

## ოპერაციების კვლევა Data mining ტექნოლოგიის საფუძველზე

ხათუნა აბულაშვილი, ეკატერინე მალრაძე

საქართველოს ტექნიკური უნივერსიტეტი, ეკონომიკური ინფორმატიკის კათედრა

### ანოტაცია:

ნაშრომში წარმოდგენილია Data Mining ტექნოლოგიით შემუშავებული პროგნოზირება საგადასახადო სტრუქტურაში, კერძოდ გადამხდელის პირადი ბარათის საფუძველზე. Data Mining ტექნოლოგიის კლასიფიკაციის მეთოდის და ალგორითმ Naivy Bayes კრიტერიუმების შერჩევის საფუძველზე ვლდებულობთ პროგნოზს, რომლის შედეგადაც შეგვიძლია მივიღოთ სასარგებლო ინფორმაცია საგადასახადო სისტემის მუშაობისათვის.

**საკვანძო სიტყვები:** Data Mining, ინტელექტუალური ანალიზი, ცოდნის მოძიება, Oracle Data Miner, Naivy Bayes ალგორითმი, პროგნოზირება.

თანამედროვე ბიზნეს სამყაროს საქმიანობის ნებისმიერ სფეროში ინფორმაციის მოცულობა ორგანიზაციებში საკმაოდ დიდია და ორგანიზაციის წარმატებული მუშაობა დამოკიდებულია მასში არსებული, კუთვნილი ინფორმაციიდან მაქსიმუმის გამოყენებისაგან. საკმაოდ რთულია მონაცემთა მნიშვნელოვნების შეფასება, რომლებიც უწყვეტად გროვდება ადამიანების საქმიანობის პროცესში, რომლებიც მონაწილეობენ ბიზნესის მართვასა და წარმოებაში, საბანკო საქმიანობაში, სამეცნიერო, ინჟინერული თუ სამედიცინო ამოცანების გადაწყვეტაში. მძლავრი კომპიუტერული სისტემები, რომლებიც მოიცავენ ინახავენ და მართავენ უზარმაზარ მონაცემთა ბაზებს წარმოადგენენ განუყრელ ატრიბუტს მსხვილი კორპორაციების საქმიანობისათვის და არა მარტო მსხვილი კორპორაციებისათვის. მაგრამ მონაცემების მხოლოდ არსებობა არ არის საკმარისი მუშაობის მაჩვენებლების გასაუმჯობესებლად. საჭიროა არსებული მონაცემების ტრანსფორმირების ცოდნა, მათგან სარგებლიანობის მისაღწევად და გადაწყვეტილებების მისაღებად.[1] ამისათვის არის წარმოდგენილი ტექნოლოგია **Data Mining**.

მომხმარებლის სათანადო მონაცემების ანალიზის მეშვეობით შესაძლოა ვიწინასწარმეტყველოთ კლიენტის მოქმედება და აგრეთვე მოვახდინოთ მასზე ზეგავლენაც კი.

Data Mining-ის გამოყენების სფერო არ არის განსაღვრული - ის ყველგან არის სადაც კი არსებობს მონაცემები. მაგრამ პირველ რიგში Data Mining -ის მეთოდებმა დააინტრიგა კომერციული საწარმოები, რომლებიც თავიანთ პროექტებს ახორციელებენ ინფ. მონაცემთა საცავის (Data Warehousing) საფუძველზე ასეთი საწარმოების მაგალითი გვიჩვენებს, რომ Data Mining-ის გამოყენების ეფექტიანობა აღწევს 100%.

Data Mining-ი წარმოადგენს დიდ ფასეულობას ანალიტიკოსებისათვის მათ ყოველდღიურ საქმიანობაში. საქმიანმა ხალხმა კარგად გაითავისა, რომ Data Mining-ის მეთოდების მეშვეობით მათ შეუძლიათ მიიღონ საგრძნობი უპირატესობა კონკურენტულ ბრძოლაში.

