Korišćenje asocijativnih pravila za istraživanje edukacionih podataka

Snježana Milinković
Elektrotehnički fakultet
Univerzitet u Istočnom Sarajevu
Istočno Sarajevo, Bosna i Hercegovina
snjeza@etf.unssa.rs.ba

Sadržaj—**Osnovni cilj dubinskog istraživanja podataka je pronalaženje i ekstrahovanje korisnih šablona i implicitnih znanja iz velikih skupova podataka. Kada se tehnike i metode dubinske analize podataka primenjuju na podatke prikupljene kroz informacione sisteme obrazovnih institucija onda se govori o posebnoj oblasti tzv. dubinskom istraživanju edukacionih podataka. Jedna od metoda koja se veoma često koristi pri dubinskoj analizi podataka je tehnika asocijativnih pravila. Tehnike asocijativnih pravila mogu da se koriste u istraživanju edukacionih podataka sa ciljem identifikovanja bitnih parametara i njihovih međusobnih veza koje utiču na proces učenja i postizanje željenih ishoda učenja.**

Ključne riječi - istraživanje podataka; edukacioni podaci; asocijativna pravila

I. UVOD

Osnovni cilj dubinskog istraživanja podataka (eng. *data mining*) je pronalaženje i ekstrahovanje korisnih šablona i implicitnih znanja iz velikih skupova podataka. Dubinsko istraživanje podataka obuhvata metode i tehnike kojima je moguće efikasno analizirati velike količine podataka. Pronalaženjem i opisivanjem strukturnih šablona među podacima generišu se nova znanja pomoću kojih se mogu otkriti i objasniti implicitne veze među podacima i kreirati prediktivni modeli bazirani na njima [1]. Pri tome, poželjno je da otkriveni šabloni budu što razumljiviji i korisniji. Ulazne strukture podataka za primenu tehnika dubinske analize predstavljaju se u formi skupa instanci, pri čemu svaka pojedinačna instanca predstavlja jednu kombinaciju vrednosti svih atributa koji karakterišu ulazni skup. Izlazne informacije dobijene nakon okončanja procesa analize ulaznih podataka najčešće se predstavljaju u prediktivnoj ili deskriptivnoj formi. Dubinska analiza podataka je multidisciplinarna oblast koja uključuje tehnike mašinskog učenja, statistike, baza podataka, veštačke inteligencije, prikupljanja informacija i vizuelizacije [2].

Kada se tehnike i metode dubinske analize podataka primenjuju na podatke prikupljene kroz informacione sisteme obrazovnih institucija onda se govori o posebnoj oblasti tzv. dubinskom istraživanju edukacionih podataka (eng. *educational data mining*). Dubinsko istraživanje edukacionih podataka je disciplina koja se poslednjih godina intenzivno razvija a fokusirana je na istraživanje i razvijanje metoda za analiziranje specifičnih podataka koji se vezuju za obrazovni kontekst [3].

Podaci koji se prikupljaju u obrazovnim institucijama mogu da budu podaci koji se generišu kroz interakciju studenata sa elektronskim sistemima koje institucija koristi kao podršku procesu učenja ali tu također spadaju i podaci o prethodnom školovanju studenata, demografski i socijalni podaci o studentima, itd. Osnovni cilj analize ovih podataka je da se otkrije korisno znanje o načinima na koji studenti uče, da se identifikuju faktori koji utiču na proces njihovog učenja, na stepen motivacije, njihove radne sposobnosti i postignuti uspeh. Dobijeno znanje koristi se da davanje smernica nastavnom osoblju kako bi što lakše i efikasnije proveli studente kroz proces učenja i postigli što bolje ishode učenja. Na osnovu dobijenog znanja nastavnici mogu imati bolji uvid u sposobnosti i potrebe svojih studenata i u skladu s tim prilagoditi nastavni proces.

