

NASTAVNO-NAUČNOM VEĆU ELEKTROTEHNIČKOG FAKULTETA U BEOGRADU

Na sednici održanoj 15.07.2014. godine Komisija za II stepen studija Elektrotehničkog fakulteta u Beogradu imenovala nas je u Komisiju za pregled i ocenu master rada Ertana Ljajića pod naslovom „Sentiment analiza tekstualnih komentara sportskih događaja na engleskom jeziku”.

Komisija je pregledala priloženi rad i dostavlja Nastavno-naučnom veću Elektrotehničkog fakulteta u Beogradu sledeći

IZVEŠTAJ

1. Biografski podaci

Ertan Ljajić je rođen 27.10.1982. godine u Novom Pazaru. Osnovnu školu i Gimnaziju je završio u Novom Pazaru, obe sa odličnim uspehom, a osnovnu i sa Vukovom diplomom. Elektrotehnički fakultet u Beogradu upisao je 2001. godine, na odseku za Računarsku tehniku i informatiku. Diplomirao je u februaru 2009. godine sa prosečnom ocenom na ispitima 7.38, na diplomskom 10. Master studije upisao je na Elektrotehničkom fakultetu u Beogradu, na odseku za Računarsku tehniku i informatiku. Položio je sve ispite sa prosečnom ocenom 9.2.

Od 2009. godine radi u preduzeću za puteve u Novom Pazaru. Prvobitno je radio kao sistem administrator, a nakon toga kao šef službe za Informacione Tehnologije. Radi na poslovima postavljanja i održavanja računarske mreže i Informacionih Sistema. Uža oblast njegovog interesovanja i rada su Informacioni Sistemi klase ERP, kako u domenu razvoja tih sistema, tako i u njihovom korišćenju.

2. Predmet, cilj i metodologija istraživanja

Predmet istraživanja je određivanje sentimenta tekstualnih komentara sportskih događaja prikupljenih sa više izvora (internet mediji, socijalne mreže, internet forumi). U radu su date definicije osnovnih pojmova, kao i pregled postojećih rešenja za određivanje sentimenta tekstualnih komentara, korišćenjem metoda mašinskog učenja.

Cilj rada je bio da se kreira nova metodologija za određivanje sentimenta tekstualnih komentara sportskih događaja. Predloženo rešenje rešava neke od problema sa kojim se susreću postojeća i daje bolje rezultate, što je pokazano kroz analizu podataka.

Da bi se došlo do novog rešenja, koje daje bolje rezultate, izvršeno je istraživanje i analiza postojećih rešenja, koja su u radu opisana. Na osnovu toga, predloženo je novo rešenje, koje daje bolje rezultate od postojećih. Da bi bilo pokazano da je novo rešenje zaista bolje, u toku izrade rada kreiran je korpus sportskih komentara i specijalizovan sportski rečnik, na osnovu kojih je bilo moguće sprovesti treniranje, testiranje modela i analiza rezultata.

3. Sadržaj rada

U uvodnom poglavlju je data definicija analize sentimenta, kao podoblasti procesiranja prirodnih jezika.

U drugom poglavlju su date oblasti gde je moguća primena analize sentimenta. Objašnjen je osnovni algoritam za analizu sentimenta, konkretno za određivanje polariteta komentara sportskih događaja.

U trećem poglavlju je dat pregled postojećih rešenja za određivanje polariteta tekstova, koja se uglavnom oslanjaju na gotove rečnike ili rečnike kreirane procesom delimično nadgledanog učenja. Opisane su osnovne metode mašinskog učenja koje se koriste za analizu sentimenta.

Predložena metoda za određivanje sentimenta sportskih komentara opisana je u četvrtom poglavlju. Ona se zasniva na nadogradnji nekih od osnovnih rečnika sa pozitivnim i negativnim rečima. U radu je korišćen rečnik „Opinion Lexicon“, autora Bing Liua. Opisan je postupak prikupljanja sirovih podataka (sportskih komentara) i njihovo sređivanje, kreiranje specijalizovanog sportskog rečnika. Objašnjeno je i na kojim način su se prikupljeni podaci koristili za testiranje i treniranje modela, kroz proces odabira i primene metoda klasifikacije.

U petom poglavlju su dati rezultati u određivanju sentimenta sportskih komentara koje daje predložena metoda. Predstavljen je i komparativni prikaz tih rezultata sa rezultatima koje daju postojeće metode gde se vidi da predložena metoda daje bolju procenu sentimenta sportskih komentara u odnosu na postojeće.

U šestom poglavlju su navedeni neki gotovi open source alati za analizu sentimenta.

4. Rezultati rada

U master radu Ertana Ljajića je predložena nova metodologija za određivanje sentimenta sportskih komentara koja daje bolje rezultate od postojećih, koje su u radu takođe detaljno prikazane. Nova metodologija je kreirana kao kombinacija postojećih. Da je predloženo rešenje zaista i bolje od postojećih pokazano je rezultatima iz primene svih ovih metodologija nad realnim podacima. Za tu potrebu je prikupljen korpus od 1.194 sportskih komentara, kao i kreiran specijalizovan rečnik sa ukupno 467 reči fraza.

U radu je kreiran i korpus sportskih komentara sa pridruženim polaritetom, kao i specijalizovani rečnik sportskih termina koji se mogu koristiti kao osnova za dalja istraživanja.

5. Zaključak i predlog

Kandidat Ertan Ljajić je u svom master radu na jasan i pregledan način predstavio postojeća rešenja za analizu sentimenta tekstualnih komentara sportskih događaja. U radu je prikazao prednosti i mane postojećih rešenja i predložio novo rešenje koje predstavlja kombinaciju postojećih i koje daje bolje rezultate u određivanju sentimenta tekstualnih komentara sportskih događaja.

Kandidat je pokazao samostalnost, sistematičnost i zrelost u istraživačkom postupku. Rad kandidata sadrži dobru analizu postojećih rešenja iz koje se može izvesti zaključak o mogućim poboljšanjima, od kojih je jedno od njih i predstavljeno u radu kao predloženo rešenje.

Na osnovu gore navedenog, smatramo da predloženi master rad Ertana Ljajića pod naslovom „Sentiment analiza tekstualnih komentara sportskih događaja na engleskom jeziku“ zadovoljava sve potrebne uslove da bude prihvaćen kao diplomski-master rad, stoga predlažemo Nastavno-naučnom veću Elektrotehničkog fakulteta u Beogradu da isti prihvati kao master rad i odobri njegovu javnu usmenu odbranu.

U Beogradu, 15.09.2014.

Članovi Komisije


Dr. Veljko Milutinović, redovni profesor


Dr. Boško Nikolić, vanredni profesor