

2.

2.1.

(data mining and knowledge discovery),

: Data Mining and Knowledge Discovery,

mining (KDD – Knowledge Discovery in Databases)

. Data

[1, 2].

– (If – Then)

()

[3, 4].

:
scratch).

(from

[5].

2.2.

“ ” “ ”,).
(flat table)

(Universe of discourse).

(),

()

1.

2.

[6].

[7].

Chikofsky Cross,

[8].

2.3.

(Quinlan 1986.
C5.0 1997.

Inductive Dichotomizer 3 – ID3
)

:
C4.5 1992.

(Association Rule - Agrawal et al., 1993),
Pawlak, 1982, 1991; Tsumoto & Tanaka, 1995; Ziarko, 1993),

(Rough Sets -

(Pawlak 1982) [9, 10]

Data mining-a.

(split constants)

(Discriminant-based univariate splits, Discriminant-based linear
combination splits, C&RT-style exhaustive search for univariate splits,).

Quinlan [11]. ID3, C4.5 i See5 (C5.0)

{true, false}, { , }

ID3
ID3

Rosetta – *A Rough Set Toolkit for Analysis of Data*. Rosetta

() [1,
12, 13].

RSES – *Rough Set Exploration System*,
[14, 15].

C4.5, ID3 (missing values), (continuous values).
 See5 (C5.0), ID3, Quinlan-a.
 Tsumoto Tanaka 1995.

2.4.

[16]:
 U (), $A = \{a_1, a_2, \dots, a_m\}$,
 V_a a , $V = \bigcup_{a \in A} V_a$ $f = U \times A \rightarrow V$
 $f(x, a) \in V_a$ $a \in A, x \in U$. $S = \langle U, A, V, f \rangle$,
 f

(indiscernibility relation): $x, y \in U$,
 f $a \in A$.
 I_P :

$$I_P = \{(x, y) \in U \times U : f(x, a) = f(y, a), \forall a \in P\} \dots\dots\dots(1)$$

I_P $U|I_P$, x
 $I_P(x)$.
 $C \subset A$ $D \subset A$ $C \cap D = \emptyset$, A C
 (condition attributes), D (decision attributes).
 (decision system)

[17]:
 $\bigwedge_{a \in C} a = a(x) \Rightarrow \bigwedge_{d \in D} d = d(x)$, $a = a(x)$ x ;
 $d = d(x)$ d x .

X U $\emptyset \neq P \subseteq A$. X
 - (2) - (3) X :
 $P(X) = \{x \in U : I_P(x) \subseteq X\} \dots\dots\dots(2)$

$$\bar{P}(X) = \bigcup_{x \in X} I_P(x) \dots\dots\dots(3)$$

X U $\emptyset \neq P \subseteq A$.
 - X $Bn(X)$:

$$Bn(X) = \overline{P}(X) - \underline{P}(X) \dots\dots\dots (4)$$

[10, 16, 17].

$$\emptyset \neq P \subseteq A \quad a \in P.$$

$$P \quad I_P = I_{P-\{a\}}.$$

NP

m

$$\binom{m}{m/2}$$

(a, v_a) ,
THEN
(konsekvent)

a v_a
(antecedent)

: IF

: $a=v_a$.

[18]

- 1.
- 2.
- 3.

$$a_1 = v_{a1} \wedge a_2 = v_{a2} \wedge \dots \wedge a_n = v_{an} \Rightarrow d = v_d.$$

n
;

- 1.
- 2.
- 3.

- 4.
- 5.

[19].

$a_1 j$,

a_1 a_2 .

(depth first search).

(AND)

(-)

(,)

$$\sum_{n=1}^{|CUD|} \prod_{i=1}^n |V_i| \dots \dots \dots (5)$$

[19],
(Last In First Out - LIFO).

