ULOGA MORALNIH I EMOCIONALNIH KOMPONENTI JEZIKA U KLASIFIKACIJI KONVERZACIONIH TEKSTOVA

Milena Šošić

O AUTORU

Milena Šošić je student završne godine doktorskih studija na Matematičkom fakultetu Univerziteta u Beogradu.

Magistrirala je 2010. godine na Matematičkom fakultetu Univerziteta u Beogradu na temi "Primena klasifikacije na N-gramsku analizu genoma" na smeru Računarstvo i informatika u oblasti bioinformatika pod mentorstvom prof. dr Nenada Mitića.

Diplomirala je na Matematičkom fakultetu 2004. godine na smeru Računarstvo i informatika.

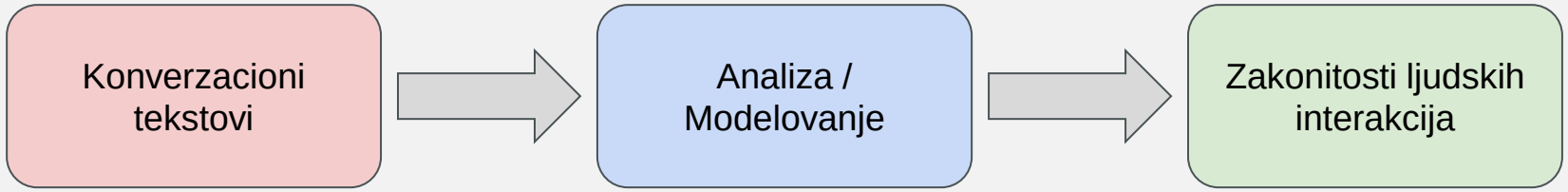
Od završetka studija radi u industriji u IT sektoru, poslednjih godina na zadacima primene metoda mašinskog učenja na rešavanje poslovnih zadataka sa posebnom pažnjom usmerenom ka zadacima koji uključuju računarsku obradu tekstova.

Sadržaj prezentacije je predstavljanje teme doktorskog rada: *"Modelovanje moralnih i emocionalnih komponenti jezika u klasifikaciji konverzacionih tekstova"* koja je prihvaćena za istraživanje.

SADRŽAJ PREZENTACIJE

MOTIVACIJA ISTRAŽIVANJA

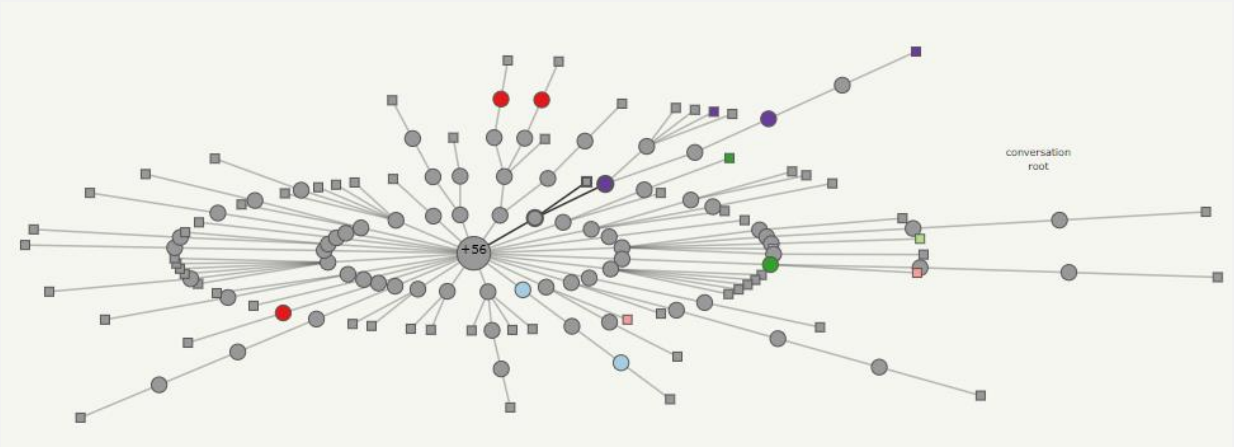
- Razvoj i dostupnost Interneta i alata za komuniciranje uticao je na pojavu velikih količina konverzacionih tekstova
- Konverzacioni tekstovi se mogu koristiti za analizu interakcija među ljudima, načina delovanja, iskazivanja sentimenta, emocija ili moralnih stavova
- Rezultati ovakvih istraživanja mogu ukazati na brojne zakonitosti i pravilnosti koje postoje u iskazima kod pojedinaca ili u međusobnim odnosima u okviru društvenih grupa

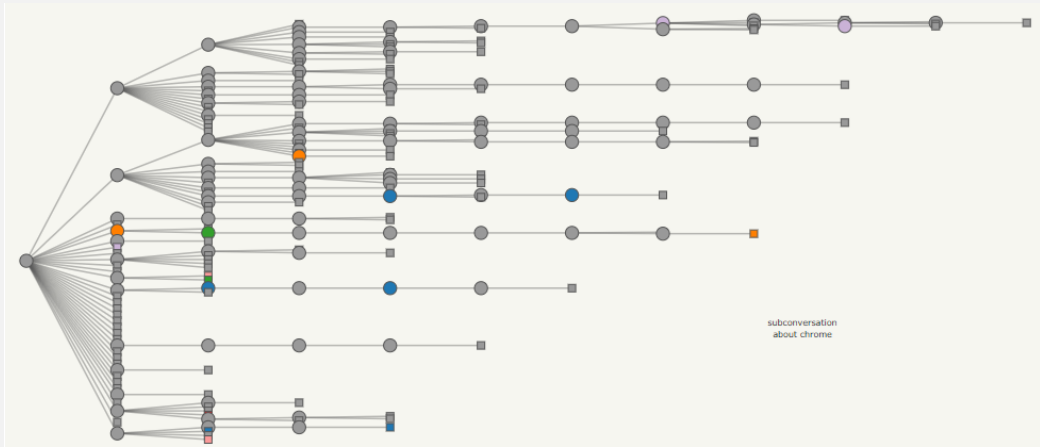


KONVERZACIONI TEKSTOVI

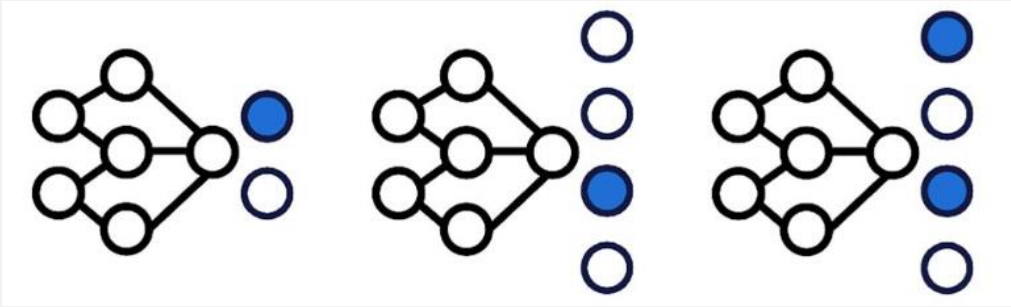
- Konverzacioni tekstovi – poruke elektronske pošte, poruke na društvenim mrežama, brze poruke, poruke u alatima za automatsko generisanje odgovora
- Analiza konverzacionih tekstova – pojedinačna poruka, parovi neposrednih poruka, luk konverzacionog niza, konverzacioni niz
- Struktura konverzacionih poruka – sadržaj i potpis poruke

- Vizuelizacija





KLASIFIKACIJA KONVERZACIONIH TEKSTOVA



- Vrste klasifikacije – binarna, višeklasna, višeznačna
- Mere – tačnost, preciznost, odziv, F-mera
- Zadaci
  - Elektronska pošta – lažne, lovac-poruke, poslovna/lična, organizaciona kategorizacija
  - Društvene mreže – sentiment, emocija, moralnost, autor, način delovanja
  - Alati za automatsko generisanje odgovora – namera, autor, imenovani entiteti (NER)
- Klasifikacija individualnih poruka ili konverzacionog niza/grane
- Klasifikacija poruka elektronske pošte na poslovni i lični kontekst:

Šošić M., Graovac J.: Effective Methods for Email Classification: Is it Business or Personal Email?

