AMSE JOURNALS-2016-Series: Advances A; Vol. 59; N°1; pp 129-144
Submitted July 2016; Revised Oct. 15, 2016, Accepted Dec. 10, 2016

Reform on Mathematical Modelling Teaching Contents in the Era of Big Data

*Li Wei, **Li Qiong

* School of Mathematics & Information Science, Jiang Xi Normal University,
Nan Chang, 330022, China (liweijxnu@163.com)
**Department of nursing, Jiang Xi Health Vocational College,
Nan Chang, 330052, China (16906165@qq.com)

Abstract

The reforms of teaching content about mathematical modelling are discussed in this paper. On the basis of the existing teaching content, with the data preprocessing methods used in the teaching content can solve current hot "big data" problem. The Rough Set Theory are introduced which is a data pre-processing method through the case study. In the big data era, it is necessary to pretreat the redundant information in data. It can help students solve practical problems better by use the mathematics knowledge if the teacher increase the content of redundant information's pretreatment in the process of mathematical modelling teaching.

Key words

Big data, pre-processing, mathematical modelling, rough set

1. Introduction

The rapid development of computing technology and mobile internet applications has opened up a new era of big data. Schonberger points out that the "big" in big data does not refer to the absolute number but the processing mode. After the collection of comprehensive, synthetic and complete data extracted from practical issues, suitable mathematical methods are required for its analysis and modelling. By revealing the internal laws between those data, the probability of events happening behind can be ascertained.

It is particularly important to process the plethora of redundant network data so as to realize usage of effective data. Brodie holds that the use of multi-disciplinary, multi-technology methods is necessary for meaningful data integration in the real, often schema-less, and complex big data world of databases and the semantic web [1]. The large quantity of data which cannot be processed by conventional data processing software demands new methods for pre-processing and processing. Wang et al. presents an overview on the research paradigms and ways of thinking on the development of big data. In the research they pointed out the current challenges and possible future direction of big data [2]. David Gil and Il-Yeol Song discussed the research and experience in modelling and developed systems and techniques to deal with big data. They pointed out that big data creates new requirements based on complexities in data processing [3].

Over the past twenty years or so, there have been students from the university that I was enrolled at attending the China Undergraduate Mathematical Contest in Modelling (CUMCM). Research has accordingly been conducted on instructing students in analysis by mathematical methods, application of mathematical tools and computers, as well as practical capability. This paper started from real life conditions to discuss reform on teaching contents of data processing during mathematical modelling, and illustrated data processing application by a case study.

For undergraduate students that have mastered rudimental knowledge of advanced mathematics and/or linear algebra, it is one of our largest concerns as mathematical modelling workers that what mathematical knowledge best fits them next so that they can better apply mathematical knowledge or related models to practical situations.

Regular mathematical courses for undergraduates merely involve fundamental mathematical knowledge and theories, but fail to address the issue of how to conduct research on mathematical application from simple to complex. Jiang et al. point out that a mathematical modelling contest is a condensed scientific practice [4]. Specialized courses on mathematical modelling are a necessity to discuss the way to flexibly solve practical problems by mathematical theories. Philippe J. et al. formed a set of design principles for teaching computational modelling, and independently implemented these principles in two public research universities [5]. These experiences can help us in mathematical modelling. Raul Ramirez-Velarde presented the Mathematical Modelling Learning strategy. To carry out the strategy students must follow a procedure consisting of 4 stages, and the students can create a model that will predict behavior of existing phenomena using real data [6].