Data Mining-ის აპლიკაციები ვრცლად არის გავრცელებული: საბითუმო ვაჭრობაში, მარკეტინგში, ფინანსებში, მედიცინაში, წარმოებაში და სხვა სფეროებში. [2]

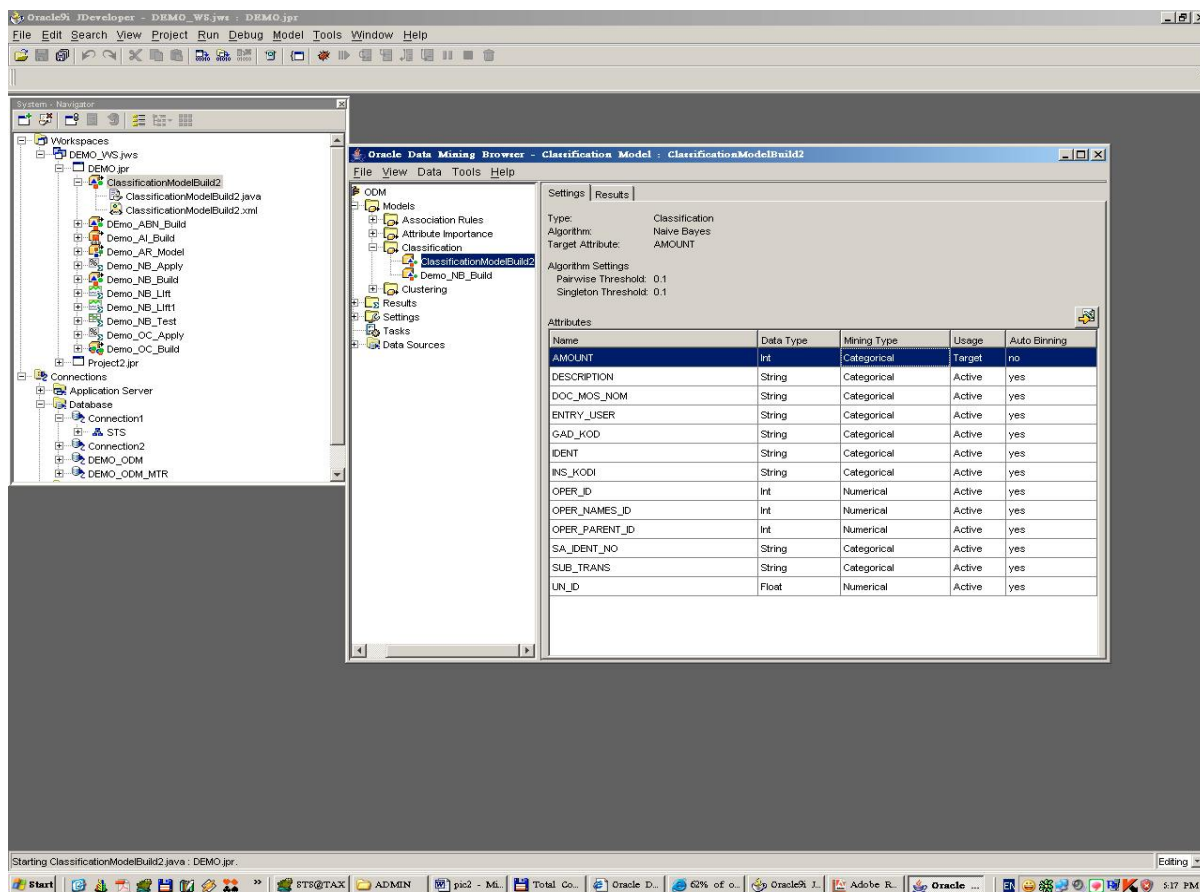
საქართველოში საბაზრო ეკონომიკის დამყარებასა და განვითარებასთან ერთად რა თქმა უნდა იზრდება მფლობელი სუბიექტების (მცირე თუ მსხვილი საწარმოების და ა.შ.), ანუ გადამხდელების რაოდენობა. საგადასახადო სტრუქტურაში დასამუშავებელი ინფორმაციის მოცულობა ყოველდღიურად იზრდება, რაც მოითხოვს ახალი თანამედროვე და ეფექტური ინსტრუმენტებისა და გადაწყვეტილებების დანერგვას. ერთ-ერთი ყველაზე ძირითადი სახელმწიფო ორგანოს ფუნქციონირების ასპექტი, რომელიც არეგულირებს გადასახადების მოგროვების პროცესს, წარმოადგენს ინფორმაციული ტექნოლოგიების განვითარება, რომელიც იძლევა საშუალებას უფრო ოპერატიულად და ზუსტად დავამუშავოდ ინფორმაციის დიდი ნაკადები. გადამხდელისაგან მიღებული ინფორმაციის დამუშავების პრობლემის გადაწყვეტის და მისგან ცოდნის მოპოვების საშუალებას იძლევა Data Mining ტექნოლოგიის გამოყენება. ცოდნის მოპოვება ამ ტექნოლოგიის და კერძოდ პროდუქტ Oracle Data Miner გამოყენება ჩვენი თვალსაზრისით მოქცევს საშუალებას შევამციროდ დროითი და შრომითი დანახარჯები გადამხდელთა პირადი აღრიცხვის ბარათის ინფორმაციის დამუშავებისას, რაც გამოიწვევს სამუშაოს საკმაო ოპტიმიზაციას და საგადასახადო სტრუქტურის მუშაობის ეფექტიანობას. უნდა აღინიშნოს, რომ არსებული ტექნოლოგია როგორც წესი წარმოდგენილია და აღწერილია შემდეგი სფეროებისათვის: საბანკო საქმე, საბითუმო ვაჭრობა, ტელეკომუნიკაცია, დაზღვევა, მედიცინა, მაგრამ ამ დროისათვის ამ ტექნოლოგიის გამოყენება საგადასახადო სისტემისათვის არ არის გავრცელებული.[3]

