

#RAČUNALNE ZNANOSTI#

Računalni znanstvenici Instituta Ruđer Bošković razvili su novu metodu grupiranja podataka koju odlikuje poboljšana točnost te je ocijenjena fundamentalnim doprinosom problemu grupiranja podataka.

Kako je priopćeno iz Ruđera, znanstvenici Laboratorija za reprezentacije znanja i strojno učenje Zavoda za elektroniku Maria Brbić i Ivica Kopriva razvili su novu metodu za grupiranje podataka temeljenih na modelu linearnih potprostora kao generatora odgovarajućih funkcionalnih skupina.

Rad je objavljen u jednom od najutjecajnijih znanstvenih časopisa u području računalnih znanosti i umjetne inteligencije 'IEEE Transactions on Cybernetics', koji se po faktoru odjeka 8.803 svrstava na treće mjesto u tome području, rečeno je.

Razvijena metoda je rezultat istraživanja doktorandice Marije Brbić u okviru istraživačkog projekta Hrvatske zaklade za znanost voditelja Ivice Koprive.

Strojno učenje je grana umjetne inteligencije koja se bavi oblikovanjem algoritama za automatsku obradu podataka, odnosno to je proces otkrivanja znanja iz velike količine podataka pri čemu računarski sustavi sami automatski poboljšavaju svoje procese kroz iskustvo.

Strojno učenje je temelj podatkovne znanosti, a dijeli se na nadzirano, nenadzirano i polunadzirano, napominje se.

To je jedno je od danas najuzbudljivijih područja računarske znanosti zbog brojnih mogućnosti primjena od raspoznavanja uzoraka i dubinske analize podataka do robotike, računalnog vida, bioinformatike i računalne lingvistike do medicine, kaže se.

Neke od najčešćih primjena grupiranja podataka u medicini odnose se na segmentaciju slike. Primjerice, kod CT slike skupine predstavljaju organi, kod PET slike skupine su tkiva, kod mikroskopske slike histopatoloških preparata skupine su tkiva i/ili stanice, a kod slike optičke koherentne tomografije oka, primjerice, skupine su slojevi unutar mrežnice, objašnjava se u priopćenju.

'Primjene koje smo ilustrirali u ovom novom radu odnose se na prepoznavanje lica odnosno grupiranje slika lica u skupine koje odgovaraju osobama, zatim prepoznavanje govornika, odnosno grupiranje značajki govora u skupine koje odgovaraju osobama, te prepoznavanje rukom pisanih brojeva, odnosno grupiranje slika u skupine koje

odgovaraju znamenkama od 0 do 9.', objašnjava Ivica Kopriva.

Razvijene metode grupiranja podataka kod navedenih primjera temelje se na modelu prema kojem su podaci unutar svake skupine generirani iz pripadajućeg linearnog potprostora.

Temeljem tog modela razvijeni su algoritmi koji daju vrlo kompetitivne rezultate na grupiranju zahtjevnih skupova podataka, kaže se.

'Ključan element u ovom pristupu je učenje matrice reprezentacije koja je rijetka i ima nizak rang. Umjesto konveksnih mjera ranga i rijetkosti koje se standardno koriste, u radu su predložene mjere koje bolje procjenjuju rang i rijetkost. Te mjere su s jedne strane egzaktne mjere temeljene na L_0 i Schatten-0 kvazi normama, a druge glatka surogat funkcija L_0 i Schatten-0 kvazi normi", kaže Maria Brbić.

Uvođenjem tih mjera odgovarajući optimizacijski problemi su postali nekonveksni, što dokaz konvergencije algoritma čini zahtjevnim. Unatoč tim poteškoćama, u radu je dan teorijski dokaz globalne konvergencije navedenog optimizacijskog problema za učenje matrice reprezentacije, kaže se u priopćenju.

'Novi algoritmi su značajno poboljšali točnost u usporedbi s postojećim metodama na svim testiranim primjenama.', zaključuje Maria Brbić.

(Hina)

Ruderovci razvili novu metodu grupiranja podataka

Kategorija: MAGAZINAžurirano: Petak, 11. Siječanj 2019. 12:18

Objavljeno: Petak, 11. Siječanj 2019. 12:18

