

Cloud Based Teacher's Assessment Data in Educational Data Mining

Than Htike Aung ¹⁺ and Nang Saing Moon Kham ²

University of Computer Studies, Yangon, Myanmar

Abstract. Questionnaire of teacher 's assessment survey in higher educational system is very important and especially cost effective method which is one of survey type from three such as face-to-face interviews and telephone interviews. Teacher assessment survey for educational system is very important in developing countries to improve the education system can be well motivated from evaluation of teacher assessment survey data by analyzing knowledge discovery data mining process. Teacher assessment survey data can be collected from mobile phone android application and which data stored on cloud storage database. In this system the real data is collected for teachers from the University of Computer Studies, Yangon, Myanmar. Association rule mining is one of the approaches to extract hidden frequent pattern and association rules between the items of dataset with varying levels of strength. The proposed system also used simple k means clustering to organize the levels of groups of data set. In this paper presents the relevant studies in the EDM field and KDD process steps.

Keywords: Association rule, K means, Educational Data Mining, Knowledge Discovery from Data.

1. Introduction

Data mining is also known as knowledge-discovery from Data (KDD), is the practice of automatically searching large stores of data for patterns. It has interested a lot of absorption from both research area and commercial communities for finding interesting pattern hidden in large datasets. Association rule mining is one of the most important areas of data mining field, its work is to find all subsets of items which frequently occur and the relationship between them. It has two steps, first is finding frequent item sets and second is generation association rules. K means clustering method is to find group of data sets or information. The whole process produced results which are taken into consideration by experts in the future decisions making in various fields. Educational Data Mining (EDM) is an emerging interdisciplinary research area that approaches the development of methods to explore data originating in an educational context.

The results can be used not only to learn the model for the performance process or instructor modelling but also to evaluate and to improve learning management systems by discovering useful learning information from learning portfolios[1].

The proposed system discovered the educational domain of data mining using a case study from the teacher assessment data collected by android mobile application from the University of Computer Studies Yangon and include three years in period [2016 - 2018]. The main goal of this proposed system is to establish a model based data mining that calculates the performance of teachers, apply on real data sets and EDM field of KDD process steps to discover hidden pattern for future decision making.

2. Related Work

There are many researchers whom work in this area. Surjeet Kumar's research paper presented the classification method for the prediction of teacher's performance. This method used the lazy IBK, j48, and Meta Bagging data mining technique implemented in WEDA 3.6.9 data mining software tool to carry out the

⁺ Corresponding author.
E-mail address: thanhtikeaung@ucsy.edu.mm

prediction and performance process. Stefanos Ougiaroglou and Giorgos Paschalis presented the KDD process which includes the application of the Apriori algorithm for the association rules mining from the educational data of ESOG Web-based application (Electre in School Occupational Guidance). A prior algorithm produced 127 association rules that could help managers to make educational decisions. Randa Kh. Hemaïd1 and Alaa M. El-Halees propose a model to evaluate their performance through the use of association, classification rules (Decision Tree, Rule Induction, K-NN, Naïve Bayesian (Kernel)) techniques to determine teacher performance.

The proposed system that will help educational administrators in decision making, which provided the basis for teacher’s performance improvement that will optimize student’s academic result and improve education. Consequently, this will contribute to successful fulfilment of the goals.

3. Methodology

This proposed system presented the apriori method of Data mining for extracting hidden pattern from teacher’s assessment data set and also used k means clustering method. These models implemented in R and their output results illustrate by plot. R is fashion with large data set.

4. Data Source

The data source that is used in this proposed system was collected from google cloud storage database as a google sheets data. This data consists of evaluations of teacher’s assessment over three years. The scores were divided into 5 roughly equal-sized categories (first, second, third, fourth and fifth). Google cloud storage is temporarily stored data and which data backup stored into University Server every yearly end of academic year. Figure 1 shown data source stored on google cloud and figure 2 shown raw data of teacher’s assessment data set.