X[top]

1. (AVK):

$$\bigwedge_{i=1}^k (a_i, v_{i,j}), \quad 1 \leq k \leq |C \cup D|, \text{ a } v_{i,j} \text{ j- } i\text{-tog}$$

2. $(K, K_{AVK}), \quad K \subseteq U$

3. $K_1 : \bigwedge_{i=1}^k (a_i, v_{i,j}), \quad K_2 : \bigwedge_{i=1}^{k+1} (a_i, v_{i,h}).$

$$K_1 = K_2 \quad (a_{k+1}, v_{k+1,h})$$

$$\bigwedge_{i=1}^k (a_i, v_{i,j}, s_{i,j}), \quad s_{i,j} = \begin{cases} \text{true,} & \text{ako je } (a_i, v_{i,j}) \text{ suvišan} \\ \text{false,} & \text{u suprotnom} \end{cases}$$

(AVKR).

4. $(K, K_{AVKR}), \quad K \subseteq U$

K_{AVKR} .

```

procedure RuleSynthesis;
begin
  X := []; // X je prazan stek
  i := 1; // odabran je atribut a1 (prvi atribut)
  for j := 1 to |Vi| do // za sve vrednosti atributa a1
  begin
    Generate(Ki,j, U, (ai, vi,j));
    new_element.K := Ki,j;
    new_element.AVKR[i] := (ai, vi,j, false);
    new_element.last_attr := i;
    Push(X, new_element);
  end;
  while X <> [] do
  begin
    top_element := X[top];
    Pop(X);
    if objekti klase top_element.K imaju iste vrednosti atributa odluke then
    begin
      GenerateRule(top_element);
      goto 1;
    end;
    if top_element.last_attr = |C ∪ D| then GenerateRule(top_element) else
    begin
      i := top_element.last_attr + 1;
      for j := 1 to |Vi| do // za sve vrednosti atributa ai
      begin
        Generate(Ki,j, top_element.K, (ai, vi,j));
        if Ki,j ≠ ∅ then
          begin

```

```

        if  $K_{i,j} = top\_element.K$  then  $s_{i,j} := true$  else
             $s_{i,j} := false$ ;
        new_element.K :=  $K_{i,j}$ ;
        new_element.AVKR :=  $top\_element.AVKR \wedge (a_i, v_{i,j}, s_{i,j})$ ;
        Push(X, new_element);
    end;
end;
1: end;
end.

```

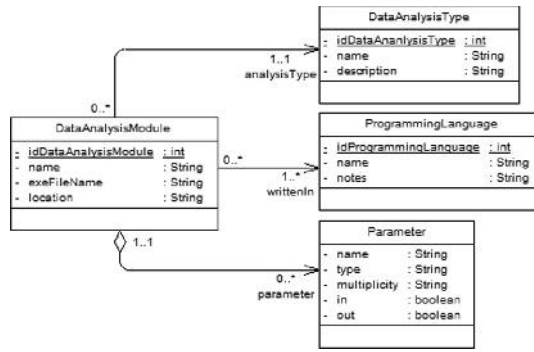
U. while (), X :

a_1
 \vdots
 a_j

$e_{i,j} = 0$
 If $i-$ a_j a_j .
 $e_{i,j}$ a_j (superfluous)
 If $i-$ a_j .
 Prolog GenerateRule () $i-$: If Then (-)
 OR Then (,, “ ”).
 LHS “ If Left Hand Support – LHS.
 Right Hand Support – RHS. “ Then $\frac{RHS}{LHS}$.

[20]. o ()

Separated Values) [21]. m Pascal CSV (Comma-
 - (IfThen) [22]. [24]. 1 UML [23]



1.

