Consumers' Attitude toward Insurance Companies: A Sentiment Analysis of Online Consumer Reviews

Submission Type: Full papers

Maryam Ghasemaghaei
Mcmaster University
ghasemm@mcmaster.ca

Seyed Pouyan Eslami
Mcmaster University
eslamisp@mcmaster.ca

Ken Deal
Mcmaster University
deal@mcmaster.ca

Khaled Hassanein
Mcmaster University
hassank@mcmaster.ca

Abstract

This study considers consumer online reviews for auto, home, and life insurance services to provide insights regarding understanding consumers' attitude toward different insurance types. We apply the sentiment analysis technique to analyze the overall sentiment of each insurance type since 2012. Our findings show that since 2013 consumers had more negative attitudes toward insurance services. We also extracted consumers' emotions for each insurance type and found that consumers generally have more negative emotions than positive ones toward insurance services. Moreover, the results show that consumer reviews sentiments could accurately predict consumer review ratings and there are significant differences in review ratings between positive, neutral, and negative consumer reviews sentiments for all insurance types. The results of this study highlight the importance of sentiment analysis in understanding consumers' opinion toward different services that could help companies to reduce consumers negative attitudes and to attract more customers.

Keywords

Sentiment analysis, insurance companies, consumer attitude, consumer online reviews.

Introduction

A recent eMarketer report indicates that the number of Internet users that are using and creating online reviews will increase significantly in the next few years. Online reviews are a form of user-generated content that is becoming an important source of information for consumers in purchasing products or using services. Studies found that 78% of users trust comments from other consumers (Hu, Koh, & Reddy, 2014). Sentiments expressed through the user-generated content provide context-specific explanations of the reviewer's attitude about the service or product (Hu, Koh, & Reddy, 2014). Sentiments could be framed as negative, positive, or neutral statements with varying degrees of emotion which provide rich information for companies to explore consumers' opinions and attitudes about their products or services.

Several researchers have noted the importance of capturing the sentiments expressed in consumer online reviews. However, they also emphasized the difficulty in doing so. For example, Liu (2006) analyzed 12,000 movie review comments using human judges and indicated that it was "an extremely tedious task." The development of text mining tools has made this task more efficient than manual coding, less tedious, and has enhanced the ability to analyze large amounts of contents generated by users. Although studies have found that text mining has the limitation of being less accurate than human judges; for online reviews the accuracy levels tend to be around 80% (Netzer, Feldman, Goldenberg, & Fresko, 2012). At this accuracy level it is still deemed useful which has prompted its application in marketing, Information Systems (IS), and other applied disciplines (Archak, Ghose, & Ipeirotis, 2011; Berger, Sorensen & Rasmussen, 2010; Duan, Gu, & Whinston, 2008).

Recent studies show that many auto, home, and life insurance companies have endured several years of constrained growth (Crawford, 2015). However, multiple opportunities exist for these companies to gain competitive advantage through a better understanding customers attitudes toward their services (Crawford, 2015). Along these lines, people's sentiments could provide insurance companies with valuable information. For example, sentiment analysis can help them respond with appropriate marketing and public relations strategies (Sautto, 2014). In addition, consumer sentiment has been shown to have a major impact on stock ratings (Schumaker, Zhang, Huang & Chen, 2012; Yu, Duan, & Cao, 2013).

Previous studies have found strong evidence that user-generated comments influence sales (Arnold & Vrugt, 2008; Chevalier & Mayzlin, 2006; Dellarocas, Zhang, & Awad, 2007; Paltoglou & Thelwall, 2010; Salehan & Kim, 2016). For instance, Ghose & Ipeirotis (2011) found that the extent of informativeness in consumer reviews impacts companies' revenue generation. Forman, Ghose, & Wiesenfeld (2008) also found consumers' attitudes about a particular product or service would impact other consumers' perceptions regarding purchasing products or services. For example, individuals often review other consumers' opinions about insurance services before making their purchasing decisions. This could be due to the fact that insurance services are considered as high involvement products for which consumers often gather extensive information from various sources (e.g., consumer reviews) before making their purchasing decisions (Kowatsch et al. 2011).

In the context of insurance industry, no studies have used sentiment analysis to identify consumers' attitudes toward insurance services. Therefore, there is a need for additional research in this area, particularly in terms of exploring consumers' attitudes toward auto, life, and home insurance companies. Thus, the first objective of this paper is to conduct a sentiment analysis of consumer reviews to better understand the overall attitudes and emotions of consumers toward different types of insurance services (i.e. auto, home and life). The insight gained into consumers' attitude towards insurance services is a novel addition to the literature, which particularly helps insurance companies to identify issues that matter most to the customers, and to predict emerging events.

Previous research has found that consumers' review ratings are considered to be very helpful by other consumers in making their decisions to purchase products or services (Salehan & Kim, 2016). Thus, the second objective of this paper is to understand if consumer reviews' sentiments could predict the consumer overall ratings for the different insurance types. In addition, this paper explores if there is a significant difference in consumers' review ratings for those who provided positive comments, neutral comments, or negative comments. By investigating such differences, this study provides insights into the consistency between consumers' comments and their review ratings.