Jedna od metoda dubinske analize podataka koja se veoma često koristi je tehnika asocijativnih pravila. Osnovni zadatak istraživanja podatka primenom tehnika asocijativnih pravila je da se pronađu međusobne veze kojima se mogu opisati pojavljivanja pojedinih instanci unutar velikih skupova podataka tzv. asocijacije [2]. Te asocijacije među podacima su najčešće veoma složene i implicitne. Prvi algoritam za otkrivanje asocijativnih pravila, Apriori algoritam, objavljen je 1993. godine i odmah po objavljivanju izazvao je ogromnu pažnju [2]. Od tada je razvijen veliki broj tehnika koje su predstavljale nadogradnju i modifikaciju kao i primenu tog algoritma [4] – [6]. Tehnike asocijativnih pravila mogu da se koriste u istraživanju edukacionih podataka sa ciljem identifikovanja bitnih parametara koji utiču na proces učenja, postizanje različitih performansi i željenih ishoda učenja, kao i njihovih međusobnih korelacija kroz tradicionalnu asocijativnu analizu. Jedan model unapređenog Apriori algoritma za dubinsko analiziranje edukacionih podataka je opisan u [7]. Jedna od mogućnosti za evaluaciju procesa učenja kroz primenu asocijativnih pravila opisana je u [8]. Autori su predložili jedan model za prikupljanje podataka relevantnih za uspešnu evaluaciju nastavnika i načina izvođenja nastave. Na osnovu predloženog modela, korišćenjem asocijativnih pravila dobijene su značajne informacije koje mogu pomoći

unapređenju nastavnog procesa u budućnosti. U [9], primenom Apriori algoritma generisan je skup pravila na osnovu kojih se mogu grupisati studenti u skladu sa njihovim akademskim performansama. Studenti su grupisani na osnovu njihovog učešća u izradi zadatka, testova, stepenu pohađanja nastave, itd. Korišćenje asocijativnih pravila za analiziranje podataka o načinu na koji studenti koriste društvene mreže predstavljeno je u [10]. Generisana su interesantna pravila koja povezuju način na koji studenti uče sa načinom na koji koriste društvene mreže.

U ovom radu istražene su mogućnosti primene algoritama asocijativnih pravila na odabranom skupu podataka o studentima. Korišćenjem Weka alata za dubinsko analiziranje podataka, [11], izvršeni su eksperimenti kojima su se generisali skupovi pravila na osnovu kojih se mogu kreirati preporuke za grupisanje studenata.

Rad je organizovan na sledeći način: u poglavlju 2 opisuju se tehnike asocijativnih pravila. 3. poglavlje opisuje specifičnosti analiziranog skupa ulaznih podataka. U 4. poglavlju opisana je implementacija odabranih algoritama asocijativnih pravila u Weka alatu, a u 5. su prikazani i analizirani rezultati izvedenih eksperimenata. 6. poglavlje je rezervisano za zaključak.

II. TEHNIKE ASOCIJATIVNIH PRAVILA

Tehnike asocijativnih pravila spadaju među najčešće korišćene tehnike deskriptivne analize podataka. Korišćenjem tehnika asocijativnih pravila otkrivaju se međusobne veze i potencijalno interesantne strukture među vrednostima atributa unutar skupova podataka [12]. Identifikovane veze se predstavljaju u formi *if-then* pravila koja su veoma pogodan metod za predstavljanje znanja zbog svoje jednostavnosti i razumljivosti. Algoritmi za generisanje asocijativnih pravila svode se na pronalaženje tzv. frekventnih skupova parova atribut-vrednost (eng. *frequent item sets*). Proces započinje formiranjem podskupova koji se sastoje od po jednog elementa atribut-vrednost a u narednim iteracijama broj elemenata u podskupovima se povećava. U svakoj narednoj iteraciji podskupovi se kreiraju kombinujući samo elemente podskupova iz prethodne iteracije koji su se pokazali kao frekventni. Kao metrika na osnovu koje se zaključuje da li je neki podskup frekventan ili ne koristi se veličina nazvana podrška (eng. *support*). Podrška se definiše kao odnos broja instanci u kojima postoje elementi jednog podskupa, parovi atribut-vrednost, u odnosu na ukupan broj instanci analiziranog skupa. U frekventne skupove spadaju samo oni podskupovi za koje je podrška veća ili jednaka od vrednosti korisnički definisane vrednosti minimalne podrške, *minsup*. Konačan skup asocijativnih pravila određuje se korišćenjem druge metrike, poverenja (eng. *confidence*). Od svih generisanih frekventnih podskupova podataka za kreiranje asocijativnih pravila odabiru se samo oni za koje je vrednost poverenja veća od korisnički definisanog minimalnog praga poverenja, *minconf*. Asocijativna pravila predstavljaju se u formi $X \rightarrow Y$, pri čemu implikacija znači istovremeno događanje a ne uzročnost. Skupovi X i Y se sastoje od jednog ili većeg broja kombinacija parova atribut-vrednost i disjunktni su. Poverenje se definiše kao verovatnoća da X implicira Y i izračunava se kao odnos broja instanci ukupnog skupa podataka unutar kojih postoje elementi *X U Y* i broja instanci unutar kojih postoje samo elementi skupa X. Na ovakav način definisano poverenje određuje prediktivnu moć generisanog pravila.

