Abstract

This research aimed to develop experts' opinions analysis system using Latent Semantic Analysis (LSA) and text clustering techniques as a tool to help Delphi technique researchers. As the Delphi technique is a research methodology that provides reliable foresight for good decision making, it is popularly used in many research areas. As its process of collecting data and creating questionnaires requires a lot of time, there are difficulties with the use of the Delphi technique. In the second phase of data collection, it takes a lot of time for text analysis. Since the questionnaires must sent back and forth between the researcher and the domain experts, they could be lost or not returned. Therefore, this research developed an opinions analysis system to reduce the time required to analyze answers, decrease lost questionnaires, lower expenses for the collection of questionnaires for each phase, and reduce errors or bias which may occur during the analysis of the answers.

As for experts' opinions analysis steps, the researcher conducted comparative analysis on the opinion of each expert, cut repetitive messages, grouped similar opinions for creating questions, and removed different opinions for each question. Two steps used for system development were as follows: The first step was to analyze similarity of opinions using LSA technique which can analyze and identify contents without using grammar. The second step was to group similar opinions using a clustering technique.

The developed system was evaluated on two aspects as follows: 1) There was an evaluation of system appropriateness according to the principles of Delphi technique research. 2) There was an evaluation of its practical use.

For the first evaluation, the researcher evaluated clustering opinions by performing a comparative study of opinions analysis using latent semantic analysis and text clustering. K-Means, Fuzzy C-Means, and Bisecting K-Means were used to find the most appropriate clustering technique to use in the development of the system. The three clustering techniques were examined to determine which techniques produced

the highest appropriateness rating score in clustering. Fifteen experts in educational research were asked to evaluate the two sets of outputs from these three clustering algorithms. The 3 x 2 within-subject design was used. The result of repeated measures ANOVA indicated that there were significant differences ($P < 0.05$) among these three techniques. Further analysis of multiple comparisons showed that rating for the appropriateness score in clustering for the K-Means technique was significantly lower than both the Fuzzy C-Means and the Bisecting K-Means. The Bisecting K-Means technique was chosen for implementation of the tool with respect to its highest rating score of appropriateness.

In the second evaluation, the researcher evaluated scores for satisfaction with the use of the system. 37 evaluators divided into two groups comprised of 15 people who have done research using the Delphi technique and 22 people who have not done research using the Delphi technique. Both groups' evaluation results in general show both groups were very satisfied with the system. Benefits of using this system are it can help reduce problems due to bias or a lack of prudent analysis. The system can be used to analyze opinions if evaluators collecting data by open-ended questions. Results derived from the use of the system were significant. Moreover, both groups' scores of satisfaction with the system in general was different at the statistical significance. This indicates that the system developed in this research, comprising use models and opinions clustering analysis research results are appropriate and consistent with the principles of Delphi technique research. This system can be used because evaluators who used the Delphi technique before had higher satisfaction scores than evaluators who had not previously used the Delphi technique.

The research results can be used for developing other systems for opinion clustering and analysis such as those that analyze opinions for survey research and group-focused research with the collection of data by interview or open-ended questions. Moreover, the research results can be applied to a web board for the analysis and clustering of similar opinions by automatic classification for easier collection or searching.