Using a panel of data on 156 insurance companies from Insureye.com, we extracted sentiments from consumer online reviews. From the online business managers' perspectives, the findings of this study are important steps toward understanding online consumers' attitudes and emotions toward insurance companies to help these companies to identify issues that matter most to customers and to consequently provide better marketing and public relations strategies.

Related work

The abundance of user-generated content in the online environment has resulted in a surge of research interest in using sentiment analysis techniques to extract consumers' opinions to understand their attitudes toward products and services (Hogenboom et al. 2014). In this section, with the objective of using sentiment analysis to extract consumers' attitudes and emotions toward insurance services, we briefly explain sentiment analysis, consumers' ratings and reviews, and the insurance market.

Sentiment Analysis

The roots of sentiment analysis are in the computational linguistics, natural language processing, and text mining fields (Hogenboom et al. 2014). Sentiment analysis is typically used in opinion mining to identify sentiment, subjectivity, affect, and emotional states in online text. The main objective of text sentiment analysis is to identify the attitude of a writer with respect to some specific topic (Li & Wu, 2010). In general, sentiment analysis has three main advantages: (i) it helps build models to aggregate the opinions of people to gain useful insights into individuals' attitudes; (ii) it converts large unstructured text into a form which allows for predictions about particular outcomes; and (iii) it is useful in gathering information on people's reaction to a particular object to design marketing and advertising campaigns (Yu et al. 2013).

The emergence of Web 2.0 has paved the way for the usage of sentiment analysis. It has created plentiful and exciting opportunities for understanding the opinions of consumers and the general public regarding company strategies, social events, marketing campaigns, and product and service preferences (Chen, Chiang, & Storey, 2012). The use of sentiment analysis has become popular in the past decade because of the availability of large datasets for machine learning algorithms to be trained on, advances of machine learning methods in information retrieval and natural language processing, and the development of many commercial artificial intelligence applications (Das & Chen, 2007; Tong & Koller, 2002; Yu et al. 2013).

Consumers' Ratings and Reviews

Consumers' ratings and reviews could be helpful for those who want to understand other consumers' experiences with particular products or services (Chen et al., 2012). According to cognitive load theory (Oviatt 2006), consumers have limited information processing capacity. Thus, they are likely to attempt to decrease the amount of effort they expend in making decisions (Hu, Koh, & Reddy, 2014). Accessibility to consumers' reviews has changed the way consumers shop and helped them to make decisions faster (Ha & Hoch, 1989). With increasing demands on their attention and time, consumers try to decrease the amount of effort they spend on making decisions by obtaining information about others' experiences or sentiments through reading their reviews and associated ratings to make decisions which ultimately could impact companies' sales (Hu et al. 2014). Consumers' ratings and reviews could also be useful for companies to understand consumers' attitudes toward their products and/or services. In this paper, we focus on consumers' ratings and reviews toward insurance companies to explore their emotions regarding insurance services which could help insurance companies to develop better marketing strategies.

The Insurance Market

The insurance market in North America is extremely competitive which makes it very difficult for some companies to attract customers (Kelly, Kleffner, & Li, 2013). Particularly, some life, home, and auto insurance companies have endured several years of constrained growth. However, despite several years of slow-to-no growth for providers of insurance, multiple opportunities exist to gain a competitive advantage in the market (Crawford, 2015). Studies found that consumers' attitudes would impact significantly on insurance companies' revenue growth (Rawson, Duncan, & Jones, 2013). For instance, Nwankwo & Ajemunigbohun (2013) found that an effective relationship with consumers would have a positive impact on consumers' attitudes and emotions which would influence companies' revenues. Moreover, Huber, Gatzert, & Schmeiser (2015) found that consumers' experiences in using life insurance services impact their satisfaction and motivate them to provide positive recommendations regarding those services to other customers.

In order to effectively understand consumers' attitudes and needs toward insurance services, recent studies suggest that companies should incorporate new approaches using data analytics tools (Crawford, 2015). Thus, in this study, we explore consumer attitudes and emotions toward different types of insurance services through sentiment analysis and provide some recommendations for insurance companies that can help them to reduce consumers' negative attitudes and to attract more customers.

Methodology

Sentiment Mining

There are different types of sentiment analysis methods based on the problem examined (Liu, 2010). For example, Archak et al. (2011) conducted feature-based sentiment analysis to extract sentiments related to product attributes. In this study, we used sentiment analysis at the consumers' review level (Liu, 2010) (i.e. document-level sentiment classification) to study consumers emotions and the inter-relationship among sentiments and ratings. Figure 1 outlines the procedure used to collect and analyze the customer reviews in our sample. With reference to the Figure, all the steps including HTML scraping were executed using an R code program. The program was run to collect the relevant reviews' information from the website Insureye.com[1]. The dataset obtained included 1584 reviews for the auto, home and life insurance

[1] This website is selected to extract consumer reviews as it provides reviews for all main insurance types (home, auto, and life insurance) and for all major Canadian insurance providers.

services of 156 insurance companies. Extracted information for each review included: the review text, date, and reviewer rating of the service in question (Figure 2 shows an example of a typical review).

