**INDIVIDUAL PAPER REVIEW**

**Title :** Data Mining Applied to Insurance Twitter Posts

**Author(s) :** Roosevelt C. Mosley Jr

**Document Date :** Casualty Actuarial Society E-Forum, Winter 2012-Volume 2

1. **What is the main problem discussed in this paper?**

The use of social media has grown significantly in recent years. With the growth in its use, there has also been a substantial growth in the amount of information generated by users of social media. Insurers are making significant investments in social media, but many are not systematically analyzing the valuable information that is resulting from their investments.

1. **What is the main related studies that worked on this problem?**

Twitter Study Reveals Interesting Results About Usage—40% is ‘Pointless Babble.

Twitter Study–August 2009,” August 12, 2009, San Antonio, TX: Pear Analytics.

1. **What are the dataset(s) used in this study?**

dataset of insurance company tweets was identified for analysis. Twapperkeeper.com was a web service that tracked and archived Twitter posts based on archives that were set up by users.

- Number of Tweets by Hour of the Day.

- Hashtag Frequency

1. **What are the approach/model proposed in the paper?**

One approach is the comparison of two strings by computing the Levenshtein edit distance (LED).

Another approach to comparing strings is to calculate the generalized edit distance (GED).

The next approach that can be applied to social media analysis is a cluster analysis. The clustering procedure is based on calculating distances between observations and is used to segment databases.

1. **How the authors evaluated the performance of the methods used in this paper?**

As can be seen all around us, the use of social media has grown significantly, and has transformed the way that people interact. This growth in social media has led to an increase in the amount of information that is being generated, and this information provides insight to companies, including insurers, related to their business. Analytics can be applied to social media to identify key words and phrases that are being expressed, and these findings can be used by insurers to assist in managing their business, and to interact more effectively with current and potential customers.

1. **Reflection: What have I learnt?**

I learned from this study that the application of correlation, clustering, and association analyses to social media which results in these analyses help identify keywords and concepts in the social media data, and can facilitate the application of this information by insurers.