Previously, undergraduate students in the university had no access to an in-depth development and application of some frequently-used mathematical methods. For recent years, to satisfy practical requirements, a new optional course, i.e. mathematical modelling, has been included for students of some majors. By educating them about basic mathematical models including ODE model, mathematical programming model, optimization model, and difference equation model, their awareness of mathematical application, mathematical analytical ability and problem-solving competent have mounted up to a great degree. Be it as it may, along with computer technology development, network popularization and the coming of the big data era, present-day mathematical modelling teaching contents have already failed to meets students' demand for data processing. Through years of mathematical content practice, we found that an increasing amount of data pre-processing knowledge has been applied to the process of mathematical modelling. Goldina Ghosh et al. discussed a novel algorithm deploying Fuzzy methodology to deal with big data paradigm [7]. Emine Ozdemir and Devrim Uzel studied the student opinions about teaching based on mathematical modelling. They pointed out Mathematical modelling can help students' better understand the world, support mathematics learning, contribute to develop various mathematical competencies and proper attitudes [8]. In reality more and more data processing knowledge are needed to deal with the rapid progress of the development of society. Therefore, apart from regular teaching contents of mathematical modelling, a series of related reforms have been undertaken in the college that is closer to reality and that can fit well for mathematical contents pertaining to popular issues at present. Big data pre-processing is one of these mathematical contents.

2. Frequent methods for data pre-processing

2.1 Data mining methods

The precursor of the analytical technology of big data in the process of mathematical modelling is the conventional data mining technology, but the essence remains as data mining. Since 1980s, research into data mining has achieved fruitful results. The data mining technology finds widespread application in plenty of cutting-edge fields. The adoption of methods in the data mining system rests on factors ranging from the type of problems to the scale and type of data. Existing major data-mining methods contain rough set theory, association rules mining, neural network, Bayesian network, decision tree, and visualization technology. After a consideration of the mathematical basis for CUMCM participants in the university as well as the faculty, teachers assembled and devised a mathematical modelling teaching scheme that conformed to the actual campus conditions. In terms of data processing, knowledge concerning the rough set theory and the neural network theory are predominated the class. Below is an explanation of the application of the rough set theory to data processing in mathematical modelling.

The theory of rough sets is first described in 1982 as a data analytical theory by Polish computer scientist Zdzisław I. Pawlak [9]. It provides an effective way to analyze and deal with imprecise, inconsistent and incomplete data in complex systems. A rough set can be used to classify vague data or noisy data as a way to reveal internal relationships. Knowledge reduction is one of the essential contents of the rough set theory.

It is well known that knowledge (attributes) in a given knowledge base is of unequal difference, let alone redundant one. Knowledge reduction refers to the deletion of unrelated or unimportant knowledge without affecting knowledge representation ability. As the quintessence of the rough set theory, knowledge reduction provides an access to knowledge in an information system. In this theory, knowledge reduction can be divided into attribute reduction and attribute-value reduction. The latter one is much simpler by comparison, thus attracting rare academic attention. The broad sense of knowledge reduction amounts to attribute reduction.

2.2 Introduction to the classical knowledge reduction algorithm

(1) Discernibility matrix method

In this method, a discernibility matrix is firstly constructed to obtain a discernibility function. Then, with logical operations, the function is simplified as a disjunctive normal form [10], for which each item corresponds to a reduction. Despite the ability to obtain all reductions in an information system, the discernibility matrix method is of high time-complexity.

(2) Information entropy method

From the perspective of information, Miao Duoqian, a Chinese scholar, measures the significance of attributes in a decision table. His proposal of a transinformation-based heuristic algorithm pertaining to relative knowledge reduction is polynomial. By a case study, he illustrates the point that the information entropy method can spawn the minimum reduction in a decision table, albeit unsure about the completeness [11, 12, 13]. Wang Guoyin et al. devised a heuristic algorithm that reduced the knowledge of a decision table, and verified its effectiveness and usability by a simulation test [14]. Liang Jiye et al. introduced the concept of information quantities to the information system, and accordingly put forward an information-quantity-based attribute reduction algorithm. However, they neglected its application to decision tables [15].

(3) A heuristic algorithm on the basis of attribute significance

This algorithm begins to solve the reduction problem with the core. The purpose is to seek for the best reduction or the user-specified minimum reduction (i.e. with the least number of attributes) [16]. The heuristic principle of this algorithm is attribute significance. First, attributes are included according to their importance in descending order, until the classification ability or decision-making ability of the corresponding subset agrees with that of the entire (conditional) attribute set. Then, a reversion elimination method is used to check the necessity of each attribute in the subset, removing unnecessary ones. In this way, it is guaranteed that the final result is a single reduction. This algorithm has low time complexity, but cannot ensure the finally obtained reduction is the minimum.