ამ შემთხვევაში კლავის ობიექტს წარმოადგენს ტექნოლოგია Data Mining. ხოლო კლავის საგანს წარმოადგენს გადამხდელთა პირადი აღრიცხვის ბარათი საქართველოს რესპუბლიკის საგადასახადო დეპარტამენტის ინფორმაციული ტექნოლოგიების სამართველოში არსებული მონაცემთა საცავების საფუძველზე.

გამოიყოფა ხუთი კანონზომიერების სტანდარტული ტიპი, რომლებიც იძლევა საშუალებას გამოვავაშკაროდ Data Mining- ის მეთოდები.

- 1 ასოციაცია
- 2 თანმიმდევრობა
- 3 კლასიფიკაცია
- 4 კლასტერიზაცია
- 5 პროგნოზირება

ჩვენ ვემთხვევაში ვიყენებთ Data Mining- კლასიფიკაციის მეთოდს, რომელიც იყენებს Naivy Bayes ალგორითმს (ნახ.1).



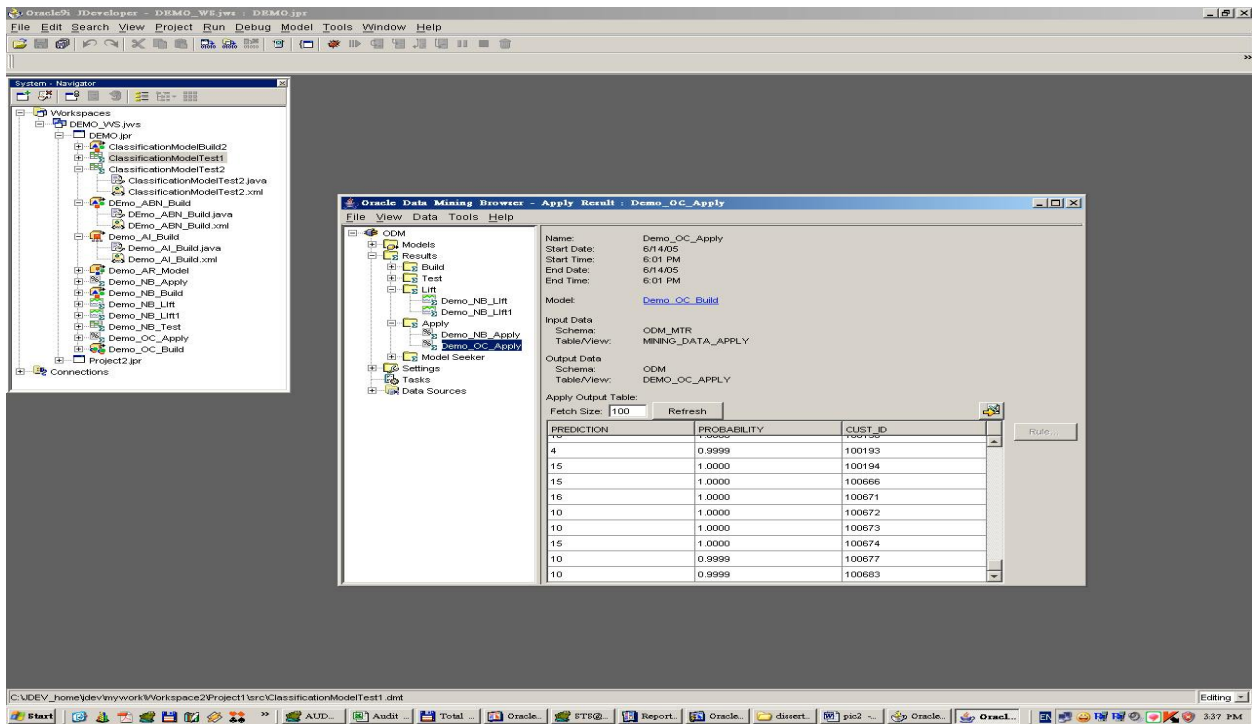
ნახ.1 წარმოდგენილია კლასიფიკაციის მოდელის შექმნის პროცესის ფრაგმენტი NB ალგორითმის საფუძველზე

