

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 238

**AUTOMATSKA KLASIFIKACIJA
DISKRETNIH EMOCIONALNIH STANJA NA
OSNOVU LINGVISTIČKIH ZNAČAJKI
GOVORA**

Antonio Kolak

Zagreb, lipanj 2011.

Zahvala:

Zahvaljujem mentoru prof.dr.sc. Davoru Petrinoviću na strpljenju i pomoći pri izradi ovog diplomskog rada, te vodstvu kroz diplomski studij. Srdačnu zahvalu također upućujem i Branimiru Dropuljiću, dipl.ing., na susretljivosti i zajedničkoj suradnji.

Najveće hvala mojim roditeljima na razumijevanju i podršci tijekom studija.

Diplomski rad posvećujem sestri Maji.

Sadržaj

1	Uvod	1
2	Problematika klasifikacije emocionalnih stanja na osnovu značajki govora	2
3	Korpus Hrvatskog emocionalnog govora.....	5
3.1	Faza prikupljanja materijala	5
3.2	Faza anotacije.....	8
3.3	Ekstrakcija neartikuliranih zvukova	13
4	Aplikacija za provođenje eksperimenta anotacije—„KEG procjena“	14
4.1	Opis funkcionalnosti aplikacije „KEG procjena“	15
5	Gradnja lingvističkog modela.....	20
5.1	Zadatak klasifikacije.....	20
5.2	Razdvajanje transkripta u klase.....	21
5.3	Tokenizacija	22
5.4	Frekvencija pojavljivanja tokena	24
5.5	Relativna frekvencija pojavljivanja tokena.....	27
5.6	Type-Token ratio	29
5.6.1	Statistika TTR-a primijenjenog na KEG	30
5.7	Skup jedinstvenih tokena.....	31
5.7.1	Skup jedinstvenih tokena sa njihovim pripadajućim relativnim frekvencijama	33
5.8	Princip rada primitivnog sustava temeljenog na unigram modelu	35
5.8.1	Sučelje za demonstraciju rada primitivnog unigram sustava.....	36
5.9	Principi redukcije dimenzionalnosti.....	36
5.9.1	Frekvencija dokumenata	37
5.9.2	Hi kvadrat χ^2 test.....	37
5.9.3	Značajka zajedničke informacije	39
5.10	Naivni Bayesov tekstualni klasifikator	40
5.10.1	Implementacija naivnog Bayesovog tekstualnog klasifikatora.....	41
6	Povezivanje sa HTK alatom	42
6.1	Priprema podataka	43
6.1.1	Priprema podataka za trening	44
6.1.2	Priprema podataka za test	47
7	Analiza rezultata.....	49
7.1	Analiza točnosti primitivnog unigram sustava na cjelokupnom korpusu	49

7.2	Analiza točnosti naivnog Bayesovog tekstualnog klasifikatora na cjelokupnom korpusu	50
7.3	Analiza točnosti prepoznavanja govora HTK alatom	52
8	Zaključak	54
9	Literatura	55
10	Sažetak	59
11	Summary	60
12	Privitci	61
12.1	Privitak — Statistika za pojedinu snimku prema spolu, dobi, ID-u govornika, ekspresiji i procijenjenoj emociji	61
12.2	Privitak — Transkript za pojedinu snimku	77
12.3	Privitak — VBA kôd aplikacije korištene za anotaciju korpusa	90
12.3.1	VBA kôd—Modul1	90
12.3.2	VBA kôd—UserForm1	90
12.4	Privitak—Matlab kôdovi	103
12.4.1	tokenizer.m	103
12.4.2	fp.m	105
12.4.3	rfp	106
12.4.4	ttr.m	107
12.4.5	unique1.m	108
12.4.6	rfp_unique.m	110
12.4.7	differenceBtwCell.m	112
12.4.8	loadModels.m	113
12.4.9	compareWithModels.m	114
12.4.10	KEGto5class.m	115
12.4.11	bezkvacice.m	117
12.4.12	bezkvacice2.m	118
12.4.13	sakvacicom.m	120
12.4.14	txt2lower.m	122
12.4.15	tokenize.m	123
12.4.16	tokenize2.m	125
12.4.17	fpStruct.m	127
12.4.18	rfpStruct.m	128
12.4.19	upis_u_unknown.m	129
12.4.20	KEG_randperm.m	130

12.4.21	KEGto2class_G_S.m.....	136
12.4.22	KEGto_all_TXT.m.....	137
12.4.23	TXToModels.m.....	138
12.4.24	loadDAT.m.....	140
12.4.25	trainMNNB.m.....	141
12.4.26	compare_with_NBM.m.....	142
12.4.27	naive_bayes_text_classifier.m.....	143
12.4.28	auto_prim_model.m.....	144
12.4.29	auto_prim_model2.m.....	145
12.4.30	auto_prim_model3.m.....	146
12.4.31	auto_prim_model4.m.....	147
12.4.32	HTK_podjela_snimaka.m.....	148

1 Uvod

Ljudski mozak u stanju je na temelju audio informacije afektivnog govora prepoznati potencijalne emocije u tome govoru. Analiza govornog signala ovim radom zalazi u jednu sasvim novu sferu, a to je prepoznavanje emocija. Ideja je razviti takav sustav koji bi uz korištenje gotovih funkcija iz alata HTK, za govornu sekvencu na ulazu, davao estimaciju najvjerojatnije emocije na izlazu. Estimacija bi se vršila kroz dva sustava. Prvi estimator bi se bavio akustičkim parametrima spomenute sekvence, dok bi se drugi estimator bavio lingvističkim značajkama. Cilj ovog diplomskog rada je razviti sustav koji bi na osnovu lingvističkih značajki tako izgovorenog teksta mogao izvršiti estimaciju emocije.

Analiza putem akustičkih parametara je opisana u drugom diplomskom radu koji se pisao paralelno, a nosi naziv „Automatska klasifikacija diskretnih emocionalnih stanja na osnovu akustičnih značajki govora“ (Milosz, 2011).

Lingvistički estimator podrazumijeva analizu verbalnog iskaza u obliku zapisanog teksta koji je dobiven kao izlaz iz sustava za automatsko prepoznavanje govora. Fuzija rezultata obaju estimatora bi trebala dati konačan izlaz iz sustava za estimaciju emocija. Sustav se temelji isključivo na pet diskretnih emocija: sreća, tuga, strah, ljutnja i neutralno. Korišteni korpus Hrvatskog emocionalnog govora je sačinjen upravo od tih emocija. Za razvoj sustava bilo je nužno prikupljanje i anotiranje jednog takvog relevantnog korpusa emocionalnog govora koji će se koristiti za postavljanje i uhadavanje sustava, te testiranje i analizu rezultata. Da bi se obavila klasifikacija diskretnih emocionalnih stanja na osnovu lingvističkih značajki govora, potrebno je izraditi transkript korpusa Hrvatskog emocionalnog govora. Transkript pored tekstualnog zapisa u obliku izgovorenih rečenica za svaku pojedinu snimku treba sadržavati i posebne glasove koji ne ulaze u uobičajene vokabulare jezika, kao što su zamuckivanja, zastajkivanja, nesuvisle riječi i sve ostale glasovne ekspresije koje su povezane uz neko emocionalno stanje, a nalaze se u snimkama korpusa. Na osnovu svega iznesenog izraditi će se jezični model za prikupljeni korpus, poput unigrama, bigrama, odnosno n-grama, čime će se pokušati postići bolja točnost u odnosu na do sada razvijene sustave.

2 Problematika klasifikacije emocionalnih stanja na osnovu značajki govora

Pri dizajniranju sustava umjetne inteligencije glavni problem je osigurati relevantni skup visoko kvalitetnih materijala koji mogu biti korišteni pri gradnji sustava, odnosno kao podaci za treniranje sustava. Stoga, prvi korak u razvoju sustava za prepoznavanje emocija na osnovu značajki govora je prikupljanje skupa značajki koje ćemo koristiti za treniranje. U ovom poglavlju je opisano kako je tekla gradnja i anotacija prvog Hrvatskog korpusa emocionalnog govora (KEG). U narednom tekstu su opisani procesi prikupljanja i anotacije emocionalnog korpusa kao i njegova statistička svojstva. Biti će objašnjeno tko su sve bili govornici, koji su izvori prikupljanja materijala te koja diskretna emocionalna stanja se nalaze u korpusu.

Pitanje prepoznavanja emocija plijeni sve više i više pažnje u različitim znanstvenim disciplinama. Rastuće zanimanje na području interakciju između čovjeka i računala zasigurno značajno pridonosi razvoju sustava prepoznavanja emocija. Razlog tome je činjenica da se modeliranje naprednih korisničkih sučelja između čovjeka i računala mora barem u nekoj mjeri osloniti na točnu identifikaciju emocionalnog stanja. Neki od uspješno modeliranih sustava prepoznavanja emocija su predstavljeni u (Picard, 1995), (Fernandez, 2003), (Lee, 2005), (Lisetti, 2005), (Devillers, 2006), (Nagel, 2007), i (Datcu, 2009). Trenutačno najpopularnije primjene ovakvih sustava su: sustavi sigurne vožnje, sustavi e-učenja, telemedicina, poboljšanja udaljenje kućne njege itd..

Emocionalna stanja se mogu manifestirati kroz različite modalitete, poput psihofiziologije, govora, izraza lica, gesta itd. Ovaj diplomski rad se bavi opisom razvoja sustava prepoznavanja diskretnih emocija na osnovu lingvističkih parametara glasa. S obzirom na prednosti i mane svih modaliteta možemo zaključiti kako govor ima veliki potencijal, prvenstveno stoga što predstavlja metodu praćenja koja zahtjeva neinvazivne senzore, ali u prilog mu ide i prirodnost izražavanja emocije kroz govor bez oklijevanja. Sve to pojednostavljuje akviziciju emocionalnog materijala.