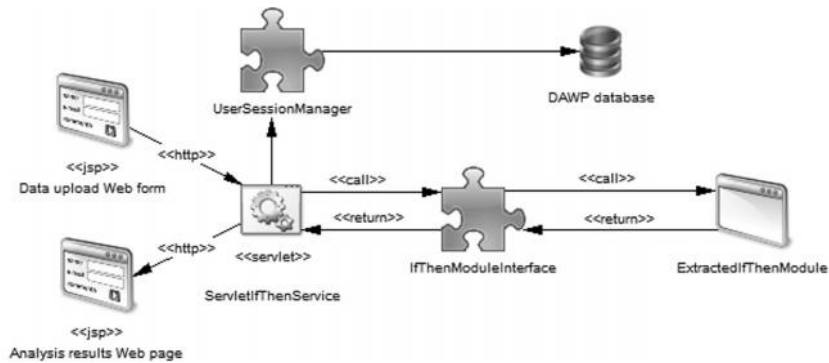
Parameter.

CSV

2):

- (ExtractedIfThenModule),
- (IfThenModuleInterface),
- (ServletIfThenService),

(Data upload Web form),
(Analysis results Web page),
(



2.

m -
(upload)

(download)

2.5.

Java Enterprise Edition [25], Apache Tomcat Java Servlet [27] [26].
 JavaServer Pages (JSP) [28]. JSP
 ServletIfThenService (upload)
 Apache Commons FileUpload [29] HTTP ()
 RFC 1867 (Form-based File Upload in HTML) [30].
 HTML JSP
 (Web browser) HTML [31], JSP
 [28], CSS [32] Java Script [33]

(3.)

(4.)



3.



4.

2.6.

2.7.

- [1] A. Øhrn, "Discernibility and Rough Sets in Medicine: Tools and Applications", *PhD thesis*, Department of Computer and Information Science, Norwegian University of Science and Technology, Trondheim, Norway, 1999.
- [2] G. Piatetsky-Shapiro, W. J. Frawley, editors: *Knowledge Discovery in Databases*, AAAI/MIT Press, 1991.
- [3] F. Bellas, "Standards for Second-Generation Portals", *IEEE Internet Computing*, vol. 8, no. 2, pp. 54-60, Mar./Apr. 2004.
- [4] C. Wege, "Portal Server Technology", *IEEE Internet Computing*, vol. 6, no. 3, pp. 73-77, May/June 2002
- [5] M. Jazayeri, "Some Trends in Web Application Development". *Future of Software Engineering* (May 23 - 25, 2007), *International Conference on Software Engineering*, IEEE Computer Society, 2007.
- [6] A. Ginige, S. Murugesan, "Guest Editors' Introduction: Web Engineering—An Introduction", *IEEE MultiMedia*, vol. 8, no. 1, pp. 14-18, Jan.-Mar. 2001.
- [7] Ž. Stojanov, V. Brtko, D. Dobrilovi, B. Odadžić, "Data Analysis Application — Migration to the Web". *INFOTEH Jahorina 2009*, Jahorina, Bosnia and Hercegovina. March 2009.
- [8] E. J. Chikofsky, II J. H. Cross, "Reverse Engineering and Design Recovery: A Taxonomy", *IEEE Software*, vol. 7, no. 1, pp. 13-17, Jan./Feb. 1990.
- [9] Z. Pawlak, J. Grzymala-Busse, R. Slowinski, W. Ziarko, "Rough sets, Association for Computing Machinery", *Communications of the ACM*, vol. 38, no. 11, ABI/INFORM Global, pg. 89, Nov 1995.