To assess the sentiments of the reviews, we looked at the polarity and the emotions expressed in these reviews. We used an opinion lexicon (Hu & Liu, 2004) as a dictionary to extract review polarities (positive, neutral, and negative) based on the words found in these reviews for each insurance type. Each "positive" word in a review (e.g. helpful, good, etc.) was assigned a polarity of 1, "neutral" words (e.g. vehicle, house, etc.) were assigned a polarity of 0, and "negative words (e.g. horrible, terrible, etc.) were assigned a polarity of −1. Irrelevant words (e.g. numbers, and, the, etc.) were excluded. For each review, the number of specific sentiment word occurrences was computed. The sentiment of each review was assessed as the sum of the polarities of all the considered words in the review.

To test the accuracy of the reviews sentiment algorithm described above, two of the authors independently read a sample of 250 reviews randomly chosen from the sample. For each review, each person determined if the sentiment score assigned through the above algorithm was accurate and reflected the overall sentiments expressed in each review. The 250 reviews were selected randomly and both raters rated the reviews in the same order. We computed Cohen's kappa to examine the strength of agreement between the two raters (Reynolds & Reynolds, 1977). Since Cohen's Kappa offers agreement beyond what could be due to chance, it provides a confidence in the ratings obtained (Reynolds & Reynolds, 1977). The raters achieved a 0.87 inter-rater reliability which is higher than the 0.60 threshold required for substantial agreement between raters (Landis & Koch, 1977). The results of this assessment indicated an accuracy of 91% for our sentiment analysis algorithm which was deemed appropriate given that it is higher than the typical 80% performance reported for sentiment analysis algorithms (Cao et al. 2013; Yu et al. 2013).

To explore more deeply consumers' attitudes and emotions embodied in reviews of each type of insurance, we used the "sentiment library" in R to assign one of six emotions (sadness, anger, joy, disgust, surprise, fear) to each review based on its content. This library contains a set of algorithms for identifying emotions embodied in text using a variety of approaches including the Naïve Bayes algorithm which we used in this analysis. The Naïve Bayes algorithm is an effective and simple classifier that has been used in various information processing applications such as information retrieval and image recognition (Escudero, Màrquez, & Rigau, 2000; Lewis, 1998; Nigam & Ghani, 2000). Naïve Bayesian classifiers often perform well for sentiment polarity classification (Cao, Thompson, & Yu, 2013; Pang & Lee, 2008). The underlying theorem for Naïve Bayesian text classification is the following Bayes Rule:

$$\mathrm{P(A|B)} = \frac{(\mathrm{P(A|B)} * \mathrm{P(A)})}{\mathrm{P(B)}}$$

The Bayes Rule calculates the likelihood of event A given that event B has occurred. By looking at the frequencies of specific words in the document (e.g. horrible), the Naïve Bayes algorithm is used in text classification to understand the probability that a review B containing a specific word (e.g., horrible) is associated with emotion A (e.g., disgust).

For visualization purposes, we used the R Wordcloud package from CRAN to classify consumers' reviews' content based on the extracted emotions (Fellows, 2015). The result of this procedure is a term document matrix which contains the various emotions, and frequencies of each word associated with each of these emotions in all the reviews. We used this term document matrix to create the word cloud (Yarowsky, 1992) which is a graphical depiction of the frequency of different words in customers' reviews and their associated emotions. In addition, we conducted regression analysis to explore if consumers' reviews' sentiment can predict their review ratings. We also conducted ANOVA analyses to understand if there is a significant difference between the sentiments (positive, neutral, and negative) in terms of review ratings.