Prva istraživanja na ovome području su započela u 1930-im godinama. U tome razdoblju su po prvi puta uspostavljene znanstveno mjerljive relacije između govora i emocije. Godine 1936., Cowan je napravio prvu analizu akustičkih značajki ljudskoga glasa

snimljenog tokom javnog govora. Nekoliko godina poslije, Fairbanks i Pronovost su otišli korak dalje i napravili analizu na širem spektru emocija. Ta početna istraživanja su bila značajna u pogledu otkrivanja potencijala ovog područja. S vremenom je ova interdisciplinarna tema počela zahvaćati sve širi interes znanstvene javnosti. Interdisciplinarnost ove teme krije se u grupaciji stručnjaka koje ona ujedinjuje, a to su: psiholozi, lingvisti, neurolozi, te u zadnje vrijeme računalni stručnjaci. Uvođenje računala i računalnih stručnjaka značajno je doprinijelo razvoju ovoga područja, putem razvoja sustava automatskog prepoznavanja emocija kao i primjene posebnih statističkih metoda i analize odgovarajućih karakteristika glasa.

Informacija o emociji koju nosi ljudski glas, može se izvući iz dva glavna svojstva glasa. Prvo svojstvo je akustička informacija koju nosi glas, a drugo lingvistička informacija. Emocije mogu biti iskazane neverbalno kroz prozodične strukture iskaza, što odgovara akustičkoj informaciji, ali također mogu biti iskazane i verbalno tako da se osjećaji direktno izraze kroz izgovorene riječi, što onda predstavlja lingvističku informaciju. Većina autora stavlja naglasak na akustična svojstva glasa zato što su više podložna utjecaju emocionalnog stanja stoga ih je teže prikriti i kontrolirati. Najrelevantnije akustičke značajke se mogu izvući iz: perioda oscilacije vokalnih akorda, energije i snage govornog signala, mjere prolaska signala kroz nulu te distribucije spektra govornog signala. Neki od autora se ipak fokusiraju na izvlačenje emocija iz lingvističkih značajki (Wu, 2006). Informacije značajne za prepoznavanje emocije mogu biti izvučene iz jednostavnih n-gram jezičnih modela, opisujući statističke vjerojatnosti pojave emocionalnih riječi i fraza. Lingvističko prepoznavanje emocija može biti napravljeno na tri razine. To su leksička, sintaktička te semantička. Akustičke kao i lingvističke značajke su dva važna aspekta afektivnog govora koje nisu nužno u korelaciji jedne sa drugima. Iz ovoga proizlazi da bi fuzija sustava dala najbolju informaciju o estimiranoj emociji, čime se proširuju svojstva svakoj pojedinog sustava.

Početak gradnje sustava automatskog prepoznavanja emocija iz značajki govora je uglavnom bazirana na prikupljanju audio materijala afektivnog govora od strane emocionalno nabijenih govornika. Iz prijašnjih istraživanja je vidljivo kako neki od istraživača preferiraju stvarne emocije dok drugi preferiraju glumljene. Međutim prema Sherer-u (Banse, 1996) teško je napraviti razliku između ta dva pristupa. „Prirodni govor je

također dijelom glumljen, kao što je i glumljeni djelomično stvaran“. Kvaliteta korpus ovisi o različitim faktorima, poput različitosti govornika (razlika u godinama, spolu, govornom stilu i dijalektu), kvaliteti prikupljenog audio zapisa, kao i o kvaliteti prikupljenih emocija i balansiranosti između prikupljenih i ciljanih emocija. Drugi faktori koji trebaju biti razmotreni pri gradnji korpusa su kvaliteta audio zapisa, razumljivost, pojave dvostrukog govora (zapis gdje dvoje ili više govornika priča istovremeno), kao i pojave šuma, glazbe ili drugih zvukova u pozadini audio zapisa koji nisu povezani sa zapisom emocionalnog govora.

Korištenje već izgrađenog korpusa je generalno poželjno pri razvoju sustava, međutim trenutačno za Hrvatski jezik postoji samo korpus VEPRAD (Ipsic, 2007), (Ipsic, 2008). Korpus VEPRAD je sačinjen od audio zapisa vijesti i vremenskih prognoza izgovorenih od strane radio spikera. Takvi audio zapisi su uglavnom neafektivni odnosno sastoje se od neutralnog govora, što nije dovoljno kako bismo taj korpus primijenili na sustav automatske klasifikacije emocionalnog stanja na osnovu lingvističkih parametara glasa. Iskorištenje već postojećih engleskih korpusa (Banse, 1996), (Bunger, 2002), (Cowie, 2003), kao i drugih (Jovicic, 2004) nije bila opcija pošto su ciljani sustavi primarno namijenjeni Hrvatskom jeziku. Ciljane aplikacije primjene ovog korpusa su: SIT (engl. *Stress Inoculation Training*), VRET (engl. *Virtual Reality Exposure Therapy*), i druge slične aplikacije koje kao cilj imaju istraživanja metoda redukcije stresa na Hrvatskim psihoterapijskim institucijama. Sustav prepoznavanja emocija bi trebao biti dio projekta—Adaptivno upravljanje virtualnim scenarijima u terapiji PTSP-a, koji se razvija na Fakultetu elektrotehnike i računarstva, Sveučilišta u Zagrebu. Prvi korpus emocionalnog govora za Hrvatski jezik je građen kroz dvije faze, fazu prikupljanja materijala i fazu anotacije. Tijek radnji i opis rada u tim faza će biti opisan kroz sljedeća poglavlja.

3 Korpus Hrvatskog emocionalnog govora

Trenutačna verzija korpusa emocionalnog govora (KEG) sadrži segmente afektivnog govora u trajanju od približno 40 minuta. Snimke su anotirane putem diskretne i kontinuirane raspodjele. Pored anotacije emocije svaka od snimaka sadrži pripadajući ID govornika, spol govornika, dobnu skupinu kojoj pripada govornik, te informaciju da li se radi o glumljenoj ili stvarnoj emociji. Ono što je osobito bitno za sustav prepoznavanja diskretnih emocija na osnovu lingvističkih parametara glasa je da svaka snimka ima i odgovarajući transkript na razini riječi. Unutar transkripta se osim standardnih riječi Hrvatskog jezika nalaze i posebni zapisi za još nekoliko specijalnih, neartikuliranih zvukova, koji su učestali pri afektivnom govoru.

3.1 Faza prikupljanja materijala

Postoji nekoliko pristupa fazi prikupljanja korpusa afektivnog govora. Jedan od pristupa je organiziranje snimanja afektivnog audio materijala od strane profesionalnih glumaca. U tome slučaju glumci izgovaraju glumljene rečenice prema specifičnom scenariju koji predstavlja određeno emocionalno stanje. Drugi pristup bi bio organizirati sustav pobude emocija u kontroliranim uvjetima, pri čemu se ispitanika nekom vrstom sustava (npr. računalnom igrom ili specijalnim ispitivanjem po već unaprijed utvrđenim verificiranim psihološkim pitanjima) želi dovesti u određeno emocionalno stanje i bilježi se njegova reakcija, bez da on zna koja je stvarna funkcija sustava, odnosno ispitivanja. Treći pristup bi bio sakupljanje afektivnog govora iz već postojećih izvora snimaka poput Interneta i filmova. Svaki od gore nabrojenih pristupa ima svoje prednosti i mane. Obično je teško izdvojiti stvarne emocije sa zadovoljavajućom kvalitetom zvuka iz već postojećih izvora. U drugu ruku, studijski snimljeni materijali obično nemaju tu razinu prirodnosti afektivnog stanja koja se može naći kod neglumljenih snimaka.

U dosada razvijanim sustavima je vidljivo kako je prva metoda izrade korpusa dominantna. Dakle, radi se o slučaju kada je korpus izgrađen od snimaka snimljenih u studijskim uvjetima, od strane profesionalnih glumaca, koji su odglumili određeni unaprijed definirani emocionalni scenarij. Međutim, u našem slučaju ipak smo se odlučili za treću metodu i to ponajprije zbog ograničenog budžeta kojeg smo imali na

raspolaganju za projekt, pa tako i za izgradnju KEG-a. U inicijalnoj verziji prikupljanja isječaka sa Interneta i iz filmova težilo se k tome da se izdvoje čiste emocionalne ekspresije. Dio korpusa koji je klasificiran kao „stvarne emocije“ uglavnom je prikupljen iz Internet isječaka „reality show“-ova, kao i različitih dokumentarnih filmova. Drugi dio korpusa nazvan „glumljene emocije“ je prikupljen iz Hrvatskih filmova, TV emisija i audio priča. Ovakva početna distribucija u izvorima korpusa omogućuje daljnju analizu korpusa s obzirom na skup glumljenih odnosno stvarnih emocija.