PREDMET I CILJ ISTRAŽIVANJA

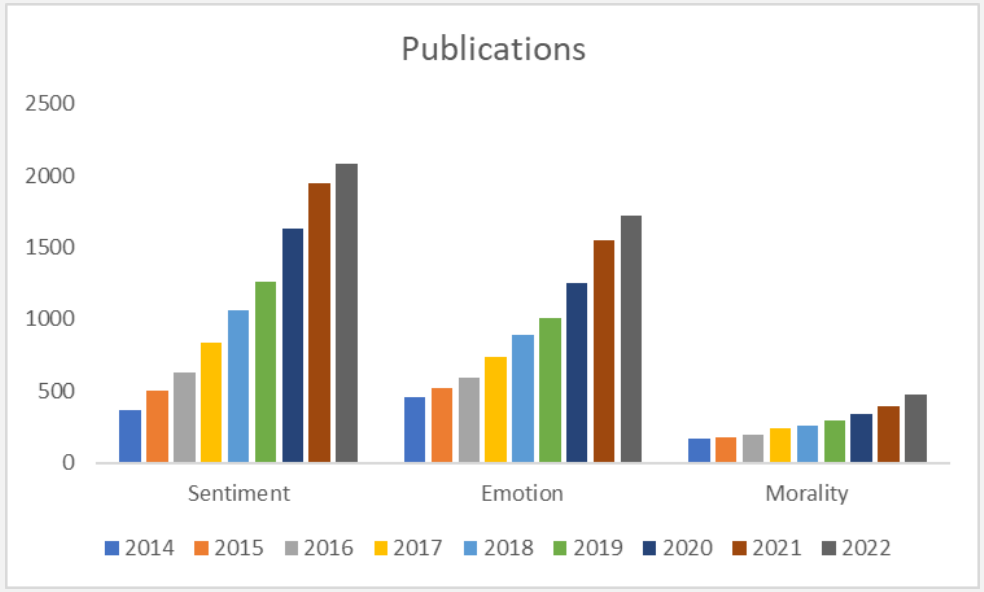
- Analiza uticaja moralnih i emocionalnih komponenti jezika u zadacima klasifikacije konverzacionih tekstova
- Tekstovi na engleskom i srpskom jeziku
- Matematičke metode - metode mašinskog učenja
  - Tradicionalni algoritmi
  - Algoritmi dubokog učenja
  - Algoritmi za utvrđivanje značaja atributa i međusobnih korelacija
- Razvoj uopštenog hibridnog pristupa za klasifikaciju konverzacionih tekstova
- Izgradnja novih resursa za srpski jezik - rečnici i korpusi konverzacionih tekstova sa oznakama za emocionalne i moralne kategorije

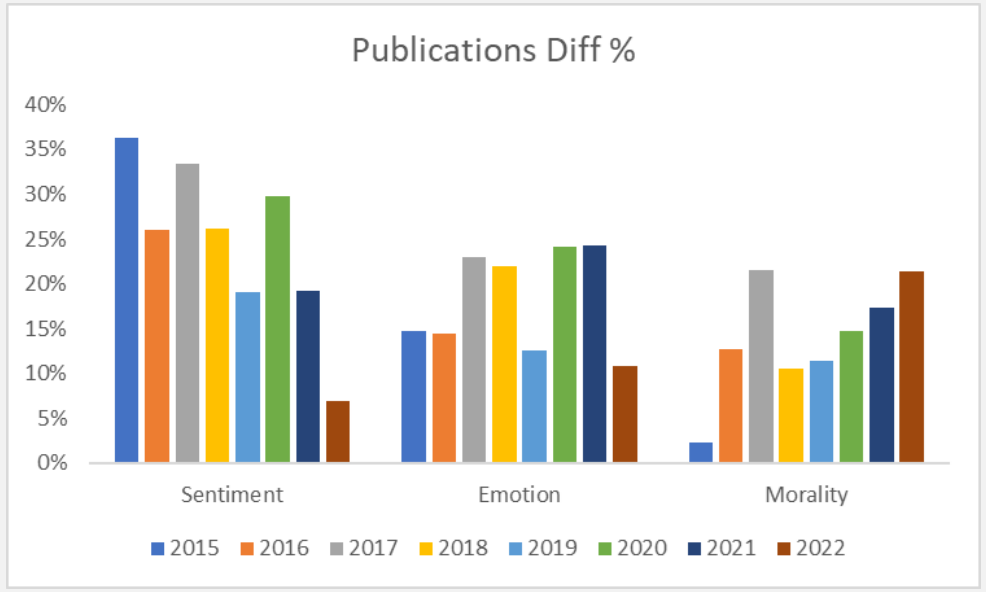
HIPOTEZE ISTRAŽIVANJA

1. Uključivanje moralnih i emocionalnih komponenti jezika u klasifikaciji konverzacionih tekstova može unaprediti tačnost modela, ukazujući da ovi atributi teksta imaju značaja u interpretaciji i razumevanju konverzacionih tekstova/poruka
2. Međusobna interakcija moralnih i emocionalnih komponenti jezika može doprineti razvoju preciznijih modela u zadacima klasifikacije konverzacionih tekstova
3. Upotreba određenih moralnih i emocionalnih komponenti jezika je specifična za pojedine klase u zadacima klasifikacije teksta
4. Razvijeni rečnici moralnih i emocionalnih afekata reči, kako za engleski, tako i za srpski jezik, mogu da doprinesu prepoznavanju ovih aspekata jezika
5. Na osnovu upotrebe reči iz svake od moralnih i emocionalnih kategorija definisanih ovim rečnicima može da se izračuna karakteristični moralni i emocionalni numerički pokazatelj tekstualne sekvence
6. Numerički atributi teksta dobijeni iz rečnika mogu se koristiti kao dodatni atributi u klasifikacionim modelima ili kao osnova za poređenje drugih (i složenijih) modela mašinskog učenja
7. Kvantitativnom analizom numeričkih atributa može se odrediti pojedinačna i međusobna korelacija između moralnog rasuđivanja i emocionalnih reakcija u jednoj ili nizu neposrednih konverzacionih poruka
8. Moguće je napraviti modele prihvatljive tačnosti koji bi na osnovu karakteristika tekstualnih sadržaja procenjivali tip moralne i emocionalne kategorije (ili više njih), kao i kategorije reakcija na poruku, za tekstove napisane na srpskom jeziku, a prema unapred definisanim kategorijama ovakvih klasifikacija

AKTUELNOST ISTRAŽIVANJA

- Izvor https://app.dimensions.ai
- Vreme pretrage: maj, 2023.
- Kriterijum:
  - polja istraživanja:
    - Information and Computing Sciences (46)
    - Language, Communication and Culture (47)
  - ključne reči: sentiment, emotion, morality





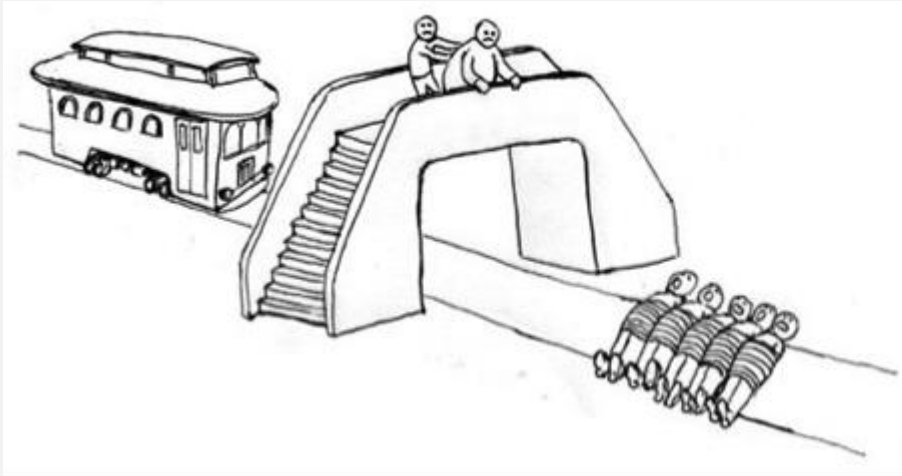
SRODNE PUBLIKACIJE

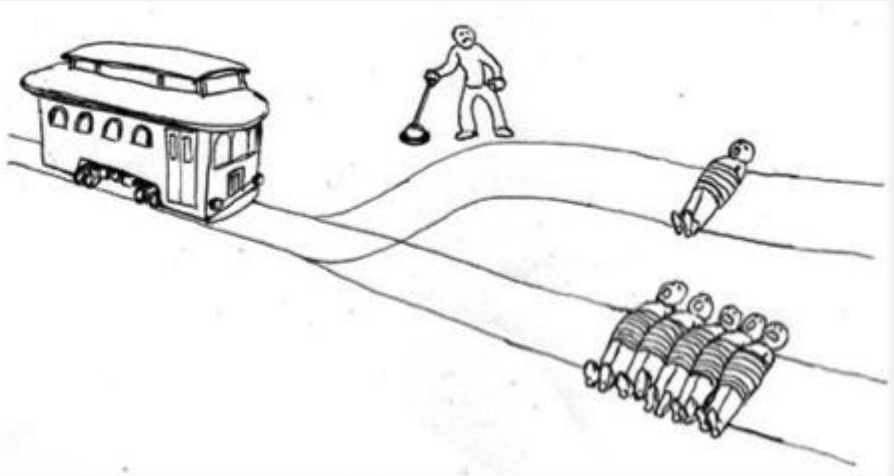
- Višeznačna anotacija tvitova u moralne kategorije: *Hoover J.: Moral Foundations Twitter Corpus: A Collection of 35k Tweets Annotated for Moral Sentiment, 2020*
- Kreiranje leksikona moralnosti: *Hopp, F.R.: The extended moral foundations dictionary (emfd): Development and applications of a crowd-sourced approach to extracting moral intuitions from text, 2021*
- Klasifikovanje tekstova prema moralnim vrednostima: *Bulla L.: Detection of Morality in Tweets Based on the Moral Foundation Theory, 2023*
- Kreiranje leksikona emocionalnog afekta (više radova istog autora): *Mohammad S.: Sentiment Analysis: Automatically Detecting Valence, Emotions, and Other Affectual States from Text, 2021*
- GoEmotions korpus - 27 emocionalnih kategorija: *Demszky D. (Google Research): GoEmotions: A Dataset of Fine-Grained Emotions, 2020*

