon of internet development makes the behavior of people express opinions also change. People today tend to use social media as their platform to share reviews, opinions, and opinions, including about their work [6]. Therefore, public opinion and sentiment are some of the most important factors for a company to take an advantage of. This can be used as a source to be analyzed and extracted to valuable insights. If used appropriately, employee sentiment analysis can provide more effective tools for determining key factors such as job satisfaction than internal surveys or other conventional methods.

As companies with the shortest tenure, the author picks Google and Amazon as a case study. The data of Google and Amazon will be crawled through the reviews contained in job site called Indeed (www.indeed.com) and it will be analyzed by using Naïve Bayes Classifier. It is one of algorithms which is well-known and used to mining the opinion and offering ease of use as well as quick processing time [7].

2. Literature Review

2.1 Sentiment Analysis

Sentiments analysis is in charge of detecting, extracting and/or summarizing opinion, polarity and/or emotion, usually based on the presence of sentiment features [8]. In most cases, sentiment analysis is treated as a classification problem, with categories of opinions and facts, positive and negative, or levels in the emotion category [8].

2.2 Naive Bayes Classifier

Naïve Bayes Classifier is one of the simplest and most commonly used classification methods in research. This classification calculates the category probabilities based on the number of word distributions in the dataset. This classification uses Bayes' theorem to predict the likelihood of features that have been labeled according to predetermined categories [9].

3. Methodology

We divided the study into 4 stages, starting with data collection, data classification, data preprocessing, and evaluation.

Figure 1: Research Stages

Data collection describes the process of data crawling. Preprocessing data describes the process of compiling the data. Data classification describes the process of categorizing data by dimension and sentiment. The evaluation describes the classification results by measuring performance.

3.1 Data Collection

Data collection from the www.indeed.com job site is obtained by the crawling process with a 1-year data period from January 2017 - December 2017. The total number of data collected is 7,070 with a proportion of 377 Google reviews and 6,693 Amazon reviews. To clear the data, all text will go through the preprocessing stage first. Some text that does not fit into any category will be deleted.

3.2 Data Classification

Prior to the establishment of training data and test data, irrelevant data such as sentiments or opinions that are not suitable for any category will be deleted. The results of review selection considered relevant will be classified by its category and sentiment as written in the following table:

Table 1: Category Classification

Text	Category
The culture is intense	Culture
Ample opportunity for advancement	Career
Management was fair and respectable	Management
Free food, snacks, incredible benefits	Benefits
Work/life balance was pretty flexible	Work/Life Balance

Table 2: Sentiment Classification

Text	Sentiment
High salary and great work benefits	Positive
No career advancement opportunities	Negative

3.3 Data Preprocessing

After categorizing the review, this research will continue with the preprocessing phase of data to prepare the input data to be analyzed afterwards. The preprocessing activities are done through 3 stages, including tokenization, filtering and stemming. Tokenize serves to remove all non-letter characters and to divide a text into small, meaningful elements such as sentences and words. Filter stopwords serve to eliminate unnecessary words. Stem is a process to convert the word to its original word.

After going through the data filtering process stage, the amount of data that has been filtered as much as 2094 data with total net data of 403 reviews for Google and 4573

reviews for Amazon.

3.4 Evaluation

The classification of texts is done using the Naive Bayes method. The value of performance evaluation must be taken into account in this classification method so that the text in this study can be trusted at the time of its implantation. Measurement can be done with the performance matrix method as follows:

Table 3: Performance Matrix

		Predicted Label	
		Positive	Negative
Actual Label	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Positive (TP)

1) TP (True Positive): class classification that is predicted to be positive and predicted by a positive classification system.

2) TN (True Negative): class classification that is predicted to be negative and predicted by a negative classification system.

3) FP (False Positive): class classification is negative but predicted by a positive class classification system.

4) FN (False Negative: class classification is positive but predicted by a negative class classification system.