Primarna ideja je sagraditi bazu podataka za izgradnju sustava automatskog prepoznavanja emocija za Hrvatski jezik. Takav sustav bi bio baziran na akustičkim svojstvima emocionalnog govora na Hrvatskom jeziku, kao i na lingvističkim svojstvima Hrvatskog jezika za iskaz emocija. To znači da gramatika govornih iskaza nije nužno točna niti povezana sa pravilima pisanja. Korpus je sačinjen od takvih govornih iskaza da su uglavnom svi Hrvatski dijalekti pokriveni, a većina govora iz segmenta „stvarnih emocija“ je spontana po svojoj definiciji. Također, mali dio KEG-a je konstruiran i iz govornih iskaza govornika koji su porijeklom iz Srbije i Bosne i Hercegovine. Osim velike sličnosti ovih dvaju jezika sa Hrvatskim jezikom takva nehomogenost je opravdana činjenicom da je Hrvatska populacija značajno pomiješana sa ove dvije nacionalne manjine.

Tri su glavne mjere opreza o kojima se vodilo računa pri sakupljanju audio materijala.

1. Unutar jednog iskaza (snimke) smije biti prisutan samo jedan govornik sa samo jednom emocijom.
2. Drugi zvukovi (okolni zvukovi i drugi glasovi) kao i pozadinska buka su morali biti minimalni kako bi snimka ušla u odabir za korpus.
3. Kvaliteta zvuka je morala biti dovoljno visoka da slušatelj snimke može prepoznati neku vrstu informacije iz govora kao i potencijalnu emociju govornika

Ovo vodi situaciji u kojoj po jednom satu obrađenog audio materijala imamo samo dvije minute korisnog emocionalnog zapisa. Glavni problem vađenja glumljenih emocija je pronalazak dijelova snimke bez pozadinskih zvukova (glazbe). Kod snimaka stvarnih emocija glavni problem je predstavljala loša kvaliteta audio zapisa koja je značajno niža

nego kod stvarnih situacija. Također kod snimaka stvarnih emocija teško je pronaći vremenski prozor unutar kojeg samo jedan govornik daje svoj iskaz (priča).

Sve snimke su normalizirane i pohranjene u wav format sa frekvencijom uzorkovanja od 11025 Hz, i 16 bps (*bit per sample*), mono. U fazi prikupljanja interesantni zvučni iskazi su izdvojeni iz snimaka i klasificirani u jednu od pet diskretnih emocionalnih kategorija, a to su: sreća, tuga, ljutnja, strah i neutralno. Svakom iskazu je dodijeljena samo jedna klasa.

Nakon prve faze prikupljanja KEG je imao 714 snimaka, ukupne duljine trajanja 56 minuta i 22 sekunde. Nakon ponovnog pregleda i procesa filtracije na osnovu glavnih mjera opreza navedenih gore, izbačeno je 40 snimaka te statistika cjelokupnog KEG-a nakon faze prikupljanja izgleda kao što je prikazano u tablici (Tablica 1).

Tablica 1-KEG statistika u fazi prikupljanja

Emocija	Snimaka (broj.)	Spol po snimci (M/Ž)	Vrsta emocije (D/Ado./Odra.)	Iskaz emocija (S/G)	Govornici (broj.)	Spol po govorniku (M/Ž)	Dobna skupina (C/Ado./Odra.)	Ukupno trajanje (mm:ss)	Prosječno trajanje (s)
Sreća	145	53/92	9/1/135	77/68	49	27/22	3/1/45	7:21	3.04
Tuga	105	28/77	2/0/103	50/55	30	11/19	1/0/29	9:39	5.51
Ljutnja	287	157/130	3/2/282	145/142	61	31/30	1/1/59	14:21	3
Strah	72	19/53	22/0/50	13/59	28	13/15	4/0/24	4:19	3.6
Neutralno	65	38/27	3/0/62	46/19	36	22/14	2/0/34	11:15	10.39
Ukupno	674	295/379	39/3/632	331/343	204	104/100	11/2/191	46:55	4.18

Iz tablice (Tablica 1) je vidljivo kako je korpus nakon faze prikupljanja sačinjen od 674 snimke ukupnog trajanja 46 minuta i 55 sekundi. Zanimljivo je primijetiti kako ljutnja sa svojim najvećim brojem iskaza njih 287, i najvećim ukupnim trajanjem u odnosu na ostale emocije, ima zapravo najmanje prosječno trajanje uzorka u odnosu na ostale emocije. Nasuprot tome neutralni iskazi imaju najmanji broj snimaka, a najveću duljinu trajanja od 10,39 sekundi. U KEG imamo nešto više muških govornika nego ženskih (104 naspram 204), ali gledano sa stajališta iskaza tada imamo više snimaka sa ženskim govornicima (379 naspram 674). Pored ostale klasifikacije iskazi su klasificirani i prema tri dobne skupine govornika, a to su: djeca (2-14 godina), adolescenti (14-19 godina), i odrasli (20+ godina). Većina govornika su ipak odrasli, dok imamo samo tri govornika koji spadaju u adolescentsku skupinu. Što se tiče distribucije snimaka po vrsti emocionalnog iskaza, tj. da li su emocije glumljene ili stvarne, omjer je 331 stvarnih naspram 343 glumljene. Takva jednaka distribucija pogoduje karakteristikama korpusa.

Pored statističkih podataka, svaki iskaz ima i pripadajući transkript koji je ručno izrađen za svaku snimku, koristeći se pri tome Hrvatskim riječima iz raznih dijalekata. Pored standardnih riječi transkript bilježi pojedine specijalne, neartikulirane zvukove, poput plača, smijeha, disanja, oklijevanja, vikanja, kašljanja, nekoliko vrsta uzvika, itd.

3.2 Faza anotacije

Kako bi se poboljšala statistička značajnost korpusa emocionalnog govora, proces anotacije je osmišljen u suradnji sa kolegama, sa Filozofskog fakulteta odsjeka za psihologiju. Pet studenata sveučilišta u Zagrebu je sudjelovalo u eksperimentu. Njihovi podatci su vidljivi iz tablice (Tablica 2).

Tablica 2-Podaci recenzenata

ID	Spol	Starost (god.)	Studij
Student_1	Žensko	22	Psihologija
Student_2	Žensko	24	Psihologija
Student_3	Žensko	23	Psihologija
Student_4	Žensko	22	Elektrotehnika i informacijska tehnologija
Student_5	Muško	24	Medicina

Njihov zadatak je bio preslušati cijelokupni korpus i dati njihovu osobnu ocjenu emocionalnog stanja za svaku pojedinu snimku. Svi studenti su prošli edukaciju u kojoj im je objašnjena osnovna problematika i što se točno od njih očekuje. Studenti su dali svoje ocjene za svaku od 674 snimke koje su nalaze u korpusu. Predviđeno trajanje anotacije je bilo procijenjeno na 15 sati. Za taj posao studenti su plaćeni prosječnim honorarom po satu, predviđenim za studente u Hrvatskoj. Kako bi se studenti dodatno motivirali za što bolju anotaciju, obećan je i bonus za onoga tko izvrši najbolje anotiranje. Visina bonusa je bila nešto manja od standardne isplaćene pristojbe po završetku obavljenog posla.

Za potrebe anotacije razvijena je aplikacija pomoću koje su studenti davali ocjene za svaku pojedinu snimku. Screenshot aplikacije je vidljiv na slici (Slika 1), a sama funkcionalnost će biti opisana u daljnjem tekstu. Svaki od studenata je dobio svoju permutiranu verziju snimaka kako međusobno nebi mogli uspoređivati rezultate. Naputak recenzentima je bio da u jednom segmentu nipošto ne anotiraju više od 100 snimaka kako nebi došlo do zasićenja, te da rade pauzu od minimalno 3 sata, između dva procesa anotacije.

Ocjenjivane su diskretne i kontinuirane emocije odnosno dva različita zapisa istog emocionalnog stanja.

Diskretne emocije (sreća, tuga, strah i ljutnja) su ocjenjivane na skali od 0 do 10, pri čemu je nula označavala odsustvo emocije. Ako su sve diskretne emocije (sreća, tuga, strah i ljutnja) ocijenjene sa niskim intenzitetom odnosno s intenzitetom bliskim nuli, tada se smatralo da se radi o diskretnoj emociji (neutralno). Jedna snimka nije nužno morala biti ocijenjena samo sa jednom diskretnom emocijom, već je bilo moguće dati distribuciju ocjena po emocijama. Odnosno, u jednoj snimci su mogle biti prisutne sve četiri emocije sa različitim intenzitetom. Teoretski je bilo moguće odabrati da su na jednoj snimci prisutne sve četiri emocije sa intenzitetom 10.

Kontinuirane emocije su opisane pomoću skale pobuđenosti i ugone. Za ugodu kao i za pobuđenost smo imali 9 razina. Razine ugone su se kretale od 1 što je predstavljalo maksimalnu neugodu, pa do 9 što je predstavljalo maksimalnu ugodu. Sličnu situaciju imamo i sa razinama pobuđenosti gdje je razina 1 predstavljala nepobuđenost (uspavanost), a razina 9 maksimalnu pobuđenost. Razina 5 se smatrala neutralnom i na skali ugone i na skali pobuđenosti.

Pored anotacije emocija recenzenti su mogli aktivirati četiri različite zastavice u slučajevima kada zatečeno emocionalno stanje nije dobro opisano sa nekim od ponuđenih opcija za diskretnu ili kontinuiranu ocjenu. Ovakvi slučajevi su bili rijetki jer se pozornost takvim iznimkama davala već pri procesu prikupljanja.

Zastavice za četiri spomenuta stanja su opisna u slijedećim točkama:

1. Pojavila se druga diskretna emocija. Potrebno je navesti o kojoj emociji/ama se radi, te navesti intenzitet svake pojedine emocije.
2. Nije moguće odrediti bilo kontinuiranu bilo diskretnu emociju, neovisno o razlogu (velika pozadinska buka, smetnje drugih zvukova, loša kvaliteta govora itd.)
3. Pojava nehomogenosti emocionalnog stanja. Tokom jedne snimke odnosno iskaza, i diskretna i kontinuirana emocija može varirati. Svake ekstremne varijacije treba zabilježiti. Primjer ekstremne varijacije bi bio da imamo unutar jedne snimke naglu promjenu iz čiste sreće u čistu tugu.