MORALNOST

- Moral predstavlja jedan od osnovnih koncepata u ljudskom društvu koji obuhvata principe i vrednosti koji oblikuju etičko ponašanje i donošenje odluka svakog pojedinca

- Problem trolejbusa i njegove varijacije - alat za prikazivanje moralnih dilema





TEORIJA O MORALNIM OSNOVAMA

- Socio-psihološka teorija o moralnim vrednostima – Moral Foundations Theory (MFT)
  - *Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S., & Ditto, P. H.: Moral foundations theory: The pragmatic validity of moral pluralism, 2012*
- Upitnik o moralnim vrednostima - Moral Foundations Questionarie (MFQ)
- 5 univerzalnih moralnih vrednosti su predstavljene dihotomnim parovima:

  care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, sanctity/degradation [liberty/oppression]

  briga/povreda, pravednost/varanje, lojalnost/izdaja, autoritet/subverzija, svetost/degradacija [sloboda/ugnjetavanje]
- Prema MFT kombinacija različitih moralnih uverenja određuju stav pojedinca na osetljive teme

MORALNOST - LEKSIKONI

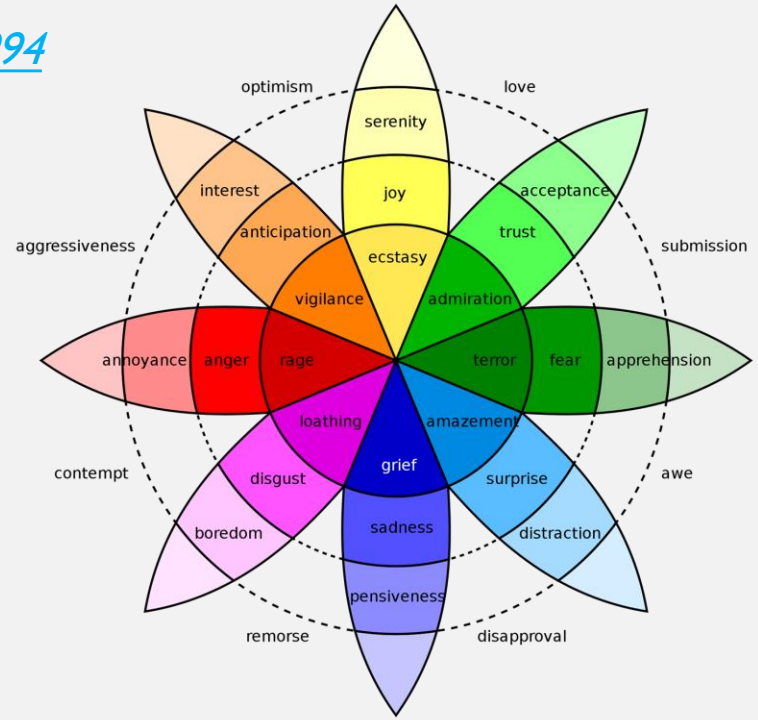
- MFD - ~300/2000 reči, binarni indikator, ekspertska klasifikacija izolovanih reči: *Graham, J., & Haidt, J.: The moral foundations dictionary, 2012; Graham, J., Haidt, J., & Nosek, B. A.: Liberals and conservatives rely on different sets of moral foundations, 2009*
- MoralStrength - ~1000 reči, indikator intenziteta: *Areque O.: MoralStrength: Exploiting a Moral Lexicon and Embedding Similarity for Moral Foundations Prediction, 2019*
- eMFD - ~3270 reči, indikator verovatnoće pripadnosti kategoriji, klasifikacija reči iz označenih korpusa korišćenjem tehnika obrade teksta, polaritet kategorije (vice/virtue): *Hopp, F.R.: The extended moral foundations dictionary (emfd): Development and applications of a crowd-sourced approach to extracting moral intuitions from text. 2021*

word	care	fairness	loyalty	authority	sanctity
close	0.073359	0.031373	0.115385	0.07722	0.084821
like(d)	0	0.047619	0.172414	0.086957	0.117647
open	0.099291	0.132075	0.118081	0.109635	0.090909
targeting	0.185185	0.261905	0.103448	0.106061	0.085106

EMOCIONALNOST

- Emocije su složena psihološka i fiziološka iskustva koja se pokreću različitim unutrašnjim i spoljašnjim stimulansima
- Teorija o osnovnim emocijama
  - Ekmanove kategorije emocija - 5 kategorija,
    Ekman, P: The nature of emotion: fundamental questions, 1994
  - Plučikove kategorije emocija - 8 kategorija,
    Plutchik, R.: The nature of emotions, 2001

    - strah (eng. fear)
    - ljutnja (eng. anger)
    - radost (eng. joy)
    - poverenje (eng. trust)
    - iznenađenje (eng. surprise)
    - iščekivanje (eng. anticipation)
    - odvratnost (eng. disgust)
    - tuga (eng. sadness)

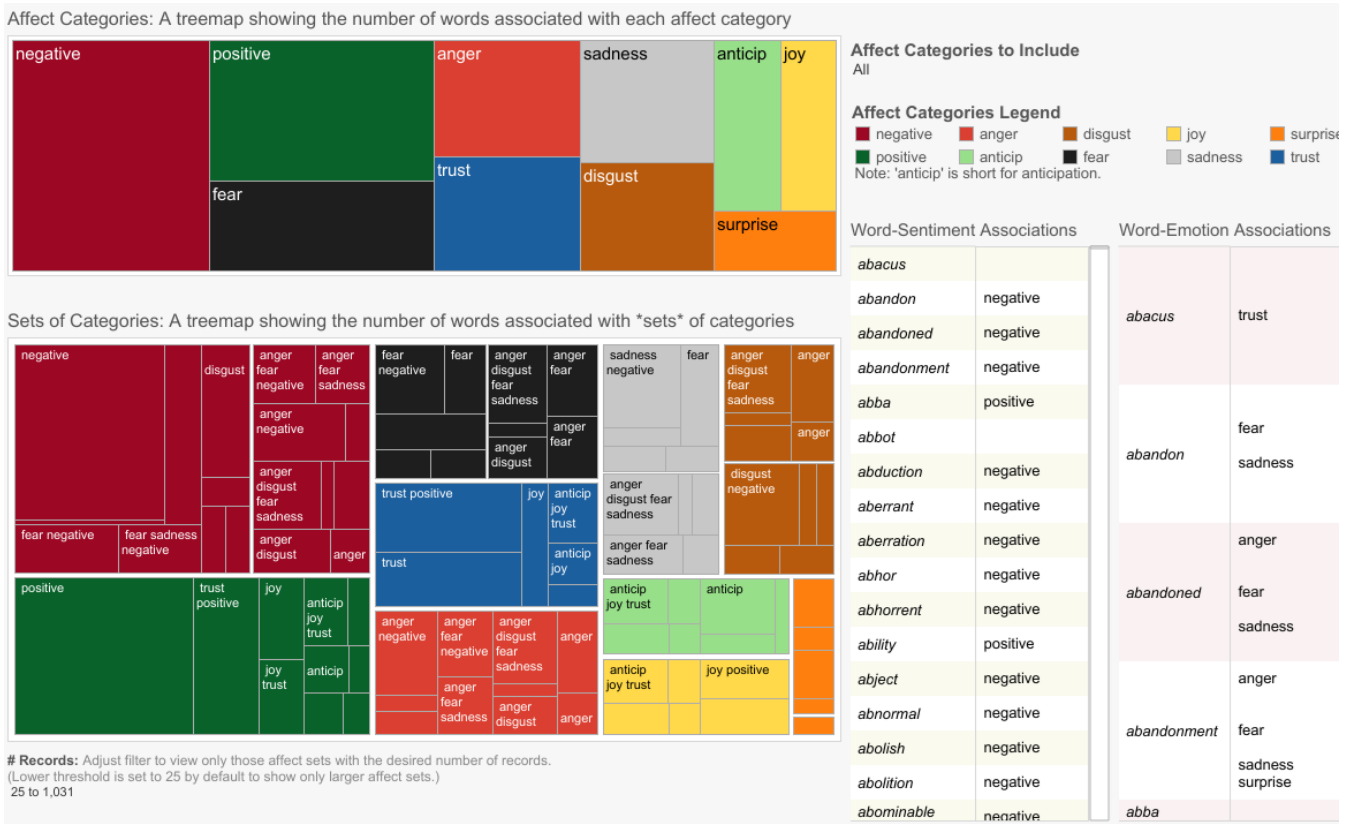


EMOCIONALNOST - LEKSIKONI

- WordNet-Affect - *Strapparava C.: Wordnet affect: an affective extension of wordnet, 2004*
- Multilingual WordNet-Affect - *Bobicev V.: Emotions in words: Developing a multilingual wordnet-affect, 2010*
- ANEW - ~1300 reči, Valence/Arousal/Dominance, *Bradley MM., Lang PJ.: Affective norms for English words (ANEW): Instruction manual and affective ratings, 1999*
- XANEW - ~14000 reči, *Warriner, A.B.: Norms of valence, arousal, and dominance for 13,915 English lemmas, 2013*
- LIWC - *Tauscyik Y.R.: The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods, 2009*
- ScenticNet - *Cambria E.: Senticnet: A publicly available semantic resource for opinion mining, 2010*
- NRC leksikoni
  - Word-Emotion Association Lexicon (EmoLex)
  - Valence, Arousal and Dominance Lexicon - Valence(positive--negative), Arousal(excited--calm), Dominance (powerful--weak)
  - Emotion Intensity Lexicon
  - Hashtag Emotion Lexicon
- Tendencija ‘prevođenja’ rečnika na različite jezike i provera njihove validnosti u drugim socio-lingvističkim okruženjima