(4) Genetic algorithm

Generally, binary encoding helps realize reduction in this method, where “1” means to choose a corresponding attribute for a point while “0” means the opposite [17, 18]. Fitness function is largely represented by the length of an attribute set, or by the ability of an attribute set to classify or decide. The former one hope to minimum the number of attribute sets as much as possible, whilst the latter one hope to classify attribute sets as many.

(5) Other feasible algorithms

The above heuristic algorithms promote the development of many derivatives, such as feature selection method [19] and characteristic matrix method [20].

3. Data pre-processing cases

Here is a case to deal with redundant data by means of rough sets.

Below is a table of actual meteorological conditions in a city, with four reference data (weather, temperature, humidity, wind) [21]. Please judge whether it contains any redundant data.

Table 1 The observed meteorological conditions in a fortnight in a city

U	weather (a_1)	temperature (a_2)	humidity (a_3)	windy (a_4)	Weather category (d)
1	sunny	high	high	No	N
2	sunny	high	high	Yes	N
3	cloudy	high	high	No	P
4	rainy	warm	high	No	P
5	rainy	cool	moderate	No	P
6	rainy	cool	moderate	Yes	N
7	cloudy	cool	moderate	Yes	P
8	sunny	warm	high	No	N
9	sunny	cool	moderate	No	P
10	rainy	warm	moderate	No	P
11	sunny	warm	moderate	Yes	P
12	cloudy	warm	high	Yes	P
13	cloudy	high	moderate	No	P
14	rainy	warm	high	Yes	N

3.1 The basic concepts in the rough set theory

Here are some basic concepts in the rough set theory.

Proposition 3.1 [9] The tetrad $S=(U,A=C\bigcup D,V,f)$ is an information system, where U denotes a finite nonempty set of objects, also known as the domain of discourse. $A=C\bigcup D$ is a finite nonempty set of attributes. C is the conditional attribute set, D is the decision attribute set, $C\bigcap D=\varnothing$. V is the attribute range, $f:U\times A\to V$ represents a mapping. we will say that the information system is a data table or a knowledge base, if $D=\varnothing$, otherwise, the information system is a decision table. If there exists an unknown number $f(x,a)=*$, $x\in U,a\in C,f(x,a)$, the information system will be incomplete, or otherwise it will be complete. The information system is in short $S=(U,A)$.

In the above case, the discourse domain is $U=\{1,2,...,14\}$, the conditional attribute set is $C=\{a_1,a_2,a_3,a_4\}$,and the decision attribute is d. Therefore, the information system is complete and is seen as a decision table.

Proposition 3.2[9] $IND(R)=\{(x,y): f(x,c)=f(y,c), \forall c \in R \subseteq A\}$ is an indiscernible relation controlled by *R*, and is in short *R* or knowledge *R*. Evidently, *R* can be reflexive, symmetric and transitive, as a family of equivalence relations in *R*. *R* divisions are represented by $U/R=\{[x]_R : x \in U\}$,where $[x]_R=\{y:(x,y)\in R\}$ denotes the equivalence family of *R* that contains $x \in U$. $[x]_R$ is also called as a granularity of *R*.

Proposition 3.3[9] We assume that $R \subseteq A, Y \subset U,$ then the upper and lower approximation of the set *Y* are $\underline{R}(Y)=\{x\in U \mid [x]_R \subseteq Y\}$, $\bar{R}(Y)=\{x\in U \mid [x]_R \cap Y \neq \varnothing\}$ (1)

Obviously, $\underline{R}X$ is the set of all elements of *U* which can be with certainty classified as elements of *X*, while $\bar{R}X$ is the set of elements of *U* which can be possibly classified as elements of *X*.

