

ტექსტების კლასიფიკაცია ნეირონული ქსელებით

გ.ბესიაშვილი,პ.ქარჩავა

ელ-ფოსტა: { [gela.besiashvili.papuna.qarchava](mailto:gela.besiashvili.papuna.qarchava@tsu.ge) }@tsu.ge.

კომპიუტერულ მეცნიერებათა დეპარტამენტი,
ზუსტ და საბუნებისმეტყველო მეცნიერებათა
ფაკულტეტი, ივ.ჯავახიშვილის სახელობის
თბილისის სახელმწიფო უნივერსიტეტი
თბილისი. ჭავჭავაძის გამზ. 3

ტექსტების კლასიფიკაცია (სენტემენტ ანალიზი ნეირონული ქსელებით), სხვა ტიპის ინფორმაციის კლასიფიკაციისგან განსხვავდება იმით, რომ თითოეულ სიტყვას, სხვადასხვა კონტექსტში განსხვავებული მნიშვნელობა აქვს, რაც ართულებს ამოცანას. სტატიაში განხილულია ორი ტიპის ნეირონული ქსელი, კონვოლუციური და რეკურენტული ნეირონული ქსელი.სტატიაში,მოდელის ასაგებად გამოყენებულია ფილმებზე კომენტარების ბაზა. აღნიშნული მოდელის მეშვეობით შეგვიძლია მაღალი სიზუსტით(-97% სიზუსტით) დავადგინოთ დადებითია ფილმის შეფასება თუ უარყოფითი. ტექსტების დასამუშავებლად (ანალიზი, კატეგორიზაცია, კლასიფიკაცია)ძირითადად რეკურენტული ნეირონული ქსელები გამოიყენებოდა, განსხვავებით კონვოლუციური ნეირონული ქსელისა, რომლებსაც აქამდე ძირითადად მულტიმედიური ტიპის ინფორმაციის კლასიფიკაციისთვის იყენებდნენ. თუმცა ამ კვლევებმა გვიჩვენა, რომ კონვოლუციები წარმატებით შეგვიძლია გამოვიყენოთ ტექსტების კლასიფიკაციისთვისაც. კონვოლუციები საშუალებას გვამლევს დასწავლის პროცესში გავითვალისწინოთ სიტყვების კონტექსტი, რაც ჭრის ტექსტის კლასიფიკაციის მთავარ პრობლემას, კონტექსტის დაკარგვას და ამასთანავე აუმჯობესებს მიღებულ შედეგებს.

ნაშრომის მიზანია შეიქმნას სისტემა, რომლის გამოყენებითაც მომხმარებელი ააგებს მისთვის სასურველ მოდელს და ამ მოდელის გამოყენებით დაადგენს შეტანილი ტექსტური ინფორმაციის სენტემენტს (დადებითია შეფასება თუ უარყოფითი).