პროექტის სათაური - უძრავი ქონების ფასების პროგნოზირება მანქანური სწავლების საშუალებით

პროექტის შემსრულებლები

1. გიორგი ბერიძე
2. დავით კობალაძე

პროექტის აღწერა

პრობლემის განსაზღვრა (არსებული სიტუაციის ანალიზი)

პროგნოზირება მრავალი წელია აქტიური კვლევის საკითხია. პროგნოზირება მრავალ სფეროს მოიცავს და მრავალი გამოყენება აქვს. პროგნოზირებისთვის იყენებენ მრავალ მეთოდს დაწყებული სტატისტიკური მოდელირებით და დამთავრებული მანქანური სწავლების მეთოდებით.

ჩვენ გადავწყვიტეთ გვეჩვენებინა მანქანური სწავლების მეთოდების ეფექტურება უძრავი ქონების ფასების პროგნოზირების მაგალითზე. გადავწყვიტეთ ეს გვეჩვენებინა ქართული უძრავი ქონების ბაზრის მაგალითზე რადგან ამის მაგალითი ქართული ბაზრისთვის არ არსებობს.

ამჟამად ფასების დასადგენად და პროგნოზირებისთვის გამოიყენება ადამიანის ექსპერტიზა რაც ხშირ შემთხვევაში ძვირი და არაზუსტია რადგან ადამიანი ყველაფერს ცივი გონებით ვერ განსჯის.

მიზანი

პირველ რიგში ჩვენ მიზნად დავისახეთ მოგვეგროვებინა და დაგვეჯგუფებინა მონაცემები ქართულ ბაზარზე უძრავი ქონების ფასების შესახებ. ამისთვის გადავწყვიტეთ უძრავი ქონების გაყიდვის ვებ-საიტებიდან ინფორმაციის მოგროვება ჩვენი შექმნილი პროგრამული უზრუნველყოფის მეშვეობით.

შემდეგი მიზანი არის ის რომ მანქანური სწავლების რამდენიმე მეთოდით გვეცადა მოცემული პროგნოზირების ამოცანის გადაწყვეტა და გვენახა რომელი მეთოდი მოგვცემდა ყველაზე მაღალ სიზუსტეს. ხოლო შემდეგ ამორჩეული მეთოდის პარამეტრების ცვლილების გზით მაქსიმალური სიზუსტის მიღწევას.

ამოცანები

ჩვენი უპირველესი ამოცანა იყო მონაცემების მოგროვება რადგან ქართულ ბაზარზე უძრავი ქონების ფასების მონაცემთა ბაზა არ არსებობს.
შემდეგი მიზანი იყო მანქანური სწავლების მეთოდების მეშვეობით ფასების პროგნოზირება.

განხორციელების გზები

მონაცემების მოსაგროვებლად გადავწყვიტეთ Python-ის და Python-ის ბიბლიოთეკის Scrapy-ს გამოყენება. მოგროვებული ინფორმაციის MongoDB მონაცემთა ბაზაში შენახვა შემდგომი ანალიზისათვის. ამ დავალებაზე პასუხისმგებელი იყო დავითი.

პროგნოზირებისთვის მოვსინჯეთ როგორც ღრმა სწავლების ასევე კლასიკური მანქანური სწავლების მეთოდები. კლასიკური ღრმა სწავლების მეთოდებიდან ვცადეთ გამოგვეყენებინა Support Vector Machine, Gradient Boosting, Logistic Regression, Random Forest. ამ უკანასკნელზე პასუხისმგებელი იყო გიორგი.

მოსალოდნელი შედეგები

მოცემული ტექნოლოგიის გამოყენება შესაძლებელია როგორც არსებული უძრავი ქონების გაყიდვის საიტების გასაუმჯობესებლად, მაგალითად ბინის დამატებისას ფასის შეთავეზებისთვის, ასევე ამ სფეროში მომუშავე ექსპერტების მიერ უძრავი ქონების ფასის შეფასებისას როგორც დამხმარე ხელსაწყო.

მოცემული ინფორმაციის მოგროვების და შემდგომი პროგნოზირების მეთოდების განვრცობა ასევე მარტივად შეიძლება სხვა ტიპის ინფორმაციის პროგნოზირებისთვის მაგალითად ისეთი როგორიცაა ამინდის პროგნოზი ან ნებისმიერი ტიპის ბაზრის კვლევა.

პროექტის შესრულების განრიგი

ამოცანები და აქტივობები	კვირეები														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
მონაცემების მოგროვება															
მონაცემების მოსაგროვებლად პროგრამული უზრუნველყოფის შექმნა															
მონაცემების მოგროვება															
მონაცემების გაწმენდა და მომზადება															
მანქანური სწავლება															
სხვადასხვა მანქანური სწავლების მეთოდის იმპლემენტაცია და სიზუსტის გაზომვა															
საუკეთესო სიზუსტის მანქანური სწავლების მეთოდის არჩევა და															

პარამეტრების ცვლილების მეშვეობით მაქსიმალური შედეგის მიღწევა															