

## მონაცემთა მეცნიერების გავლენა მსოფლიოში მიმდინარე პროცესებზე

ცირა ბერაძე

ქუთაისის აკაკი წერეთლის სახელმწიფო უნივერსიტეტი

**აბსტრაქტი** - სტატიაში ნაჩვენებია დიდ მონაცემთა თანამედროვე ტექნოლოგიები. აღნიშნულია რომ, მონაცემთა გიგანტური ნაკრებების დამუშავების ხერხებით, მონაცემთა მეცნიერება ნათლად განმარტავს თავიანთ დასკვნებს, ეხმარება მკვლევარებს ოპტიმალური გადაწყვეტილებების მიღებაში და ზუსტი პროგნოზების გაკეთებაში. მონაცემთა მეცნიერება ინფორმაციის დამუშავების ალტერნატიულ ტექნოლოგიას წარმოადგენს და მათი რეალიზება შესაძლებელია თანამედროვე კომპიუტერებზე, რომლებიც გამოიყენება ნეირონული ქსელის სამართავად. მაქსიმალური ეფექტურობით პრობლემების გადაწყვეტა და მოვლენათა წინასწარი პროგნოზირება საზოგადოების უდიდესი მოთხოვნილება და აუცილებლობაა.

**საკვანძო სიტყვები** - მონაცემთა მეცნიერება, მონაცემთა ინჟინერია, მონაცემთა ტექნოლოგიები, ხელოვნური ინტელექტი, ნეირონული ქსელები.

ინტენსიური ეკონომიკურ-ტექნოლოგიური განვითარების შედეგად ირღვევა ბიოსფეროში ნივთიერებათა ბუნებრივი წრებრუნვა, რაც საფრთხეს უქმნის თანამედროვე და მომავალი თაობების ჯანმრთელობას.

მონაცემთა მეცნიერება რომელიც წარმოადგენს კომპიუტერული მეცნიერების დარგს, ეხმარება მკვლევარებს გადაწყვეტილებების მიღებაში და პრობლემების მოგვარებაში. თანამედროვე მსოფლიო სრულად მოიცვა ინოვაციებმა და სიახლეებმა, მონაცემთა ეპოქა უკვე დაწყებულია და ყოველდღიურად განვითარების ახალ საფეხურზე გადადის. კვლევებზე დაყრდნობით, მსოფლიოს მონაცემთა 90% ბოლო 2 წლის განმავლობაში შეიქმნა. მონაცემთა შექმნისა და დაგროვების რიტმი დღითიდღე იზრდება, ხელოვნური ინტელექტი და მონაცემთა მეცნიერება სწრაფად ცვლის ჩვენს გარშემო არსებულ სამყაროს. საზოგადოება, ნელ-ნელა, სულ უფრო მეტს იგებს ხელოვნური ინტელექტის შესაძლებლობების შესახებ და აღიარებს მის როლს თანამედროვე სამყაროში. მონაცემთა მეცნიერებას შეუძლია დაეხმაროს კომპანიებს ცვლილებების პროგნოზირებაში და სხვადასხვა გარემოებებზე ოპტიმალურად რეაგირებაში, მათ შეუძლიათ გამოავლინოს ხარვეზები და პრობლემები, რომლებიც სხვაგვარად შეუმჩნეველი დარჩებოდა. მონაცემთა მეცნიერება რეკოლუციას ახდენს კომპანიების ფუნქციონირებაში, პროგნოზირებადი ანალიზი იყენებს ისტორიულ მონაცემებს, რათა გააკეთოს ზუსტი პროგნოზები მონაცემთა შაბლონების შესახებ, რომლებიც შეიძლება მომავალში მოხდეს. მას ახასიათებს ისეთი ტექნიკები, როგორცაა მანქანათმცოდნეობა, პროგნოზირება, შაბლონების შედარება და პროგნოზირებადი მოდელირება. თითოეულ ამ ტექნიკაში კომპიუტერები არიან გაწვრთნილნი, რათა შეცვალონ მიზეზობრიობის კავშირები.