EMOLEX LEKSIKON

- [Word-Emotion Association Lexicon](#) (EmoLex) – 14182 reči
- Prema Plučikovoj kategorizaciji 8 osnovnih emocija + 2 kategorije za sentiment
- Binarni indikator pripadnosti kategoriji
- Višestruko označavanje



LEKSIKON EMOCIONALNIH INTENZITETA

- Emotion Intensity Lexicon – 9829 reči
- Prema Plučikovoj kategorizaciji 8 osnovnih emocija
- Numerički/realni indikator pripadnosti kategoriji u rangu [0, +1]
- Višestruko označavanje
- Označavanje korišćenjem efektivne metode - skaliranje poređenjem sa najboljim/najgorim u grupi (best-worst scaling)

Word	Anger	Word	Fear	Word	Joy	Word	Sadness
outraged	0.964	*horror*	0.923	*sohappy*	0.868	*sad*	0.844
brutality	0.959	*horrified*	0.922	*superb*	0.864	*suffering*	0.844
satanic	0.828	*hellish*	0.828	*cheered*	0.773	*guilt*	0.750
hate	0.828	*grenade*	0.828	*positivity*	0.773	*incest*	0.750
violence	0.742	*strangle*	0.750	*merrychristmas*	0.712	*accursed*	0.697
molestation	0.742	*tragedies*	0.750	*bestfeeling*	0.712	*widow*	0.697
volatility	0.687	*anguish*	0.703	*complement*	0.647	*infertility*	0.641
eradication	0.685	*grisly*	0.703	*affection*	0.647	*drown*	0.641
cheat	0.630	*cutthroat*	0.664	*exalted*	0.591	*crumbling*	0.594
agitated	0.630	*pandemic*	0.664	*woot*	0.588	*deportation*	0.594
defiant	0.578	*smuggler*	0.625	*money*	0.531	*isolated*	0.547
coup	0.578	*pestilence*	0.625	*rainbow*	0.531	*unkind*	0.547
overbearing	0.547	*convict*	0.594	*health*	0.493	*chronic*	0.500
deceive	0.547	*rot*	0.594	*liberty*	0.486	*injurious*	0.500
unleash	0.515	*turbulence*	0.562	*present*	0.441	*memorials*	0.453
bile	0.515	*grave*	0.562	*tender*	0.441	*surrender*	0.453
suspicious	0.484	*failing*	0.531	*warms*	0.391	*beggar*	0.422
oust	0.484	*stressed*	0.531	*gesture*	0.387	*difficulties*	0.421
ultimatum	0.439	*disgusting*	0.484	*healing*	0.328	*perpetrator*	0.359
deleterious	0.438	*hallucination*	0.484	*tribulation*	0.328	*hindering*	0.359

item_1	item_2	item_3	item_4	best_item	worst_item

```
$scores{$item} = ($count_best{$item} - $count_worst{$item}) / $count_item{$item}
```

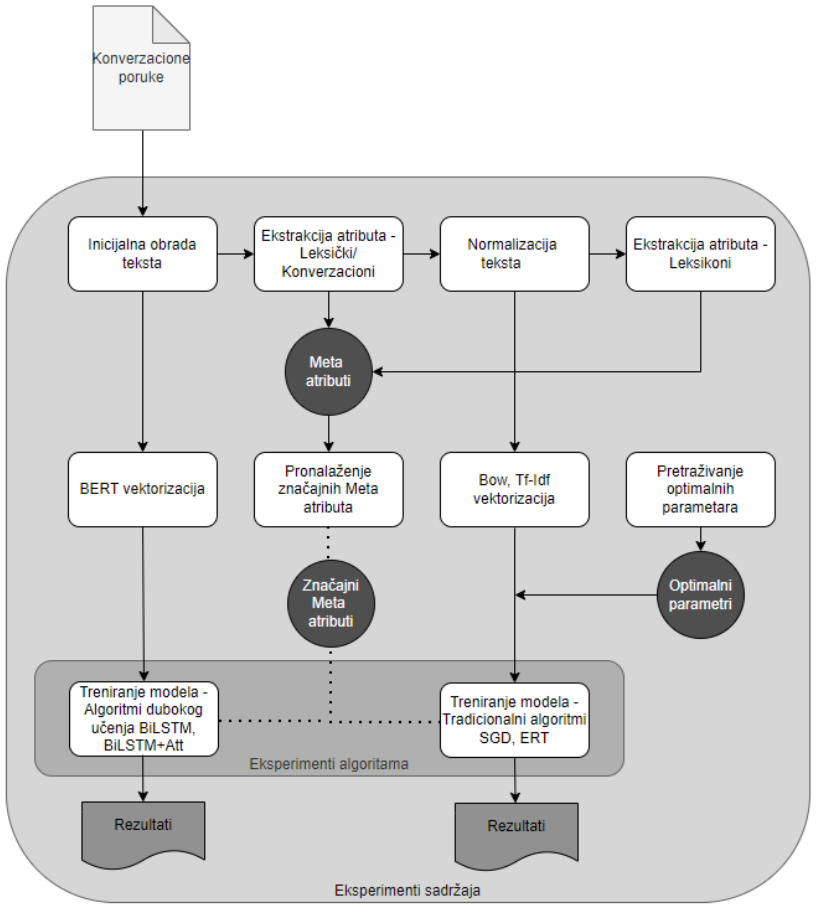
EMOCIONALNOST NASUPROT SENTIMENTA

- Sentiment – koji je moj stav prema određenoj pojavi – pozitivan/negativan/neutralan
- Emocije – koji su mogući emocionalni pokretači takvog stava – granularniji pristup

Kontekst	Sentiment	Emocije
Jako mi se dopao film	*+1 (+58)*	*dopadati - radost, poverenje*
Film je na mene ostavio loš utisak, iako sam dobio preporuke da ga pogledam	*-1(-0.45)*	*loš - ljutnja, odvratnost, strah; utisak - radost; pogledati - očekivanje*
I pored dobrih vizuelnih efekata, ovaj film ne zaslužuje visoku ocenu	*+1 (+0.17)*	*dobar - očekivanje, radost, iznenađenje, poverenje; zaslužiti - očekivanje, poverenje; ocena - ljutnja, strah, tuga*

PREDLOŽENA METODOLOGIJA

- Izdvajanje segmenata poruke
- Predstavljanje sadržaja u vektorskom obliku (BoW, Tf-Idf, BERT ugnježdeni vektori reči)
- Ekstrakcija dodatnih atributa – leksičko-sinktatički, konverzacioni, atributi izražajnosti, emocionalnosti i moralnosti ~60 atributa
- Kreiranje modela klasifikacije
- Evaluacija rezultata
- Evaluacija značajnosti atributa