იმისათვის, რომ წარმოდგენა შეგვექმნას თუ რა პროგნოზირების ჩატარება არის შესაძლო Data Mining ტექნოლოგიის მეშვეობით თავდაპირველად ვაყალიბებთ კვლევის ობიექტის აღწერას ანუ გადამხდელის პირადი აღრიცხვის ბარათის აღწერას.[2]

გადამხდელთა პირადი აღრიცხვის ბარათი (გპაბ) წარმოადგენს მკაცრი აღრიცხვის დოკუმენტს, რომელშიც აისახება გადამხდელების/საგადასახადო აგენტების ან სხვა ვალდებული პირების მიერ დეკლარირებული და გადახდილი თანხები, აგრეთვე საგადასახადო და სხვა უფლებამოსილი ორგანოების მიერ საქართველოს კანონმდებლობით გათვალისწინებული გადასახდელად დარიცხული ან შემცირებული თანხები, ვადაგადაცილებული დავალიანებები, ზედმეტად გადახდილი თანხები, სანქციები, ჯარიმები და საურავები.

გადასახადის გადამხდელის/საგადასახადო აგენტის ან სხვა ვალდებული პირის მიერ საგადასახადო ვალდებულების შესრულების აღრიცხვის წარმოების მიზნით საგადასახადო ორგანოებში იხსნება გპაბ, სადაც შესაძლოა კონკრეტული პერიოდისათვის ბარათზე ჩატარებული ოპერაციების დათვალიერება და შესაბამისად შედეგების ნახვა. ნახ2.





ნახ.3 ალგორითმ NB-ს კრიტერიუმების შერჩევის საფუძველზე მიღებული პროგნოზი

კრიტერიუმების შერჩევის საფუძველზე ვლდებულობთ პროგნოზს, რომლის შედეგადაც შეგვიძლია გამოვაგლინოთ ის დარგები რომლებიც განვითარების სტატიაზეა, განვითარებულ საფეხურზეა და ის დარგები, რომლებიც არიან ცუდად განვითარებულნი. ამ ინფორმაციის საფუძველზე შეგვიძლია დავადგინოთ, თუ რომელ დარგს შეუძლია მოუტანოს ბიუჯეტს მეტი შემოსავალი და რომელ დარგებს უნდა მიექციოს ყურადღება.

**ლიტერატურა:**

1. ნ. კრეჩეტოვი. პროდუქტები მონაცემთა ინტელექტუალური ანალიზისათვის. პროგრამული საშუალებების ბაზარი, №14-15-97, გვ. 32-39
2. Knowledge Discovery Through Data Mining: What is Knowledge discovery? – Tandem Computers Inc., 1996
3. მ. კისელიოვი, ე. სოლომატინი. ცოდნის მოპოვების საშუალებები ბიზნესსა და ფინანსებში. - ღია სისტემები, №14, 1997, გვ. 41-44.
4. საქართველოს ფინანსთა სამინისტროს ბრძანება №228 “გადამხდელთა პირადი აღრიცხვის ბარათების წარმოების წესის შესახებ”//2005 წ. 7 აპრილი, ქ. თბილისი. მუხლი1-5.
5. Oracle Documentation: “Solving Business Problems with Oracle Data Mining” Oracle Inc., 2004.

**ნახატების რაოდენობა: 3**