დიაგნოსტიკური ანალიზი არის მონაცემთა დეტალური გამოკვლევა იმის გასაგებად, თუ რატომ მოხდა რაღაც, იგი ხასიათდება ისეთი ტექნიკით, როგორცაა საბურღი, მონაცემთა აღმოჩენა, მონაცემთა მოპოვება და კორელაციები. მონაცემთა მრავალრიცხოვანი ოპერაციები და ტრანსფორმაციები შეიძლება შესრულდეს უკვე არსებულ მონაცემთა ნაკ-

რებზე, რათა აღმოჩენილი იქნას უნიკალური ნიმუშები თითოეულ ამ ტექნიკაში. ხელოვნურმა ინტელექტმა და მანქანათმცოდნეობის ინოვაციებმა მონაცემთა დამუშავება უფრო სწრაფი და ეფექტური გახადა.

Big Data აღწერს მონაცემთა გიგანტურ ნაკრებებს, რომლებიც გროვდება ინტერნეტის მომხმარებლებისგან. Big Data გადაწყვეტილებები გვთავაზობენ ღრუბლოვან ჰოსტინგს, მაღალ ინდექსირებულ და ოპტიმიზებულ მონაცემთა სტრუქტურებს, ავტომატურ არქივისა და ამოღების შესაძლებლობებს და ანგარიშგების ინტერფეისებს, რომლებიც შექმნილია უფრო ზუსტი ანალიზების უზრუნველსაყოფად, რაც საშუალებას იძლევა მიღებული იქნას უკეთესი გადაწყვეტილებები. ამ მხრივ, ბოლო დროის ერთ-ერთი მნიშვნელოვანი მიღწევა ხელოვნური ინტელექტია, რომელიც უკვე იქცა ჩვენი ყოველდღიურობის მნიშვნელოვან ნაწილად. ხელოვნური ინტელექტი კომპიუტერული სისტემების მიერ ადამიანის ინტელექტუალური პროცესების სიმულაციაა. ხელოვნური ინტელექტის აპლიკაციები მოიცავს საექსპერტო სისტემებს, ბუნებრივი ენის დამუშავებას, მანქანურ ხედვას. ადამიანები მუდმივად მიისწრაფვიან სიახლეებისკენ, ტექნოლოგიური განვითარებისა და შესასრულებელი სამუშაოების შემსუბუქებისკენ. მართლაც, თანამედროვე მეცნიერება ქმნის და ავითარებს ისეთ მექანიზმებს, რომლებიც ადამიანის შრომას გაცილებით ეფექტიანს ხდის, თანაც, გაცილებით მოკლე დროში. დღეისათვის აქტუალური ხდება მათი გამოყენება ისეთ სისტემებში, რომლებშიც შეიცავენ მრავალ დამოუკიდებელ ცვლადს, რომლებიც დეტერმინისტულად არიან დამოკიდებულნი მიზნობრივ ფუნქციაზე, თანამედროვე კომპიუტერები იყენებენ შემეცნებით მიდგომას პრობლემის გადასაწყვეტად. მათ შორის მონაცემთა არქიტექტორები, ბიზნეს ინტელექტის შემქმნელები, მანქანათმშენებლობის მეცნიერები, მანქანათმშენებლობის ინჟინრები, სტატისტიკოსები, ინფრასტრუქტურის არქიტექტორები, მონაცემთა ინჟინრები და მონაცემთა ანალიტიკოსები. რომლითაც პრობლემა უნდა იყოს გადაწყვეტილი, უნდა იყოს რეალიზებული დაპროგრამების ენის არაორაზროვან ინსტრუქციებში და გადაყვანილი მანქანურ კოდში, რომელსაც შემდეგ კომპიუტერი შეასრულებს. ასეთი მანქანები არიან ტოტალურად წინასწარ გამოცნობადი. თუ რაიმე კეთდება არასწორად, იგი გამოწვეულია პროგრამული უზრუნველყოფის ან ტექნიკური მოწყობილობის შეცდომებით.