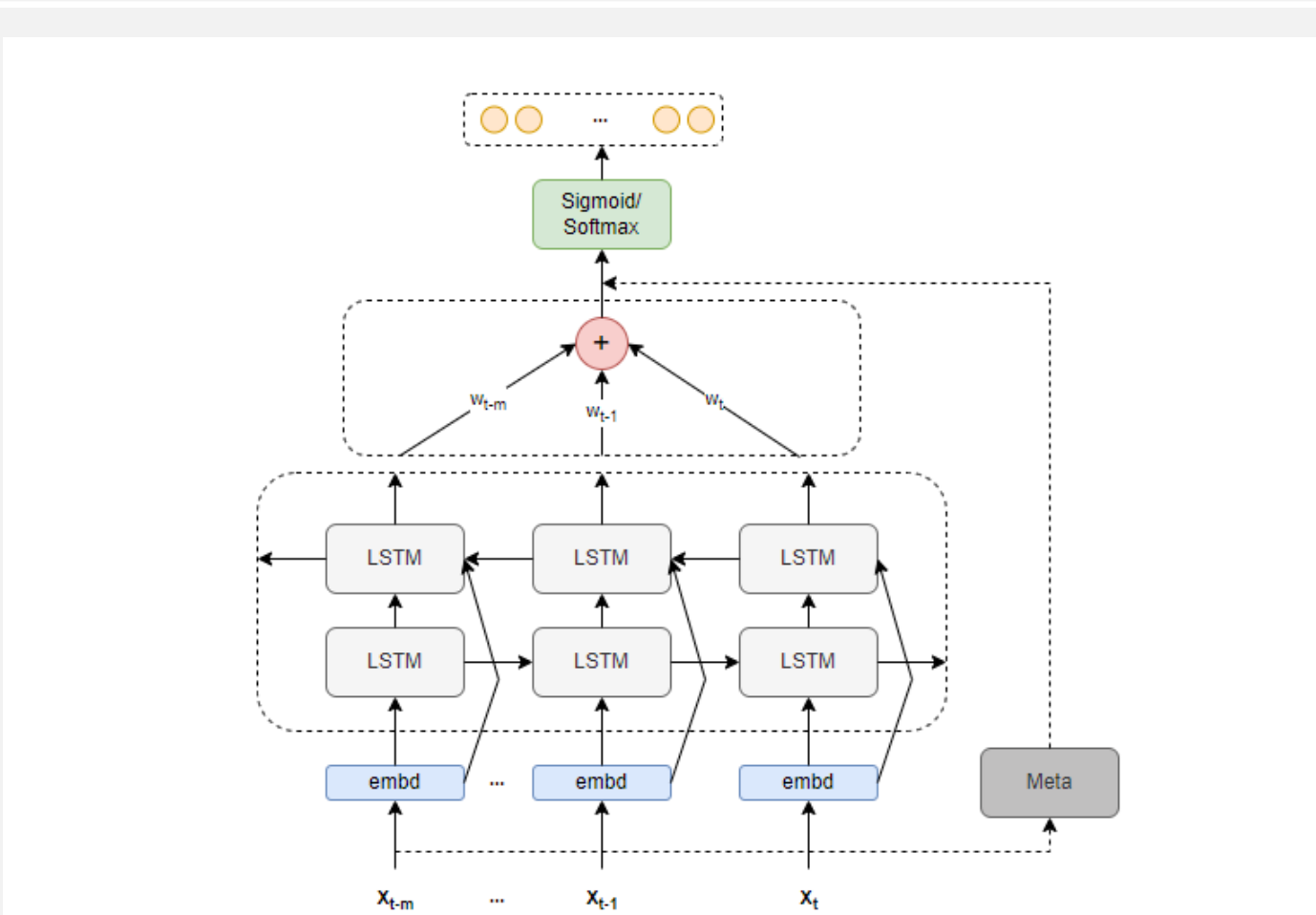
PREDLOŽENA METODOLOGIJA - DODATNI ATRIBUTI

- Leksičko-sintaktički
- Konverzacioni - zavisni od zadatka
- Izražajnosti (razumljivost i čitljivost, polaritet i subjektivitet teksta)
  - Težinske formule koje koriste broj slova, slogova, reči, rečenica
  - ARI (Automated Readability Index) [1->14]
  - FRES (Flech Reading Ease Score)[100->0]
  - LWM (Linsear Write Metric)[80->70]
- Moralnosti
- Emocionalnosti

Feature Group		Features List	# of Features
Lexical		Number of characters and words in content and subject, sentences count, average sentence length, average word length, noun phrases, average syllables per word, average syllables per sentence, sentence and word density, difficult words, business indicator, acronyms indicator	16
	NER-based	Ratio of personal name, organization name, number, connectors, month name, day name, email and url address tags	8
	Punctuation-based	Ratio of dots, question marks, exclamation marks, hash tags, reference tags	5
Conversational		Free domains in headers ratio, number of recipients, recipients domains coherency	3
Expressional		Automated Readability Index(ARI), Flech Reading Ease Score(FRES), Linsear Write Metric(LWM), content subjectivity, content polarity	5
Moral		Probability measures of care, sanctity, authority, loyalty and fairness on word and sentence, moral/non-moral ratio	11
Emotional		Measures of trust, joy, anger, disgust, sadness, fear, surprise, positive, negative	9
All features (Meta)			57

PREDLOŽENA METODOLOGIJA - ALGORITMI

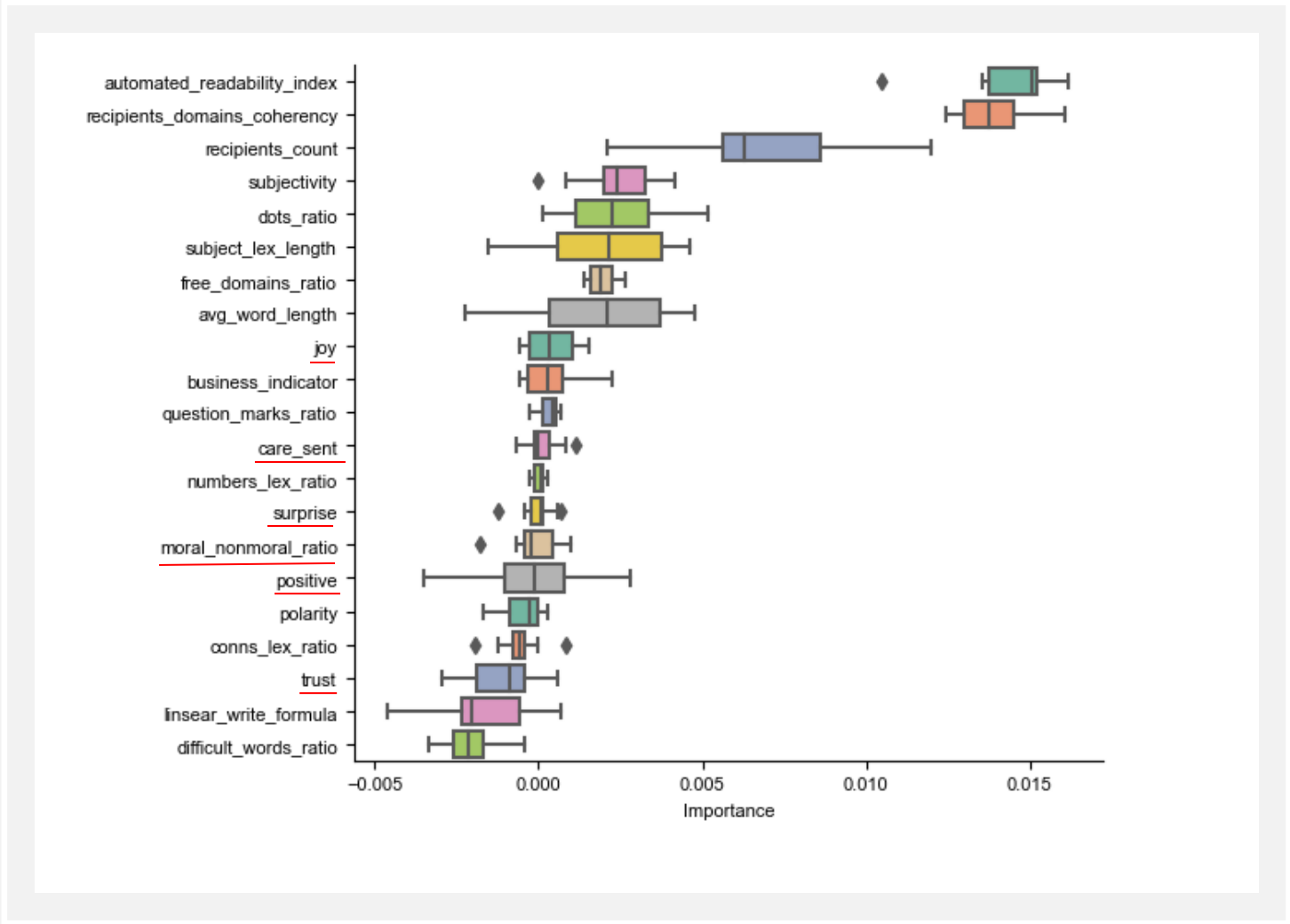
- Tradicionalni algoritmi:
  - SGD-SVM – stohastički gradijentni spust
  - Ansambli/Skupovi drveta odlučivanja (Extremely Randomized Trees)
- Algoritmi dubokog učenja
  - RNN, LSTM, Bi-LSTM, Bi-LSTM sa pažnjom