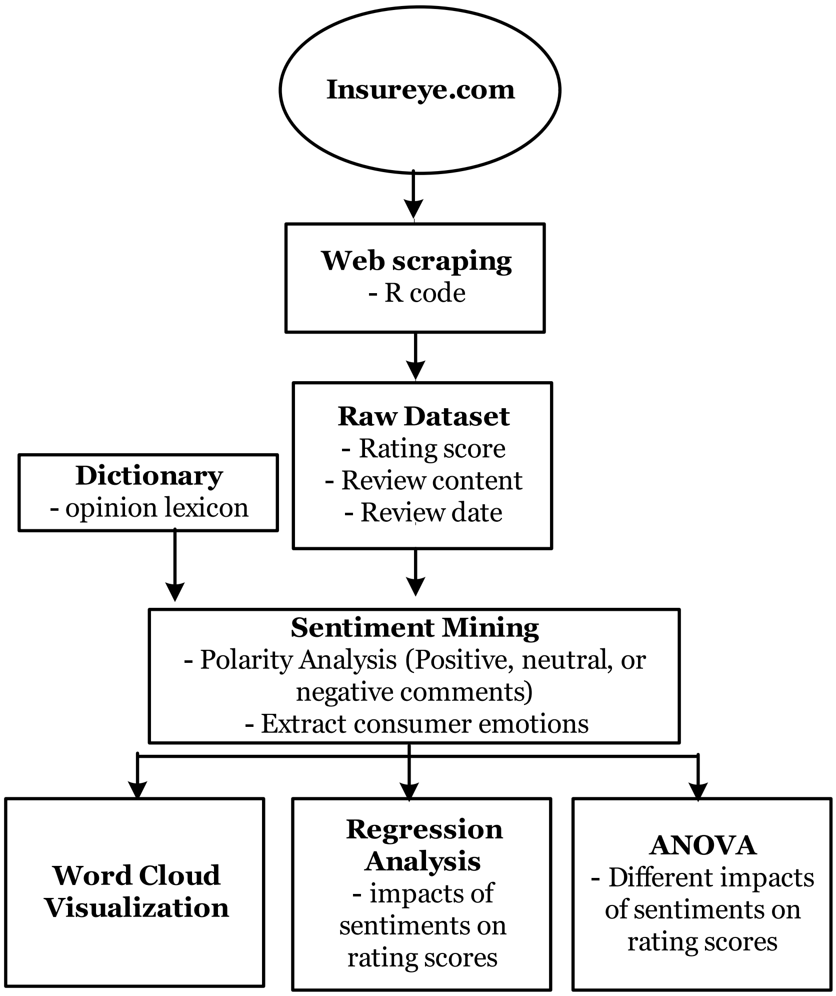

Figure 1. A Methodology for Review Sentiment Mining Process

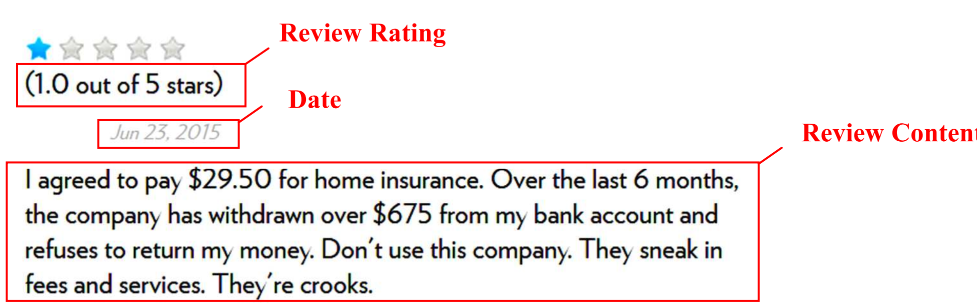

Figure 2. Example of Online Reviews

Data Analysis and Results

To conduct our study, we collected a panel data set of customer reviews for auto, life, and home insurance in Insureye.com from January 2012 to December 2015 using a web scraping technique. We extracted 796 text reviews[2] for auto insurance, 553 reviews[3] for home insurance, and 235 reviews[4] for life insurance for

[2] 300, 70, 146, and 280 reviews in 2012, 2013, 2014, and 2015, respectively.
[3] 266, 50, 115, and 122 reviews in 2012, 2013, 2014, and 2015, respectively.
[4] 126, 12, 31, and 66 reviews in 2012, 2013, 2014, and 2015, respectively.

156 insurance companies. The summary statistics of the sentiments generated for reviews as well as the ratings assigned by reviewers in our sample in their reviews are shown in Table 1. This table shows the downward shift in consumer sentiments and ratings since 2012. Further, the last column illustrates the overall means and medians for each type of insurance sentiment and review rating.

	2012		**2013**		**2014**		**2015**		**Overall**	
Variable	**M**	**Mean (SD)**	**M**	**Mean (SD)**	**M**	**Mean (SD)**	**M**	**Mean (SD)**	**M**	**Mean (SD)**
Auto review sentiment	1	0.45(.73)	-.5	-.18(.8)	-1	-.28(.8)	-1	-.35(.8)	0	-.02(.89)
Auto review rating	4	3.22(1.48)	1	1.77	1	1.78(1.34)	1	2.14(1.54)	2	2.45(1.59)
Home review sentiment	1	0.42(.78)	0	-.06(.8)	0	-.04(.9)	-1	-.1(.95)	0	.16(.89)
Home review rating	4	4.1(.85)	1	1.86(1.29)	2	2.26(1.6)	1	1.92(1.46)	3	3.03(1.6)
Life review sentiment	1	0.61(.63)	0	0(.95)	-1	-.29(.9)	0	-.13(.8)	1	.25(.85)
Life review rating	4	3.89(1.05)	2	2.33(1.49)	1	1.54(1.1)	1	1.5(.89)	3	2.83(1.55)

Note: M: Median; Review sentiment: -1, 0, or 1; Review rating: from 1 (low) to 5 (high)

Table 1. Statistics of Reviews' Sentiment and Ratings

We examined the trend of the sentiment ratio (positive and negative words to total) as an indicative measure of individuals' attitude toward each insurance for each of the years between 2012 to 2015. As can be seen in this Figure 3, for all insurance types (i.e., auto, home, life), the ratio of positive terms to total is higher in 2012. However, this ratio has gradually decreased since 2012, and in 2015 the ratio of negative terms to total is higher in all insurance types. Studies show that many auto, home, and life insurance companies have endured several years of constrained growth (Crawford, 2015). Thus, these results are consistent with the insurance literature.