4. Pojava dvoje ili više govornika iste dominacije glasa unutar jedne snimke. U ovome slučaju treba opisati emociju i vrijeme zapažanja te emocije za svakog od govornika. Popratne nedominantne zvukove se može ignorirati.

Također, svaka snimka je mogla biti komentirana ako se ukaže potreba za takvo što, u slučaju pojave nekakvog oblika neregularnosti, iako se za tim nije ukazala potreba.

KEG sučelje za ocjenu emocije

< Prethodna snimka

SNIMKA BROJ: 1

NIJE SPREMLJENO

Sljedeća snimka >

IZNIMNO	SREĆA	TUGA	STRAH	LJUTNJA	IZNIMNO	MAKSIMALNA UGODA	UGODA	POBUĐENOST	MAKSIMALNA UZBUĐENOST
VRLO	10	10	10	10	VRLO		9	9	
	9	9	9	9			8	8	
	8	8	8	8			7	7	
	7	7	7	7			6	6	
	6	6	6	6			5	5	
UMJERENO	5	5	5	5	UMJERENO	NEUTRALNO	4	4	UMJERENA UZBUĐENOST (NEUTRALNO)
	4	4	4	4			3	3	
	3	3	3	3			2	2	
BLAGO	2	2	2	2	BLAGO		1	1	
	1	1	1	1					
ODSUTSTVO	Odsutstvo sreće	Odsutstvo tuge	Odsutstvo straha	Odsutstvo ljutnje	ODSUTSTVO	MAKSIMALNA NEUGODA			NEUZBUĐENOST (USPAVANOST)

KOMENTAR

Pusti snimku (Play)

Spremi odabir za snimku

Zatvori

DRUGO

Slika 1-Aplikacija za provođenje eksperimenta—"KEG procjena"

Kako bismo izdvojili snimke koje nisu adekvatne za daljnju obradu, odnosno izdvojili samo one koje su homogene, definirane diskretnim emocijama navedenim gore, primijenili smo sljedeće kriterije:

1. Snimke koje su označene sa zastavicom 2, 3 ili 4 od strane najmanje 3 recenzenta, ukazuju na složenost situacije u zapisu te taj zapis biva filtriran.
2. Za svaku snimku je izvučena dominantna emocija. Lista jedne ili više diskretnih emocija koje su imale najviši intenzitet su bilježene za svakog od recenzenta. Emocije dodane kroz zastavicu 1 su također uključene. U slučaju da niti jedna emocija nije označena sa intenzitetom većim od 3, pretpostavilo se kako se radi o neutralnoj emociji. Konačno je snimci dodijeljena ona emocija za koju je četvoro ili više recenzenta reklo kako se radi o dominantnoj emociji.

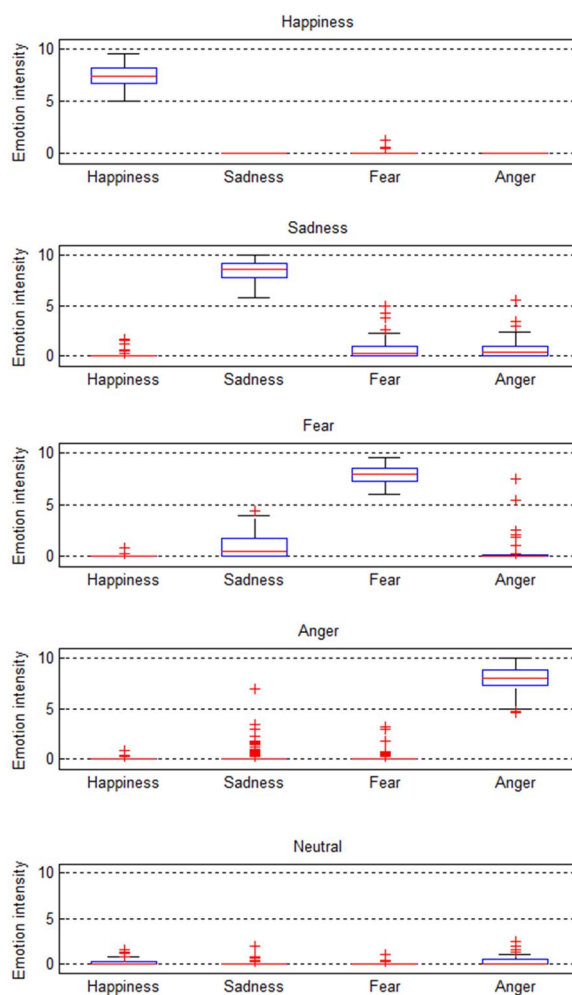
Rezultati anotacije su predloženi u tablici (Tablica 3) gdje su sve anotirane emocije uzete u obzir.

Tablica 3-KEG statistika nakon anotacije

Emocija	Snimaka (broj.)	Spol po snimci (M/Ž)	Dobna skupina (D/Ado./Odra.)	Vrsta emocija (S/G)	Govornici (broj.)	Spol po govorniku (M/Ž)	Dobna skupina (C/Ado./Odra.)	Ukupno trajanje (mm:ss)	Prosječno trajanje (s)
Sreća	77	25/52	10/1/66	34/43	34	15/19	3/1/30	5:27	4.25
Tuga	83	16/67	1/0/82	44/39	30	7/23	1/0/29	8:57	6.47
Ljutnja	231	133/98	2/2/227	105/126	60	31/29	1/1/58	13:27	3.49
Strah	33	7/26	14/0/19	3/30	16	6/10	3/0/13	2:26	4.42
Neutralno	72	43/29	1/0/71	61/11	50	30/20	1/0/49	10:59	9.15
Ukupno	496	224/272	28/3/465	247/249	146	69/77	7/2/137	41:16	4.99

Kao što je prikazano u tablici (Tablica 3) broj snimki je reduciran sa 674 na 496. Uspoređujući inicijalnu anotaciju u fazi prikupljanja i naknadno provedenu anotaciju, vidimo kako su 462 od 496 snimke jednako anotirane, odnosno nose istu emocionalnu oznaku. Konačno su 462 snimke filtrirane na principu gore navedena dva kriterija. Nakon prve faze otpale su samo četiri snimke. Omjeri prema spolu po snimci, dobi, i vrsti emocije su se promijenili minimalno, dok su se faktori ukupni broj govornika kao i spol po govorniku, značajno promijenili. Ovime je iz korpusa izbačeno gotovo 200 snimki. Ipak, ukupno vrijeme trajanja korpusa je smanjeno samo za otprilike 5 minuta. Takva promjena je povezana sa činjenicom da se prosječno trajanje snimke povećalo za gotovo sekundu.

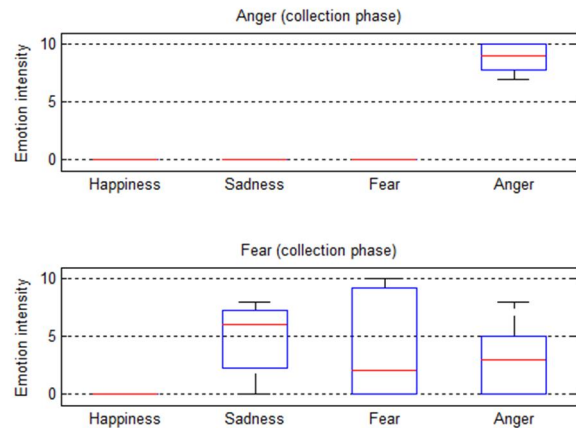
Slika (Slika 2) prikazuje rezultate za sve filtrirane snimke po svakoj emociji (sreća, tuga, strah, ljutnja, neutralno) zasebno. Snimke su klasificirane prema provedenom postupku anotacije i plotane na 5 različitih grafova za svaku emociju posebno, pomoću box-whisker dijagrama. Centralna oznaka na svakom grafu označava medijan, dok su rubovi na 25. i 75. percentili. Box-whisker dijagram obuhvaća krajnje vrijednosti, a sve grube pogreške su zasebno prikazane crvenim križićima. Iz grafova je vidljivo kako uzorci koji su anotirani kao sreća bilježe jako



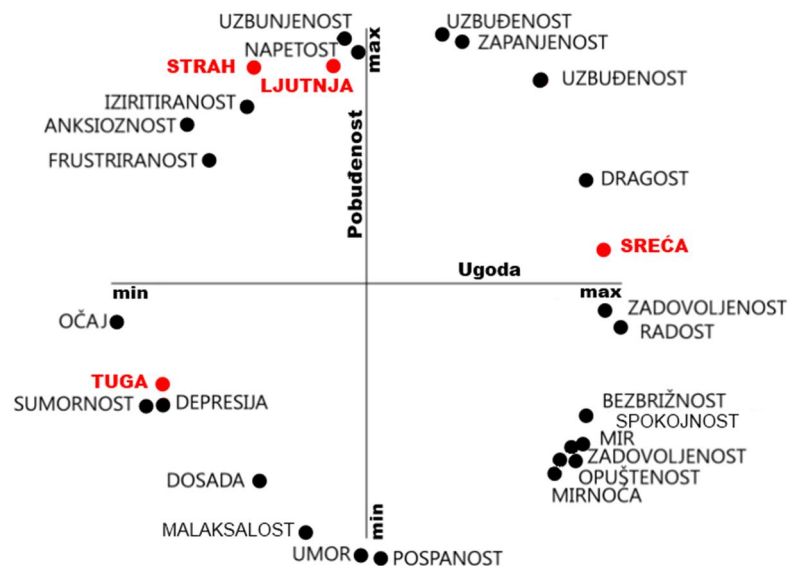
Slika 2-Rezultati po emocijama za sve snimke (Chmura, 2011)

malo ocjena neke druge emocije, dok je to nije slučaj kod tuge, ljutnje i straha. Takvi rezultati su opravdani sukladno Russellovom modelu (Russell, 1980), pošto su emocije straha i ljutnje bliže jedna drugoj nego što je to sreća naspram te dvije emocije (Slika 4).