Fig. 1:Survey Data Storage

| | A | B | C | D | E |
|---|---------------------|--------|----------|---------------------|---------|
| 1 | Timestamp | Course | Semester | InstructorsName | Section |
| 2 | 2018-08-19 14:22:18 | HCI | | 1 U Than Htike Aung | A |
| 3 | 2018-08-19 14:30:14 | HCI | | 1 U Than Htike Aung | A |
| 4 | 2018-08-19 14:39:04 | HCI | | 1 U Than Htike Aung | A |
| 5 | 2018-08-19 14:56:42 | HCI | | 1 U Than Htike Aung | A |
| 6 | 2018-08-19 14:57:38 | HCI | | 1 U Than Htike Aung | A |
| 7 | 2018-08-19 14:58:39 | HCI | | 1 U Than Htike Aung | A |
| 8 | 2018-08-19 15:04:54 | Java | | 1 U Mya | A |

Fig. 2: Teacher’s Assessment Data Set

5. Application and Data Set Collection

Android-based phones and tablets have become a reliable and useful tool in the collection of data related to the teacher’s assessment survey data collection. Mobile phone data collection saves time and money. Mobile data collection can be used for most interaction in everywhere. JSON objects are parsed into values and they are passed to google cloud storage database. Figure 3,4 shown android mobile survey application and figure 5 is survey application generated JSON objects for survey answers[2].

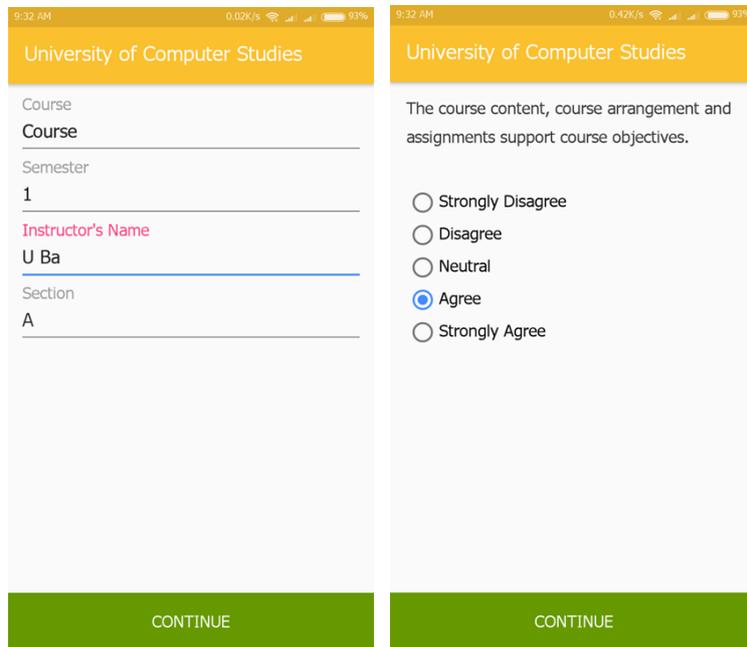


Fig. 3: Survey Form

Fig. 4: Survey Question

```
{
  "AnswerId": "ipang180709022106",
  "PersonalInfo": {
    "value": {
      "PersonalId": 0,
      "Title": "Course",
      "Value": "Data Mining",
      "PersonalId": 0,
      "Title": "Semester",
      "Value": "First",
      "PersonalId": 0,
      "Title": "Instructor's Name",
      "Value": "U Than Htike Aung",
      "PersonalId": 0,
      "Title": "Section",
      "Value": "C"
    }
  },
  "StatusSend": false,
  "SurveyAnswer": {
    "Answer": {
      "IdChoice": "3",
      "Selected": false,
      "ValueChoice": "Neutral"
    },
    "Ask": "The course content, course arrangement and assignments support course objectives.",
    "TypeAsk": "Radioboxes",
    "idAsk": "1",
    "Answer": {
      "IdChoice": "9",
      "Selected": false,
      "ValueChoice": "Agree"
    },
    "Ask": "The course materials that the instructor prepared are good enough to grasp the concept of the course chapters.",
    "TypeAsk": "Radioboxes",
    "idAsk": "2",
    "Answer": {
      "IdChoice": "15",
      "Selected": false,
      "ValueChoice": "Strongly Agree"
    },
    "Ask": "The instructor clearly explains difficult materials.",
    "TypeAsk": "Radioboxes",
    "idAsk": "3",
    "Answer": {
      "IdChoice": "20",
      "Selected": false,
      "ValueChoice": "Strongly Agree"
    },
    "Ask": "The instructor adjusts his/her presentation to the"
  }
}
```