Figure 4 illustrates the results of the word clouds for the auto, home, and life insurance reviews in our sample. The size of each term in the word cloud is based on the number of times it appeared in reviews associated with a specific emotion (i.e., sadness, anger, joy, disgust, surprise, fear) across all reviews. It should be noted that emotion associated with each word in the cloud is a context-related emotion related to the reviews that a specific word appears in as opposed to the specific meaning of the word in question. As a result, the same word in different reviews can have different associated emotions that are based on the emotions of the different reviews this word appears in.

As shown in Figure 4, there are generally more negative emotions than positive emotions for all insurance types. For auto insurance words such as "horrible", and "policies" have been classified under the *disgust* emotion; "company", "insurance", and "service" have been classified under the *fear* emotion; "high", and "despite" have been classified under the *anger* emotion; "pay", and "bad" have been classified under the *sadness* emotion; and "good" has been considered under the *joy* emotion. For home insurance words such as "increase", and "insurance" have been classified under the *disgust* emotion; "policy", and "coverage" have been classified under the *fear* emotion; "high", and "damage" have been classified under the *anger* emotion; and "good" and "like" have been considered under the *joy* emotion. For life insurance words such as "paperwork" has been classified under the *disgust* emotion; "coverage" has been classified under the *fear* emotion; "frustrated", and "claim" have been classified under the *anger* emotion; "insurance" and "policy" have been considered under the *sadness* emotion; and "good" and "great" have been considered under the *joy* emotion.

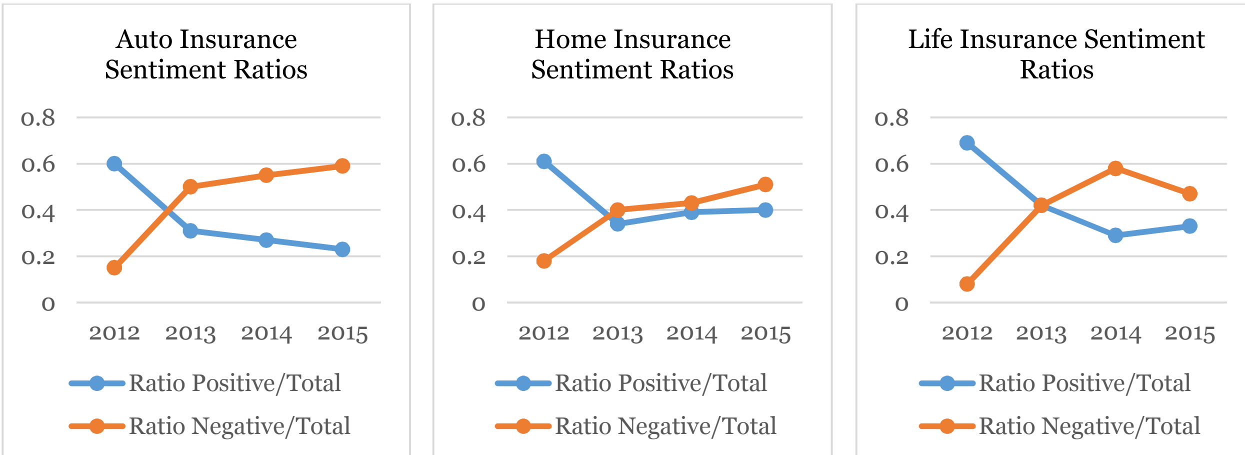

Figure 3. Annual Positive/Negative Sentiment Ratios for Each Insurance Type

The results from the word clouds could be very useful for insurance companies to identify the most frequent words in consumer comments that are associated with specific consumers' emotions toward their services. These words should be seen as flags representing customers' positive or negative attitudes toward their services. To understand the real meanings behind these words, companies should carefully examine the comments associated with the most frequent non-generic key words (e.g. happy, great, good, horrible, increase, paperwork, etc.) as revealed by the word cloud.

Regression results show that for auto insurance the relationship between sentiment reviews and review ratings was significant ($\beta = 0.254$; $p < 0.000$). Moreover, for home insurance the relationship between sentiment reviews and review ratings was significant ($\beta = 0.471$; $p < 0.000$). Likewise, for life insurance the relationship between sentiment reviews and review ratings was also significant ($\beta = 0.565$; $p < 0.000$).

ANOVA analyses were also conducted to understand whether any differences exist between negative, neutral, and positive sentiments (-1, 0, and 1) regarding consumer review ratings for all insurance types. As can be seen in Table 2 and Figure 5 there is a significant difference between reviews' sentiment (i.e., -1, 0, and 1) and review ratings such that when reviews have negative sentiments, they significantly have lower review ratings compared to when they have neutral or positive sentiments.