PREDLOŽENA METODOLOGIJA - ZNAČAJNI ATRIBUTI

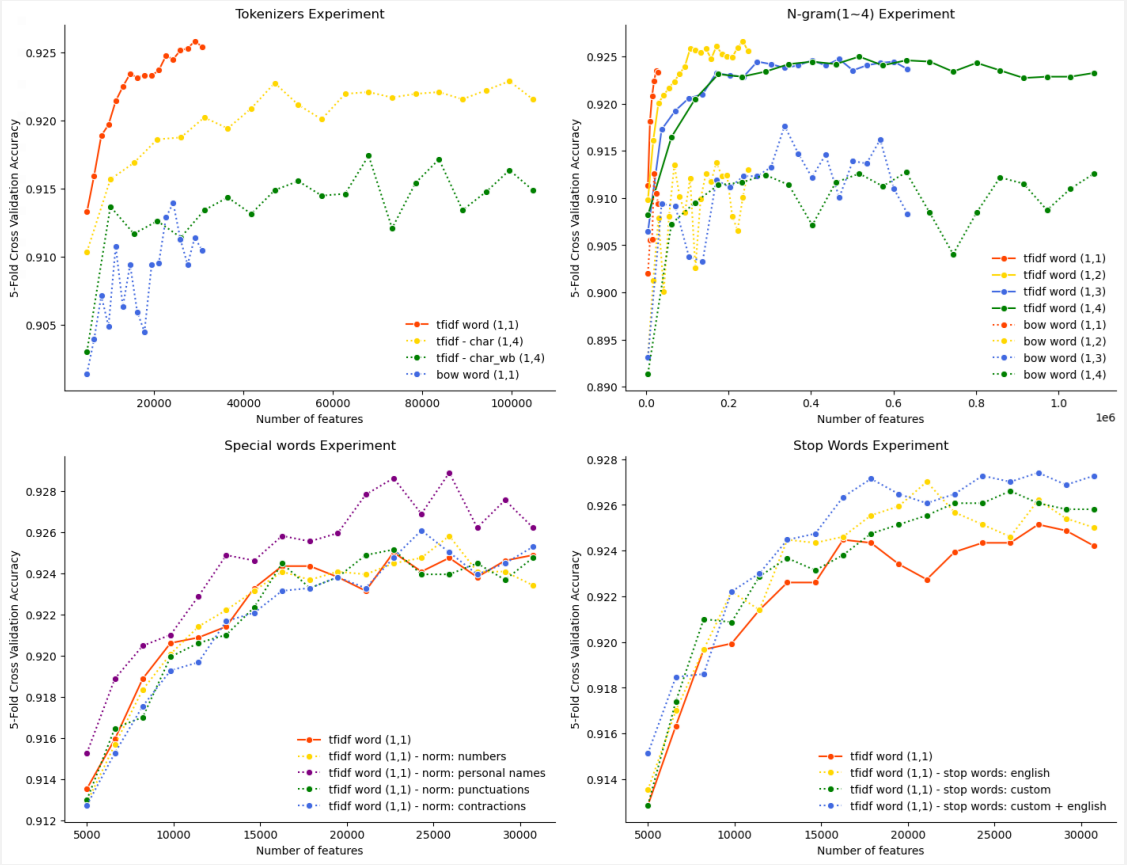
- Interpretacija i razumevanje modela
- Poređenje različitih metoda

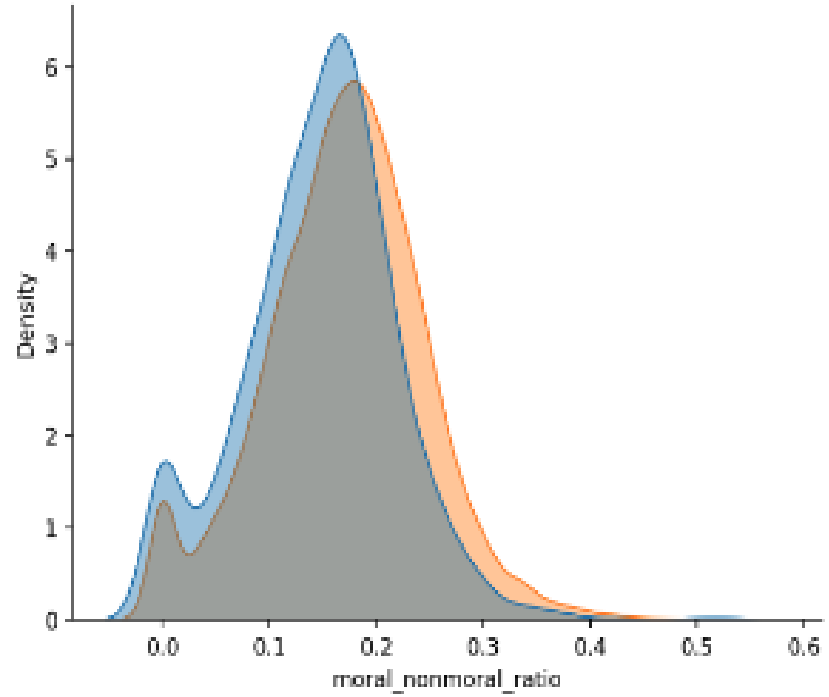
- Algoritmi prema broju posmatranih atributa:
  - Univarijabilni
  - Multivarijabilni
- Algoritmi prema zavisnosti od modela
  - Zavisni (u toku obuke)
  - Nezavisni (nakon obuke)
    - Izbacivanje jednog atributa
    - Permutacija vrednosti atributa
- Koji je značaj moralnih i emocionalnih atributa u procesu klasifikacije konverzacionih tekstova?
- Da li postoji korelacija između moralnih i emocionalnih atributa?

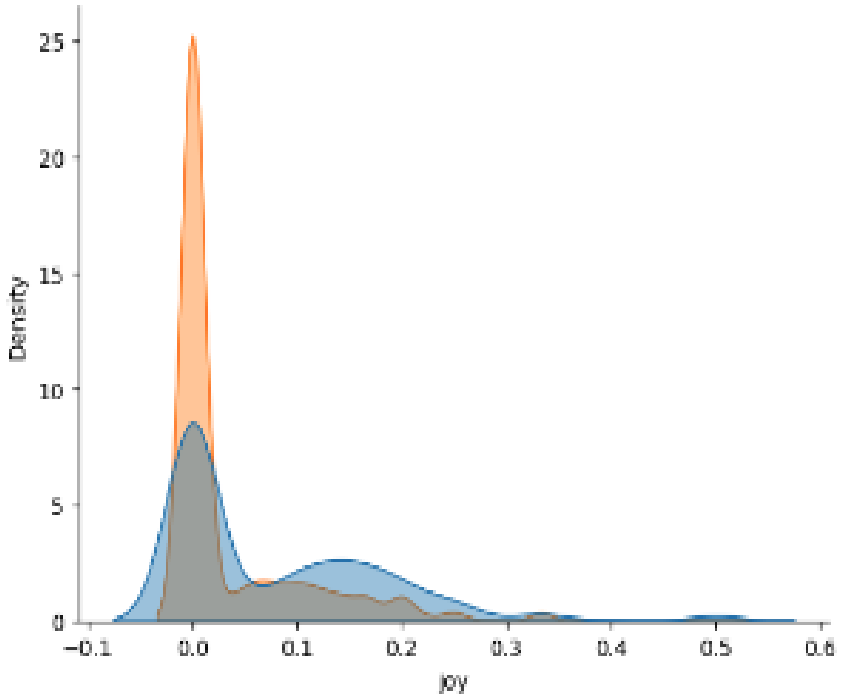


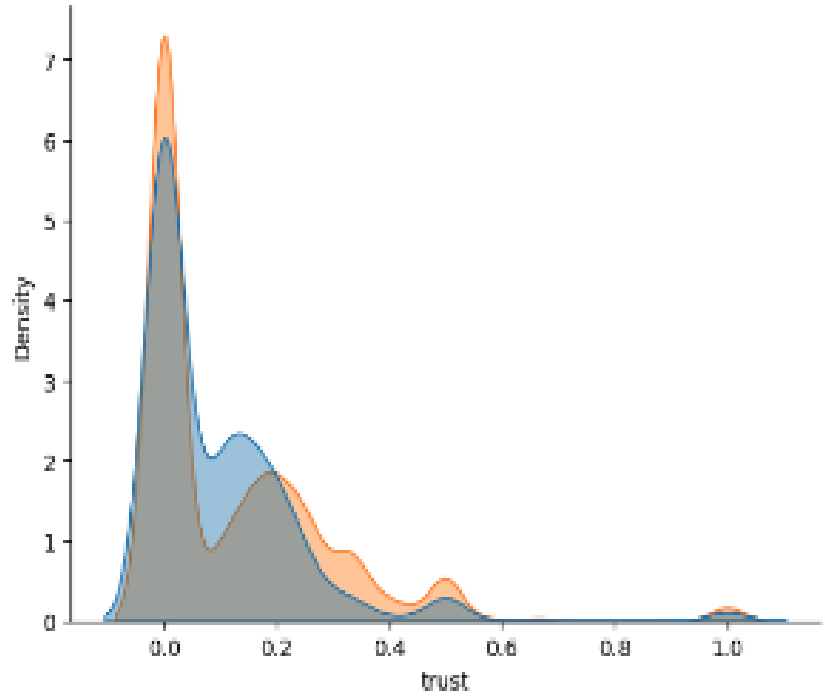
PREDLOŽENA METODOLOGIJA - TEHNIKE NORMALIZACIJE TEKSTA

- Eksperimenti sa procesiranjem teksta:
  - Tokenizator – reč, karakter, prošireni karakter
  - Dužina fraze – 1-4, 2-3
  - Specijalne reči – brojevi, interpunkcijski znakovi, lična imena, kontrakcije
  - Lista stop reči
  - Maksimalni i minimalni broj pojavljivanja reči u korpusu
  - Lematizacija
  - Dužina vektora