Fig. 5: JSON Data

6. Preprocessing of Teacher’s Assessment Data

The proposed system contains teacher’s assessment information collected from the University of Computer Studies, Yangon which included three years in period from 2016 to 2018. Teacher assessment dataset contains 3335 records and 15 attributes after removing the unnecessary attributes. Table 1 shows the attributes and their description[3].

Table 1. Teacher’s Assessment Data Set Description

| Attribute | Description | Selected |
|-----------|------------------|----------|
| Times | Time Stamp | |
| Course | Java Programming | |
| Semester | First | |

| | | |
|-------------------|--|---|
| Instructor's Name | U Mya | |
| Section | A | |
| Q1 | The course content, course arrangement and assignments support course objectives. | √ |
| Q2 | The course materials that the instructor prepared are good enough to grasp the concept of the course chapters. | √ |
| Q3 | The instructor clearly explains difficult materials. | √ |
| Q4 | The instructor adjusts his/her presentation to the Students' level of understanding. | √ |
| Q5 | The instructor uses inter active teaching Students' participation and questions. | √ |
| Q6 | The lecturer and course assignments are rightly matched. | √ |
| Q7 | The instructor is available on an individual basis outside of class when I request it. | √ |
| Q8 | I have learned a great deal in this class. | √ |
| Q9 | The course has increased my interest in this field of study. | √ |
| Q10 | The course depends upon understanding concepts rather than memorizing facts. | √ |
| Q11 | This course has improved my skills in problem solving. | √ |
| Q12 | This course enhances my critical thinking skill about the subject. | √ |
| Q13 | The key points of this course can be understandable. | √ |
| Q14 | Overall, the effectiveness of the instructor is good. | √ |
| Q15 | Overall, this course has been efficient to advance my learning. | √ |

(Strongly Disagree=e, Disagree =d, Neutral=c, Agree=b, Strongly Agree=a) option choice which is user selected.

Data preparation and pre-processing of the data set need data mining techniques to get better input data. Thus, this proposed system did some processing for the collected data before loading the data set to the data mining software and which irrelevant attributes should be removed.

7. Teacher Assessment Survey Frame Work

Please acknowledge collaborators or anyone who has helped with the paper at the end of the text. Figure 6 illustrates the proposed method framework. This framework starts from data collection, pre-processing, selection and then data mining methods which are association and clustering followed by the evaluation of results, finally generate rules [4].

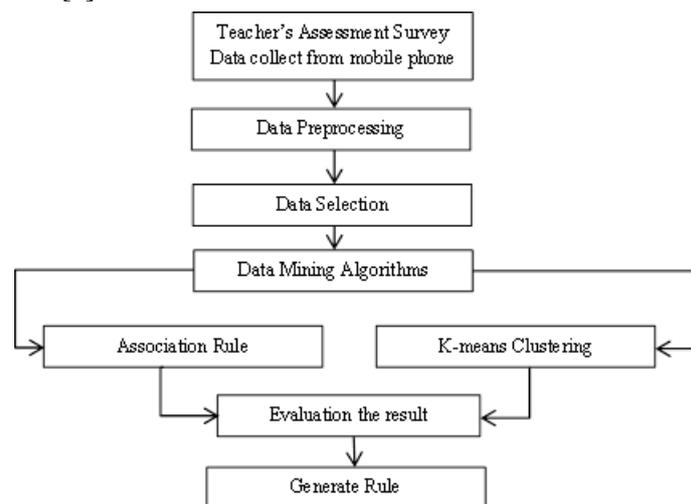


Fig. 6: Teacher's Assessment Survey Frame Work

8. Application of Results and Analysis

The proposed system model was implemented by R tool. The model was built with two machine learning algorithms: association rule mining algorithm and k means clustering algorithm. In association rule mining used teacher's assessment data set which example of items are Q3,Q5,Q6,Q8,Q12,Q13,Q14 have the Strongly Agree =a . These rules are sorted by lift metric. The result of lift value is greater than 1 indicates a

positive correlation. The proposed system used support 0.9 and confidence 1.0. Experiment result shown in figure 7, 8.

