ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ROUGH SET BASED CONTEXT SUGGESTIONS

[1] **ENAS FADHIL ABDULLAH** , [2] **GHAIDAA A. Al-SULTANY** , [3] **HUDA NAJI NAWAF**

[1]University of Kufa, Faculty of Education for Girls, Dept. Computer Science, Iraq
[2]Universty of Babylon, Faculty of Information Technology, Dept. Software, Iraq
[3]Universty of Babylon, Faculty of Information Technology, Dept. Software, Iraq

E-mail:[1] inasf.alturky@uokufa.edu.iq,[2]ghaidaa.almulla@itnet.uobabylon.edu.iq, [3]halmamory@itnet.uobabylon.edu.iq

ABSTRACT

The classification the progression from splitting the objects on the basis of some criteria. On various occurrences, the class of each object is given in progress then it becomes easy to collection the objects in to their classes. This type of classification is called supervised classification. Rule-based classifiers such as rough set classifiers provide rules that basis classify classes of items context such as (social and location).In this paper exploited rough set theory fundamental for context suggestion as contribution and comparing results with classification methods are J48, K-nearest neighbor (K-NN), and decision stump (DS).

Keywords: *CARS, Rough Set Theory, Context Suggestion*

1.INTRODUCTION

Recommender systems (Rs) can be permit to clarified programs that is enterprise in order to recommend the most appropriate items “:products or services” to exceptional users “individuals or businesses” and the prediction of a user's service in an item occur depends on the closely related information in the items, the connection among items and users and the users themselves. Generally, the two major forms of recommendations are: user recommendation and item recommendation. An example of user recommendation is “social recommender in Facebook and Twitter” while an example of item recommendation are “recommender of products in Amazon” and “recommender of movie by Netflix” [1]. By returning the most pertinent services from a huge amount of data, recommending system process is looking for to minimize information overload, that way providing individualized assistance. The most important characteristic of (Rs) is the ability to speculate a user's interests and preferences via the analyzing of user attitude or the other users attitude in order to get “personalized recommendations (PR)” [2]. The “Context-Aware Recommender Systems” are personalization systems emphasis on recommender systems, they are more discriminate in the development of conventional recommender systems [3]. The traditional systems mostly provide recommendations for the users about the products and services, the "CARS" vary from the conventional recommendation mechanisms as it takes into account the combination of users’ rating or items interactions with the case where ranking or interactions were generated. Also, it depends on user’s sides and his/her context during purchase time [4, 5].

similarly as in the conventional recommending systems, the aim of a "context-aware recommender" is to recommending of an item or user such as in case of social recommendation to a objective user. There are numerous scenarios where the final decision of a user on whether to choose an item might be count on contexts the item might be utilized or consumed [6] . Meanwhile, the existence of CARS brought chances for new recommendation context, like context proposition that recommends a list of suitable contextual situations for the users to consume the items. usually context is defined as, “ each information that will be utilized to describe the entity situation”.The context recommender is a narrative recommender system which target to propose an appropriate contexts list to the end users to increase the total test on the items consumption for example "listening to a song" [7].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.RELATED WORK

New recommendation opportunities that featured from development CARS ”such as” context suggestion which propose an appropriate contexts list to the end users for the items consumption. In [7] the research discussed the direction context suggestion or recommender that extension for context aware-recommender (CARS), the work compared between item and context recommender and used user splitting approach for context suggestion.

The work in [8] presented recommender process that focused user-oriented context suggestion based on user profile for appropriate recommender context, “such that” proposed to construction various recommenders to propose contexts at first, by the original multidimensional training set. Secondly, via the 2-dimensional testing set which is derived for evaluation goal.

The authors [9] worked in their research on context recommender by inserting the concept of best context", i.e., the contextual conditions most convenient for a particular item to be recommended. they proposed three personalized best context prediction methods that are based on the classical user-to-user collaborative filtering approach.

The context recommender has been used in [1] as set of tools for assistance decision-making for the user. The research presented the problem of context recommendation and explored the potential implementations of the concept. they specified two categories of algorithms to address the problem: "direct context prediction" and "indirect context recommendation". they presented and evaluated various direct context prediction algorithms depends on "multi-label classification (MLC)".In this paper, context suggestion has been proposed to recommend set of items with a list of contexts that are more appropriate to the items’ substances. The Rough Set Theory (RST) Method has been used for extracting the decision rules for context suggestion.

