

ივანე ჯავახიშვილის სახელობის თბილისის სახელმწიფო უნივერსიტეტი

ზუსტ და საბუნებისმეტყველო მეცნიერებათა ფაკულტეტი



Tobot - ფსევდო-ინტელექტუალური მოსაუბრე ბოტი

გოგიტა გოგოლიძე
ვარუჟან შახმურადიანი
გიორგი ბერიძე
ნოდარ ჭანკოტაძე
გიორგი ბეროზაშვილი

კომპიუტერულ მეცნიერებათა სრული პროფესორი მანანა ხაჩიძე

თბილისი
2021

სარჩევი

Tobot - ფსევდო-ინტელექტუალური მოსაუბრე ბოტი	3
შესავალი	4
ზოგადი ინფორმაციის მოძიების აუცილებლობა	4
პრობლემის გადაჭრის თეორიული და ტექნოლოგიური მეთოდები	5
განხორციელებული პროექტის შედეგი	5
დასკვნა	7
გამოყენებული ლიტერატურა	7

Tobot - ფსევდო-ინტელექტუალური მოსაუბრე ბოტი

ნაშრომში წარმოდგენილია ფსევდო-ინტელექტუალური მოსაუბრე ჩეთ-ბოტი, რომლის მთავარი დანიშნულებაა თსუს ვებ-საიტზე შემოსული მომხმარებლების დახმარება და მათ მიერ დასმულ კითხვებზე პასუხის გაცემა. პროგრამულ უზრუნველყოფას წინასწარ აქვს განსაზღვრული მონაცემები, რომლის დამუშავების შედეგადაც პოულობს მომხმარებლის დასმულ შეკითხვას და გამოაქვს შესაბამისი პასუხი. ძირითადი კონტექსტი უკავშირდება თბილისის სახელმწიფო უნივერსიტეტს, შესაბამისად მომხმარებელს საშუალება ექნება, მიიღოს ზოგადი ინფორმაცია უნივერსიტეტის შესახებ და ამ მხრივ გამარტივდეს თსუ-სთან კონტაქტი და სიახლეების გაზიარება სტუდენტებისთვის და სხვა პირებისთვის.

This paper presents pseudo-intelligent chatbot that provides answers for questions of Tbilisi State University website visitors. The software searches requested question in predefined data and responds to user with processed answer. Context of data is about Tbilisi State University so users can get any public information about university. This simplifies users' contact to call center and information sharing with students and other people.

შესავალი

დღესდღეობით უამრავი კომპანია იყენებს ჩეთ-ბოტებს თავიანთ ვებ-საიტებზე. ამის მიზეზი არის მომხმარებლისთვის ზოგადი ინფორმაციის მარტივი მიწოდება და ადამიანური რესურსის დაზოგვა. ინფორმაცია, რომელიც მომხმარებელმა ვებ-საიტზე შემოსვლის დროს შეიძლება მოითხოვოს, შეუზღუდავია, თუმცა არსებობს ძირითადი თემები და ხშირად დასმული კითხვების სიმრავლე და შესაბამისად კომპანიამ ან ბიზნემამ ყოველთვის იცის, ძირითადად რა აინტერესებს მის კლიენტს. ამიტომ ზოგადი ინფორმაციის შეგროვება და მომხმარებლების კითხვებზე საპასუხოდ ბოტის გამოყენება მნიშვნელოვნად ეხმარება ერთი მხრივ კომპანიას, ადამიანური რესურსების დაზოგვაში და მეორე მხრივ - კლიენტს, ინფორმაციის სწრაფად მოძიებაში.

Tobot-იც წარმოადგენს ფსევდო-ინტელექტუალურ მოსაუბრე ბოტს, რომლის მიზანია თსუს მომხმარებლებისთვის ინფორმაციის გაზიარება და მათ კითხვებზე საშუალების ფარგლებში პასუხის გაცემა.

თბილისის სახელმწიფო უნივერსიტეტის შესახებ მრავალ მომხმარებელს შეიძლება გაუჩნდეს კითხვა, მათ შორის აბიტურიენტებს, სტუდენტებს, მაგისტრანტებს და ა.შ. კონკრეტული და სპეციფიკური ინფორმაციის მოძიება შესაბამის სამსახურებში არის შესაძლებელი, თუმცა ზოგადი, საჯარო ინფორმაციის მოძიებისთვის ეფექტური იქნება პროგრამული უზრუნველყოფა, რომელიც წინასწარ განსაზღვრული ინფორმაციის დამუშავების გზით პოულობს დასმულ კითხვაზე პასუხს.