Slika (Slika 3) predstavlja dva tipična primjera ocjena za pojedinu snimku, koji su se pojavili tokom anotacije. Prvi dijagram predstavlja slučaj gdje se svih 5 recenzenata složilo oko ocjene emocije. Svi recenzenti su emociju ocijenili sa visokim intenzitetom „ljutnje“, što je ujedno i ocjena emocije koju je snimka dobila fazi prikupljanja. Drugi dijagram predstavlja snimku kod koje su recenzenti dali različite ocjene. Dotična snimka je u fazi prikupljanja podataka bila klasificirana kao „strah“. Međutim, samo dvoje recenzenata je snimku ocijenilo kao visoki stupanj „straha“, dok su drugi recenzenti dali drugačiju ocjenu.



Slika 3-Rezultati za dva reprezentativna primjerka (Chmura, 2011)



Slika 4-Prikaz Russellovog modela

3.3 Ekstrakcija neartikuliranih zvukova

Već smo gore spomenuli kako se u transkriptu pored standardnih iskaza riječi nalazi i cijeli niz neartikuliranih zvukova. Pošto ne postoje jezični modeli zapisa takvih zvukova u obliku riječi onda se pokušalo što bliže opisati njihov stvarni zvučnik zapis, pri čemu on ne predstavlja stvarni izgovor te riječi već samo model koji označava postojanje takvog zvuka u zapisu snimke. U tablici (Tablica 4) se nalaze svi korišteni modeli.

Tablica 4-Oznake modela neartikuliranih zvukova

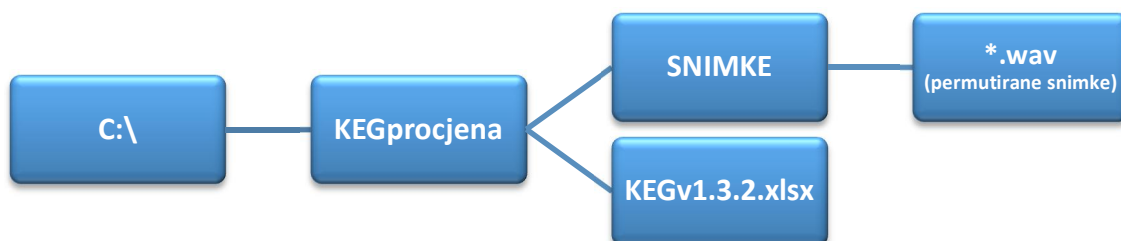
Neartikulirani zvukovi	Oznake modela
Šmrcaње	SMRC_
Plač	CMIZDR_
Zastajkivanje	ZASTAJK_
Udah	UDH_
Izdah	IDH_
zamuckivanje	$(A - Z) \times 3_$
Kašalj	KHM_
Smijeh	HIHI_, HAHA_, HOHO_, HEHE_
Režanje	GRRR_
Vikanje	DER_
Začudenost	HA_, MHM_, EA_
Uzvici	HEJ_, EJ_, OJ_, AH_, AJ_, JAO_

4 Aplikacija za provođenje eksperimenta anotacije—„KEG procjena“

Za potrebe provođenja eksperimenta anotacije bilo je potrebno razviti aplikaciju koja će omogućiti potrebnu funkcionalnost preslušavanja snimka i pohranjivanja rezultata anotacije. Zapisivanje preko 10 različitih podataka za svaku snimku zasigurno nije jednostavan zadatak za recenzenta ako taj posao mora obaviti ručno. Time bi se njegova pažnja sa procesa anotacije usmjerila na proces zapisivanja i pohrane rezultata za pojedinu snimku, što nipošto ne želimo, jer bi takvo što smanjilo točnost procesa anotacije. S obzirom da ovdje govorimo o približno 700 snimaka, tada je jasno da je taj posao gotovo nemoguće obaviti ručno, već da je potrebno na neki način pohranjivati rezultate.

Sukladno opisanim razlozima, razvijena je aplikacija u VBA (Visual Basic for Applications) programskom jeziku, koja zadovoljava gore navedene potrebe. Kôd cjelokupne aplikacije, koji uključuje kôd pomoćnog modula te kôd UserForm-a, se nalazi u poglavlju 12.3 „Privitak — VBA kôd aplikacije korištene za anotaciju korpusa“. Aplikacija se oslanja na Microsoft-ov paket Office, konkretno alat Excel.

Kako bi pokrenuli aplikaciju potrebno je otvoriti dokument pod nazivom KEGv1.3.2 i pritisnuti gumb „Pokreni sučelje“. Također u slučaju da su u sustavu onemogućeni macros-i potrebno je sniziti razinu sigurnosti i omogućiti macros-e. Sustav je potrebno postaviti u sljedeću strukturu na disku:



Slika 5-Struktura za KEG aplikaciju

U direktoriju „SNIMKE“ se nalaze audio snimke korpusa koje je potrebno permutirati za svakog pojedinog recenzenta kako bi se mogućnost uspoređivanja rezultata svela na minimum. Nakon inicijalnog pokretanja aplikacija izgleda kao na slici (Slika 6), sa standardnim vrijednostima odsustva diskretnih emocija i neutralnom položaju za kontinuirane emocije.

Slika 6-Inicijalni izgled aplikacije za anotaciju

4.1 Opis funkcionalnosti aplikacije „KEG procjena“

U ovom poglavlju je opisana funkcionalnost aplikacije kroz sve njezine segmente. Svaki od pojedinih funkcionalnih dijelova aplikacije je naveden i povezan sa popratnom slikom kako bi se zornije predložilo o čemu se radi.

Aplikacija sadrži sljedeće funkcionalnosti:

- listanje snimaka, naprijed<->natrag
- nastavak ocjenjivanja od prve neanotirane snimke
- kontrolna lampica stanja snimke (je li snimka već anotirana ili ne)
- spremanje ocjena u tablicu
- mogućnost višestrukog prolaženja kroz snimke i ocjenjivanja istih
- povlačenje podataka iz tablice natrag u sučelje tako da se direktno iz sučelja mogu pratiti prijašnje ocjene

- pokretanje snimke (puštanje audio zapisa, koristi se Windows Media Player)
- ljestvica u rasponu od 1 do 10 za ocjenu diskretne emocije
- opcije odsustva diskretne emocije
- ljestvica u rasponu od 1 do 9 za ocjenu u kontinuiranom sustavu
- textbox za komentar
- opcija „Drugo“ za specijalna stanja

KEG sučelje za ocjenu emocije

SNIMKA BROJ: 1

NIJE SPREMLJENO

< Prethodna snimka

Sljedeća snimka >

IZNIMNO

VRLO

UMJERENO

BLAGO

ODSUTSTVO

SREĆA

TUGA

STRAH

LJUTNJA

UGODA

POBUĐENOST

MAKSIMALNA UGODA

MAKSIMALNA UZBUĐENOST

NEUTRALNO

UMJERENA UZBUĐENOST (NEUTRALNO)

NEUŽUDENOST (USPAVANOST)

KOMENTAR

Pusti snimku (Play)

Spremi odabir za snimku

Zatvori

DRUGO

Slika 7-Interaktivni gumbi u aplikaciji

Aplikacija ima 5 mogućih gumba, sa intuitivnom funkcionalnošću.

- <Prethodna snimka
- >Sljedeća snimka
- Pusti snimku (Play)
- Spremi odabir za snimku
- Zatvori

Gumbi „Prethodna snimka“ i „Sljedeća snimka“ omogućavaju listanje između snimka i njihovo preslušavanje, kao i pohranjivanje ocjene. Kako bi recenzent u svakom trenutku znao o kojoj snimci se radi, postoji indikator koji se sukladno listanju mijenja i ispisuje broj snimke.

Gumb „Pusti snimku (Play)“, pokreće audio zapis u Windows Media Player-u, za onu snimku koja piše na sučelju aplikacije pod lokacijom „SNIMKA BROJ:“

Gumb „Spremi odabir za snimku“ vrši pohranjivanje svih unesenih podataka za tu dotičnu snimku. Pohranjuju se sve ocjene (i diskretne i kontinuirane), kao i komentar i zastavica „DRUGO“. Kada je pohranjivanje izvršeno, kontrolna lampice se mijenja iz crvene boje sa tekstom „NIJE SPREMLJENO“, u zelenu sa tekstom „SPREMLJENO“.

Gumb „Zatvori“ služi da zatvorimo aplikaciju, po završetku rada, što ujedno možemo napraviti i na crveni „x“ u gornjem desnom kutu. Pri ponovnom pokretanju aplikacije, ona će se locirati na prvu neanotiranu snimku u korpusu.