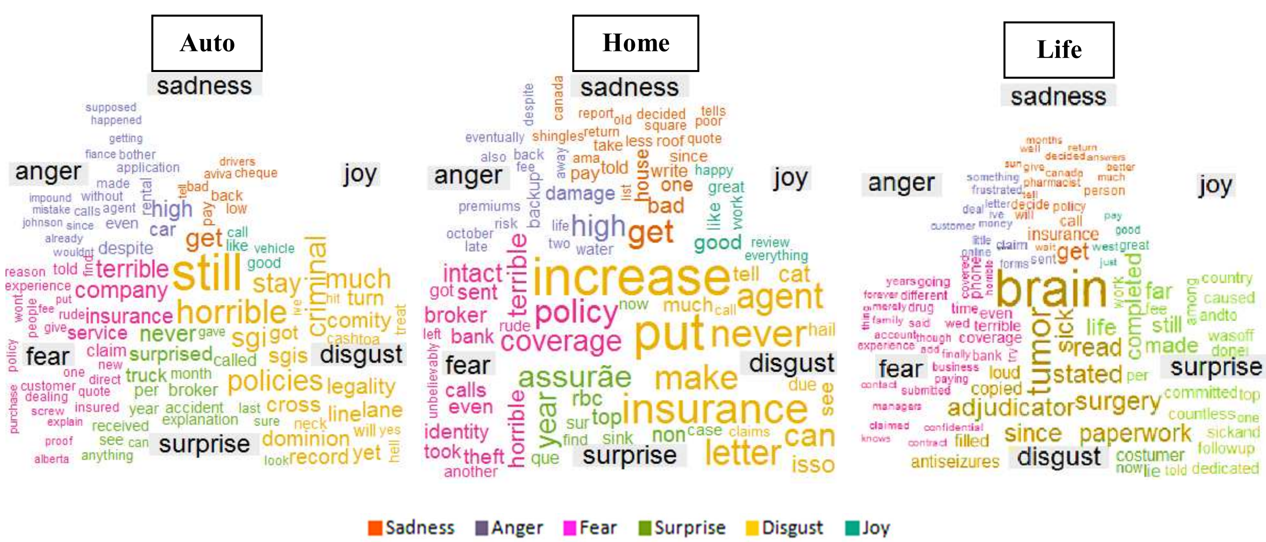

Figure 4. Word Cloud for each Insurance Type

Groups	Multiple Comparisons	Auto		Home		Life	
		Mean Difference	Sig.	Mean Difference	Sig.	Mean Difference	Sig.
Negative sentiment	Neutral	-0.550*	0.001	-0.9*	.000	-0.82*	.003
	Positive	-0.902*	0.000	-1.7*	.000	-2.01*	.000
Neutral sentiment	Positive	-0.352*	0.049	-0.8*	.000	-1.1*	.000

* The mean difference is significant at the 0.05 level.

Table 2. ANOVA Summary Table for All Insurance Types

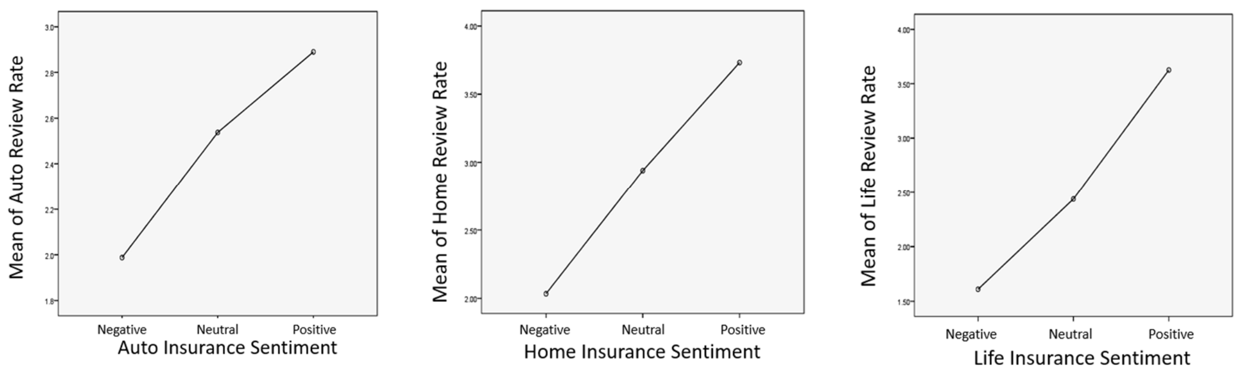

Figure 5. ANOVA Results for Each Insurance Type

Discussion

Studies have found that consumers' attitude is one of the critical factor affecting their satisfaction (Ghose, & Wiesenfeld (2008). In this study, we conducted a sentiment analysis based on consumers' reviews to understand consumers' attitudes and emotions toward insurance services. With a sample of 1584 reviews from 156 companies, we covered three insurance types including auto, home, and life insurance. We found that although in 2012 consumers had more positive attitudes toward insurance services, since then, they have had more negative attitudes toward these services. The results of this study highlight the importance of sentiment analysis in understanding consumers' opinions toward different services which could help companies to reduce consumers' negative attitudes and to attract more customers.