U opštem slučaju sa leve i desne strane implikacije generisanih asocijativnih pravila mogu se naći bilo koje kombinacije parova atribut-vrednost. Međutim, u nekim primenama se javlja potreba da se generišu pravila koja će sa desne strane implikacije imati uvek prikazanu informaciju o vrednosti samo jednog atributa. U tom slučaju radi se o posebnoj vrsti asocijativnih pravila tzv. klasna asocijativna pravila (eng. *class association rules*), a atribut čije vrednosti će se prikazivati sa desne strane pravila naziva se klasni atribut. Algoritam za generisanje klasnih asocijativnih pravila je opisan u [13]. Definicije korišćenih metrika, podrške i poverenja, za određivanje snage tj. za evaluaciju generisanih pravila, i u slučaju ovog algoritma iste su kao i kod standardnih asocijativnih pravila.

III. KARAKTERISTIKE ULAZNOG SKUPA PODATAKA

Za potrebe ovog rada metodom slučajnog uzorka prikupljeni su podaci o dvema generacijama studenata sva tri studijska program koja se izvode na Elektrotehničkom fakultetu u Istočnom Sarajevu. Osnovni cilj istraživanja prikupljenih podataka metodama asocijativnih pravila je da se istraži da li postoji korelacija između podataka koji se mogu prikupiti o studentima pri njihovom upisu na fakultet i uspeha koji će da ostvare pri polaganju ispita Uvod u programiranje koji se izvodi u letnjem semestru na prvoj godini studija. Podaci do kojih se može doći kroz informacioni sistem studentske službe su: srednja škola koju su kandidati završili, prosek koji su imali u srednjoj školi, grad u kojem su završili srednju školu, broj bodova koji su osvojili na polaganju prijemnog ispita i ocena koju su dobili pri polaganju ispita Uvod u programiranje. Podaci koji se ne unose u informacioni sistem, ali do kojih se može doći ručnom pretragom studentskih dosijea, su ocene koje su studenti imali iz predmeta matematika i informatika u pojedinim razredima srednje škole. Na osnovu njih izračunavaju se prosečne ocene iz tih predmeta.

Proces dubinske analize podataka podrazumeva tri osnovna koraka: priprema ulaznih podataka – njihova predobrada da bi se prilagodili formi koja se zahteva algoritmom koji će se primenjivati za istraživanje, primena samog algoritma za dubinsku analizu podataka i obrada i analiza dobijenih rezultata. Jedan od preduslova za primenu algoritama asocijativnih pravila je da ulazni podaci moraju biti kategorijskog, nominalnog, tipa. Većina ulaznih podataka prikupljenih za ovaj eksperiment je numeričkog tipa i zbog toga će biti potrebno izvršiti njihovu transformaciju radi konvertovanja u nominalne vrednosti. Odabir načina konvertovanja podataka veoma često u velikoj meri utiče na dobijene rezultate provedenog eksperimenta. Zbog toga se tom poslu mora posvetiti posebna pažnja i on najčešće iziskuje dosta vremena.

IV. WEKA IMPLEMENTACIJA ALGORITAMA ASOCIJATIVNIH PRAVILA

Weka je jedan od najčešće korišćenih alata za dubinsku analizu podataka. Spada u *open-source* softvere i ima implementiran veliki broj tehnika mašinskog učenja. Za učenje asocijativnih pravila Weka ima implementiranih 6 algoritama a za potrebe eksperimenata u ovom radu koristit će se dva algoritma: Weka implementacija standardnog Apriori algoritma i modifikovana verzija Apriori algoritma nazvana Prediktivni Apriori algoritam (eng. *Predictive Apriori*). Oba algoritma imaju mogućnost podešavanja parametara za učenje klasnih asocijativnih pravila.