KEG sučelje za ocjenu emocije

< Prethodna snimka

SNIMKA BROJ: 1

NIJE SPREMLJENO

Sljedeća snimka >

IZNIMNO	SREĆA	TUGA	STRAH	LJUTNJA	IZNIMNO	MAKSIMALNA UGODA	UGODA	POBUĐENOST	MAKSIMALNA UZBUĐENOST
VRLO	10	10	10	10	VRLO	9	9	9	MAKSIMALNA UZBUĐENOST
UMJERENO	9	9	9	9	UMJERENO	8	8	8	UMJERENA UZBUĐENOST (NEUTRALNO)
BLAGO	8	8	8	8	BLAGO	7	7	7	NEUŽBUĐENOST (USPAVANOST)
ODSUTSTVO	7	7	7	7	ODSUTSTVO	6	6	6	
	6	6	6	6		5	5	5	
	5	5	5	5		4	4	4	
	4	4	4	4		3	3	3	
	3	3	3	3		2	2	2	
	2	2	2	2		1	1	1	
	1	1	1	1					

KOMENTAR

Pusti snimku (Play)

Spremi odabir za snimku

Zatvori

DRUGO

Slika 8-Prikaz funkcionalnosti indikatora ID-a i pohrane 1/2

KEG sučelje za ocjenu emocije

< Prethodna snimka

SNIMKA BROJ: 25

SPREMLJENO

Sljedeća snimka >

IZNIMNO	SREĆA	TUGA	STRAH	LJUTNJA	IZNIMNO	MAKSIMALNA UGODA	UGODA	POBUĐENOST	MAKSIMALNA UZBUĐENOST
VRLO	10	10	10	10	VRLO	9	9	9	MAKSIMALNA UZBUĐENOST
UMJERENO	9	9	9	9	UMJERENO	8	8	8	UMJERENA UZBUĐENOST (NEUTRALNO)
BLAGO	8	8	8	8	BLAGO	7	7	7	NEUŽBUĐENOST (USPAVANOST)
ODSUTSTVO	7	7	7	7	ODSUTSTVO	6	6	6	
	6	6	6	6		5	5	5	
	5	5	5	5		4	4	4	
	4	4	4	4		3	3	3	
	3	3	3	3		2	2	2	
	2	2	2	2		1	1	1	
	1	1	1	1					

KOMENTAR

Pusti snimku (Play)

Spremi odabir za snimku

Zatvori

DRUGO

Slika 9-Prikaz funkcionalnosti indikatora ID-a i pohrane 2/2

Na slikama (Slika 8 i Slika 9) je vidljiva izmjena stanja indikatora sukladno broju snimke koja se trenutačno anotira i ovisno o njezinom stanju pohranjenosti.

Slika (Slika 10) prikazuje dva područja za unos ocjene emocionalnog stanja. Područje označeno sa brojem jedan predstavlja prostor za unos diskretne ocjene emocije. Ponuđene su četiri emocije (sreća, tuga, strah i ljutnja) sa intenzitetom od 1 do 10 i njihova odsustva. Kako bi se recenzent lakše snašao, sa lijeve i desne strane skale, nalazi se tekstualni ekvivalenti intenziteta emocionalnog stanja. Gotovo identičan zapis se nalazi i u području označenim brojem 2, samo što se kroz taj dio aplikacije daje kontinuirana ocjene emocije preko ugone i pobuđenosti, te je skala u zapisu od 1 do 9, pri čemu se 5 smatra neutralnim stanjem.

Slika 10-Područja za unos diskretne i kontinuirane ocjene emocionalnog stanja

KEG sučelje za ocjenu emocije

< Prethodna snimka

SNIMKA BROJ: 1

NIJE SPREMLJENO

Sljedeća snimka >

IZNIMNO	SREĆA	TUGA	STRAH	LJUTNJA	IZNIMNO	MAKSIMALNA UGODA	UGODA	POBUĐENOST	MAKSIMALNA UZBUĐENOST
VRLO	10	10	10	10	VRLO	9	9	9	9
UMJERENO	9	9	9	9	UMJERENO	8	8	8	8
BLAGO	8	8	8	8	BLAGO	7	7	7	7
ODSUTSTVO	7	7	7	7	ODSUTSTVO	6	6	6	6
	6	6	6	6		5	5	5	5
	5	5	5	5		4	4	4	4
	4	4	4	4		3	3	3	3
	3	3	3	3		2	2	2	2
	2	2	2	2		1	1	1	1
	1	1	1	1					
	Odsutstvo sreće	Odsutstvo tuge	Odsutstvo straha	Odsutstvo ljutnje		MAKSIMALNA NEUGODA			MAKSIMALNA UZBUĐENOST (NEUTRALNO)

KOMENTAR

Pusti snimku (Play)

Spremi odabir za snimku

Zatvori

DRUGO

Slika 11-Područje za unos komentara i postavljanje zastavice „DRUGO“

Slika (Slika 11) prikazuje prostor za unos komentara i prostor za aktivaciju zastavice „DRUGO“. Opciju „DRUGO“ moguće je i deaktivirati klikom na žutu okolinu, čime opcija biva neaktivna. Svi podatci iz sučelja se pohranjuju u tablicu pri kliku na gumb „Spremi odabir za snimku“, kao što se i već spremljeni podaci ponovno učitavaju u sučelje pri listanju kroz snimke. Pohranjivanje se automatski vrši i u *.x/sx (KEGv1.3.2.xlsx) datoteku tako da nije potrebno vršiti pohranu izvan sučelja.

5 Gradnja lingvističkog modela

Da bismo krenuli sa gradnjom sustava moramo za početak na nekoj apstraktnoj razini definirati što želimo da nam dotični sustav radi. Iz prijašnjeg teksta, jasno je kako bismo željeli za pojedinu rečenicu, na osnovu njezinoga transkripta, reći kojoj od gore navedenih 5 diskretnih emocija (sreća, tuga, strah, ljutnja, neutralno) pripada. Ako sagledamo malo šire ovaj problem onda uviđamo kako se ovdje radi o sustavu raspoznavanja uzoraka odnosno određivanja pripadnosti. Kod takvih sustava razlikujemo:

- sustave učenja pod nadzorom (klasifikacija)
- sustave učenja bez nadzora (grupiranje)

Sustavi učenja pod nadzorom se treniraju kroz skup ulaznih uzoraka koji su unaprijed klasificirani, na osnovu čega se sustav nauči kako raditi klasifikaciju nepoznatog uzorka. Za razliku od njih sustavi učenja bez nadzora imaju unaprijed očekivani broj klasa te na osnovu znanja ugrađenog u sustav za razdvajanje tih klasa vrši se klasifikacija ulaznih uzoraka.

5.1 Zadatak klasifikacije

Općenito zadatak klasifikacije je da skup dokumenata $D = \{D_1, D_2, \dots, D_n\}$ razvrsta u klase $C = \{C_1, C_2, \dots, C_m\}$ na osnovu skupa odluka $A = \{A_{1,1}, A_{1,2}, \dots, A_{n,m}\}$.

Takav sustav možemo vizualizirati kroz tablicu (Tablica 5), gdje stupci predstavljaju dokumente, a redovi klase. Presjek svakog reda i stupca je odgovarajuća odluka od sortiranju pojedinog dokumenta u određenu klasu.

Tablica 5-Tablični prikaz klasifikacije

	D_1	D_2	D_3	...
C_1	$A_{1,1}$	$A_{1,2}$	$A_{1,3}$	
C_2	$A_{2,1}$	$A_{2,2}$	$A_{2,3}$	
C_3	$A_{3,1}$	$A_{3,2}$	$A_{3,3}$	
...				

Kod rada sa sustavima učenja treba biti svjestan svih ograničenja koji sustavi raspoznavanja uzoraka imaju, poput preklapanja klasa i pogrešaka klasifikacije.

Postoje dvije strategije procesiranja:

- Kategorijski orijentirana klasifikacija (CPC- *Category-pivoted classification*)
- Klasifikacija orijentirana na dokumente (DPC- *Document-pivoted classification*)

Kod kategorijski orijentirane klasifikacije radi se o dodjeli svakoga dokumenta kategoriji, bazirano na retku. Ovakav sustav je primjenjiv kada imamo uglavnom nepromjenjiv skup dokumenata, te je kreiranje novih kategorija učestalo.

Kod klasifikacije orijentirane na dokumente imamo suprotnu situaciju kada se radi o dodjeli svakog dokumenta kategoriji bazirano na stupcu. Ovakav sustav je primjenjiv kada imamo uglavnom nepromjenjiv skup kategorija.

U nekim slučajevima obje metode klasifikacije mogu biti primjenjivanje, a dodijeljene vrijednosti za parove dokument–klasa, mogu biti logičke vrijednosti (nula ili jedan) ili vjerojatnost. Sukladno tome izlaz sustava može biti jedna klasa po dokumentu odnosno n klasa po dokumentu.

Iz svakog dokumenta je potrebno izvući elemente bitne za sadržaj dokumenta odnosno za njegovo klasifikaciju. Taj proces je opisan u sljedećem poglavlju.

5.2 Razdvajanje transkripta u klase

Kako bismo mogli provesti treniranje sustava potrebno je izvršiti pripremu podataka u obliku razdvajanja cjelokupnog transkripta u pojedine klase. Ovime je informacija o vrsti emocije pojedine rečenice (snimke) iskorištena za ugradnju toga znanja u sustav. Sukladno činjenici da su snimke, a time i transkript tih snimaka anotirani, potrebno je iskoristiti tu informaciju i ugraditi je u sustav kako bi nakon završetka faze učenje sustav samostalno mogao izvršiti procjenu ulaznih uzoraka.