Proposition 3.4[9] We assume that $R \subseteq A, X \subseteq U$,then $BN(X)=\bar{R}X-\underline{R}X$ is the *R*-boundary of *X*. $BN(X)$ is the set of elements which cannot be classified either to *X* or to –*X*, employing knowledge *R*.

Proposition 3.5[9] We assume that $P,Q \subseteq A$,then $POS_P(Q)=\bigcup_{X\in U/Q} \underline{P}(X)$ is defined as the positive region *P* of *Q*.

Proposition 3.6[9] We assume that $R \subseteq A$, $\varnothing \neq X \subseteq U$,then $\alpha_R(X)=\frac{|\underline{R}X|}{|\bar{R}X|}$ is the approximation quality of the set *X* described by *R*, where $|\square|$ represents the cardinality.

Precision $\alpha_R(X)$ is *R*-definable degree in the set *X*. Evidently, for each *R* and $X \subseteq U$, there is $0 \le \alpha_R(X) \le 1$. In fact, a rough set is imprecise due to *R*-boundary. The broader the *R*-boundary is, the less accurate the rough set is, and the smaller $\alpha_R(X)$ becomes. When $\alpha_R(X)=0$, *X* cannot be defined in the knowledge base *R*. When $\alpha_R(X)=1$,*X* can be exactly defined in this knowledge base.

Proposition 3.7[9] We assume that P,Q is a pair of equivalence relations in *U*. If $P \subseteq Q$, U/P is finer than U/Q, which is expressed as $U/P \preceq U/Q$. It also means *P* is finer than *Q*. If $P \subset Q$, then $U/P \prec U/Q$.

Apparently, for the information system $S=(U, A=C \bigcup D, V, f)$, if $P \subseteq A, Q \subseteq A$ and $P \subseteq Q$, then $U/P \preceq U/Q$.

Finer equivalence relations produce smaller equivalence classes. The finer the classification is, the more knowledge the classification element has, and the higher the granularity is.

3.2 Concepts of relative reduction

First is the introduction of two fundamental concepts in reduction of knowledge — a reduction and the core.

Proposition 3.8[9] $S=(U, C \bigcup D, V, f)$ is an information system. For an arbitrary attribute $a \in C$, a is dispensable in C if $IND(C)=IND(C-a)$, otherwise a is indispensable in C. C is independent if each $\forall a \in C$ is indispensable in C, or otherwise C is dependent. Suppose $P \subseteq C$ and *P* is independent, such that $IND(P)=IND(C)$, then *P* is a reduction of C, i.e. $RED(C)$. *P* will be called the core of C, if it represents the set of all indispensable relations in C, and will be denoted $core(C)$.

Obviously $core(C)=\bigcap RED(C)$.

Proposition 3.9[10] $S=(U, C \bigcup D)$ is a decision system, $R \subseteq P \subseteq C, Q \subseteq D$. *R* is the *Q*-dispensable element of *P*, if $POS_P(Q)=POS_{(P-R)}(Q)$, otherwise *R* is *Q*-indispensable. *P* is *Q*-independent if each attribute of *P* is *Q*-indispensable, or otherwise *P* is *Q*-dependent. *R* is the *Q*-reduction in *P*, if $R \subseteq P$ is a *Q*-dependent subset in *P* and $POS_P(Q)=POS_R(Q)$. In this case, *R* is briefed as dependent reduction and is denoted $RED_Q(P)$. The set of all *Q*-indispensable attributes in *P* is called the *Q*-core of *P*, and will be denoted $core_Q(P)$.

Obviously $core_Q(P) = \bigcap RED_Q(P)$. There is not necessarily a single dependent reduction in the decision system $S = (U, C \bigcup D)$.