| | lhs | rhs | support | confidence | lift |
|------|----------------|------------|-----------|------------|----------|
| [1] | {q3=a, q13=a} | => {q12=a} | 0.9043191 | 1.0000000 | 1.101421 |
| [2] | {q14=a} | => {q13=a} | 0.9007199 | 1.0000000 | 1.091683 |
| [3] | {q13=a} | => {q14=a} | 0.9007199 | 0.9833006 | 1.091683 |
| [4] | {q3=a, q12=a} | => {q13=a} | 0.9043191 | 1.0000000 | 1.091683 |
| [5] | {q12=a, q13=a} | => {q3=a} | 0.9043191 | 1.0000000 | 1.088475 |
| [6] | {q12=a} | => {q13=a} | 0.9043191 | 0.9960357 | 1.087355 |
| [7] | {q13=a} | => {q12=a} | 0.9043191 | 0.9872299 | 1.087355 |
| [8] | {q5=a} | => {q13=a} | 0.9010198 | 0.9940437 | 1.085181 |
| [9] | {q13=a} | => {q5=a} | 0.9010198 | 0.9836280 | 1.085181 |
| [10] | {q12=a} | => {q3=a} | 0.9043191 | 0.9960357 | 1.084160 |
| [11] | {q3=a} | => {q12=a} | 0.9043191 | 0.9843291 | 1.084160 |
| [12] | {q5=a} | => {q3=a} | 0.9025195 | 0.9956982 | 1.083793 |
| [13] | {q3=a} | => {q5=a} | 0.9025195 | 0.9823702 | 1.083793 |
| [14] | {q6=a} | => {q3=a} | 0.9019196 | 0.9875205 | 1.074892 |
| [15] | {q3=a} | => {q6=a} | 0.9019196 | 0.9817173 | 1.074892 |

Fig. 7: strong rules results

Graph for 34 rules

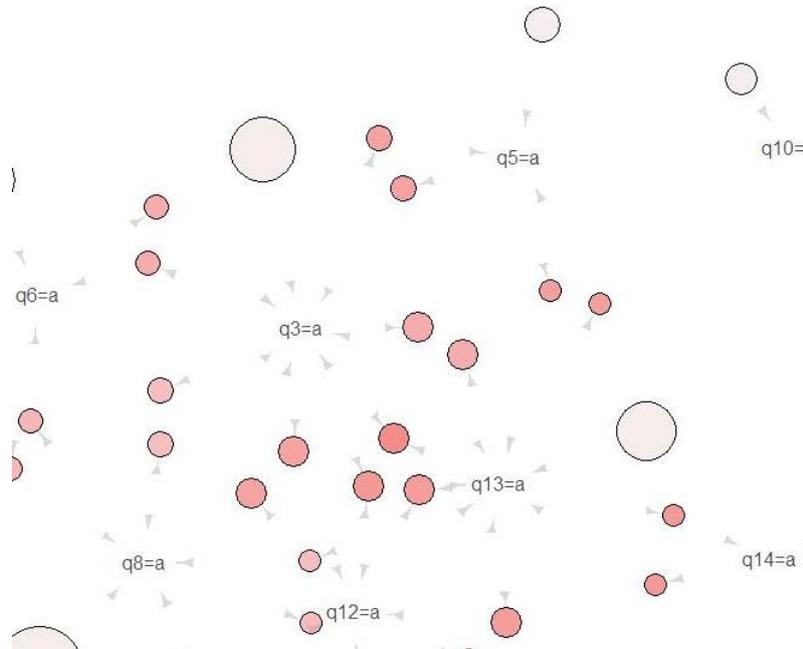


Fig. 8: Strong rules with graph

k means clustering algorithm of number of cluster k=5, and resulted output clusters are shown in figure 9. Five clusters consist of question answers with most Agree, Strongly Agree for all questions. This result is excellent of teacher's assessment survey evaluation result.