The mathematical equations and explanations of the above parameters are described below:

1) Recall: this measurement is used to identify how many methods can be remembered by correct text groupings.
$$\text{Recall} = \frac{TP}{TP+FN} \qquad (1)$$

2) Precision: this measurement ratio is used to see how precisely the textual classification method in predicting text.
$$\text{Precision} = \frac{TP}{TP+FP} \qquad (2)$$

3) Accuracy: this measurement is used to test the accuracy of text classification methods that have been made.
$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN=FN} \qquad (3)$$

4) F-Measure: this measurement is a combination of precision and recall measurements.
$$\text{F-Measure} = 2x \frac{Precision \; x \; Recall}{Precision + Recall} \qquad (4)$$

5) Kappa: this measurement is used to measure the agreement between each of the anotator pairs in which the annotator is used to assess the making of a text classification method.
$$K = \frac{P(A)-P(E)}{1-P(E)} \qquad (5)$$

The table below shows the interpretation of kappa values in the classification method which shows the quality and performance of text classification methods that have been constructed by the researchers. If the kappa value is more than 0.75 then the text classification method can be concluded very well [10].

Table 4: Kappa Score Interpretation

Kappa Score	Description
<0	Less than Chance Agreement
0.01-0.20	Slight Agreement
0.21-040	Fair Agreement
0.41-0.60	Moderate Agreement
0.61-0.80	Substantial Agreement
0.81-0.99	Almost Perfect Agreement

4. Results and Analysis

Accuracy, Kappa, Precision, Recall and F-Measure are methods used for evaluating the performance of sentiment analysis. Performance evaluation results of category and sentiment classification model using Naive Bayes Classifier are as follow:

Table 5: Performance Evaluation Results of Category Classification

Company	Accuracy	Kappa	Recall	Precission	F-Measure
Google	91.88%	0.890	86.23%	85.62%	85.92%
Amazon	87.33%	0.822	85.54%	81.75%	85.00%

Table 5 above shows the results of category classification using Naive Bayes Classifier. For Google, the accuracy shows about 91.88% with kappa score of 0.890. This means that the classifier has a high percentage of accuracy as evidenced by the kappa score stands above 0.80 which classified as almost perfect agreement (Landis & Koch, 1977). For Amazon, the accuracy shows about 87.33% with kappa score of 0.822. This means that the classifier has a high percentage of accuracy as evidenced by the kappa score stands above 0.80 which classified as almost perfect agreement (Landis & Koch, 1977). The F-Measure of both companies is 85.92% for Google and 85.00% for Amazon, this shows the harmonization of precision and recall values from the classification process works well.

Table 6: Performance Evaluation Results of Sentiment Classification

Company	Accuracy	Kappa	Recall	Precission	F-Measure
Google	96.86%	0.932	96.44%	97.26%	96.84%
Amazon	92.06%	0.822	93.43%	89.78%	91.56%

Table 6 above shows the results of sentiment classification using Naive Bayes Classifier. For Google, the accuracy shows about 96.86% with kappa score of 0.932. This means that the classifier has a high percentage of accuracy as evidenced by the kappa score stands above 0.80 which classified as almost perfect agreement (Landis & Koch, 1977). For Amazon, the accuracy shows about 92.06% with kappa score of 0.822. This means that the classifier has a high percentage of accuracy as evidenced by the kappa score stands above 0.80 which classified as almost perfect agreement (Landis & Koch, 1977). The F-Measure of both companies is 96.84% for Google and 91.56% for Amazon, this shows the harmonization of precision and recall values from the classification process works well.

The interaction that related to Google and Amazon is mapped to understand the sentiment classification. Figure below will explain the sentiment classification result of each category. To review issues that exist in Google companies, this will be analyzed through the negative sentiments of each category with the following proportions:
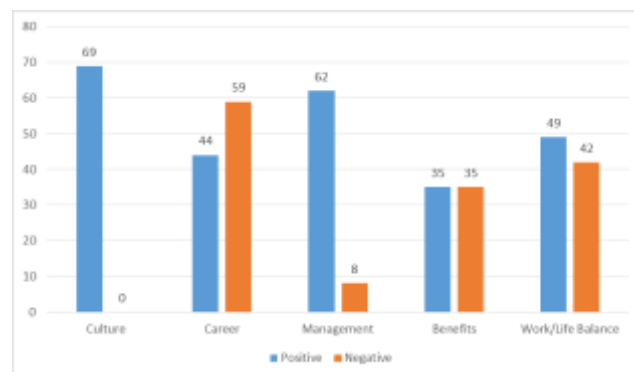

Figure 2: Google Sentiment Classification Results based on Category

As shown in Figure 2, this indicates that career category has the most dominant negative sentiment in Google when compared to other categories.