Za početak je potrebno izdvojiti samo one emocije koje su dobile pozitivnu ocjenu pri prolasku kroz fazu anotacije. Sve ostale snimke koje nisu dobile prolaznu ocjenu nisu pogodne za gradnju sustava. Tako inicijalno filtrirani korpus je potrebno podijeliti u 5 klasa (sreća, tuga, strah, ljutnja, neutralno). Pri tome ćemo dobiti tekstualne datoteke u kojima će se nalaziti transkripti samo onih snimaka koje pripadaju dotičnoj klasi odnosno

emociji. Svaka od tako odbijenih tekstualnih datoteka će se zatim podložiti daljnjoj obradi pri gradnji sustava.

Cjelokupno gore opisani postupak obavlja se unutar funkcije `KEGto5class.m`. pri čemu funkcija kao ulaz prima: `*.x/sx` datoteku sa cjelokupnim transkriptom korpusa, stazu do direktorija u kojem želimo da funkcija kreira klase te zastavicu koja označava da li je potrebno da funkcija kreira novu `*.x/sx` datoteku sa samo onim snimkama koje su prošle anotaciju.

Kao što se može i zaključiti iz gore opisanog postupka, za izgradnju 5 klasa, funkciju je potrebno pozvati dva puta. Prvi put za izgradnju nove `*.x/sx` datoteke sa samo onim snimkama koje su dobile pozitivnu ucjenu pri anotaciji, za daljnju upotrebu u gradnji sustava, te drugi put za kreiranje 5 klasa. Pri pozivu funkcije drugi put, ulazna datoteka sa transkriptom je novo dobivena datoteka iz prvog koraka, koja sadrži samo snimke sa pozitivnom ocjenom anotacije. Zastavice u drugom pozivu se postavlja na nulu.

Pozivi:

```
KEGto5class('Staza do *.xlsx datoteke',...  
  
            'Staza do direktorija za pohranu klasa',...  
  
            zastavica_za_gradnju_nove_xlsx_datoteke)
```

5.3 Tokenizacija

Tokenizacija je postupak kojim iz određenog dokumenta želimo izvući lingvističke značajke tog dokumenta odnosno tokene. U našem slučaju radi se isključivo o riječima, ali zavisno o vrsti dokumenata za koje se sustav radi to ne mora biti slučaj. Cilj tokenizacije je niz znakova, riječi i paragrafa razbiti u listu jedinstvenih tokena, čuvajući distributivna svojstva tokena.

Primjer tokenizacije na sljedećem tekstu dan je na slici (Slika 12).

Tablica 6-Primjer teksta za tokenizaciju

Tekst za tokenizaciju	Danas je lijep i sunčan dan. Uvezivanje diplomskog rada košta 100,00 kn. ZESOI, je, najbolji, zavod, na, FER-u
------------------------------	--

```
ans =

token1: 'Danas'
token2: 'je'
token3: 'lijep'
token4: 'i'
token5: 'sunčan'
token6: 'dan'
token7: '.'
token8: 'Uvezivanje'
token9: 'diplomskog'
token10: 'rada'
token11: 'košta'
token12: '100,00'
token13: 'kn'
token14: '.'
token15: 'ZESOI'
token16: ','
token17: 'je'
token18: ','
token19: 'najbolji'
token20: ','
token21: 'zavod'
token22: ','
token23: 'na'
token24: ','
token25: 'FER'
token26: '-'
token27: 'u'
```

Slika 12-Primjer tokenizacije

Pitanja koja postavlja tokenizacija se odnose na postupak, identifikacije, delimitera, riječi i rečenica. Delimiteri s kojima radi implementirani sustav su navedeni u tablici (Tablica 7). Treba biti svjestan da u nekim slučajevima pojava delimitera ne označava generiranje novog tokena već da se radi o jedinstvenoj cjelini tokena. Kao najjednostavniji primjer možemo navesti brojeve sa decimalnim zarezom. U ovome primjeru moram biti svjesni dvosmislenosti koju ovaj token može značiti, jer se na primjer broj 100,55, može

protumačiti kao sto cijelih pedeset pet, ali i kao dva broja, odnosno kao 100 i kao 55, pri čemu se između njih nalazi još jedan token koji je u biti delimiter. Stoga sve iznimke treba uključiti u kôd lingvističkog automata, tako da se napravi pravilno parsiranje niza.

Tablica 7-Delimiteri

Vidljivi	; . , ! ? \ " () ' : [] / + - = ? ? { } > < * ' " ^ _
Nevidljivi	Space, Tab, New Line, Carriage return

S obzirom da je gotovo nemoguće uključiti sve iznimke, sustav treba prilagoditi namjeni. Gore navedeni sustav delimitera je primjenjiv na naš emocionalni korpus. Pored toga radila se još i normalizacija velikih i malih slova, tako da je ulazni transkript prebačen u mala slova.

Za tokenizaciju su razvijene dvije funkcije u matlabu. Prva funkcija nosi naziv `tokenizer.m`, te omogućuje tokenizaciju ulazne datoteke i njezino pohranjivanje. Dok funkcija `tokenize.m`, samo učitava datoteku sa tekstom i tokene pohranjuje u matlab strukturu.

Pozivi:

```
tokenizer('Staza do ulazne datoteke ', ...
          'Staza do izlazne datoteke\*.lst')

TokensStruct=tokenize('Staza do ulazne datoteke')
```

5.4 Frekvencija pojavljivanja tokena

U prethodnom poglavlju smo opisali kako je potrebno ulaznu datoteku razbiti na niz tokena. Jednom kada imamo takav niz moguće je vršiti daljnju obradu i izvlačenje karakterističnih podataka za takav niz tokena. Jedan od parametara je i frekvencija pojavljivanja pojedinog tokena u nizu. Na osnovu toga možemo zaključiti koji su tokeni (u našem slučaju riječi) najzastupljeniji za pojedini skup dokumenta odnosno za pojedinu klasu. Visoka frekvencija pojavljivanja će označavati funkcijske riječi koje se često koriste, dok niska frekvencija pojavljivanja označava specifične izraze. Za osrednje pojavljivanje ili ono pojavljivanje koje je veće od n možemo reći da se radi o nekim uobičajenim izrazima

za pojedinu klasu odnosno dokument. U specifičnim tekstovima postojat će specifična raspodjela uobičajenih izraza kao i nisko frekventnih izraza.

Za potrebe izvlačenja parametra frekvencije tokena napisane su dvije funkcije u matlabu pod nazivom, **fp.m** i **fpStruct.m**. Kao što je vidljivo iz naziva, jedna funkcija služi za direktno pohranjivanje frekvencije pojavljivanja pojedinih tokena u određenu datoteku, dok druga služi za učitavanje tih podataka u strukturu matlaba i daljnju obradu. Bitno je reći kako funkcije ne vrše sortiranje tokena po frekvenciji pojavljivanja, već samo vrše ispis frekvencije pojavljivanja pojedinog tokena.

Pozivi:

```
fp('Staza do ulazne datoteke ','Staza do izlazne datoteke')
```

```
FpTokensStruct=fpStruct('Staza do ulazne datoteke')
```

Kao primjer ispisa na slici (Slika 13) prikazane su sortirane strukture frekvencije pojavljivanja tokena u svih pet klasa (sreća, tuga, strah, ljutnja i neutralno). Za funkcije sortiranja može se pogledati matlab skripta **auto_prim_model2.m**

Tablica (Tablica 8) prikazuje ukupni broj tokena po pojedinoj klasi odnosno emociji. Bitno je reći kako ovaj broj predstavlja isključivo broj riječi i neartikuliranih zvukova, pošto su delimiteri izbačeni kao nerelevantni tokeni.

Tablica 8-Broj tokena po emociji

	Sreća	Tuga	Strah	Ljutnja	Neutralno
Broj tokena	672	1205	319	2552	1745
Ukupno	6493				