- [10] J. Komorowski, Z. Pawlak, L. Polkowski, A. Skowron, "Rough Sets: A Tutorial", <http://citeseer.ist.psu.edu/komorowski98rough.html>, 1998.
- [11] J. R. Quinlan, *C4.5: Programs for Machine Learning*, Elsevier, 1999.
- [12] J. Komorowski, A. Øhrn, "Modelling prognostic power of cardiac tests using rough sets", *Artif. Intell. in Medicine* 15, pp. 167–191, 1999.
- [13] A. Øhrn, J. Komorowski, A. Skowron, P. Synak, "A Software Systems for Rough Data Analysis", *Bulletin of International Rough Set Society*, vol. 1, no 2, 1997.
- [14] J. G. Bazan, M. S. Szczuka, J. Wróblewski, "A new version of rough set exploration system", In: James J. Alpigini, James F. Peters, Andrzej Skowron, Ning Zhong, redaktorzy, *Third International Conference on Rough Sets and Current Trends in Computing RSCTC*, volume 2475, Lecture Notes in Artificial Intelligence, pp. 397–404, Malvern, PA, October 14–16, Springer-Verlag, 2002.
- [15] J. Bazan, M. Szczuka, "RSES and RSESLib – A Collection of Tools for Rough Set Computations", *Extended version of paper presented at RSCTC*, 2000.
- [16] S. Greco, M. Benedetto, R. Slowinski, "New Developments in the Rough Set Approach to Multi-Attribute Decision Analysis", in *Bulletin of Int. Rough Set Society*, vol. 2, no 2/3, pp. 57–87, 1998.
- [17] Z. Pawlak, A. Skowron, "Rudiments of rough sets", *An International Journal of Information Sciences* 177, pp. 3–27, 2007.
- [18] J. Bazan, H. S. Nguyen, S. H. Nguyen, P. Synak, J. Wróblewski, "Rough set algorithms in classification problems", in L. Polkowski, T. Y. Lin, S. Tsumoto, eds., *Rough Set Methods and Applications: New Developments in Knowledge Discovery in Information Systems*, volume 56 of Studies in Fuzziness and Soft Computing, pp. 49–88. Physica-Verlag, Heidelberg, Germany, 2000.
- [19] G. Luger, W. Stubblefield, *Artificial Intelligence - structures and strategies for complex problem solving*, University of New Mexico, Albuquerque, The Benjamin/Cummings Publishing Company Inc, 1993.
- [20] Ž. Stojanov, D. Dobrilovi, V. Brtka, "Technological Infrastructure of Web Portal for Data Analysis", *Proceedings of 6th International Conference on Informatics, Educational Technology and New Media in Education*, pp. 389-400, Sombor, Serbia. March 28-29, 2009.
- [21] RFC 4180, Common Format and MIME Type for CSV Files, <http://tools.ietf.org/html/rfc4180>, [accessed February 2009].
- [22] . , " " , , , 2008. ,
- [23] G. Booch, J. Rumbaugh, I. Jacobson, *The Unified Modeling Language User Guide*, First Edition, Addison-Wesley Professional, 1998.
- [24] J. Conallen, *Building Web Applications with UML*, Addison-Wesley Longman Publishing Co., 2000.
- [25] Java Platform Enterprise Edition. <http://java.sun.com/javaee> [accessed February 2009].
- [26] Apache Tomcat, open-source application server. <http://tomcat.apache.org> [accessed February 2009]
- [27] Java Servlet technology, <http://java.sun.com/products/servlet/> [accessed February 2009].
- [28] JavaServer Pages (JSP) technology, <http://java.sun.com/products/jsp/> [accessed February 2009].
- [29] Apache Commons FileUpload package. <http://commons.apache.org/fileupload/> [accessed February 2009].
- [30] RFC 1867, Form-based File Upload in HTML, <http://datatracker.ietf.org/doc/rfc1867/> [accessed February 2009].
- [31] HTML. <http://www.w3.org/html/wg/> [accessed February 2009].
- [32] Cascading Style Sheets, CSS. <http://www.w3.org/Style/CSS/> [accessed February 2009].
- [33] Java Script, <https://developer.mozilla.org/en/JavaScript/> [accessed February 2009].

Мишљење о техничком решењу:

Веб сервис за аутоматску анализу података базиран на систематској синтаксној класификацији

Са циљем да се обезбеди доступност софтвера за аутоматску анализу података, која је базирана на систематској синтаксној класификацији, великом броју корисника развијен је веб сервис за анализу података. Сервис је интегрални део веб портала. Сервис омогућује кориснику да учита своје податке и да након извршене анализе преузме резултате анализе. Потребно је да подаци буду у форми табеле што је веома чест случај, а резултати анализе су у форми Ако – Онда (*If Then*) правила, која су лако читљива и подесна за даљу анализу.