In order to further understand consumers' attitudes toward the different insurance services, we also investigated consumers' emotions. Results reveal that in general, consumers have more negative emotions than positive ones toward insurance services. The results of our word cloud analysis can help insurance companies to better understand review words that are most associated with specific emotions towards their services. For example, for auto insurance, consumers provided terms such as "horrible" and "policies" for *disgust* frequently. For home insurance, consumers provided terms such as "policy" and "coverage" for *fear* frequently. For life insurance, consumers provided terms such as "frustrated" and "claim" for *sadness* frequently. Understanding consumers' emotions is a major step toward determining consumers' attitudes toward different insurance services. The results of our regression and ANOVA analyses lend further support for the validity of our reviews' sentiment analysis algorithm.

Limitation and Future Research

Although sentiment analysis is a useful technique in analyzing large textual information, it still has limitations. For example, the software we used in this study lacks processing capability for alternate styles of writing such as sarcasm. There are many areas for improvement in the natural language processing field and future research can provide better insights on the information contained in online consumer reviews using more advanced technology.

Second, the sample used in this study was collected only from the Insureye.com website which has consumer reviews from 2012 to 2015. Thus, future research may further investigate consumers review toward insurance companies by analyzing reviews from other platforms.

Third, in this study we focused only on auto, home, and life insurance services. Thus, our conclusions may remain limited to these types of insurance. Future research can extract reviews for other insurance types (e.g., travel insurance) and other industries to determine if the results of this study could be generalized to them.

Finally, the samples lack language and cultural diversity. The reviews were collected from the Insureye.com website which contains customer reviews of different insurance companies in Canada and all reviews were written in English. Future research can extract reviews from other websites that also contains customer reviews from companies in other countries to determine if the results of this study could be generalized to companies in other countries. Moreover, future studies could utilize reviews written in other languages to include the effect of language on the results of this research.

Conclusions

This research used sentiment analysis to provide insights regarding understanding consumers' attitude toward different insurance services. We found that in 2012 consumers had more positive attitudes toward insurance services. However, since then they have had more negative attitudes toward auto, home, and life insurance services. This could be one of the reasons that insurance companies have endured several years of constrained growth. We also extracted consumer emotions for each insurance type and found that consumers generally have more negative emotions toward insurance services. Moreover, we found the most frequent words in consumers' reviews that shape their emotions for different types of insurance services. The results of this study also showed that consumer reviews' sentiment can accurately predict review ratings. Moreover, we found that significant differences exist between positive, neutral, and negative sentiments regarding review ratings for all insurance types. This research can be used by insurance companies to develop better strategies regarding their relations with their customers to reduce their negative attitudes and increase positive ones.

REFERENCES

Archak, N., Ghose, A., & Ipeirotis, P. G. 2011. "Deriving the pricing power of product features by mining consumer reviews," *Management Science* (57:8), pp. 1485-1509.

Arnold, I. J., & Vrugt, E. B. 2008. "Fundamental uncertainty and stock market volatility," *Applied Financial Economics* (18:17), pp. 1425-1440.

Berger, J., Sorensen, A. T., & Rasmussen, S. J. 2010. "Positive effects of negative publicity: When negative reviews increase sales," *Marketing Science* (29:5), pp. 815-827.

Cao, Q., Thompson, M. A., & Yu, Y. 2013. "Sentiment analysis in decision sciences research," *Decision Support Systems* (54:2), pp. 1010-1015.

Chen, H., Chiang, R. H., & Storey, V. C. 2012. "Business Intelligence and Analytics: From Big Data to Big Impact," *MIS Quarterly* (36:4), pp. 1165-1188.

Chevalier, J. A., & Mayzlin, D. 2006. "The effect of word of mouth on sales: Online book reviews," *Journal of Marketing Research* (43:3), pp. 345-354.

Crawford, S. 2015. *Global insurance outlook*, Available at http://www.ey.com/Publication/vwLUAssets/ey-2015-global-insurance-outlook/$FILE/ey-2015

Das, S. R., & Chen, M. Y. 2007. "Yahoo! for Amazon: Sentiment extraction from small talk on the web," *Management Science* (53:9), pp. 1375-1388.

Dellarocas, C., Zhang, X. M., & Awad, N. F. 2007. "Exploring the value of online product reviews in forecasting sales: The case of motion pictures," *Journal of Interactive Marketing* (21:4), pp 23-45.

Duan, W., Gu, B., & Whinston, A. B. 2008. "Do online reviews matter?—An empirical investigation of panel data," *Decision Support Systems* (45:4), pp. 1007-1016.

Escudero, G., Màrquez, L., & Rigau, G. 2000. *Naive Bayes and exemplar-based approaches to word sense disambiguation revisited. arXiv Preprint cs/0007011.*

Fellows, Ian 2015. *Package 'wordcloud'. CRAN*, February 20, 2015.