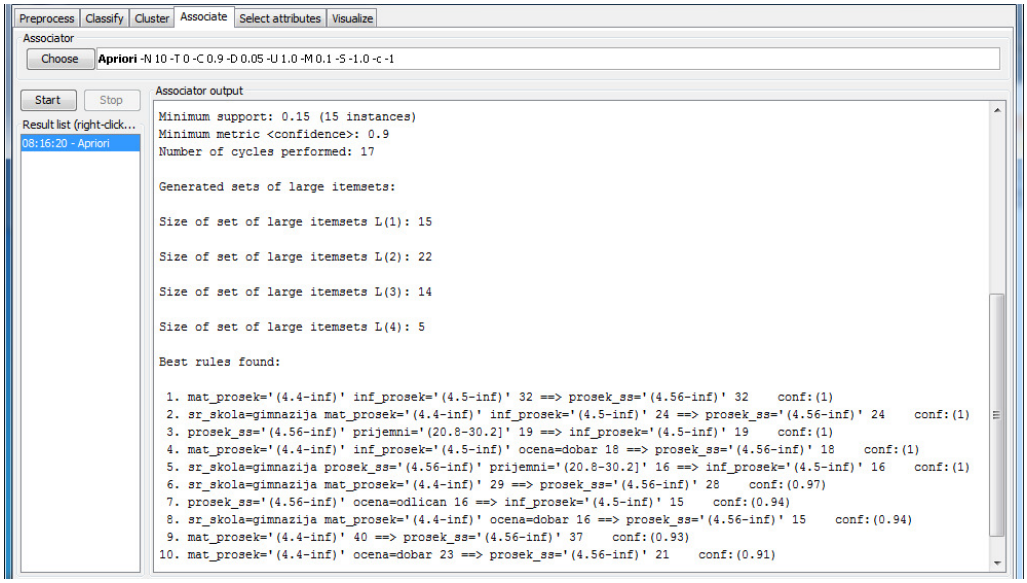
Weka implementacija Apriori algoritma započinje proces učenja sa predefinisanom vrednošću *minsup* = 1, tj. sa minimalnom podrškom od 100% i vrši njeno iterativno smanjivanje za korak 5% sve dok se ne generiše najmanje 10 pravila sa minimalnim poverenjem *minconf* = 0,9 ili dok vrednost podrške u procesu njenog iterativnog smanjivanja ne dostigne definisanu donju granicu minimalne podrške od 10%, *minsup* = 0,1. Ove predefinisane vrednosti za *minsup* i *minconf,* kao i korak iterativnog smanjivanja podrške se mogu promeniti podešavanjem parametara Weka Apriori algoritma ukoliko to korisnik želi. Pored toga, korisnik može odabrati da umesto poverenja koristi neku drugu metriku za evaluaciju snage generisanih pravila [1]. Generisana pravila mogu da budu u standardnoj formi ili u formi klasifikacijskih pravila podešavanjem vrednosti parametra *car* na *True*.

Prediktivni Apriori algoritam kombinuje vrednosti poverenja i podrške u jedinstvenu meru tzv. prediktivne tačnosti (eng. *predictive accuracy*) [1]. Algoritam pronalazi N željenih asocijativnih pravila poredanih po vrednostima prediktivne tačnosti. Podešavanem parametara algoritam se može transformisati u algoritam za predviđanje vrednosti jednog, klasnog, atributa tj. u algoritam za generisanje klasifikacijskih pravila.

V. REZULTATI IZVOĐENJA EKSPERIMENATA

Kao što je već pomenuto, za primenu algoritama asocijativnih pravila potrebno je da ulazni fajl ima sve nominalne atribute. Weka nudi veliki broj filtera za transformisanje vrednosti ulaznih podataka u željeni format [1]. Za potrebe ovog rada korišćen je Weka *Discretize* filter iz skupa nenadgledanih (eng. *unsupervised*) atributskih filtera. *Discretize* omogućava transformisanje numeričkog opsega vrednosti u nominalni kroz definisanje parametara filtriranja kojima se podešava broj intervala na koje će se podeliti kontinualni opseg, na koji način će se ta podela realizovati, nad kojim atributima će se fitriranje primeniti itd. Pored korišćenja filtera, ulazni podaci mogu da se transformišu i ručno oslanjajući se pri tome na intuitivne zahteve samog korisnika tih podataka. Za potrebe ovog rada pripremljena su dva fajla sa ulaznim podacima. U oba fajla kombinovan je ručni metod sa Weka *Discretize* filterom. Prvo su konačne ocene studenata ručno transformisane na sledeći način: ocena 5 – nepoloženo, ocene 6, 7 i 8 se konvertuju u nominalnu vrednost dobar, ocene 9 i 10 u odličan. U prvom eksperimentu, algoritmi asocijativnih pravila su primenjeni nad ulaznim fajlom u kojem su svi ostali numerički atributi konvertovani korišćenjem *Discretize* filtera sa odabranih 5 intervala diskretizacije. U drugom eksperimentu, numerički atributi *mat_prosek*, *inf_prosek* i *prosek_ss* su ručno trasformisani u 5 intuitivno kreiranih intervala: manje od 3, od 3 do 3.5, od 3.5 do 4, od 4 do 4.5 i od 4.5 do 5. Numeričke vrednosti atributa *prijemni* su diskretizovane na 5 intervala jednake širine korišćenjem *Discretize* filtera kao i u prvom slučaju.