Као основа за развој веб сервиса је искоришћена десктоп апликација, која је методама софтверског реинжењеринга адаптирана и интегрисана у веб окружење. Модел сервиса је развијен помоћу језика за моделовање UML (*Unified Modeling Language*) који је данас де-факто прихваћени стандард у софтверској индустрији за моделовање софтверских система. Приказани модел сервиса је генерички у смислу да се може применити уз одговарајуће адаптације и за реализацију сервиса који би понудили услугу анализе података и неком другом методом.

Сервис је тестиран на скупу података које је припремио корисник, а такође и на скупу података који су из области едукације и медицине.

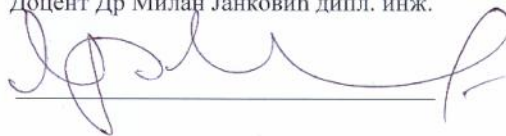
Реализовано решење омогућује приступ сервису за аутоматску анализу података базираном на систематској синтаксној класификацији како кориснику решења Пословном инкубатору Зрењанин, БИЗ д.о.о из Зрењанина тако и осталим заинтересованим корисницима.

Техничко решење спада у категорију М85 “Софтвер”, према Правилнику о поступку и начину вредновања, и квантитативном исказивању научноистраживачких резултата истраживача, који је донео Национални савет за научни и технолошки развој 2008. године.

Београд,

26. април 2010. године

Доцент Др Милан Јанковић дипл. инж.



Електротехнички факултет Београд

Мишљење о техничком решењу:

Веб сервис за аутоматску анализу података базиран на систематској синтаксној класификацији

Са циљем да се обезбеди доступност софтвера за аутоматску анализу података, која је базирана на систематској синтаксној класификацији, великом броју корисника развијен је веб сервис за анализу података. Сервис је интегрални део веб портала. Сервис омогућује кориснику да учита своје податке и да након извршене анализе преузме резултате анализе. Потребно је да подаци буду у форми табеле што је веома чест случај, а резултати анализе су у форми Ако – Онда (*If Then*) правила, која су лако читљива и подесна за даљу анализу.

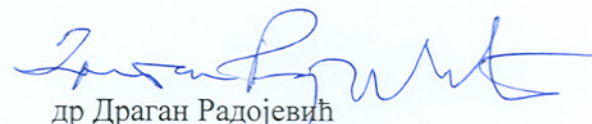
Као основа за развој веб сервиса је искоришћена десктоп апликација, која је методама софтверског реинжењеринга адаптирана и интегрисана у веб окружење. Модел сервиса је развијен помоћу језика за моделовање UML (*Unified Modeling Language*) који је данас де-факто прихваћени стандард у софтверској индустрији за моделовање софтверских система. Приказани модел сервиса је генерички у смислу да се може применити уз одговарајуће адаптације и за реализацију сервиса који би понудили услугу анализе података и неком другом методом.

Сервис је тестиран на скупу података које је припремио корисник, а такође и на скупу података који су из области едукације и медицине.

Реализовано решење омогућује приступ сервису за аутоматску анализу података базираном на систематској синтаксној класификацији како кориснику решења Пословном инкубатору Зрењанин, БИЗ д.о.о из Зрењанина тако и осталим заинтересованим корисницима.

Техничко решење спада у категорију М85 “Софтвер”, према Правилнику о поступку и начину вредновања, и квантитативном исказивању научноистраживачких резултата истраживача, који је донео Национални савет за научни и технолошки развој 2008. године.

Београд, 20. 04. 2010. године



др Драган Радојевић

Институт Михајло Пупин

Волгина 15, 11000 Београд



poslovni inkubator zrenjanin
business incubator zrenjanin
biz d.o.o

Kralja Aleksandra I Karađorđevića 2, Zrenjanin 23000
Tel. +381(0)23 512 260 Fax. . +381(0)23 512 264
Mob. tel. . +381(0)62 365 063
E-mail : inkubator@biz-zr.co.rs , www.biz-zr.co.rs
PIB 104431016, Matični br. 20127325
Broj poslovnog računa kod banke:
Societe Generale Srbija 275-0010220484307-62

POSLOVNI INKUBATOR ZRENJANIN
BIZ d.o.o. Zrenjanin

Del. br. 70/10

Datum 20.04.2010.