ნეირონული ქსელები და თანამედროვე ალგორითმული კომპიუტერები არ არიან ურთიერთ შეჯიბრებაში, არამედ ავსებენ ერთიმეორეს. არსებობენ ამოცანები, რომელთათვისაც უფრო სასურველია ალგორითმული მიდგომა და არსებობენ ამოცანები, რომელთათვისაც სასურველია ნეირონული ქსელები. უფრო მეტიც, არსებობენ დიდი რაოდენობა ამოცანებისა, რომლებიც მოითხოვენ სისტემებს, სადაც გამოყენებულია კომბინირებულ მიდგომა (ჩვეულებრივ, თანამედროვე კომპიუტერი გამოიყენება ნეირონული ქსელის სამართავად), რომ მაქსიმალური ეფექტურობით გადაწყდეს პრობლემა.

**დიდი მონაცემები** არის ის, რომ კომპანიებს ექნებათ ბევრად მეტი ინტელექტი, რათა მიიღონ ზუსტი გადაწყვეტილებები და პროგნოზები.

**მოცულობა:** დიდი რაოდენობით მონაცემები გენერირდება სხვადასხვა წყაროდან, როგორცაა სოციალური მედია, IoT მოწყობილობები და ბიზნეს ტრანზაქციები.

**სიჩქარე:** სიჩქარე, რომლითაც ხდება მონაცემების გენერირება, დამუშავება და ანალიზი.

**მრავალფეროვნება:** სხვადასხვა ტიპის მონაცემები, მათ შორის სტრუქტურირებული, ნახევრად სტრუქტურირებული და არასტრუქტურირებული მონაცემები, სხვადასხვა წყაროდან მოდის.

**სიზუსტე:** მონაცემების ხარისხი და სიზუსტე, რომელზეც შეიძლება გავლენა იქონიოს შეუსაბამობამ, გაურკვეველობამ ან თუნდაც დეზინფორმაციამ.

**ღირებულება:** სარგებლობა და პოტენციური მონაცემებიდან ინფორმაციის ამოღების მიზნით, რომელსაც შეუძლია უკეთესი გადაწყვეტილების მიღება და ინოვაცია. ძირითადი სტატისტიკის შეჯამება TechJury დიდი მონაცემების ტენდენციებსა და პროგნოზებზე:

**მონაცემთა მოცულობის ზრდა:** 2025 წლისთვის, გლობალური მონაცემთა სფერო, სავარაუდოდ, მიაღწევს 175 ზეტაბაიტს, რაც აჩვენებს მონაცემთა ექსპონენციალურ ზრდას.

- **IoT მოწყობილობების გაზრდა:** პროგნოზირებულია, რომ IoT მოწყობილობების რაოდენობა 64 წლისთვის 2025 მილიარდს მიაღწევს, რაც შემდგომში ხელს შეუწყობს დიდი მონაცემების ზრდას.
- **დიდი მონაცემთა ბაზრის ზრდა:** გლობალური დიდი მონაცემთა ბაზრის ზომა 229.4 წლისთვის 2025 მილიარდ დოლარამდე გაიზრდება.
- **მზარდი მოთხოვნა მონაცემთა მეცნიერებზე:** 2026 წლისთვის მონაცემების მეცნიერებზე მოთხოვნა 16%-ით გაიზრდება.
- **AI და ML-ის მიღება:** 2025 წლისთვის, ხელოვნური ინტელექტის ბაზრის ზომა 190.61 მილიარდ დოლარს მიაღწევდა, რაც გამოწვეულია AI და ML ტექნოლოგიების მზარდი გამოყენებით Big Data ანალიზისთვის.
- **ღრუბელზე დაფუძნებული Big Data გადაწყვეტილებები:** მოსალოდნელი იყო, რომ ღრუბლოვანი გამოთვლები შეადგენდა მთლიანი დატვირთვის 94%-ს 2021 წლისთვის, რაც ხაზს უსვამს ღრუბელზე დაფუძნებული გადაწყვეტილებების მზარდ მნიშვნელობას მონაცემთა შენახვისა და ანალიტიკისთვის.
- **საცალო მრეწველობა და დიდი მონაცემები:** საცალო მოვაჭრეები, რომლებიც იყენებდნენ Big Data-ს, მოსალოდნელი იყო, რომ გაზარდონ თავიანთი მოგების მარჟა 60%-ით.
- **დიდი მონაცემების მზარდი გამოყენება ჯანდაცვაში:** ჯანდაცვის ანალიტიკის ბაზარი 50.5 წლისთვის 2024 მილიარდ დოლარს მიაღწევდა.
- **სოციალური მედია და დიდი მონაცემები:** სოციალური მედიის მომხმარებლები ყოველდღიურად აწარმოებენ 4 პეტაბაიტ მონაცემს, რაც ხაზს უსვამს სოციალური მედიის გავლენას დიდი მონაცემების ზრდაზე.