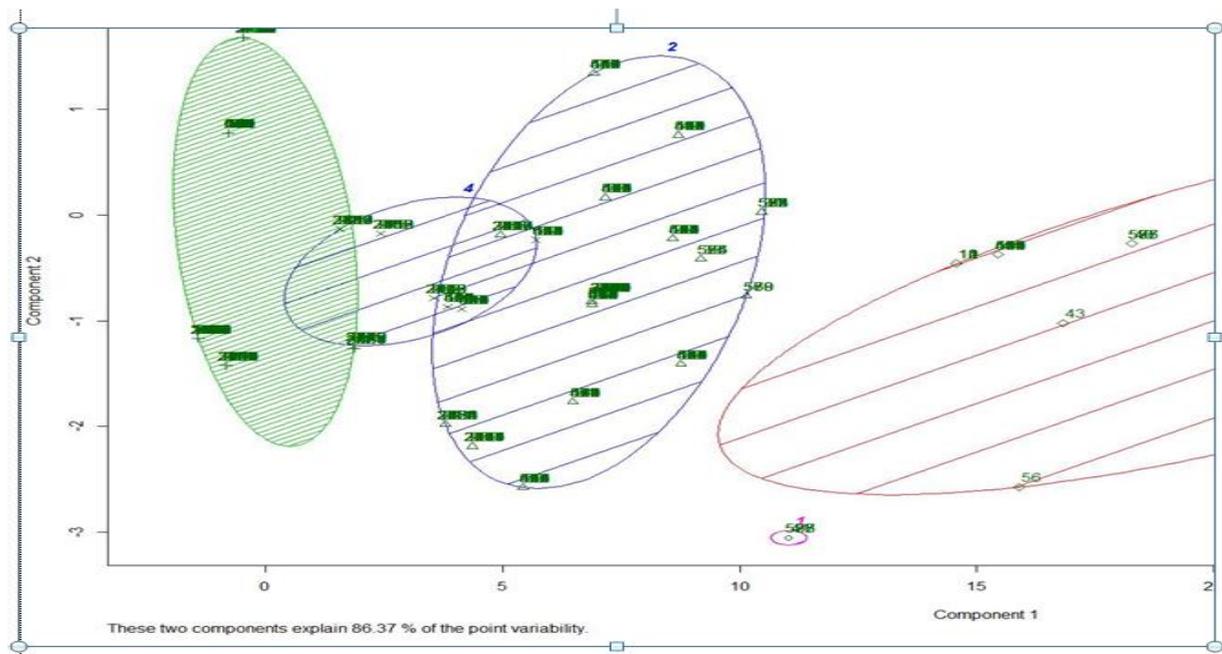


Fig. 9: Group of survey answers

9. Conclusion

In this study analyzed the teaching performance assessment, prediction of training course , lab resources and interesting on learning that will be provide to reach the high quality of education. This proposed system produced 34 association rules that could help and guide Educators to make educational decision, learning management according their student's answers survey data and efficiently manage the University admin work. During this work many questions another Universities arose that indicated directions for future research. One of these directions is modelling the other types of data mining algorithms in University of Computer Studies cloud storage.

10. References

- [1] S.Ougiaroglou and G.Paschalis. Association Rules Mining from the educational data of ESOG web-based application. Artificial Intelligence Applications and Innovations. 10.1007/978-3-642-33412-2_11. hal-01523060. Submitted on 16 May 2017
- [2] J.T.Wassan.Discovering Big Data Modelling for Educational World. Procedia - Social and Behavioral Sciences 176 (2015) 642 – 649, IETC 2014
- [3] S. Kumar. A Modern Data Mining Method for Assessment of Teaching Assistant in Higher Educational Institutions. ISSN:0975-9646. International Journal of Computer Science and Information Technologies, Vol. 8 (3) , 2017,424-429
- [4] R. K. Hemaïd, A. M. El-Halees. Improving Teacher Performance using Data Mining. ISSN (Online) 2278-1021. ISSN (Print) 2319-5940. International Journal of Advanced Research in Computer and Communication Engineering.Vol. 4, Issue 2, February 2015