3.3 Attribute reduction based on harmonic mean of attribute importance

3.3.1 Propositions and properties

Proposition 3.10[9] Suppose $S = (U, A = C \bigcup D)$ is a decision system, and

$U/_{IND(D)} = \{Y_1, Y_2, ..., Y_k\}$ □ $D, P \subseteq C, Y \subseteq U$, then $\alpha_P(D) = \dfrac{\sum_{i=1}^{k} |\underline{P}(Y_i)|}{\sum_{i=1}^{k} |\overline{P}(Y_i)|}$ is the rough approaching approximation of classification *D* measured by the attribute set *P*, and $\gamma_P(D) = \dfrac{\sum_{i=1}^{k} |\underline{P}(Y_i)|}{|U|}$ is the approaching approximation of classification *D* measured by the attribute set.

$\alpha_P(D)$ is the percentage of correct decisions in all possible decisions, employing *P*-classification. $\gamma_P(D)$ is the percentage of *P*-classifications that can be with certainty included in *D* class.

Property 3.1[9] suppose P, R partition $U/_P, U/_R$ satisfies $U/_R \preceq U/_P$, then $\gamma_R(D) \geq \gamma_P(D), \alpha_R(D) \geq \alpha_P(D)$.

Property 3.2[9] We assume that P, R partition $U/_P, U/_R$ satisfies $U/_R \preceq U/_P$. *R* is a dependent reduction of the decision attribute *D*, if $\gamma_P(D) = \gamma_R(D)$ and *R* is the minimum set that satisfies this equation.

We proposed new definition about the importance of the attribute as follows:

Proposition 3.11 Suppose there is a decision system $S = (U, C \bigcup D)$, $U/_{IND(D)} = \{Y_1, Y_2, ..., Y_k\}$, $B \subseteq C$,

then $$sig_B^{(1)} = \frac{\sum_{j=1}^{k} \alpha_B(Y_j) + \alpha_B(D) + \gamma_B(D)}{k+2} \quad (2)$$

is the importance of the attribute set *B*. Particularly, if $B = \{c_i\} \subseteq C$,then $sig_{\{c_i\}}^{(1)}$ is denoted

c_i importance.

Obviously $0 \le sig_B^{(1)} \le 1$.

Property 3.3 Suppose there is a decision system $S = (U, C \cup D)$, $U/IND(D) = \{Y_1, Y_2, ..., Y_k\}$, $P \subseteq R \subseteq C$. If P, R partition $U/P, U/R$ satisfies $U/R \preceq U/P$, then $sig_R^{(1)} \ge sig_P^{(1)}$.

Proposition 3.12 For the decision table $S = (U, C \cup D)$, $B \subseteq C$, $U/IND(B) = \{X_1, X_2, ..., X_n\}$, we have

$$I(B) = 1 - \frac{1}{|U|^2} \sum_{i=1}^{n} |X_i|^2 \tag{3}$$

which represents the information volume provided by the attribute set *B*.

Evidently, the information volume of a single attribute c_i is $I(c_i)$. In a general sense, a larger $I(c_i)$ is more helpful to decision-making.

Property 3.4 $c_i \in B \subseteq C$ is B-indispensable $\Leftrightarrow I(B \setminus \{c_i\}) < I(B)$.

Property 3.5 $B \subseteq C$ is independent $\Leftrightarrow \forall c_i \in B, I(B \setminus \{c_i\}) < I(B)$.

Property 3.6 $core(B) = \{c_i \in B \mid I(B \setminus \{c_i\}) < I(B)\}$.

Property 3.7 $R \subseteq B \subseteq C$ is a reduction of B $\Leftrightarrow I(R) = I(B)$, and $\forall c_i \in B, I(B \setminus \{c_i\}) < I(B)$.

Proposition 3.13 For decision table $S = (U, C \cup D)$, $B \subseteq C$,we have $I(D \mid B) = I(B \cup D) - I(B)$ as the conditional information volume of the attribute set *B*.

Proposition 3.14 For decision table $S = (U, C \cup D)$, $P \subseteq B \subseteq C$,we have $sig_B^{(2)}(P) = I(D \mid B \setminus P) - I(D \mid B)$ as the *D*-importance of the attribute set *P* in *B*. Particularly, if *P* is a single point set, i.e. $P = \{c_i\}$,then $sig_B^{(2)}(\{c_i\}) = I(D \mid B \setminus \{c_i\}) - I(D \mid B)$. If a decision table only has a single attribute $C = \{b\}$, then the *C*-importance of $\{b\}$ is $sig_C^{(2)}(\{b\}) = I(D \mid \{b\})$.