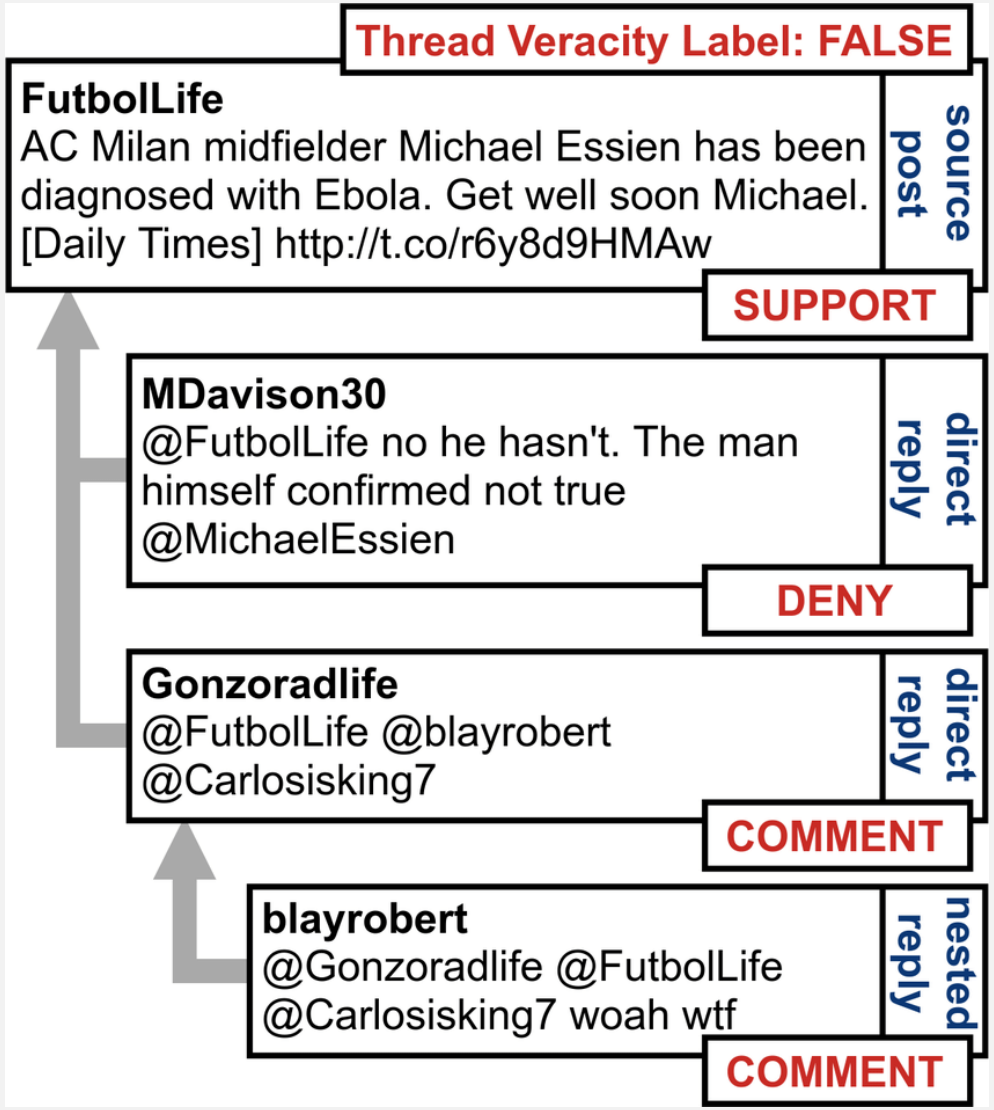


KLASIFIKACIJA PORUKA ELEKTRONSKE POŠTE NA POSLOVNI I LIČNI KONTEKST

- Zahtevan i nedovoljno istraživan zadatak - više klasa koje se semantički preklapaju; nebalansirani podaci
- Označeni podaci dobijeni u saradnji sa istraživačima sa Kolumbija Univerziteta
- Uočena važnost uticaja moralnih i emocionalnih komponenti jezika (engleski) na tačnost klasifikacije

KLASIFIKACIJA GLASINA I TIPA DELOVANJA NA GLASINU

- Važan zadatak pravovremenog identifikovanja glasina i njihovog razlikovanja od istinitih tvrdnji
- Tipovi delovanja na glasinu: podrška, odbijanje, pitanje, komentar
- Označeni podaci preuzeti iz SemVal-2019 internacionalnog takmičenja na zadacima (zadatak 7) iz semantičke analize tekstova
  - Kategorizacija glasina – tačno/netačno – mali broj označenih poruka
  - Kategorizacija tipa delovanja – nebalansirana višeklasna klasifikacija
- Uočena važnost uticaja leksičkih i emocionalnih atributa na tačnost klasifikacije na ovim zadacima
- Koji je uticaj moralnih atributa?

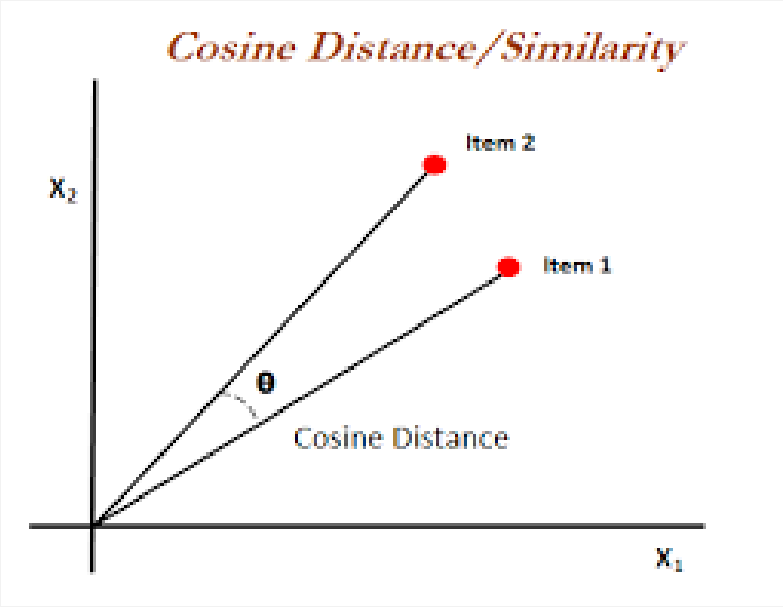


KLASIFIKACIJA GLASINA I TIPA DELOVANJA NA GLASINU

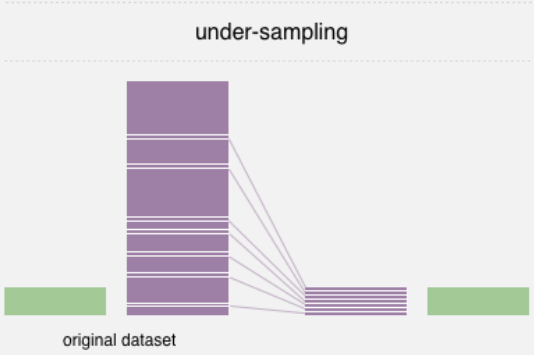
- Označene poruke
- Struktura konverzacionih nizova
- Zadaci u kojima moralni i emocionalni atributi mogu imati značajnu ulogu u klasifikaciji
- Testiranje u kojoj meri je predložena metoda primenljiva na druge skupove podataka
- Prijavljena rešenja/rezultati su dobijeni korišćenjem SoA tehnika

Izazovi u podacima:

- Metode za do-označavanje podataka - sličnost tekstova
- Metode za balansiranje podataka (under/over sampling)



$$\text{cosine similarity} = S_C(A,B) := \cos(\theta) = \frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}},$$



PRIMENA METODE NA SRPSKI JEZIK

- Konverzacioni tekstovi na srpskom jeziku – Tviter i Redit nalozi na srpskom jeziku – objave i komentari (konverzacioni nizovi)

Korpus	Naloga/ Grupa	Poruka	Nizova
Tviter	76	44835	7211
Redit	8	140608	9161

- Konstruisanje i validacija rečnika
  - Emocije: NRC.SR – inicijalno (eng) 14182 reči / 8 emocionalnih kategorija + 2 kategorije sentimenta / binarni ili indikator verovatnoće
    [anger, trust, fear, sadness, joy, anticipation, surprise, disgust | positive, negative]
    [ljutnja, poverenje, strah, tuga, radost, iščekivanje, iznenađenje, odvratnost | pozitivno, negativno]
  - Moralnost: eMFD.SR – inicijalno (eng) 3270 reči / 5 moralnih kategorija * 2 dihotoma / binarni ili indikator verovatnoće
    [care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, sanctity/degradation]
    [briga/povreda, pravednost/varanje, lojalnost/izdaja, autoritet/subverzija, svetost/degradacija]
- Označavanje tekstova – moralna i emocionalna kategorija (ili više njih)
- Pravljenje modela za predviđanje moralne i emocionalne kategorije poruke - višeklasni i višeznačni modeli

PRIMENA METODE NA SRPSKI JEZIK - IZAZOVI

- Tokenizator, PoS tager, lematizator – eksperimenti sa različitim razvijenim modelima (JeRteh, RELDI, nltk)
- Priprema teksta – restauracija dijakritika, stop reči
- Ekstrakcija atributa:
  - Polaritet i subjektivitet teksta – SRPOL alat + SentiWords.SR leksikon

    Šošić M.: SRPOL - A Lexicon Based Framework for Sentiment Strength of Serbian Texts
  - Imeničke fraze, 'teške' reči – podela reči na slogove
  - Mere čitljivosti teksta – predlog algoritma za procenu čitljivosti teksta na srpskom jeziku
  - Atributi moralnosti i emocionalnosti – izgrađeni leksikoni
- BERT ugnježdeni vektori – višejezični (multilingual-bert), jednojezični za srpski jezik (RELDI:bertic; JeRteh: bertovic i dr.)
- Dovoljne količine **označenih** podataka za izgradnju modela
- Prikupljanje dodatnih arhivnih podataka sa Tviter platforme