3.Feature Selection

Basically, the purchased items are represented in terms of their content and context. However, those items may have many contents (features or attributes) in addition to the context. Accordingly, the selection procedure for the most effective and significant features is essentially required. One of the methods that are widely used for features selection is the Information Gain as illustrated in the Algorithm “(1)”.

Algorithm (1) : Feature Selection using Information Gain

Input: *k-features*
Output: *gain of feature (gof)*
Begin
1: *Read class and values of class* // *class:rating*
2: $f_c \leftarrow$ *calculate frequency for each value in class* //*c :number of value in class*
3: *Calculate probability of frequency* $p_c\,(f_c) = \frac{fc}{N}$ // *N total number of instances*
4: *Calculate entropy of class E(C)* $=-\sum_{j=1}^{c} pc\ log2(pc)$
5: while k > 0 do
6: *for each value in feature do*
7: $f_V \leftarrow$ *calculate frequency for each value in feature*
8: *Calculate probability of frequency* $p_V(f_V) = \frac{fV}{N}$
9: *Calculate entropy of feature E(F)* $=-\sum_{j=1}^{v} pv\ log2(pv)$
10: *end for*
11: *Calculate gain of feature by Gain (F)=E(C)-E(F)*
12: *End while*
13: *Descending arrange for gain values*
14: *Select value with high rank*
15: *Return gof*
End

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. CONTEXT RECOMMENDER

Context recommender is the development of CARS, the first effort was contrived in [9]. They strained to guess the better contexts for users to music listen. There are different types of context recommender are context suggestion (CS), bundle suggestion, and rich suggestion. Items prediction is the typical outcome of traditional recommendation and CARS, whereas, context recommendation output is lists of predicted contexts. The Differences between the two recommendation mechanisms can be depicted through the different inputs and outputs that clarify author in [7], as shown in Table "(1)".

Table1. Recommender Summary Of Item & Context

Recommender		**Inputs**	**Outputs**
Item Recommender		user contents, user contexts	Item lists
Context Recommender	**Context Suggestion**	user	Context lists
		item	**Context lists**
		user, item	Context lists
	Bundle Suggestion	user	items and contexts
		item	users and contexts
	Rich Suggestion	user, item	contexts and items
		user, item	contexts and users

5. ROUGH SET THEORY (RST)

RST is specifically for discover relationships in data. This series of actions are regularly named "knowledge discovery". rough sets are inner importance approximation sets and established on "ordinary sets" and It is adequate to reasoning around unclear or incomplete data [10]. in more than thirty years RST has been explored and many achievements has been made on numerous areas such as decision analysis, knowledge discovery in database. furthermore, the rough sets can be used to optimize many existing soft computing methods not only to work on new vague information systems [11]. The rough set analysis goal is to induction approximations of meanings. It can be utilized for decision rule generation, feature extraction , feature selection. The RST basic concepts are described as below:

A. Information System (IS)

The term of IS that constructed by the 4-tuples (quadruple system) S = (C,A,E,f), where C is a finite set of objects(O), A is a finite set of attributes, E=Cq ∈ A Eq ,where Eq is field of the attribute q and

f: C × A ⟶ E is total function such that p(m,q) ∈ Eq for every q ∈A, Ý ∈ C called information function. Any pair (q,v),q ∈ A, E ∈ Cq is called descriptor in S.

B. Indiscernibility Relation (IR)

Indiscernibility Relation(IR) is central concept in Rough Set theory, and is considered as a relation between two objects or more, where all the values are identical in relation to a subset of considered attributes, each subset B □ A of attributes determines a binary relation INDC(B) called indiscernibility relation, defined as follows in Eq. "(1)":

$$INDC(B) = \{(m, n) \;\square\; O^2 | \forall a \in B,\ f(m, a) = f(n, a)\} \quad (1)$$

Where m and n refer to rows has same values. Given any B □ A, relation INDC(B) induces a partition of object (O), which is denoted by Ý | IND(B), where an element from Ý | INDC(B) is called an equivalence class or elementary set.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

C. Decision tables

The decision tables include two categories of features or attributes described as the condition attribute (CA) and decisions attribute (DA).

