

Croatian Sentiment Analysis

UPUTE ZA PROVJERU OZNAČAVANJA

U ovome su dokumentu opisane upute za provjeru označavanja sentimenta u rečenicama iz hrvatskih tekstova.

Sentiment kao stručni naziv obuhvaća stavove, emocije, mišljenja i vrijednosne sudove. Kada rečenici pridajemo sentiment, zapravo pokušavamo odrediti subjektivan doživljaj autora prema izrečenom u rečenici iz perspektive prosječnoga govornika hrvatskoga jezika

Sentiment je u ovoj fazi automatski označen na razini rečenice, ali kako ćete dobiti rečenice koje pripadaju istome tekstu, pri provjeri automatskoga označavanju sentimenta pojedine rečenice možete u obzir uzeti i širi **kontekst** tj. okolne rečenice.

Proces provjere označavanja sentimenta sastoji se od dvaju potprocesa:

1. razumijevanja – pri čemu čovjek-označivač razumijeva sadržaj rečenice
2. procjene sentimenta – pri čemu čovjek-označivač provjerava je li sentiment ispravno automatski dodijeljen rečenici.

Sentiment se određuje prema 5-stupanjskoj ljestvici:

- 1 – **negative**: negativan sentiment
- 2 – **neutral**: neutralan sentiment
- 3 – **positive**: pozitivan sentiment
- 4 – **mixed**: dio rečenice je pozitivan, a dio negativan
- 5 – **sarcasam**: sarkastični i ironični komentari.

Uputa, kojom se vodite pri provjeri označavanja sentimenta rečenice, je sljedeća:

Molimo označite sentiment sljedeće rečenice iz perspektive prosječnoga hrvatskoga korisnika. Kako ste se osjećali nakon što ste pročitali ovu vijest / rečenicu?

1. Pozitivne rečenice označavaju se oznakom **positive**.

- govornik se koristi pozitivnim izrazima, primjerice, izražava potporu, divljenje, pozitivan stav, opraštanje, poticanje, uspjeh, pozitivne osjećaje
 - *To! Francuska je pobijedila Njemačku 3:1.* – označit ćemo kao pozitivno zbog pozitivnoga izraza „To!“
 - *Vodafone ima još mnogo prostora za rast.*

- Izvoz se povećao za 4,5 %.

2. Negativne rečenice označavaju se oznakom **negative**.

- govornik se koristi negativnim izrazima, primjerice, izražava kritiku, osudu, negativan stav, propituje valjanost/vještine, upućuje na greške i neuspjehe, izražava negativne osjećaje
 - *Rusija je igrala očajno protiv Finske.*
 - *Tanašićevo se ime pojavilo u javnosti nakon pljačke poštanskoga ureda u Kresnicama u lipnju prošle godine, u kojoj je Blaž Ereiz ubio poštansku službenicu Anicu Liković.*
 - *U usporedbi s dugoročnim prosjekom pokazatelj se smanjio za 11 postotnih poena, objavio je danas Državni zavod za statistiku.* (tu se treba uzeti u obzir kontekst kako bi se odredilo da je pad pokazatelja negativan)

3. Činjenice i objektivne izjave označavaju se oznakom **neutral**.

- govornik objektivno izvještava o činjenicama
 - *U najnovijem izvješću Dolamić je napisao da tvrtka trenutno nema problema s likvidnošću i redovito podmiruje svoje obveze.*
 - *Današnja jutarnja temperatura u Zagrebu iznosila je 5 Celzijevih stupnjeva.*

4. Rečenice s podijeljenim sentimentom označavaju se oznakom **mixed**.

- govornik se u dijelu rečenice koristi pozitivnim izrazima, a u dijelu rečenice negativnim izrazima
- U slučaju složenih rečenica u kojima je izraženo više od jednoga sentimenta, npr. kada imamo kombinaciju novinarove izjave i izjave političara, također ćemo upotrijebiti oznaku **mixed**.
 - *Dolar je ojačao vrijednost eura na \$1,357, a cijene lakih ulja pale su ispod \$74.*

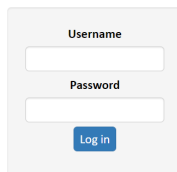
5. Rečenice u kojima je izražen sarkazam, ironija, ismijavanje ili sl. označavaju se oznakom **sarcasm**. Pri tom se obvezatno, u skladu s uputama, u obzir uzima kontekst. Rečenice su izvan konteksta pozitivne, ali se iz konteksta vidi da je mišljeno suprotno od rečenoga, tj. obično se koriste pozitivne riječi da bi se izrazile negativne konotacije. Zbog toga pri označavanju sarkazma moramo uzeti u obzir kontekst cijeloga dokumenta. Nama kao ljudskim označivačima to je jednostavno, dok je računalima to izuzetno složeno.