**დიდი მონაცემთა ტექნოლოგიებს მიეკუთვნება:**

- **განაწილებული ფაილური სისტემები:** სისტემები, როგორცაა Hadoop განაწილებული ფაილური სისტემა (HDFS) ჩართეთ მონაცემთა დიდი მოცულობის შენახვა და მართვა მრავალ კვანძში. ეს მიდგომა უზრუნველყოფს შეცდომების შემწყნარებლობას, მასშტაბურობას და საიმედოობას დიდი მონაცემების დამუშავებისას.

- **NoSQL მონაცემთა ბაზები:** მონაცემთა ბაზები, როგორცაა MongoDB, Cassandra და Couchbase შექმნილია არასტრუქტურირებული და ნახევრად სტრუქტურირებული მონაცემების დასამუშავებლად. ეს მონაცემთა ბაზები გთავაზობთ მოქნილობას მონაცემთა მოდელირებაში და უზრუნველყოფს ჰორიზონტალურ მასშტაბურობას, რაც მათ შესაფერისს ხდის დიდი მონაცემთა აპლიკაციებისთვის.
- **Map Reduce:** პროგრამირების ეს მოდელი საშუალებას იძლევა დამუშავდეს დიდი მონაცემთა ნაკრები განაწილებულ გარემოში პარალელურად. MapReduce საშუალებას იძლევა რთული ამოცანების დაყოფა უფრო მცირე ქვეამოცნებებად, რომლებიც შემდეგ დამოუკიდებლად მუშავდება და გაერთიანებულია საბოლოო შედეგის მისაღებად.
- **ApacheSpark:** ღია კოდის მონაცემთა დამუშავების ძრავას, Spark-ს შეუძლია გაუმკლავდეს როგორც ჯგუფურ, ასევე რეალურ დროში დამუშავებას. ის გთავაზობთ გაუმჯობესებულ შესრულებას MapReduce-თან შედარებით და მოიცავს ბიბლიოთეკებს მანქანური სწავლისთვის, გრაფიკის დამუშავებისა და ნაკადის დამუშავებისთვის, რაც მას მრავალმხრივს ხდის დიდი მონაცემთა გამოყენების სხვადასხვა შემთხვევისთვის.
- **SQL-ის მსგავსი მოთხოვნის ინსტრუმენტები:** ინსტრუმენტები, როგორცაა Hive, Impala და Presto საშუალებას აძლევს მომხმარებლებს გაუშვან მოთხოვნები Big Data-ზე ნაცნობის გამოყენებით SQL სინტაქსი. ეს ხელსაწყოები ანალიტიკოსებს საშუალებას აძლევს გამოიტანონ შეხედულებები დიდი მონაცემებიდან უფრო რთული პროგრამირების ენების გამოცდილების საჭიროების გარეშე.
- **მონაცემთა ტბები:** ამ საცავის საცავებს შეუძლიათ შეინახონ ნედლეული მონაცემები მშობლიურ ფორმატში, სანამ ის საჭირო იქნება ანალიზისთვის. მონაცემთა ტბები უზრუნველყოფენ მასშტაბურ და ეკონომიურ გადაწყვეტას დიდი რაოდენობით მრავალფეროვანი მონაცემების შესანახად, რომელიც შემდგომში შეიძლება დამუშავდეს და გაანალიზდეს საჭიროებისამებრ.
- **მონაცემთა შენახვის გადაწყვეტილებები:** პლატფორმები, როგორცაა Snowflake, BigQuery და Redshift, გთავაზობენ მასშტაბურ და ეფექტურ გარემოს დიდი რაოდენობით სტრუქტურირებული მონაცემების შესანახად და მოთხოვნისთვის. ეს გადაწყვეტილებები შექმნილია დიდი მონაცემების ანალიტიკისთვის და სწრაფი შეკითხვისა და მოხსენების გასააქტიურებლად.
- **მანქანათმცოდნეობის ჩარჩოები:** ჩარჩოები, როგორცაა TensorFlow, PyTorch და scikit-learn საშუალებას აძლევს ტრენინგის მოდელებს მონაცემთა დიდ ნაკრებებზე ისეთი ამოცანებისთვის, როგორცაა კლასიფიკაცია, რეგრესია და კლასტერირება. ეს ხელსაწყოები გეხმარებათ დიდი მონაცემებიდან მიღებული შეხედულებებისა და პროგნოზების მიღებაში მოწინავე AI ტექნიკის გამოყენებით.
- **მონაცემთა ვიზუალიზაციის ინსტრუმენტები:** ინსტრუმენტები, როგორცაა Tableau, Power BI და D3.js, გვხმარება დიდი მონაცემების ანალიზსა და წარმოდგენაში ვიზუალური და ინტერაქტიული გზით. ეს ხელსაწყოები საშუალებას აძლევს მომხმარებლებს გამოიკვირონ მონაცემები, დაადგინონ ტენდენციები და ეფექტურად მიაწოდონ შედეგები.

