

ZAVRŠNI ZADATAK br. 660

Pristupnik: **Franjo Mindek (0036523095)**

Studij: Elektrotehnika i informacijska tehnologija i Računarstvo

Modul: Računarstvo

Mentor: doc. dr. sc. Mario Brčić

Zadatak: **Sustav za označavanje tekstualnog sadržaja navođeno aktivnim učenjem**

Opis zadatka:

Strojno učenje je veoma uspješna, štoviše dominantna, paradigma za obradu teksta. No, efektivnost tih metoda znatno ovisi o kvaliteti ulaznih podataka. Za postojeća je rješenja uobičajeno da se utreniravaju nad podacima koji su indiskriminativno skupljeni s interneta i labelirani za neku općenitu ili specifičnu svrhu. To rezultira potrebom za velikim skupovima da bi se ostvarila željena performansa. Za uspjeh samog model u određenoj zadaći je bitno imati kvalitetan skup podataka koji učinkovito pokriva prostor mogućnosti tako da model može pouzdano i precizno vršiti svoju namjenu. Aktivno učenje povećava uzorkovnu učinkovitost postupka utreniravanja modela jer se pronalaze čim informativniji uzorci. Time se smanjuju veličina, trošak i trajanje postupka označavanja skupa podataka. U ovom radu treba, koristeći dostupne alate, implementirati sustav za označavanje tekstualnog sadržaja koje je navođeno aktivnim učenjem. Označivačima podataka treba biti izloženo web sučelje koje ih vodi kroz postupak te im zadaje uzorke koji su visoko informativni za odabrani model.

Rok za predaju rada: 10. lipnja 2022.