- *PLENKOVIĆ EKSPRESNO I SARKASTIČNO UZVRATIO MILANOVIĆU: 'Vidi se iz aviona da se stabilizirao'*
- *Na čelu Dinama je jako ugladen čovjek. (misli se na Z. Mamića)*
- *Predsjednik radi tako dobar posao.* (u negativnome kontekstu)

Provjera označavanja obavlja se u programu WebAnno. Programu pristupate na sljedećoj poveznici:

<http://faust.ffzg.hr/inception/login.html>

INCEpTION



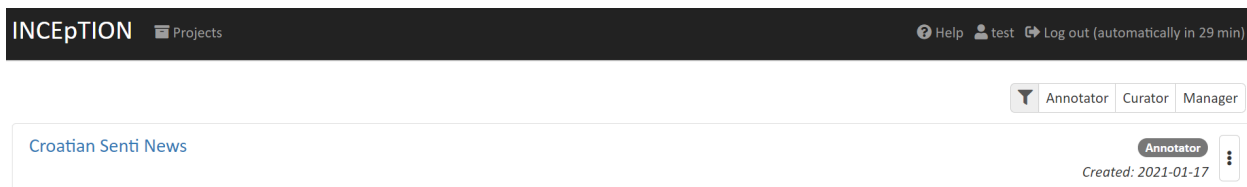
A login form with two input fields labeled 'Username' and 'Password', and a 'Log in' button below them.

Prijavni podatci:

Username: adresa vaše e-pošte

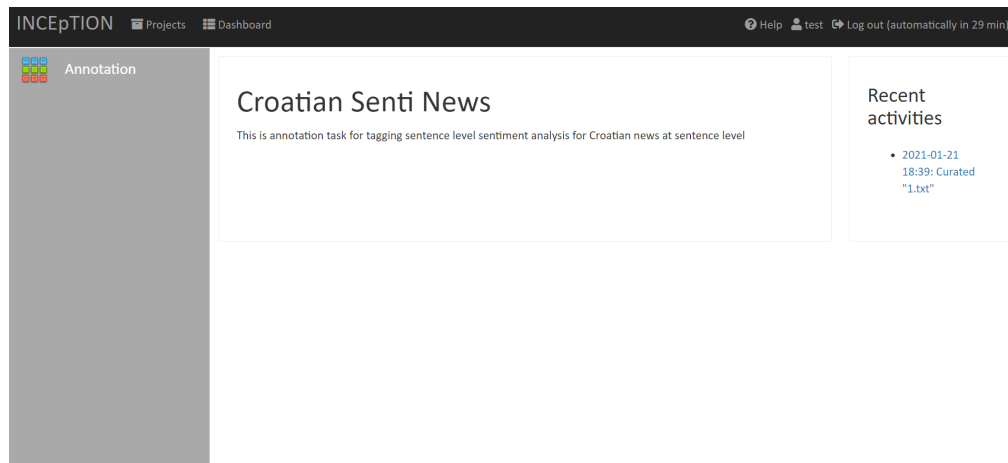
Password: adresa vaše e-pošte

Nakon što se prijavite, prikazat će Vam se projekt: Croatian SentiNews – X, gdje će X označavati broj vašeg paketa dokumenata.

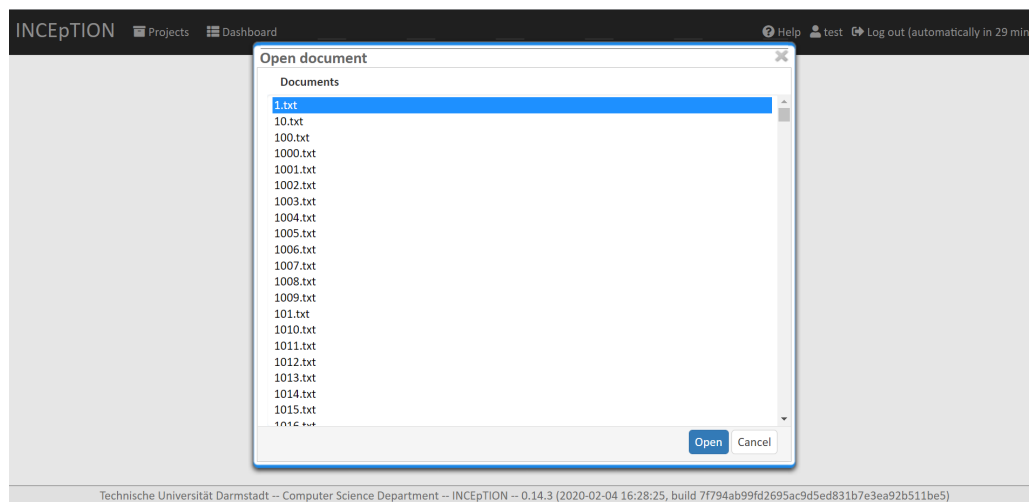


The INCEpTION interface shows the project 'Croatian Senti News' with a filter icon, roles 'Annotator', 'Curator', and 'Manager', and a status 'Annotator' with a dropdown menu. The creation date is 'Created: 2021-01-17'. The top bar includes 'INCEpTION', 'Projects', 'Help', 'test', and 'Log out (automatically in 29 min)'.

