

# ქართული ენა ხმის ამომცნობ სისტემებში

ლაშა გოგიჩაიშვილი

*სამაგისტრო ნაშრომი წარდგენილია ილიას სახელმწიფო უნივერსიტეტის  
ბიზნესის, ტექნოლოგიისა და განათლების ფაკულტეტის,  
ტექნოლოგიის სკოლის, პროგრამული ინჟინერის მაგისტრის აკადემიური ხარისხის  
მინიჭების მოთხოვნების შესაბამისად*

სამაგისტრო ნაშრომი

სამეცნიერო ხელმძღვანელი: მიხეილ თუთბერიძე, ასოცირებული პროფესორი

ილიას სახელმწიფო უნივერსიტეტი

თბილისი, 2021

## განაცხადი

როგორც წარდგენილი სამაგისტრო ნაშრომის ავტორი, ვაცხადებ, რომ ნაშრომი წარმოადგენს ჩემს ორიგინალურ ნამუშევარს და არ შეიცავს სხვა ავტორებისადმი აქამდე გამოქვეყნებულ, გამოსაქვეყნებლად მიღებულ ან/და დასაცავად წარდგენილ მასალებს, რომლებიც ნაშრომში არ არის მოხსენიებული ან ციტირებული სათანადო წესების შესაბამისად.

ლაშა გოგიჩაიშვილი

20 მაისი 2021

## აბსტრაქტი

ადამიანსა და კომპიუტერს შორის ურთიერთობის ყველაზე ხშირი მეთოდები გრაფიკულ ინტერფეისზე დამყარებული, სამართავი მოწყობილობებით მათი გამოყენებაა. დღეს ეს მეთოდები თუმცა სტანდარტია მაგრამ, პრიმიტიულ მეთოდებად მოიაზრება. HMI (human machine interaction) დღითიდღე იხვეწება და დღეს ჩვენ გვაქვს ისეთი მეთოდები, როგორსაც ადამიანები ერთმანეთთან ურთიერთობისთვის ვიყენებთ, ესენია: ჟესტები, ხმოვანი ბრძანებები და ა. შ. ეს მეთოდები ქართულად მოსაუბრე ხალხში არც ისე დიდი პოპულარობით სარგებლობს. ძირითადად ენის ბარიერის გამო, თუმცა ჟესტებისა და სურათის ამოცნობას ენის ბარიერი არ ეხება, ეს ბარიერი ეხება ხმოვან ბრძანებებს. დღეს კი თითქმის არ შეინიშნება ისეთი პროდუქტების არსებობა, რომლებიც ქართულად მოსაუბრე ხალხისთვისაა განკუთვნილი და დამყარებულია ხმოვან ბრძანებებზე. ამ პრობლემის გადასაჭრელად ჩატარდა გამოყენებითი კვლევა, რომელიც გვაჩვენებს თუ რა პრაქტიკული გამოყენება შეიძლება ჰქონდეს ასეთ პროდუქტს. რას ცვლის ის ქართულად მოსაუბრე ხალხში, რა ღირებულება აქვს და აქვს თუ არა გამართლება ასეთი პროდუქტის არსებობას. მაგალითად ხმოვანი ბრძანებებით საზოგადოებრივი მოხმარების აპარატების (ბანკომატი, სწრაფი გადახდის აპარატი, ლიფტი) მართვა ამარტივებს მათ მოხმარებას, ხდის ხელმისაწვდომს უნარშეზღუდული პირებისთვის და თავიდან გვაცილებს ფიზიკურ კავშირს ბაქტერიული კოლონიზაციის წყაროებთან, რაც დღეს, პანდემიის პირობებში განსაკუთრებულად მნიშვნელოვანია. კვლევა მიზნად ისახავს პროტოტიპის შექმნას, რომელსაც ექნება უნარი ამოიცნოს ქართული მოკლე ხმოვანი ბრძანებები. დასახული მიზნის მისაღწევად შეირჩა ისეთი, რომელიც დღეს ხელოვნების ნიმუშად არის აღიარებული, ტრადიციული მეთოდებისგან განსხვავებით მაღალი სიზუსტით და შესასწავლი მასალიდან საჭირო მახასიათებლების ავტომატური ამოცნობით გამოირჩევა და დამყარებულია ადამიანის ტვინის შესწავლის უნარის მეთოდზე. ეს მეთოდი კონვოლუციური ნეირონული ქსელებია (CNN). CNN-ი ისეთ უსტრუქტურო მონაცემების დასამუშავებლადაა განკუთვნილი, როგორიცაა

სურათი, მაგრამ მიზანი არა სურათის, არამედ ხმის ამოცნობაა. იმისთვის, რომ შესაძლებელი იყოს CNN-ის ამ მიზნით გამოყენება, საჭიროა ხმოვანი ტალღები გარდაიქმნას სპექტროგრამად.

ხმოვანი ბრძანებების ამოცნობის მოდელის შექმნა, მითუმეტეს ისეთი უცნობი და იშვიათი ენის მოდელის შექმნა როგორცაა ქართული, საკმაოდ ბევრ შესაძლებლობას იძლევა. პოტენციურად შეუძლია ყველა სფეროს შეეხოს და შეცვალოს უკეთესობისკენ. ადამიანის პროდუქტიულობა კი საგრძნობლად გაზარდოს.

**ძირითადი საძიებო სიტყვები:** ხმის ამომცნობი სისტემები, ქართული ენის ხმის, ამომცნობი სისტემები, ნეირონული ქსელები, კონვოლუციური ნეირონული ქსელები, ხმოვანი ბრძანებები ქართულ ენაზე.

## **Abstract**

There are a lot of methods for interaction between human and computer, the most popular one is based on graphical interface and depends on external devices (mouse, keyboard, touchscreen). Today these methods are standard, but they are considered as primitive methods. HMI (human machine interaction) is being improved day by day and today we have methods, which are used between humans, these are: hand gestures, speech commands and others. This concept dates back to computer's early days, but only later became practical application possible. These methods' popularity increases exponentially. Although, it's not as popular in Georgian speaking people, mostly because of the language barrier, this barrier is not applicable for hand gestures but it applies to speech recognition. Today Georgian based speech recognition systems are almost nonexistent. In order to fix this problem, applied research was conducted, which will show us what practical usage can be found for speech recognition systems, what does it change

for Georgian speaking people and if its existence can be justified. For example, speech controlled public usage objects (ATMs, elevators, etc.) make them more accessible, easier to use and makes physical contact obsolete, this means less contact with bacterial colonization sources, which is especially important in pandemic. Research aims to create a product which will be capable of identifying short speech commands in Georgian. To reach the goal, searching for speech recognition methods was necessary. From existing methods, state of art, for today's standards, method was chosen. This method differs from traditional ones with its high accuracy and automatic feature extraction, which is modelled on an actual biological brain's capability to learn. It's called neural networks (NN), from its different variations the most suitable for small speech commands recognition, convolutional neural networks (CNN) was chosen. CNN is specifically designed to work with unstructured data, like images and it has high accuracy in doing so. But the aim is to recognize speech and not images. In order to use CNN for this, first it is necessary to convert sound waves into an image, which will have all the features needed for recognizing spoken words. For this audio waves go through couple transformation techniques. In the end we will have a spectrogram, that holds all the features needed and it will have a format of an image, which is necessary for training CNN. Creation of speech recognition model, especially for a rare and unique language such as Georgian, gives us endless possibilities. It has potential to apply to most of the fields of work, change them for the better and increase productivity.

**Key Words:** Speech recognition systems, Georgian speech recognition systems, Neural networks, Convolutional neural networks, Georgian speech commands.