This means that the importance of c_i in B is measured by the change of information volume when c_i is removed. For an attribute set, the larger such change is, the greater importance the attribute has, and vice versa.

Proposition 3.15 For decision $S = (U, A = C \cup D)$,the harmonic mean *D*-importance of $B \subseteq C$ in

C is expressed as

$$sig(B)=\begin{cases}\dfrac{1}{\frac{1}{2}\left(1/sig_B^{(1)}+1/sig_C^{(2)}(B)\right)} & sig_B^{(1)}\neq 0, sig_C^{(2)}(B)\neq 0\\ 0 & else\end{cases} \tag{4}$$

Property 3.8 Suppose P,R partition $U/P, U/R$ satisfies $U/R \preceq U/P$, then $sig(P)\leq sig(R)$

3.3.2 Attribute reduction algorithm

According to Proposition 3.15, an attribute with the maximum harmonic mean importance is the core of the decision table. We view it as the starting point of reduction solution, and accordingly construct the attribute reduction algorithm.

Input: A complete decision table $S=(U,A,V,f)$.

Output: A reduction of $S=(U,A,V,f)$.

Step 1: Calculate $\gamma_C(D)$ of the conditional attribute set C.

Step 2: According to Preposition 3.15, we calculate the C-importance $sig(\{c\})$ of all attributes. $c\in C$, and denote it as $R=\max_{c\in C} sig(\{c\})$. If $\gamma_C(D)=\gamma_R(D)$, then R is a reduction and the algorithm stops. If not, go to step 3.

Step 3: For the attribute set $C\setminus R$, we repeat the following steps:

1) calculate $sig(\{R\cup\{c\}\})$ for each $c\in C\setminus R$,according to Preposition 3.15.

2) select c_0 that satisfies $sig(\{R\cup\{c_0\}\})=\max_{c\in C\setminus R} sig(\{R\cup\{c\}\})$, and let $R\cup\{c_0\}=R$.

3) judge whether $\gamma_C(D)=\gamma_R(D)$. We output R if $\gamma_C(D)=\gamma_R(D)$, otherwise we go back to step1.

3.4 Solution to problems

Intuitively, we convert the observation results into numbers.

a_1 : sunny—0 cloudy—1 rainy—2. a_2 : hot—0 warm—1 cool—2. a_3 : high—0 moderate—1. a_4: yes—0 no—1 .

Below is the solution to the minimum related reduction using the said attribute reduction method.

Step 1: After calculation, we have $\gamma_C(D) = \sum_{i=1}^{2} |\underline{C}(Y_i)| \Big/ |U| = 14/14 = 1$.

Step 2: $I(C) = 1 - \frac{14}{14^2} = \frac{13}{14}$. According to Preposition 3.6, we calculate $sig(\{a_1\})$, $sig(\{a_2\})$, $sig(\{a_3\})$, $sig(\{a_4\})$, and the corresponding result is $sig(\{a_1\}) = 0.0668$, $sig(\{a_2\}) = 0$, $sig(\{a_3\}) = 0$, $sig(\{a_4\}) = 0$, let $R = \{a_1\}$, calculate $\gamma_R(D) = \frac{\left|\sum_{i=1}^{2} \underline{R}(Y_i)\right|}{|U|} = \frac{4}{14} \neq \frac{14}{14}$, which means R in this case is not the minimum.