Предмет:

Извештај о пројекту **“Веб портали за анализу података и саветодавство”** (бр. пројекта TP – 13013)

Пословни Инкубатор Зрењанин је задовољан сарадњом и оствареним резултатима рада на пројекту технолошког развоја кога финансира Министарство науке Републике Србије под називом **“Веб портали за анализу података и саветодавство”** (бр. пројекта 13013, област 13) у периоду од 1.04.2009.-31.03.2010.године. Такође, Партиципант је задовољан постигнутим резултатима пројекта, а посебно оствареним техничким решењима у чијем тестирању и валидацији је и сам учествовао.

Пословни Инкубатор Зрењанин је у протеклој истраживачкој години извршио све своје обавезе у погледу партиципирања, а које се односе на опрему, стручни кадар и податке потребне за реализацију пројекта. Остварени су сви услови за сарадњу на Пројекту у наредном периоду, од 1.04.2010.-31.12.2010.године. Партиципант ће и даље стављати на располагање своје пословне податке, опрему и кадар како би валидирани новедефинисане процедуре и софтверска решења.

У Зрењанину, 20.04.2010.године.

Директор
Бојан Љутић, дипл. Инж.



Република Србија – АП Војводина
Универзитет у Новом Саду
Технички факултет «Михајло Пупин»
Зрењанин, Буре Ђаковића бб
www.tfzg.uns.ac.rs
Тел.023/550-515 факс: 023/550-520
ПИБ: 101161200



ИЗВОД ИЗ ЗАПИСНИКА

са 19.седнице Наставно-научног већа Техничког факултета «М.Пупин» у Зрењанину одржане 28.04.2010.године.

Непотребно изостављено!

АД:2.

Након поднетог Извештаја Катедри гласањем, једногласно је донета

ОДЛУКА

УСВАЈА СЕ предлог Катедри о верификацији техничких решења као резултата пројекта „Веб портали за анализу података и саветодавство“ ИТ13013, под називом:

- ♦ „Веб сервис за аутоматску анализу података базиран на систематској синтаксној класификацији“, аутори: доц. др Владимир Бртка, мр Жељко Стојанов, мр Далибор Добриловић, проф. др Ивана Берковић, др Весна Јевтић и мр Елеонора Бртка, са ТФ „Михајло Пупин“ у Зрењанину, за корисника Пословни инкубатор БИЗ Зрењанин,

Решење је прихваћено од рецензената: проф. др Драгана Радојевића са Института „Михајло Пупин“ у Београду и доц. др Милана Јанковића са Електротехничког факултета у Београду;

- ♦ „Аугментовани видео запис стања пута“, аутори: проф. др Драган Иветић, мр Дину Драган и Срђан Михаић, са Факултета техничких наука у Новом Саду, за корисника Центар за путеве Војводине АД, Нови Сад,

Решење је прихваћено од рецензената: проф. др Ивана Луковића и проф. др Милана Тривунића са Факултета техничких наука у Новом Саду;

- ♦ „Прикушљање података ефеката производње Газдинства“, аутор: др Тихомир Зорановић, са Пољопривредног факултета у Новом Саду, за корисника ПИК Бечеј АД.

Решење је прихваћено од рецензената: проф. др Бојана Срђевића и проф. др Мирослава Плавшића са Пољопривредног факултета у Новом Саду.

Образложење

Наставно-научно веће констатује да су техничка решења ваљана и квалитетна и врши њихову верификацију.



ПРЕДСЕДАВАЈУЋИ ВЕЋА
Проф. др Вјекослав Сајферт