

## Referee's Comments

on

“Textual Sentiment of Chinese microblog toward the Stock Market”

(Manuscript Number: IJITDM-D-17-00138)

In this paper, the authors employed text mining and statistical models to explore the relationship between the Shanghai Stock Exchange Composite Index and the collective emotions of individual investors. The emotions of individual investors are quantified by extracting and aggregating investor online posts that contain finance-related keywords. Viewpoints are quantified by analyzing and aggregating the online posts of individual investors that are related to stock markets, in which the major challenge lies in how to automatically identify a set of suitable keywords. To overcome this difficulty, a keyword set that contains  $k$  finance-related words is determined semi-automatically. Text segmentation is also a challenge when dealing with Chinese text mining. The authors proposed a specialized procedure for breaking Chinese sentences into words should be developed. The proposed procedure is based on two concepts, namely, stickiness and entropy. This procedure is reasonable.

Statistical results of posts from Sina Weibo suggest a relationship between collective viewpoints reflected on Sina Weibo and movements of the SSECI. This study contributes to the existing literature by demonstrating that the microblog sentiment level reports can be quantitatively incorporated as a proxy to provide valuable support to portfolio decision making.

Considering the number of the keywords is large, the authors showed the VIF for each variable (the frequency of each keyword). For rigorous logic reasoning, the correlation matrix of these variables is suggested to be presented here. And the authors are suggested to using the principal component analysis to the variables before they are incorporated into the models.

In sum, I think this paper is well organized, and be suitable for publication in *International Journal of Information Technology & Decision Making*