PRIMENA METODE NA SRPSKI JEZIK - IZAZOVI

- Označavanje i validacija leksikona - MFD.SR

word	word_sr	lemma_sr	PoS	care	fairness	loyalty	authority	sanctity
close	*blizu/zatvoriti*	*blizu/zatvoriti*	*ADV/VERB*	0.073359	0.031373	0.115385	0.07722	0.084821
like(d)	sviđati (se)	sviđati (se)	VERB	0	0.047619	0.172414	0.086957	0.117647
open	*otvoren/otvoriti*	*otvoren/otvoriti*	*ADJ/VERB*	0.099291	0.132075	0.118081	0.109635	0.090909
targeting	*ciljanje*	*ciljanje*	*NOUN*	0.185185	0.261905	0.103448	0.106061	0.085106

- Upitnik za moralne vrednosti na srpskom govornom području

- Označavanje i validacija leksikona - NRC.SR

word	word_sr	lemma_sr	PoS	anger	anticipation	disgust	fear	joy	sadness	surprise	trust
friendship	prijateljstvo	prijateljstvo	NOUN	0	0	0	0	1	0	0	1
abundance	*obilnost/izobilje/obilje/zastupljenost*	*obilnost/izobilje/obilje/zastupljenost*	NOUN	0	1	1	0	1	0	0	1
abundant	*obilan/izobilan/zastupljen*	*obilan/izobilan/zastupljen*	ADJ	0	0	0	0	1	0	0	0

PRIMENA METODE NA SRPSKI JEZIK - IZAZOVI

- Označavanje i validacija tekstualnih korpusa

Lančani sudar na Banovom brdu, povređene majka, dve devojčice i beba – harm (care)

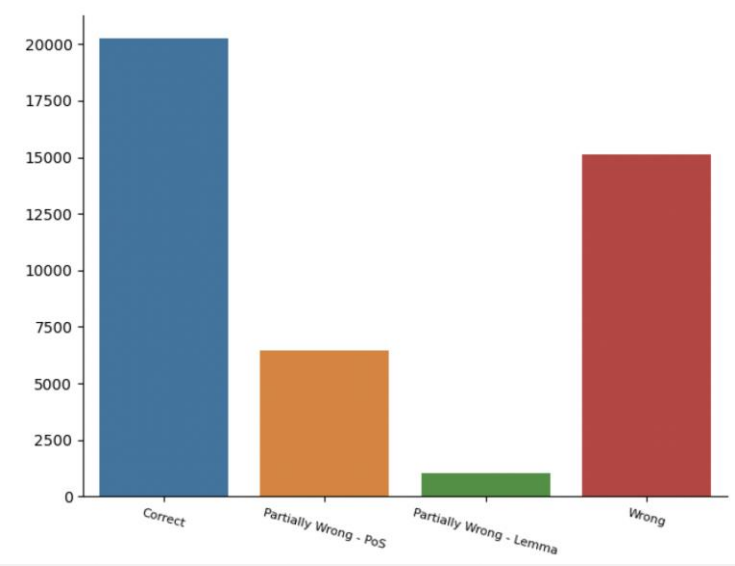
Poljski premijer optužio Rusiju i Belorusiju za hakerski napad i izmenu imejlova - harm,cheating,authority

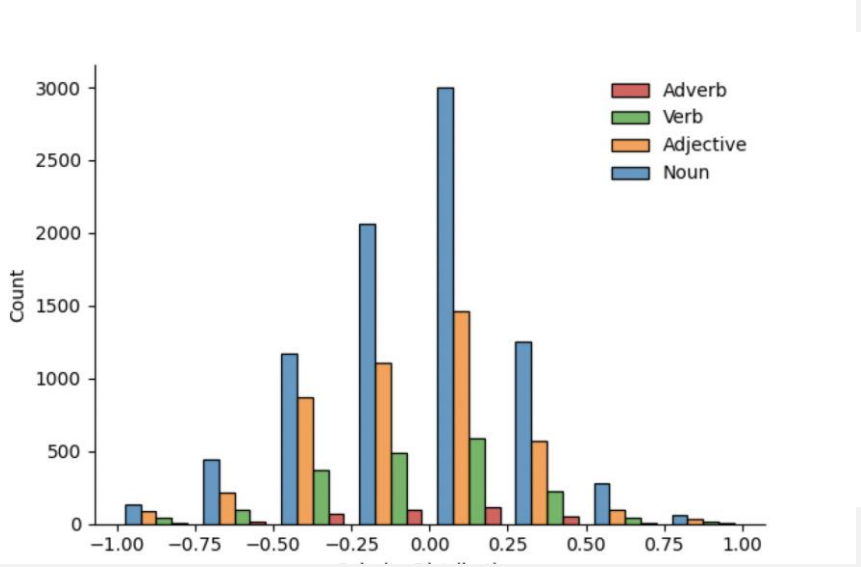
UNDP: Devet odsto ljudi u svetu siromašno, sa prihodima do 1,90 dolara dnevno – care, harm, fairness, degradation

- Pitanja:
  - Kako definisati anotacione šeme da bi se neusaglašenosti prilikom označavanja svele na minimum?
  - Primeniti jednoznačno ili višeznačno označavanje?
  - Algoritam za konsenzus oznaku između više anotatora
  - Koji broj anotatora je prihvatljiv? 2 + super-evaluacija razlika?
  - Istraživači ukazuju na varijacije u moralnim vrednostima među različitim socio-demografskim grupama

SRPOL - ALAT ZA IZRAČUNAVANJE INTENZITETA SENTIMENTA

- Kreiranje SentiWords.SR leksikona ~15000 lema/PoS parova sa pridruženim vrednostima polariteta u rangu [-1,+1]
  - Po uzoru na analogni leksikon na engleskom jeziku *Gatti L., Guerini M., and Turchi M.:"SentiWords: Deriving a high precision and high coverage lexicon for sentiment analysis, 2015*
  - Translacija: korektna~48.3%, parcijalno korektna~18%, nekorektna~33.7%





Algorithm 1: Creation of SentiWords.SR from the English SentiWords lexicon

FindPolarity
inputs : $SentiWords; GoogleTranslate$
output: $SentiWords.SR$
foreach $lemma, PoS \in SentiWords$ **do**
  $score \leftarrow score(lemma, PoS)$;
  $lemmas_{Sr} \leftarrow clean(GoogleTranslate(lemma, PoS))$;
  $lemmas_{Sr}, PoS \leftarrow score(lemma, PoS)$;
  $SentiWords.SR \leftarrow evaluate(lemmas_{Sr}, PoS, score)$;
foreach $lemma_{Sr}, PoS \in SentiWords.SR$ **do**
  $score \leftarrow mean(lemma_{Sr}, PoS, score)$;
  $std \leftarrow std(lemma_{Sr}, PoS, score)$;
  $count \leftarrow count(lemma_{Sr}, PoS, lemma)$;
return $SentiWords.SR$;

SRPOL - ALGORITAM

- SRPOL algoritam pored polariteta reči uključuje kontekstualne informacije:
  - Prilozi kao modifikatori intenziteta
  - Negacije
  - Uzvičnici
  - Produžene reči
  - Emotikoni i emoji

$$\text{"Veoma } (\rightarrow MOD=1.2) \text{ dobar } (p=+0.43) \text{ film..."} \xrightarrow{1.2\times(+0.43)} +0.52$$

$$\text{"Very } (\rightarrow MOD=1.2) \text{ good } (p=+0.43) \text{ movie..."} \xrightarrow{1.2\times(+0.43)} +0.52$$

  - Segmentacija teksta u rečenice ili delove rečenica
  - Intenzitet polariteta tekstualnog sadržaja je težinska sredina intenziteta polariteta njegovih segmenata

$$\text{"Film nije } (\rightarrow NEG) \text{ zanimljiv } (p=+0.53)\text{"} \xrightarrow[\times(-1)]{+0.53} -0.53$$

$$\text{"The movie is not } (\rightarrow NEG) \text{ interesting } (p=+0.53)\text{"} \xrightarrow[\times(-1)]{+0.53} -0.53$$

ZAKLJUČAK I DALJI RAD

- Predstavljen je rad na zadatku klasifikacije konverzacionih tekstova u kome su korišćeni atributi koji mere emocionalni i moralni afekt poruka
- Na ovom zadataku se pokazalo da ovi atributi imaju značajan uticaj na tačnost klasifikacije

Dalji rad:

- Proveriti metodu na drugim zadacima/korpusima
- Izvršiti proveru značajnosti/korelacija atributa korišćenjem drugih matematičkih metoda

- Srpski jezik - cilj je razvoj resursa koji bi omogućili da se ovi aspekti jezika bolje razumeju i koriste za dalja istrazivanja

Dalji rad:

- Označavanje i validacija rečnika i korpusa
- Definisanje zadatka za rešavanje - predviđanje moralne kategorije/emoconalne kategorije poruke ili odgovora na poruku, neke druge kategorije korišćenjem moralnih i emocionalnih komponenti sadržaja
- Rešavanje pitanja atributa koji su korišćeni za engleski jezik, a nemaju svoj ekvivalent u srpskom jeziku - predlaganje metode/algoritma za imeničke fraze, subjektivitet, čitljivost, ‘teške’ reči
- Eksperimentisanje za različitim razvijenim tokenizatorima/lematizatorima/PoS tagerima

KRAJ

Hvala na pažnji!

Pitanja i sugestije

kontakt: milena.sosic@gmail.com