To review issues that exist in Amazon companies, this will be analyzed through the negative sentiments of each category with the following proportions:
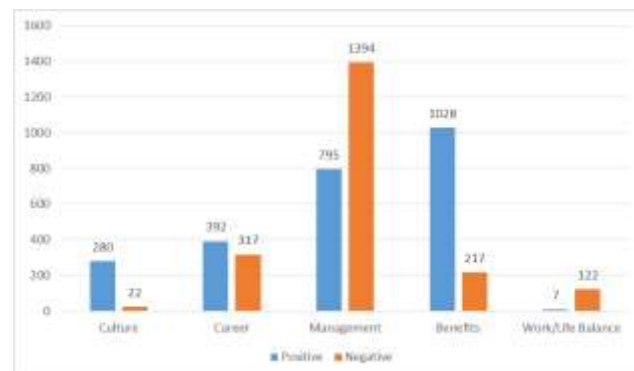

Figure 3: Amazon Sentiment Classification Results based on Category

As shown in Figure 3, this indicates that management category has the most dominant negative sentiment in Amazon when compared to other categories.

5. Conclusion

This paper presents a method of sentiment analysis that is accomplished based on Naive Bayes algorithm. In our opinion, the method we use in this research is better in the term of real-time processing capability compared to conventional way.

In this study, we found that data classification was able to separate sentiments correctly. The classification of categories and sentiments using the Naive Bayes Classifier has a high percentage of accuracy. This is evidenced by the Kappa score of each company stands above 0.80 which is categorized as

Almost Perfect Agreement (Landis & Koch, 1977). F-Measure of both companies shows the harmonization of precision and recall values from the classification process works well.

This method can be applied in other sector. Our suggestions for future research, in order to increase the accuracy and kappa of sentiment, stop word dictionary must be enriched and reduce the text ambiguity because one sentence may contain more than one meaning.

References

[1] Moeheriono. (2009). Pengukuran Kinerja Berbasis Kompetensi. Bogor: Ghalia Indonesia.

[2] PayScale. (2017). Full List of Most and Least Loyal Employees. Retrieved from PayScale Survey: https://www.payscale.com/data-packages/employee-loyalty/full-list

[3] Dell Technologies. (2017, November 28). How AI is Helping HR Gauge Employee Sentiment in Real Time. Retrieved from Dell Technologies: https://www.delltechnologies.com/en-us/perspectives/how-ai-is-helping-hr-gauge-employee-sentiment-in-real-time/

[4] Rowh, M. (2011, November 9). *How to Find Out What Employees Really Think*. Retrieved November 28, 2017, from CIO: https://www.cio.com/article/2402428/enterprise-software/how-to-find-out-what-employees-really-think.html

[5] Hootsuite & We Are Social. (2017). *Digital in 2017 Global Overview*. Retrieved from We Are Social.

[6] Costa, A., & Veloso, A. (2015). Employee Analytics through Sentiment Analysis. *Conference: Brazilian Symposium on Databases.* Petrópolis - RJ - Brazil: SBBD'15 Proceedings.

[7] Hadna, N. M., Santosa, P. I., & Winarno, W. (2016). Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen di Twitter. *Seminar Nasional Teknologi Informasi dan Komunikasi 2016.* Yogyakarta: (SENTIKA 2016).

[8] Moenz, M. F., Li, J., & Chua, T. S. (2014). *Mining User Generated Content.* Boca Raton: CRC Press.

[9] Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment Analysis Algorithms and Applications: A Survey. Ain Shams Engineering Journal. Science Direct.

[10] Zhao, Y., & Cen., Y. (2014). Data Mining: Practical Machine Learning Tools and Techniques (Third Edition). Burlington: Elsevier.