D. Approximation sets

Pompous concept in RST, being connected with the signifies of the approximations topological operations, discussed in three concepts [12].

1) *Lower Approximation(LA)*: A explanation of the purview objects that are recognized with certainty belong to the subset of attention. Let C be a "finite set" of objects and $R \subseteq C \times C$ be an equivalence relation on C. Let $R^* = \{Ý_1, Ý_2, ..., Ý_n\}$ indicate the partition convinced by R, where term equivalence class of (RST) represented in relation (R^*). $Ý_i$ is also called an elementary set of A. Term a definable set relized for any finite union of elementary sets. Let Ý be any subset of C. Then define lower approximation of Ý as show in Eq." (2)":

$$\underline{LA}\,(Ý) = \cup Ý_i \subseteq Ý\; Ý_i \qquad (2)$$

In words, $\underline{LA}$ (Ý) is definable set "the union of all the elementary sets of A", where each elementary set is totally included (i.e., subset) in Ý.

2) *Upper Approximation(UA):* The depiction of the objects that perhaps belong to the subset of attention. The upper approximation set of a set Ý regarding R is the set of all of objects which can be perhaps classified with Ý apropos R, then describe upper approximation of Ý in A as clarify in Eq. "(3)":

$$\overline{UA}\,(Ý) = \cup Ý_i \cap Ý \neq \emptyset Ý_i \qquad (3)$$

$\overline{UA}$ (Ý) is definable set "the union of all the elementary sets of A", where each elementary set contains at least one element in Ý.

3) *Boundary Region(BR):* Is the set of all the objects of a set Ý apropos R, which cannot be classified neither as Ý nor - Ý apropos R. If the BR is a set Ý =∅ , then the set is "considered Crisp", that is, exact in relation to R; otherwise, if the BR is a set Ý ≠ ∅ the set Ý "Rough is considered". In that the boundary region is BR = *UA* - $\underline{LA}$.

6. ROUGH SET BASED CONTEXT RECOMMENDER (RSCR)

Building context predict model for items under address context suggestion, inputs are items and outputs lists of contexts as show in "Table 1" process based on principle of rough set theory as contribution in research. Challenge in this method is uncertainty of items context, therefor based on item contents (attributes such as genres of item) to conclude classes of decisions (such as class home of decision location) for items. Details of RSCR method are presented as Algorithm"(2)".

Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Algorithm (2) : Rough Set based Context Recommender (RSCR)

```
Input:    Number of instance (O ), Number of Attributes (A ), Number of
          decision (D )
Output:   LoC (List of Item Context)
Begin
   LoC        F                        // F:  save file for item context
   while n > 0 do                      // n: number of decisions
      for each attribute do
         Calculate indiscernibility attribute sets
         INDC (B) = {(m, n) ∈ O²|∀a ∈ B, f (m, a) = f (n, a)}
      end for
      Apply disjoint between first two indiscernibility  attribute
      sets to calculate intermediate relation
      for each intermediate relation do
         Apply disjoint between intermediate relation and new
         indiscernibility attribute sets to calculate final
         indiscernibility relation R*
      end for
      for each class of decision do
         Compute indiscernibility set class of decision R+
      end for
      for each class indiscernibility decision sets do
         Calculate the intersection between final relation R* and
         R+ to find lower and upper set (rules)
      find lower and upper sets            //equation (2) and (3)
         redact rules
      end for
      Matching rules with items content
   end  while
   return LoC
End
```

7. DATA DESCRIPTION

In this paper, context suggestion (CS) has been performed on the LDOS-CoMoDa dataset [13, 14] and InCarMusic dataset, in which they both have been produced for the purposes of prediction applications. the first dataset, comprises of a rated 1232 movies by 121 users Whereas the second dataset, comprises of a rated 139 music by 42 users. “Table 2” illustrates the description of the LDOS_CoMoDa and InCarMusic.