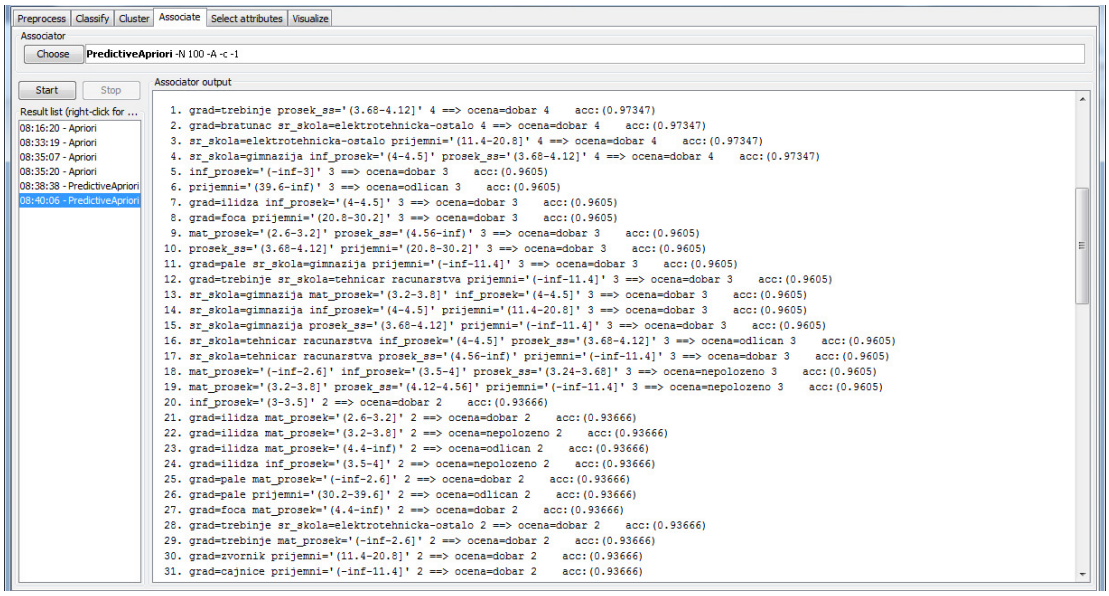
Rezultat primene Apriori algoritma na prvi skup podataka sa predefinisanim vrednostima parametara algoritma je prikazan na sl. 1.



Slika 1. Pravila generisana Apriori algoritmom, 1. eksperiment

Sa sl. 1 se može videti da od 10 generisanih pravila samo u 4 figuriše atribut *ocena* i to svaki put sa leve strane implikacije. Takva pravila, iako imaju veoma visok stepen pouzdanosti, ne mogu da budu od koristi u konkretnom slučaju gde se pokušava zaključiti kako ostali atributi utiču na vrednost atributa konačna ocena. Stoga je eksperiment ponovljen nad istim skupom podataka ali sa podešenom vrednošću parametra *car* = *True* čime se algoritam iz svoje standardne forme konvertovao u klasifikacijsku. Međutim, za predefinisanu minimalnu vrednost metrike *minconf* = 0.9 algoritam nije uspeo da pronađe nijedno pravilo. Smanjivanjem vrednosti pouzdanosti moguće je generisati pravila ali je pitanje kakva je korist od njih ako je njihova prediktivna moć tako mala (generisano pravilo sa najboljom pouzdanošću u ovom slučaju ima pouzdanost od 63%).