ზოგადი ინფორმაციის მოძიების აუცილებლობა

თბილისის სახელმწიფო უნივერსიტეტში რამდენიმე ფაკულტეტი და უამრავი მიმართულებაა, შესაბამისად სწავლობს ათიათასობით სტუდენტი და მაგისტრანტი. გარდა ამისა, ეროვნული გამოცდების პიკის დროს უამრავ აბიტურიენტს და მათთან დაახლოებულ პირს უჩნდება კითხვა უნივერსიტეტის შესახებ, ამიტომ ზოგადი ინფორმაციის მიღებისთვის ისინი მიმართავენ ქოლ-ცენტრს ან პირადად სტუმრობენ უნივერსიტეტს. ამდენი მომხმარებლისთვის ინფორმაციის გასაზიარებლად არსებობს მრავალი ვებ-საიტი და უნივერსიტეტი ხშირად მართავს ღია კარის დღეებს, თუმცა დაინტერესებული პირების რაოდენობა მაინც ძალიან დიდია, ამიტომ საჭიროა, რომ ინფორმაციის გაზიარებისთვის ახალი გზები გამოჩნდეს თანამედროვე ტექნოლოგიების დახმარებით.

პრობლემის გადაჭრის თეორიული და ტექნოლოგიური მეთოდები

მომხმარებლებთან ინფორმაციის გასაზიარებლად კომპანიები რამდენიმე საშუალებას იყენებენ, მათ შორის არის ქოლ-ცენტრის სამსახური, სარეკლამო აქციები, ვებ-საიტები და ა.შ. თუმცა ზოგჯერ მომხმარებელს ძალიან მარტივი და ზოგადი ინფორმაცია აინტერესებს, რომელსაც ვებ-საიტზე სწრაფად ვერ პოულობს მონაცემების დიდი რაოდენობის გამო. ასეთ დროს საკმაოდ მომგებიანია ძიების ველი, რომელშიც მომხმარებელი წერს საკვანძო სიტყვას და პოულობს შესაბამის სტატიას, თუმცა ესეც არ არის საკმარისი ზოგიერთი ინფორმაციისთვის, რადგან ძიების ველი მხოლოდ ვებ-საიტზე არსებულ მონაცემების საკვანძო სიტყვებში ეძებს, რაც ხშირ შემთხვევაში არ არის საკმარისი.

დღეს მრავალი კომპანია იყენებს მოსაუბრე ბოტებს, რომლებიც ისეთ შთაბეჭდილებას ქმნიან, ვითომ კომპეტენტურ პირთან ხდება საუბარი, ეს ძალიან საინტერესო და მიმზიდველია მომხმარებლისთვის. ჩეთ-ბოტს გარდა ინფორმაციის მოძიებისა, მომხმარებელთან კომუნიკაციის უნარები აქვს, რის დახმარებითაც მომხმარებელს ისეთ განწყობას უქმნის, რომ რეალურ ადამიანს ესაუბრება და ეს ძალიან დიდი კომფორტია.

ჩეთ-ბოტების შესაქმნელად იყენებენ მონაცემების დასამუშავებელ და დასწავლად, მათ შორის მანქანური სწავლების მეთოდებს. სტატისტიკური დამუშავებით და მრავალი თეორიის გამოყენებით პროგრამული უზრუნველყოფა ახერხებს მიღებული ინფორმაციის სწრაფ ანალიზს და არსებულ მონაცემებში სასურველი პასუხის ძებნას, შედეგად კი მომხმარებელთან გააქვს ისეთი პასუხი, რომელიც არსებული მონაცემების ფარგლებში საუკეთესოა.

განხორციელებული პროექტის შედეგი

პროექტის მიზანი არის მსგავსი ჩეთ-ბოტის შექმნა, რომელიც დაეხმარება აბიტურიენტებს, სტუდენტებს და სხვა დაინტერესებულ პირებს, მიიღონ უნივერსიტეტის შესახებ ზოგადი ინფორმაცია.

პროგრამული უზრუნველყოფა მონაცემების გადასამუშავებლად სხვა ჩეთ-ბოტებისგან განხვავებით არ იყენებს მანქანური სწავლების მეთოდებს და ნეირონულ ქსელებს. მონაცემთა ბაზაში წინასწარ გაწერილია კითხვა-პასუხი. ინფორმაციის დამუშავება და ორი ტექსტის შედარება ხდება მიახლოებითი ძებნის ალგორითმებით და რამდენიმე ალგორითმის კომბინაციის შედეგად დგინდება მომხმარებლის კითხვასთან ფონეტიკურად ყველაზე ახლოს მდგომი კითხვა, შედეგად კი ამ კითხვის პასუხი გამოდის ჩეთში.