- **მონაცემთა ინტეგრაცია და ETL:** ისეთი ინსტრუმენტები, როგორცაა Apache NiFi, Talend და Informatica, იძლევა სხვადასხვა წყაროებიდან მონაცემების მოპოვებას, ტრანსფორმაციას და ჩატვირთვას ცენტრალურ შენახვის სისტემაში. ეს ხელსაწყოები ხელს უწყობს მონაცემთა კონსოლიდაციას, რაც საშუალებას აძლევს ორგანიზაციებს შექმნან ერთიანი ხედვა მათი მონაცემების ანალიზისა და მოხსენებისთვის.

ხელოვნური ინტელექტისა და დიდი მონაცემების გადაფარვა იმაში მდგომარეობს, რომ AI ტექნიკა, განსაკუთრებით მანქანური სწავლა და ღრმა სწავლა (DL), შეიძლება გამოყენებულ იქნას დიდი მოცულობის მონაცემების ანალიზისა და ამოსაღებად. Big Data უზრუნველყოფს აუცილებელ საწვავს AI ალგორითმებისთვის, რათა ისწავლონ და მიიღონ პროგნოზები ან გადაწყვეტილებები. თავის მხრივ, AI-ს შეუძლია დაეხმაროს რთული, მრავალფეროვანი და ფართომასშტაბიანი მონაცემთა ნაკრების გაგებას, რომელთა დამუშავება და ანალიზი ტრადიციული მეთოდების გამოყენებით რთულია. აქ მოცემულია რამდენიმე ძირითადი სფერო, სადაც ხელოვნური ინტელექტი და დიდი მონაცემები იკვეთება:

1. **მონაცემთა დამუშავება:** ხელოვნური ინტელექტის მქონე ალგორითმები შეიძლება გამოყენებულ იქნას დიდი მონაცემთა წყაროებიდან ნედლეული მონაცემების გასასუფთავებლად, წინასწარ დამუშავებისა და გარდაქმნისთვის, რაც ხელს უწყობს მონაცემთა ხარისხის გაუმჯობესებას და უზრუნველყოფს ანალიზისთვის მზადყოფნას.
2. **ფუნქციის მოპოვება:** ხელოვნური ინტელექტის ტექნიკის გამოყენება შესაძლებელია დიდი მონაცემებიდან შესაბამისი ფუნქციებისა და შაბლონების ავტომატურად ამოსაღებად, რაც ამცირებს მონაცემთა განზომილებას და გახდის მას უფრო მართვადი ანალიზისთვის.
3. **პროგნოზირებადი ანალიტიკა:** მანქანათმცოდნეობის და ღრმა სწავლის ალგორითმები შეიძლება ივარჯიშონ დიდ მონაცემთა ნაკრებებზე პროგნოზირებადი მოდელების შესაქმნელად. ეს მოდელები შეიძლება გამოყენებულ იქნას ზუსტი პროგნოზების გასაკეთებლად ან ტენდენციების იდენტიფიცირებისთვის, რაც გამოიწვევს უკეთესი გადაწყვეტილების მიღებას და გაუმჯობესებულ ბიზნეს შედეგებს.
4. **ანომალიის გამოვლენა:** ხელოვნური ინტელექტის წყალობით შესაძლებელია დიდი მონაცემების უჩვეულო შაბლონების ან ამომწურავი ელემენტების იდენტიფიცირება, რაც შესაძლებელს გახდის პოტენციური საკითხების ადრეულ გამოვლენას, როგორცაა თაღლითობა, ქსელში შეჭრა ან აღჭურვილობის გაუმართაობა.
5. **ბუნებრივი ენის დამუშავება (NLP):** AI-ზე მომუშავე NLP ტექნიკა შეიძლება გამოყენებულ იქნას დიდი მონაცემთა წყაროებიდან არასტრუქტურული ტექსტური მონაცემების დასამუშავებლად და გასაანალიზებლად, როგორცაა სოციალური მედია, მომხმარებელთა მიმოხილვები ან ახალი ამბების სტატიები, რათა მიიღოს ღირებული შეხედულებები და განწყობის ანალიზი.
6. **სურათის და ვიდეოს ანალიზი:** ღრმა სწავლის ალგორითმები, განსაკუთრებით კონვოლუციური ნერვული ქსელები (CNN-ები), შეიძლება გამოყენებულ იქნას სურათებისა და ვიდეოების დიდი მოცულობის მონაცემების ანალიზისა და ამოსაღებად.

7. **პერსონალიზაცია და რეკომენდაცია:** AI-ს შეუძლია გააანალიზოს მომხმარებლების, მათი ქცევისა და პრეფერენციების დიდი რაოდენობით მონაცემები, რათა უზრუნველყოს პერსონალიზებული გამოცდილება, როგორცაა პროდუქტის რეკომენდაციები ან მიზნობრივი რეკლამა.

8. **ოპტიმიზაცია:** ხელოვნური ინტელექტის ალგორითმებს შეუძლიათ მონაცემთა დიდი ნაკრების ანალიზი რთული პრობლემების ოპტიმალური გადაწყვეტის დასადგენად, როგორცაა მიწოდების ჯაჭვის ოპერაციების ოპტიმიზაცია, ტრაფიკის მართვა ან ენერჯის მოხმარება.

AI-სა და Big Data-ს შორის სინერგია საშუალებას აძლევს ორგანიზაციებს გამოიყენონ ხელოვნური ინტელექტის ალგორითმების ძალა, რათა გააცნობიერონ მასიური მოცულობის მონაცემები, რაც საბოლოოდ გამოიწვევს უფრო ინფორმირებულ გადაწყვეტილებების მიღებას და უკეთეს შედეგებს.

### ლიტერატურა

1. კ. ოდიშარია, ს. ხოშტარია, ჟ. ებანოიძე სისტემების და პროცესების მოდელირება თბილისი 2011
2. ბ. მეფარიშვილი, გ. ჯანელიძე „საინფორმაციო სისტემების აგება MS SQL Server-ის გამოყენებით“, თბილისი, 2013წ, ISBN 978-9941-20-351-0
3. რ. სამხარაძე, ლ. გაჩეჩილაძე. SQL სერვერი. თბილისი. „ტექნიკური უნივერსიტეტი“. 2016. ISBN 978-9941-28-393-2
4. Geman S., Geman D. 1984. Stochastic relaxation, Gibbs distribution and Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6:721-41.
5. Hinton G. E., Sejnowski T. J. 1986. Learning and relearning in Boltzmann machines. In Parallel distributed processing, vol. 1. Cambridge, MA: MIT Press.

### **The influence of data science on current processes in the world Summary**

The article describes the modern big data technology. It is noted that by methods of processing giant sets of data, the data science clearly explains its findings, helps researchers in taking optimal decisions, and in making accurate predictions. Data science is an alternative information processing technology and can be implemented on modern computers that are used to control neural networks. Solving problems with maximum efficiency and predicting events in advance is the greatest need and necessity of society.