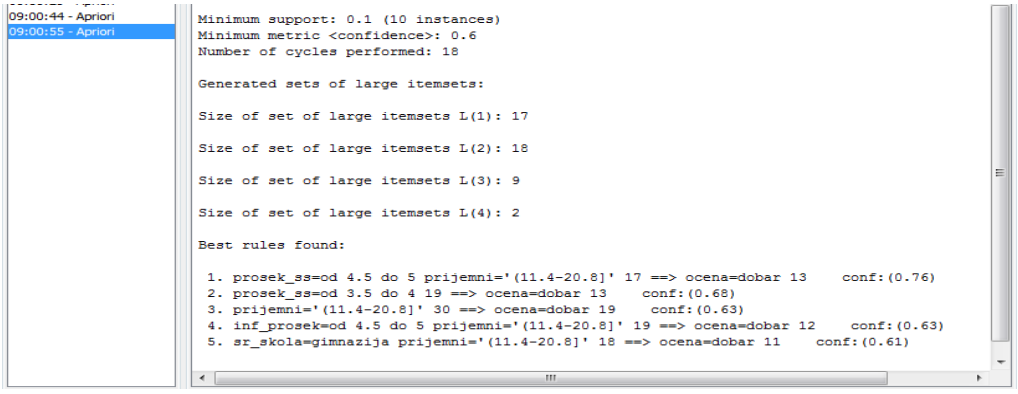
Primenom Prediktivnog Apriori algoritma na isti ulazni skup generisan je izlazni skup od 100 pravila prikazan na sl. 2.



Slika 2. Pravila generisana Prediktivnim Apriori algoritmom, 1. eksperiment

Analizom generisanih pravila uočava se da postoji veliki broj pravila koji imaju veoma veliku prediktivnu tačnost: 48 pravila sa predikcijom većom od 90%. Pri tome, u svakom od prva 4 pravila identifikovane su po 4 instance tj. ukupno 16 instanci ulaznog skupa koji imaju tako veliku verovatnoću, odnosno, narednih 15 sa po 3 pronađene instance što u ukupnom zbiru predstavlja značajan broj instanci koje imaju veliku prediktivnu moć.

Drugi eksperiment je rađen nad drugim opisanim ulaznim skupom. U ovom slučaju, primenom standardnog Apriori algoritma, u skupu generisanih 10 pravila nema nijedno u kojem se implicira vrednost atributa konačna ocena. Zbog toga je ovaj algoritm izvršen u klasifikacijskoj varijanti a rezultat njegovog izvršenja je prikazan na sl. 3.



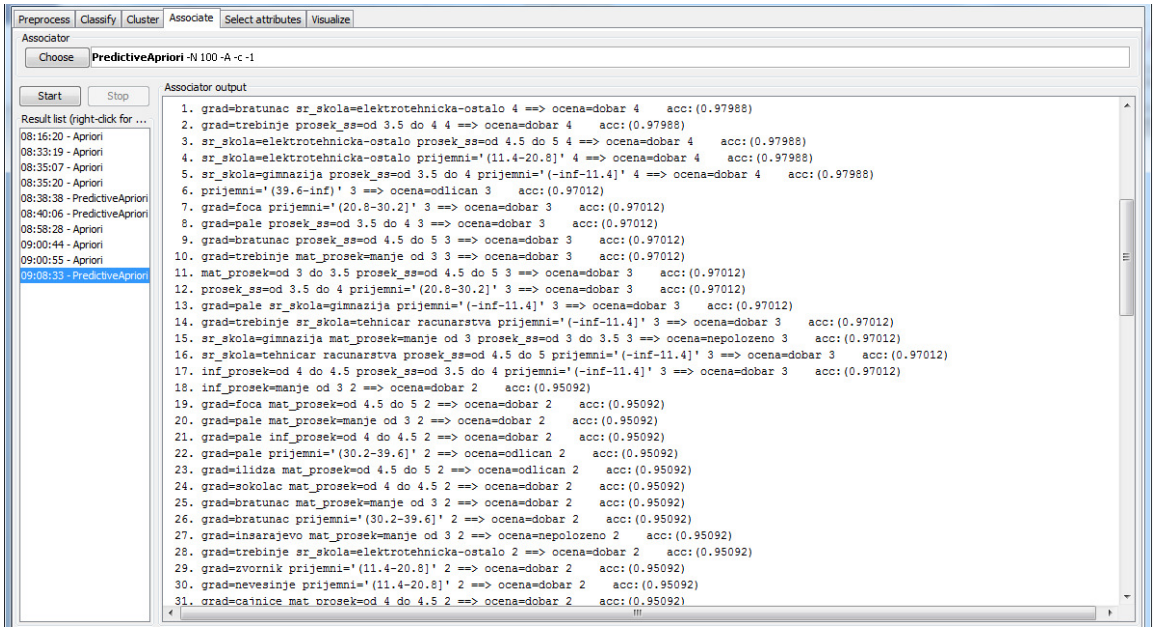
Slika 3. Pravila generisana Apriori algoritmom, 2. eksperiment