Step 3: For $C \setminus R = \{a_2, a_3, a_4\}$, according to Preposition 3.6, we calculate $sig(\{a_1, a_2\})$, $sig(\{a_1, a_3\})$, $sig(\{a_1, a_4\})$, and the result shows: $sig(\{a_1, a_2\}) = 0.1336$, $sig(\{a_1, a_3\}) = 0.1252$, $sig(\{a_1, a_4\}) = 0.155$,

Therefore, R-importance of all attributes in descending order is $sig(\{a_1, a_4\}) > sig(\{a_1, a_2\}) > sig(\{a_1, a_3\})$. We let $R = \{\{a_1\} \cup \{a_4\}\}$ and calculate

$$\gamma_R(D) = \frac{\left|\sum_{i=1}^{2} \underline{R}(Y_i)\right|}{|U|} = \frac{9}{14} \neq \frac{14}{14}$$, which means R in this case is not the minimum.

Then, we compare the R-importance of a_2 and that of a_3.

$sig(\{a_1, a_2, a_4\}) = 0.2794 > sig(\{a_1, a_3, a_4\}) = 0.2655$ We let $R = \{\{a_1, a_4\} \cup \{a_2\}\}$ and calculate

$$\gamma_R(D) = \frac{\left|\sum_{i=1}^{2} \underline{R}(Y_i)\right|}{|U|} = \frac{14}{14} = \gamma_C(D)$$ which means R in this case is the minimum reduction.

The above analysis result indicates that the redundant weather information of this city is humidity (a_3). Problems with large data volume can be tackled with the proposed algorithm in the paper. Through programming for the algorithm, attribute reduction can be realized as a way to eliminate redundant information.

3.5 Discussion

Rough Set Theory can derive the classification or decision rule from knowledge reduction. A new definition of attribute's importance is used in above research to construct the algorithm of attribute reduction. From the case study, the validity of this algorithm is verified. For the teaching of mathematical modeling, this method of data processing is a simple and practical way. It is a good content in the reform of mathematical modeling teaching. In the process of teaching, we can choose some popular method combine with the actual situation to achieve good teaching effect.

4. Conclusion

Civilization is an evolution from understanding of the real world to the creation of InfoWorld. By obtaining knowledge from data-centered big data, we make breakthroughs in sample randomness [22-24]. Big data is a complicated data set with massive, diversified and fast-changing data [25-28]. In this paper, the attribute reduction algorithm in the rough sets philosophy is used to delete redundant information as data pre-processing, whose effectiveness is verified by a case study.

The course of mathematical modelling covers a wide variety of mathematical theories and practice. In addition to imparting fundamental mathematical approaches and mathematical programming to students, teachers are supposed to introduce some in-depth, large-span mathematical theories in class, which will be much better to fit for present social focuses. The reason is that students will better tackle practical problems with a mathematical thought. Therefore, mathematical modelling instructors should intensify expertise and keep up with the time. Moreover, they are suggested to continue absorb teaching experience from teachers in other universities and colleges by stronger interschool communication, as a helpful way to explore relative teaching methods and contents of higher quality.

Acknowledgments

This research was financially supported by the Jiangxi Province Society Project Planning (No.15GL44) and Jiangxi Province educational reform project of degree and postgraduate education (No.JXYJG-2015-048)

References

[1] Bizer C, Boncz P, Brodie M L, et al. The Meaningful Use of Big Data: Four Perspectives-Four Challenges, 2012, ACM SIGMOD Record, vol.40, no. 4, pp.56-60.

[2] Hai Wang, Zeshui Xu, Hamido Fujita, et al. Towards felicitous decision making: An overview on challenges and trends of Big Data, 2016, Information Sciences, vol. 367-368, pp.747-765.

[3] David Gil, Il-Yeol Song, Modelling and Management of Big Data: Challenges and opportunities, 2016, Future Generation Computer Systems, vol.63, pp.96-99.

[4] Jiang L.P., Dong Y.C. Unique charm of Mathematical modelling contest, 2002, Mathematics in Practice and Theory, vol.32, no.2, pp.351-352.

[5] Philippe J., Giabbanelli, Vijay K., Mago. Teaching computational modelling in the data science era, 2016, Procedia Computer Science, vol. 80, pp. 1968-1977.