Kliknite na naslov projekta i otvorit će se sljedeći prozor:



Kliknite na Annotation da bi vam se otvorio popis dostupnih dokumenata koje trebate provjeriti i ispraviti.



Odaberite dokument i kliknite na Open. Otvorit će se sučelje za provjeru označavanja. Sve su rečenice prethodno automatski označene, a vi provjeravate i ako je potrebno, mijenjate oznake.

INCEpTION Projects Dashboard Help test Log out (automatically in 29 min)

test: Croatian Senti News/1.txt Showing 1-5 of 8 sentences [document 1 of 1992]

Layer Chunk

Annotation

No annotation selected

1 Pjevačica je pobjedničku strategiju sa svojim partnerom dogovarala tajnim znakovima. positive

2 Odbojkaška je utakmica privukla brojne obožavatelje kojima je ovo bila prilika da Rihannu vide negdje drugdje osim. positive

3 Ljepotica s Barbadosa prije tri tjedna krenula je na turneju 'Last Girl On Earth'. neutral

4 RIHANNA VRAĆA VUNU U MODU Da pletivo može biti i više nego seksi Rihanna je pokazala na nastupu u Ateni. neutral

5 Nosila je sivi vuneni triko sa izvezenim detaljima te bež kožne čizme iznad koljena. neutral

Technische Universität Darmstadt -- Computer Science Department -- INCEpTION -- 0.14.3 (2020-02-04 16:28:25, build 7f794ab99fd2695ac9d5ed831b7e3ea92b511be5)

Najvažnije opcije u sučelju:

1. Idi na prethodni dokument i idi na sljedeći dokument:



2. Zaključaj dokument, tj. predaj dokument kad je svaka rečenica provjerena i potvrđena.



PAZITE! Kada pritisnete taj gumb, više nećete moći mijenjati oznake u tom dokumentu. Stoga kliknite na taj gumb tek kad ste sigurni da ste provjerili sve oznake.

3. Briši sve oznake. Ta opcija nije preporučena. Pritiskom na taj gumb izbrisat će se sve oznake na stranici i resetirati dokument.



4. Otvaranje dijaloškog okvira za odabir pojedinoga dokumenta.



Proces provjere označavanja:

Pročitajte rečenicu i odlučite je li sentiment rečenice označen pravilno ili ga treba izmijeniti. Ako želite promijeniti oznaku, kliknite na nju i otvorit će se okvir s njezinom vrijednošću na desnoj strani prozora.

test: Croatian Senti News/1.txt Showing 1-5 of 8 sentences [document 1 of 1992]

1 Pjevačica je pobjedničku strategiju sa svojim partnerom dogovarala tajnim znakovima. (positive)

2 Odbojkaška je utakmica privukla brojne obožavatelje kojima je ovo bila prilika da Rihannu vide negdje drugdje osim. (positive)

3 Ljepotica s Barbadosa prije tri tjedna krenula je na turneju 'Last Girl On Earth'. (neutral)

4 RIHANNA VRAĆA VUNU U MODU Da pletivo može biti i više nego seksi Rihanna je pokazala na nastupu u Ateni. (neutral)

5 Nosila je sivi vuneni triko sa izvezenim detaljima te bež kožne čizme iznad koljena. (neutral)

Annotation: Layer: Sentiment, Text: RIHANNA VRAĆA VUNU U MODU Da pletivo može biti i, value: neutral

Technische Universität Darmstadt -- Computer Science Department -- INCEpTION -- 0.14.3 (2020-02-04 16:28:25, build 7f794ab99fd2695ac9d5ed831b7e3ea92b511be5)

Otvorite padajući izbornik i odaberite jedan od pet ponuđenih sentimenta. Promjena vrijednosti odražava se i u glavnome okviru:

test: Croatian Senti News/1.txt Showing 1-5 of 8 sentences [document 1 of 1992]

1 Pjevačica je pobjedničku strategiju sa svojim partnerom dogovarala tajnim znakovima. (positive)

2 Odbojkaška je utakmica privukla brojne obožavatelje kojima je ovo bila prilika da Rihannu vide negdje drugdje osim. (positive)

3 Ljepotica s Barbadosa prije tri tjedna krenula je na turneju 'Last Girl On Earth'. (neutral)

4 RIHANNA VRAĆA VUNU U MODU Da pletivo može biti i više nego seksi Rihanna je pokazala na nastupu u Ateni. (positive)

5 Nosila je sivi vuneni triko sa izvezenim detaljima te bež kožne čizme iznad koljena. (neutral)

Annotation: Layer: Sentiment, Text: RIHANNA VRAĆA VUNU U MODU Da pletivo može biti i, value: positive

Kada ste provjerili sve rečenice i zadovoljni ste svim oznakama, stisnite gumb s lokotom kako biste spremili svoje oznake.



Prijeđite na sljedeći dokument.



Želimo vam uspješno rješavanje zadatka!