O ovom slučaju generisano je 5 pravila koja imaju poverenje veće od 60% što je malo bolji rezultat u odnosu na prethodni eksperiment. Međutim, analizom generisanih pravila uočava se da se svih 5 pravila odnosi na kombinaciju atributa i vrednosti ulaznog skupa za predviđanje samo jedne klase, *ocena=dobar*. Pored toga, u 4 od 5 generisanih pravila figuriše atribut *prijemni* sa svojim opsegom vrednosti *(11.4- 20.8)*. Daljim analiziranjem navedenih pravila, zaključuje se da 1., 4. i 5. pravilo nisu relevantni i da ustvari predstavljaju podskup 3. pravila. Ovo je situacija koja se često dešava pri primeni tehnika asocijativnih pravila: pojava redundantnosti u skupu generisanih pravila. Odbacivanjem nerelevatnih pravila, ostaju samo dva pravila koja su relevantna u ovom slučaju i oba se odnose na predviđanje klase *dobar*. U pokušaju da se dođe do međusobnih veza atributa i njihovih vrednosti za predviđanje ostalih klasa, nad ovim ulaznim skupom je izvršen Prediktivni Apriori algoritam. Primenom Prediktivnog Apriori algoritma u klasifikacijskoj formi dobije se skup od 100 generisanih pravila delimično pokazan na sl. 4.

Sa sl. 4 se može videti da je prediktivna tačnost prvih 48 instanci povećana, sve su preko 95%, a također prediktivna tačnost je povećana i za naredne instance tako da u ovom slučaju postoji i dodatnih 10 pravila za koja je prediktivna tačnost veća od 70%. Ovo upućuje na zaključak da su ulazni podaci u drugom eksperimentu bolje pripremljeni.

Analizom generisanih pravila uočava se da postoje predviđanja za sve 3 vrednosti klasnog atributa, što daje neospornu prednost Prediktivnom Apriori u odnosu na osnovni Apriori algoritam. Predefinisani broj pravila koja se generišu Prediktivnim Apriori algoritmom je prilično veliki, 100 pravila, i tako veliki broj pravila lako može da ima za posledicu pojavu redundantnosti. Zbog toga je potrebno pažljivo izvršiti evaluaciju generisanih pravila i izdvojiti ona koja su u konkretnom slučaju relevantna. Relevantnost pravila se može određivati korišćenjem nekih od statističkih metoda ili subjektivnim razmatranjem korisnika koji na osnovu prethodnih znanja o analiziranom skupu podataka može izvesti određene zaključke. Tako npr. u ovom skupu generisanih pravila moguće je uočiti redundantnost pravila 8 i 13 jer su veze koje se njima opisuju već obuhvaćene pravilom broj 5. Subjektivni zaključak je izveden na osnovu poznavanja činjenice da su svi studenti koji su završili srednju školu u Palama, *grad=Pale*, završili gimnaziju, *sr_skola=gimnazija*. Do istog zaključka se dolazi i analiziranjem pravila generisanih pod rednim brojevima od 38 – 43. Sva pravila od 39 – 43 su samo specijalizacija pravila 38, pri čemu pravila 42 i 43 pored toga predstavljaju ustvari kombinacije parova atribut – vrednost iz pravila 39 i 41, odnosno, 40 i 41, sukcesivno.

Evaluacijom kompletnog skupa generisanih pravila može se izvesti opšti zaključak da verovatnoća da je neko pravilo redundantno raste sa rednim brojem generisanog pravila, a istovremeno važnost tih pravila opada jer opada njihova prediktivna moć. Ono što je važno za ovaj skup pravila je da većina pravila sa velikom prediktivnom moći spada u skupinu interesantnih i relevantnih pravila. Ona korisniku pružaju znanje koje do tada nije postojalo i na osnovu njih se može preduzeti neka konkretna akcija. Na ovakav način generisana pravila mogu se iskoristiti kao preporuka za grupisanje studenata na samom početku organizovanja kursa kako bi se kroz grupne aktivnosti što bolje prilagodili nastavni materijali njihovom nivou predznanja.