Table 2: Clarify Data Set Description

Data sets	No. users	No. items	No. total instance	No. context condition	No. context factor
LDOS_CoMoDa	121	1232	2296	12	45
InCarMusic	42	139	4012	8	26

Algorithm “(1)” has been applied on two data set to select contexts features (context condition), where day type, location, and social are selected for firstly dataset. Secondly data set selected traffic condition, mood, and natural phenomena.

8. EVALUATION PROCESS

A procedure of building a model of classes from a set of instance that contains class labels called classification. Decision tree algorithm is to discover out the way the attributes-vector perform for a number of instances, many ways for classification “such as” J48 is an extension of ID3, k-nearest neighbor (K-NN) is a not complex procedure that classifies new cases by stock piling all available cases based on a similarity measure and a decision stump (DS) is a machine learning model consisting of a one-level decision tree. The proposed system (RSRC) has been evaluated using performance measures such as precision, recall and F-measure. In comparison rough set classifier with three methods for classification are J48, K-NN, and (DS). The cross validation have been used for test the performance of our system with 10-fold. The precision metric (P) as shown in equation (4) calculated a proportional relationship between two different numbers, the number of hits and the total number of recommended items" (|recsetu|) while the Recall (R) computed "the ratio of hits to the total number of hits" that should be retrieved by the system (|testsetu|) as clarified in equation (5). In addition, the harmonic mean (F_ Measure) of the precision and recall has been quantified as illustrated in equation(6).

$$\text{Precision}_u = \frac{|\text{hits}_u|}{|\text{recset}_u|} \quad (4)$$

$$\text{Recall}_u = \frac{|\text{hits}_u|}{|\text{testset}_u|} \quad (5)$$

$$\text{F\_measure} = \frac{2\,(\text{Precision} \cdot \text{recall})}{\text{Precision} + \text{recall}} \quad (6)$$

Table "3" clarify applied of equations (4,5,and 6) on two datasets LDOS_CoMoDa and InCarMusic.

Table 3: Mention Precision, Recall, and F-measures for Two Data Set

LoDoS-Dataset				InCarMusic-Dataset			
D day type	**precision**	**recall**	**f-measure**	**D Traffic condition**	**precision**	**recall**	**f-measure**
Rs	44.98	49.42	**47.09**	**Rs**	44.92	48.97	**47**
J48	39.49	48.86	36.41	J48	43.27	46.63	42.21
KNN	44.05	49.21	46.17	KNN	45.1	49.3	44.9
DS	38.99	49.21	35.57	DS	19.52	44.19	27.08
D location	**precision**	**recall**	**f-measure**	**D Mood**	**precision**	**recall**	**f-measure**
Rs	68.01	80.49	**73.7**	**Rs**	60.98	64.76	**63**
J48	66.19	81.35	72.99	J48	58.19	67.44	62.23
KNN	66.14	80.87	72.77	KNN	60.69	64.55	61.35
DS	66.19	81.35	72.99	DS	40.47	59.27	47.92
D social	**precision**	**recall**	**f-measure**	**D Natural Phenomena**	**precision**	**recall**	**f-measure**
Rs	48.16	51.95	**49.98**	**Rs**	71.33	71.69	75
J48	48.89	53.87	47.75	J48	75.06	75.12	75.06
KNN	47.21	51.95	48.24	KNN	72.32	72.53	72.36
DS	38.83	52.26	42.02	DS	37.77	56.8	45.37

Experimental results on classification decisions classes are day type ($D_{\text{day type}}$), location (D_{location}),and social (D_{social}) for LDOS_CoMoDa data set. Table "3" and Fig."(1)", Fig."(2)", and Fig."(3)" show precision, recall, and f-measure for rough set (RS), J48, K-Nearest neighbor (K-NN), and decision stump (DS) respectively.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