Forman, C., Ghose, A., & Wiesenfeld, B. 2008. "Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets," *Information Systems Research* (19:3), pp. 291-313.

Ghose, A., & Ipeirotis, P. G. 2011. "Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics," *Knowledge and Data Engineering, IEEE Transactions on* (23:10), pp. 1498-1512.

Ha, Y.-W., & Hoch, S. J. 1989. "Ambiguity, processing strategy, and advertising-evidence interactions," *Journal of Consumer Research*, pp. 354-360.

Hogenboom, A., Heerschop, B., Frasincar, F., Kaymak, U., & de Jong, F. 2014. "Multi-lingual support for lexicon-based sentiment analysis guided by semantics," *Decision Support Systems* (62), pp. 43-53.

Huber, C., Gatzert, N., & Schmeiser, H. 2015. "How Does Price Presentation Influence Consumer Choice? The Case of Life Insurance Products," *Journal of Risk and Insurance* (82:2), pp. 401-432.

Hu, M., & Liu, B. 2004. "Mining and summarizing customer reviews," In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168–177). ACM.

Hu, N., Koh, N. S., & Reddy, S. K. 2014. "Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales," *Decision Support Systems* (57), 42-53.

Kelly, M., Kleffner, A. E., & Li, S. 2013. *The Impact of Regulation on the Availability and Profitability of Auto Insurance in Canada. Available at SSRN 1698042.*

Kowatsch, T., Maass, W. and Fleisch, E. 2011. "The role of product reviews on mobile devices for in-store purchases: consumers' usage intentions, costs and store preferences," *International Journal Internet Marketing and Advertising* (6:3), pp. 226-243.

Landis, J. R., & Koch, G. G. 1977. "The measurement of observer agreement for categorical data," *Biometrics*, pp. 159-174.

Lewis, D. D. 1998. "Naive (Bayes) at forty: The independence assumption in information retrieval," In *Machine learning: ECML-98* (pp. 4-15). Springer.

Li, N., & Wu, D. D. 2010. "Using text mining and sentiment analysis for online forums hotspot detection and forecast," *Decision Support Systems* (48:2), pp. 354-368.

Liu, B. 2010. "Sentiment Analysis and Subjectivity," *Handbook of Natural Lang. Processing* (2), 627-666.

Liu, Y. 2006. "Word of mouth for movies: Its dynamics and impact on box office revenue," *Journal of Marketing* (70:3), pp. 74-89.

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. 2012. "Mine your own business: Market-structure surveillance through text mining," *Marketing Science* (31:3), pp. 521-543.

Nigam, K., & Ghani, R. 2000. "Analyzing the effectiveness and applicability of co-training," *the ninth international conference on Information and knowledge management* (pp. 86-93). ACM.

Nwankwo, S. I., & Ajemunigbohun, S. S. 2013. "Customer Relationship Management and Customer Retention: Empirical Assessment from Nigeria?Insurance Industry," *Business & Economics Journal*.

Oviatt, S. 2006. "Human-centered design meets cognitive load theory: designing interfaces that help people think," *14th annual ACM international conference on Multimedia* (pp. 871-880). ACM.

Paltoglou, G., & Thelwall, M. 2010. "A study of information retrieval weighting schemes for sentiment analysis," In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (pp. 1386-1395). Association for Computational Linguistics.

Pang, B., & Lee, L. 2008. "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval* (2:2), pp. 1-135.

Rawson, A., Duncan, E., & Jones, C. 2013. "The truth about customer experience," *Harvard Business Review* (91:9), pp. 90-98.

Reynolds, H. T., & Reynolds, H. T. 1977. *The analysis of cross-classifications*. Free Press New York. Retrieved from https://www.ncjrs.gov/App/abstractdb/AbstractDBDetails.aspx?id=47084

Salehan, M., & Kim, D. J. 2016. "Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics," *Decision Support Systems* (81), pp. 30-40.

Sautto, J. M. 2014. "Decision Support Systems for Business Intelligence," *Investigación Operacional* (35:1), pp. 89-90.

Schumaker, R. P., Zhang, Y., Huang, C.-N., & Chen, H. 2012. "Evaluating sentiment in financial news articles," *Decision Support Systems* (53:3), pp. 458-464.

Tong, S., & Koller, D. 2002. "Support vector machine active learning with applications to text classification," *The Journal of Machine Learning Research* (2), pp. 45-66.

Yarowsky, D. 1992. *Word-sense disambiguation using statistical models of Roget's categories trained on large corpora*, in: Proceedings of the 14th International Conference on Computational Linguistics (COLING-92), Nantes, France, pp. 454-460.

Yu, X., Liu, Y., Huang, X., & An, A. 2012. "Mining online reviews for predicting sales performance: a case study in the movie domain," *Knowledge and Data Engineering, IEEE* (24:4), pp. 720-734.

Yu, Y., Duan, W., & Cao, Q. 2013. "The impact of social and conventional media on firm equity value: A sentiment analysis approach," *Decision Support Systems* (55:4), pp. 919-926.