Slika 4. Pravila generisana Predictive Apriori algoritmom, 2. eksperiment

VI. ZAKLJUČAK

Dubinsko istraživanje podataka primenom metoda asocijativnih pravila je veoma interesantno jer se generišu pravila u veoma jednostavnoj i razumljivoj formi i kao takva laka su za analizu i implementaciju u budućim konkretnim softverskim rešenjima koja bi koristila te rezultate. Istraživanje podataka generisanih unutar obrazovnih institucija

je veoma značajno jer se time može poboljšati proces izvođenja nastave. U eksperimentima pokazanim u ovom radu generisan je veliki broj pravila na osnovu kojih se mogu sa velikom prediktivnom tačnošću grupisati studenti na osnovu njihovih predviđenih akademskih performansi i na takav način se izvršiti prilagođenje nastavnih materijala različitim nivoima predznanja sa kojim se oni upisuju na fakultet. To prilagođenje bi u budućnosti trebalo da dovede do postizanja boljih ishoda učenja i ostvarivanja boljih akademskih rezultata.

LITERATURA

[1] I. H. Witten, E. Frank, M.A. Hall, "Data mining: practical machine learning tools and techniques", 3rd edition, Elsevier, 2011

[2] B. Liu, Web "DataMining - Exploring Hyperlinks, Contents, and Usage Data", © Springer-Verlag Berlin Heidelberg 2007

[3] C. Romero, S. Ventura, "Data mining in education", WIREs Data Mining Knowl Discov 2013, 3: 12–27 doi: 10.1002/widm.1075

[4] H. Yu, J. Wenb, H. Wangc, L. Jun, "An Improved Apriori Algorithm based On the Boolean Matrix and Hadoop", Procedia Engineering 15 (2011) 1827 - 1831, Elsevier, 2011

[5] J. Yabing, "Research of an Improved Apriori Algorithm in Data Mining Association Rules", International Journal of Computer and Communication Engineering, Vol. 2, No. 1, January 2013

[6] M. Al-Maolegi, B. Arkok, "AN IMPROVED APRIORI ALGORITHM FOR ASSOCIATION RULES", International Journal on Natural Language Computing (IJNLC) Vol. 3, No. 1, February 2014

[7] J. Jha, L. Ragha, "Educational Data Mining using Improved Apriori Algorithm", International Journal of Information and Computation Technology, ISSN 0974-2239 Volume 3, Number 5, pp. 411-418, 2013

[8] D. Jiabin, H. JuanLi, C. Hehua, W. Juebo, "An Apriori-based Approach for Teaching Evaluation", Information Engineering and Electronic Commerce (IEEC), 2nd International Symposium on, Ternopil, Ukraine, 2010

[9] D. M. D. Angeline, "Association Rule Generation for Student Performance Analysis using Apriori Algorithm", The SIJ Transactions on Computer Science Engineering & its Applications (CSEA), Vol. 1, No. 1, March-April 2013

[10] S. S. Phulari, P. U. Bhalchandra, Dr. S. D. Khamitkar, S. N. Lokhande, "Understanding Rule Behavior through Apriori Algorithm over Social Network Data", Global Journal of Computer Science and Technology Volume 12 Issue 10 Version 1.0 May 2012

[11] Weka softerski alat dostupan na: http://www.cs.waikato.ac.nz/ml/weka/

[12] O. Maimon, L. Rokach, "Data Mining and Knowledge Discovery Handbook", Springer, NewYork, 2010

[13] B. Liu, W. Hsu, Y. Ma, "Integrating Classification and Association Rule Mining", Fourth International Conference on Knowledge Discovery and Data Mining, 80-86, 1998

ABSTRACT

The main goal of data mining is finding and extracting useful patterns and implicite knowledge from large data sources. Data mining techniques can be applied to explore and analyze data that come from different types of educational environments. In that case the term educational data mining is used. Association rules is one of the most popular data mining technique. Association rules can be applied on educational data mining for identifying important parameters and their relationships that could influence teaching process and learning outcomes.

ASSOCIATION RULES APPLIED ON EDUCATIONAL DATA

Snježana Milinković