Tobot არქიტექტურულად რამდენიმე დონიანი პროგრამული უზრუნველყოფაა. პირველ დონეზე ხდება მომხმარებელთან კომუნიკაცია, კითხვების მიღება და ინფორმაციის გადაგზავნა შემდეგ დონეზე, სადაც ხდება კითხვის ტექნიკური დამუშავება. შემდეგ დამუშავებული მონაცემი იგზავნება მესამე - ალგორითმულ

დონეზე, რომელიც თავის მხრივ იწყენებს მონაცემთა ბაზის მეოთხე დონეს, მოაქვს საჭირო ინფორმაცია და ტექსტის დამუშავების ალგორითმებით ახერხებს არსებული მონაცემებიდან საუკეთესო პასუხის მოძიებას. საბოლოოდ კი შედეგი გადის მომხმარებელთან, რომელსაც სურვილის შემთხვევაში შეუძლია უკუკავშირის დაბრუნება.

დამუშავების ალგორითმებს წარმოადგენს Wagner-Fischer-ის მინიმალური მანძილის და Jaro-Winkler-ის მსგავსობის დადგენის ალგორითმების კომბინაცია (ამასთანავე ფარდობითი შედარება ხდება როგორც ასოებს შორის, ისე სიტყვებს შორითაც, ამიტომ ქართული ენის სპეციფიკიდან გამომდინარე სიტყვების გადალაგებით საერთო აზრი არ იცვლება), რომლითაც დგინდება ორ წინადადებას შორის რა არის მსგავსების კოეფიციენტი და მონაცემთა ბაზაში რომელ სხვა წინადადებასთან აქვს მოცემულ წინადადებას მაქსიმალური დამთხვევა. შედეგად გამოდის საუკეთესო პასუხი და მომხმარებელი ეცნობა ამ ინფორმაციას.

გარდა ცალსახა დამუშავებისა, პროგრამულ უზრუნველყოფას გააჩნია დასწავლადი აგენტის ელემენტები. კერძოდ, ჩეთ-ბოტი მომხმარებელთან კომუნიკაციის დროს მოძიებული პასუხის გატანის შემდეგ, გარკვეულ სტატისტიკურ შემთხვევებში მომხმარებელს სთხოვს უკუკავშირის დაბრუნებას, იმის მიხედვით მომხმარებელი დადებითად განიხილავს ბოტის პასუხს თუ უარყოფითად, ეს პასუხი აისახება მონაცემთა ბაზაში და სამომავლოდ ბოტი იღებს ახალ კითხვას და სხვა მომხმარებელთან უფრო ზუსტად შეეცდება პასუხის დაფიქსირებას. შესაბამისად საერთო ჯამში Tobot-ს გააჩნია თვითგანვითარების საშუალება და მომხმარებელთან ინტერაქციის დროს ახალი დეტალების “სწავლა”.

პროექტი დაწერილია .NET core აპლიკაციაში, რომელიც იყენებს MSSQL მონაცემთა ბაზას. ძირითადი სამუშაო დაპროგრამების ენა არის C#. პროგრამული უზრუნველყოფის ვიზუალური ნაწილი შექმნილია Angular framework-ის გამოყენებით. ხოლო ძირითადი მონაცემები მოძიებულია ზოგადი ცნობებით და თბილისის სახელმწიფო უნივერსიტეტის ვებ-გვერდზე არსებული ინფორმაციით.

პროექტის სამომავლო პერსპექტივა არის, რომ დაემატოს ქოლ-ცენტრის თანამშრომლებისთვის CMS, რათა მათ თავად შეძლონ ახალი ინფორმაციის დამატება და ხშირად დასმულ კითხვების შესახებ ინფორმაციის შეტანა.

პროექტის კიდევ ერთი უპირატესობა ის არის, რომ პროგრამული უზრუნველყოფა გამოირჩევა მოდულარული არქიტექტურით. არქიტექტურის ყოველი მოდული ერთიან სისტემას წარმოადგენს, ამიტომ რაიმე სახის ცვლილება ძალიან მარტივია. ამასთანავე აპლიკაციის ნაწილი არ არის კავშირში მონაცემთა ბაზის ელემენტების კონტექსტზე. კერძოდ, თუ მონაცემთა ბაზაში შევიყვანთ ინფორმაციას სხვა თემაზე ან სხვა კონტექსტზე, ის ყველანაირი ცვლილების გარეშე იმუშავებს, რადგან დამუშავების ალგორითმები არ არის კონტექსტურად დაკავშირებული მონაცემებთან.

დასკვნა

ჯგუფური პროექტის ფარგლებში შესრულებული ამოცანა წარმოადგენს ბაკალავრიატზე ნასწავლი მასალის და თვითგანათლებით მიღებული ცოდნის გაერთიანებას და გადაჭრის პრობლემას, რომელიც მიმდინარე პერიოდში დგას. პროგრამული უზრუნველყოფით შესაძლებელია უამრავი მომხმარებლის დახმარება და ამისთვის ადამიანური და ადმინისტრაციული რესურსის დაზოგვა.

გამოყენებული ლიტერატურა

https://en.wikipedia.org/wiki/Jaro-Winkler_distance

https://en.wikipedia.org/wiki/Wagner-Fischer_algorithm

https://en.wikipedia.org/wiki/Modular_programming