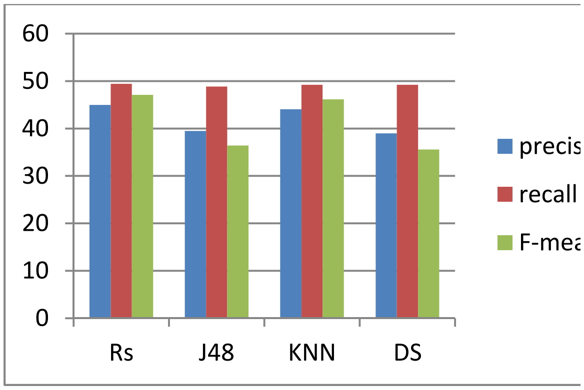

Figure 1: Represent D $_{day\ type}$ accuracy

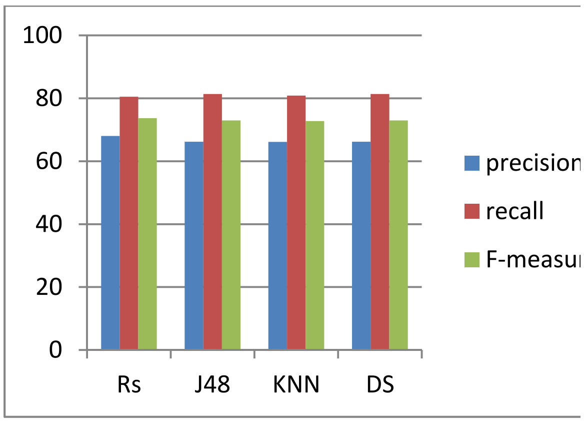

Figure 2: Represent D $_{location}$ accuracy

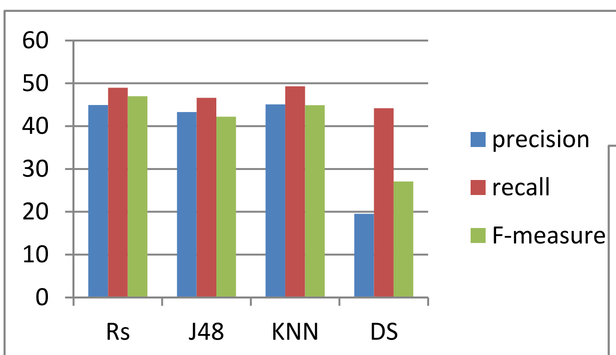

Figure 4: Clarify D $_{Traffic\ condition}$ accuracy

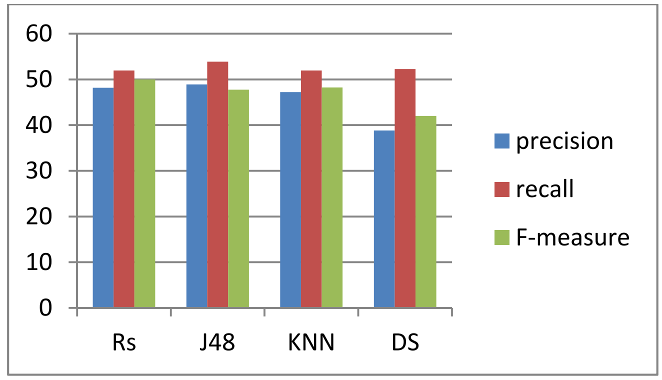

Figure 3: Represent D $_{social}$ accuracy

Classification experimental results of decisions classes that are traffic condition (D $_{Traffic\ condition}$), mood (D $_{Mood}$) ,and Natural Phenomena (D $_{Natural\ Phenomena}$) for InCarMusic data set. Fig."(4)", Fig."(5)", and Fig."(6)" show precision, recall, and F-measure.