[6] Raul Ramirez-Velarde, Nia Alexandrov, Raul Perez-Cazares et al. Mathematical Modelling Based Learning Strategy, 2015, Procedia Computer Science, vol. 51, pp. 1694-1704.

[7] Goldina Ghosh, Soumya Banerjee, Neil Y. Yen, State transition in communication under social network: An analysis using fuzzy logic and Density Based Clustering towards big data paradigm, 2016, Future Generation Computer Systems, vol. 65, pp. 207-220.

[8] Emine Ozdemir, Devrim Uzel, Student Opinions On Teaching Based On Mathematical Modelling, 2012, vol. 55, pp. 1207-1214.

[9] Pawlak Z. Rough Set theory and its applications to data analysis, 1998, Cybernetics and System, vol.29, no.7, pp.661-688.

[10] Baowen Xu, Yuming Zhou, HongminLu, An improved accuracy measure for rough sets,2005, Journal of Computer and System Sciences, vol. 71, pp.163-173.

[11] Miao D.Q., Hu G.R., A heuristic algorithm of knowledge reduction. Journal of Computer Research and Development, 1999, vol. 36, no. 6, pp. 681-684.

[12] Miao D.Q., Wang J., Presentation of Information about Concepts and Operation in Rough Set Theory, 1999, vol.10, no.2, pp.113-116.

[13] Li C.Z., Wu G.X., Yan W.F. et al., Conversion Method of Covering Rough Set Model, 2011, Journal of JiangXi Normal University(Natural Science Edition), vol.35, no.4, pp.391-395.

[14] Wang G.Y., Yu H., Yang D.Q., Reduction of decision table based on conditional information

entropy, 2002, Chinese Journal of Computers, vol. 25, no. 7, pp. 759-766.

[15] Liang J.Y., Qu K.S., Xu Z.B., Attribute Reduction of information systems, 2001, System Engineering Theory and Practice, no. 12, pp. 76-80.

[16] Skowron, A., Rauszer, C. The discernibility matrices and functions in information systems, In: Slowinski, 1991, Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, pp.331-362.

[17] Hu X. Knowledge discovery in databases: an attribute-oriented rough set approach, 1995, Doctoral dissertation, University of Regina, Canada.

[18] Li Q.H., Zhao W.D., The application of Genetic Algorithm in the decision table reduction, 2001, Computer Engineering, vol.27, no.2, pp. 80-81.

[19] Huang B., Zhou X.Z., Zhang R.R. Attribute Reduction of Incomplete Information System Based on Information Content, 2005, no. 4, pp. 55-60.

[20] Xu Y., Huai J.P., Wang ZH.J., Heuristic reduction algorithm and its application based on the size of distinguish ability, 2003, Chinese Journal of Computers, vol. 26, no. 1, pp.97-103.

[21] Wang Y.Y., Zhang CH.K., Shao H.H., Research and Application of heuristic algorithm for knowledge reduction, 2001, Control and Decision, no.11, pp. 886-889.

[22] Burstein F. HOLSAPPLE C. W. Handbook of Decision Support System, 2008, Berlin: Springer, pp. 46-48.

[23] Cressie N. statistics for Spatial Data, 1993, Revised edition. New York: John Wiley and Sons Inc, pp.107-109.

[24]Haining R. Spatial Data Analysis: Theory and Practice, 2003, Cambridge: Cambridge University Press, pp.225-226.

[25] United Nations Global Pulse. 2012, Big Data for Development: Challenges $ Opportunities, 2012.

[26] Office of Science and Technology Policy. Executive Office of the President, Fact Sheet: Big Data across the Federal Government, www.WhiteHouse.gov/OSTP, access date Dec.2012.

[27] Office of Science and Technology Policy. Executive Office of the President, Obama Administration Unveils ”Big Data” Initiative: Announces $200 Million in New R$D Investments. www.WhiteHouse.gov/OSTP, access date March 2012.

[28] Mckinsey Global Institute. 2011 Big Data the Next Frontier for Innovation, Competition, and Productivity, 2011, United States, pp. 55-57.