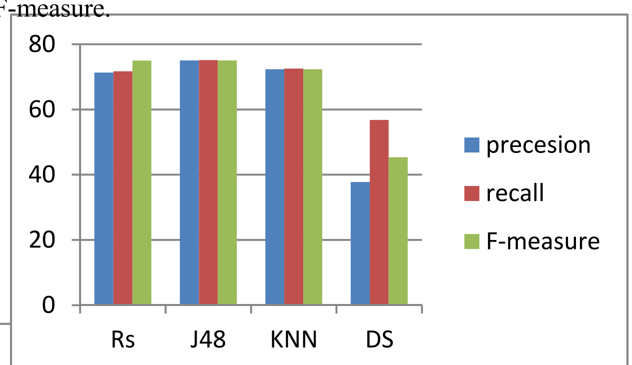

Figure 5 : Discern D $_{Mood}$ accuracy

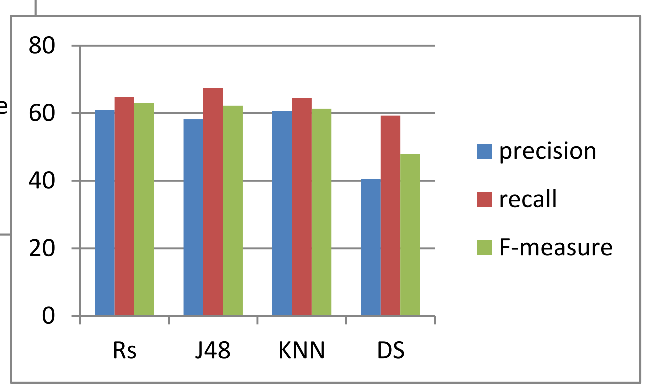

Figure 6: Represent D $_{Natural\ Phenomena}$ accuracy

9. CONCLUSION

In conclude, contexts suggestion for items based on fundamental rough set theory. Uncertainty item context (classes of decision), classification

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

classes of decisions by rule –based classification, using rough set. F-measure trade of precision and recall show accuracy of model for two datasets. Experimental results to decisions classes classification for RS is better from three methods classification J48, K-NN, and DS for LoDoS-CoMoDa dataset as show Fig.”1”,Fig.”2”, and Fig.”3”. RS classification results for InCarMuSic dataset are better, except result in Fig.”6” was 75% for RS and 75.06% for J48.

REFERENCES

[1] Y. Zheng, B. Mobasher,and R. Burke,” Context Recommendation Using Multi-Label Classification”, IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014.

[2] J. Lu, D.Wu, M. Mao,W.Wang, G. Zhang,” Recommender system application developments: A survey”, Elsevier B.V. Decision Support Systems PP. 12–32, 2015 .

[3] G. Adomavicius, A. Tuzhilin,” Context-aware recommender systems”, in: F. Ricci, L. Rokach, B. Shapira, P.B.Kantor (Eds.) Recommender Systems Handbook, Springer US., pp. 217-253, 2011.

[4] L. Baltrunas, , B. Ludwig, , S.Peer, , F. Ricci, ,” Context relevance assessment and exploitation in mobile recommender systems”, Personal and Ubiquitous Computing Vol.16, no.5, pp. 507–526 , 2012.

[5] V.Codina, , J. Mena, , L.Oliva,:” Context-Aware User Modeling Strategies for Journey Plan Recommendation”. In: User Modeling, Adaptation and Personalization, Springer, pp. 68–79, 2015.

[6] Y. Zheng, B. Mobasher,and R. Burke,” Context Recommendation Using Multi-Label Classification”, IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014.

[7] Y. Zheng,” Context Suggestion: Solutions and Challenges”, IEEE 15th International Conference on Data Mining Workshops, 2015.

[8] Y. Zheng, B. Mobasher,and R. Burke,” User-Oriented Context Suggestion”, ACM Canada, pp.13-17, 2016.

[9] L. Baltrunas, M. Kaminskas,and F. Ricci,” Be Usage Context Prediction for Music Tracks”, CARS Barcelona, Spain ,2010.

[10] D. Gries and B. Schneider,” fandemantale of new artificial intelligence”, springer, second edition, 2008.

[11] Q. Zhang , Qin Xie, Guoyin Wang ,” A survey on rough set theory and its applications”, Chongqing University
of Technology Production and hosting by Elsevier B.V, 2468-2322, 2016.

[12] S. Rissino and G. Lambert-Torres,” Rough Set Theory – Fundamental Concepts, Principals, Data Extraction, and Applications” Data Mining and Knowledge Discovery in Real Life Applications, Book edited by: Julio Ponce and Adem Karahoca, ISBN 978-3-902613-53-0, pp. 438, I-Tech, Vienna, Austria, 2009.

[13] A. Koir, A. Odi, M. Kunaver, M. Tkali, and J. Tasi,“Database for contextual personalization”, ELEKTROTEHNSKI VESTNIK, Vol.78, No.5, pp. 270–274, 2011.

[14] A. Odic, M. Tkalcic, J. Tasic, and A. Kosir, “Predicting and detecting the relevant contextual information in a movie recommender system”, Interacting with Computers, Vol.25, No.1, pp.74–90, 2013.