fp_sreca_cell =		fp_tuga_cell =		fp_strah_cell =		fp_ljutnja_cell =		fp_neutralno_cell =	
'JE'	[20]	'I'	[47]	'JE'	[8]	'TI'	[80]	'JE'	[58]
'HAHA_'	[20]	'DA'	[39]	'I'	[8]	'DA'	[66]	'I'	[57]
'SE'	[16]	'JE'	[30]	'DA'	[7]	'JE'	[61]	'U'	[50]
'I'	[15]	'NE'	[25]	'SE'	[7]	'JA'	[60]	'DA'	[41]
'A'	[13]	'TE'	[21]	'TI'	[6]	'NE'	[58]	'SE'	[39]
'DA'	[12]	'UDH_'	[21]	'TO'	[6]	'SE'	[57]	'UDH_'	[37]
'UDH_'	[11]	'SE'	[18]	'UDH_'	[6]	'I'	[57]	'NA'	[36]
'TI'	[10]	'CMIZDR_'	[18]	'TAMO'	[5]	'U'	[41]	'ZASTAJK_'	[31]
'NA'	[8]	'NA'	[17]	'SMO'	[5]	'MI'	[33]	'TO'	[29]
'MI'	[8]	'U'	[15]	'CXU'	[5]	'SI'	[33]	'NE'	[25]
'O'	[8]	'JA'	[14]	'IZH'	[5]	'SXTA'	[33]	'SU'	[24]
'TO'	[7]	'SAM'	[13]	'TUP'	[4]	'SAM'	[33]	'S'	[16]
'HEHE_'	[7]	'MI'	[12]	'TAP'	[4]	'A'	[30]	'SXTO'	[12]
'U'	[6]	'IDH_'	[12]	'MENE'	[4]	'ME'	[25]	'JA'	[11]
'HIHI_'	[6]	'TEBE'	[12]	'HOCXU'	[4]	'TE'	[25]	'ZA'	[11]
'JA'	[5]	'II'	[11]	'MAJA'	[4]	'TO'	[24]	'CXE'	[11]
'KAKO'	[5]	'BILO'	[10]	'SI'	[4]	'PA'	[23]	'A'	[10]
'SI'	[5]	'SVI'	[10]	'A'	[3]	'NA'	[23]	'BITI'	[10]
'HOHO_'	[5]	'NIJE'	[10]	'JA'	[3]	'S'	[19]	'NIJE'	[9]
'CXE'	[5]	'TO'	[10]	'EVO'	[3]	'CXU'	[16]	'KAD'	[9]
'SMO'	[5]	'SMRC_'	[10]	'SAMO'	[3]	'SVE'	[16]	'KOJI'	[9]
'NE'	[5]	'SXTO'	[9]	'DOBRO'	[3]	'ZA'	[16]	'DVIJE'	[9]
'SVE'	[4]	'SI'	[9]	'SU'	[3]	'MA'	[15]	'ILI'	[8]
'S'	[4]	'TAKO'	[9]	'STRASXNO'	[3]	'KO'	[15]	'ZXENSTVENOST'	[8]
'TE'	[4]	'KAD'	[8]	'AKO'	[2]	'MENI'	[15]	'PA'	[7]
'ME'	[4]	'ME'	[8]	'SAD'	[2]	'SXTO'	[14]	'NACYIN'	[7]
'PA'	[4]	'OD'	[8]	'BI'	[2]	'KAD'	[14]	'KOJA'	[7]
'MA'	[3]	'MAMA'	[8]	'TE'	[2]	'SAD'	[13]	'SVE'	[6]
'EJ_'	[3]	'SAD'	[7]	'SXTA'	[2]	'KURAC'	[13]	'DO'	[6]

Slika 13-Primjer frekvencije pojavljivanja pojedinih tokena za svaku emociju

5.5 Relativna frekvencija pojavljivanja tokena

Pored frekvencije pojavljivanja pojedinog tokena unutar klase, bitan podatak je i relativna frekvencija pojavljivanja pojedinog tokena. Tu se zapravo radi o vjerojatnosti pojavljivanja pojedinog tokena u klasi, pošto relativna frekvencija predstavlja frekvenciju pojavljivanja nekog tokena u odnosu na druge tokene u toj klasi.

$$RF = \frac{token_freq}{\sum token_freq} \quad (5.1.)$$

Za potrebe analize relativne frekvencije tokena u klasama, po uzoru na dosadašnje primjere, razvijene su dvije funkcije u matlabu. Jedna za direktni zapis rezultata u određenu datoteku, te druga za pohranjivanje rezultata u matlab strukturu. Funkcija za zapis rezultata u određenu datoteku nosi naziv **rfp.m**, a funkcija za zapis u strukturu **rfpStruct.m**. Primjer rezultata za nekoliko tokena sa najvećom vjerojatnošću za svaku klasu odnosno emociju dan je slici (Slika 14).

Pozivi:

```
rfp('Staza do ulazne datoteke ', 'Staza do izlazne datoteke')  
  
RfpTokensStruct=rfpStruct('Staza do ulazne datoteke')
```


rfp_sreca =	rfp_tuga =	rfp_strah =	rfp_ljutnja =	rfp_neutralno =
JE: 0.0298	I: 0.0390	JE: 0.0251	TI: 0.0313	JE: 0.0332
HAHA_: 0.0298	DA: 0.0324	I: 0.0251	DA: 0.0259	I: 0.0327
SE: 0.0238	JE: 0.0249	DA: 0.0219	JE: 0.0239	U: 0.0287
I: 0.0223	NE: 0.0207	SE: 0.0219	JA: 0.0235	DA: 0.0235
A: 0.0193	TE: 0.0174	TI: 0.0188	NE: 0.0227	SE: 0.0223
DA: 0.0179	UDH_: 0.0174	TO: 0.0188	SE: 0.0223	UDH_: 0.0212
UDH_: 0.0164	SE: 0.0149	UDH_: 0.0188	I: 0.0223	NA: 0.0206
TI: 0.0149	CMIZDR_: 0.0149	TAMO: 0.0157	U: 0.0161	ZASTAJK_: 0.0178
NA: 0.0119	NA: 0.0141	SMO: 0.0157	MI: 0.0129	TO: 0.0166
MI: 0.0119	U: 0.0124	CXU: 0.0157	SI: 0.0129	NE: 0.0143
O: 0.0119	JA: 0.0116	IZH: 0.0157	SXTA: 0.0129	SU: 0.0138
TO: 0.0104	SAM: 0.0108	TUP: 0.0125	SAM: 0.0129	S: 0.0092
HEHE_: 0.0104	MI: 0.0100	TAP: 0.0125	A: 0.0118	SXTO: 0.0069
U: 0.0089	IDH_: 0.0100	MENE: 0.0125	ME: 0.0098	JA: 0.0063
HIHI_: 0.0089	TEBE: 0.0100	HOCXU: 0.0125	TE: 0.0098	ZA: 0.0063
JA: 0.0074	TI: 0.0091	MAJA: 0.0125	TO: 0.0094	CXE: 0.0063
KAKO: 0.0074	BILO: 0.0083	SI: 0.0125	PA: 0.0090	A: 0.0057
SI: 0.0074	SVI: 0.0083	A: 0.0094	NA: 0.0090	BITI: 0.0057
HOHO_: 0.0074	NIJE: 0.0083	JA: 0.0094	S: 0.0074	NIJE: 0.0052
CXE: 0.0074	TO: 0.0083	EVO: 0.0094	CXU: 0.0063	KAD: 0.0052
SMO: 0.0074	SMRC_: 0.0083	SAMO: 0.0094	SVE: 0.0063	KOJI: 0.0052
NE: 0.0074	SXTO: 0.0075	DOBRO: 0.0094	ZA: 0.0063	DVIJE: 0.0052
SVE: 0.0060	SI: 0.0075	SU: 0.0094	MA: 0.0059	ILI: 0.0046
S: 0.0060	TAKO: 0.0075	STRASXNO: 0.0094	KO: 0.0059	ZXENSTIVENOST: 0.0046
TE: 0.0060	KAD: 0.0066	AKO: 0.0063	MENI: 0.0059	PA: 0.0040
ME: 0.0060	ME: 0.0066	SAD: 0.0063	SXTO: 0.0055	NACYIN: 0.0040
PA: 0.0060	OD: 0.0066	BI: 0.0063	KAD: 0.0055	KOJA: 0.0040
MA: 0.0045	MAMA: 0.0066	TE: 0.0063	SAD: 0.0051	SVE: 0.0034
EJ_: 0.0045	SAD: 0.0058	SXTA: 0.0063	KOJA : 0.0051	DO: 0.0034
PO: 0.0045	SXTA: 0.0058	ON: 0.0063	ONDA: 0.0051	KAO: 0.0034

Slika 14-Prikaz dijela tokena i njihovih relativnih frekvencija pojavljivanja za svaku emociju

5.6 Type-Token ratio

Type-Token ratio (TTR) je mjera varijacije vokabulara pisanog ili izgovorenog teksta. Radi se zapravo o omjeru između različitih tokena i ukupnog broja tokena. Ovim parametrom često možemo pratiti o kakvoj vrsti analiziranog teksta se radi. Naime, poznato je kako vokabular standardnog svakodnevnog govora ima značajno manju varijaciju nego li recimo pisani tekst. Sami Type-Token ratio možemo objasniti kroz sljedeći primjer.

$$\text{Type – Token Ratio} = \left(\frac{\text{broj različitih tokena}}{\text{ukupno tokena}} \right) \times 100 \quad (5.2.)$$

Tablica 9-Tekst 1

Tekst 1	JE LI OVO FER ? JE NE, PITAM VAS JE LI OVO FER ? KAŽEM DA JE FER BITI FER NA FERU.
----------------	---

Tablica 10-Tekst 2

Tekst 2	MOJ DIPLOMSKI RAD JE NAPRAVLJEN NA ZAVODU ZESOI KAO I OSTALI DIPLOMSKI RADOVI I ON ĆE IMATI KORICE TVRDOG UVEZA.
----------------	---

Tablica 11-TTR za tekst 1 i tekst 2

	Tekst 1	Tekst 2
Različitih tokena	12	18
Ukupno tokena	20	20
TTR	60 %	90 %

Vidljivo je kako je broj tokena u oba teksta jednak. Usprkos tome Type-Token ratio je znatno drugačiji u ova dva teksta (Tablica 11). Pri čemu je u tekstu 1 (Tablica 9), TTR 60 %, a u tekstu 2 (Tablica 10), TTR 90 %. Ako možemo primijetiti da je tekst 1 primjer transkripta jednog razgovora, tada je bilo i za očekivati kako će kod takvog teksta varijacija vokabulara biti manja nego kod pisanog teksta. Ovime zaključujemo kako visoki TTR ukazuje na veliku količinu leksičkih varijacija.

Glavna razlika između govornog transkripta i pisanog teksta leži u činjenici da se konverzacija odnosno govorni transkript odvija u realnom vremenu, što znači da postoji ograničeno vrijeme u kojem govornik može smisliti ono što želi reći. Pri čemu govornik obično raspolaže sa ograničenim vokabularom s obzirom na vrijeme dohvata željenih riječi i same fluidnosti govora. Nasuprot tome pri pisanju teksta imamo znatno veće mogućnosti u pogledu vremena, da izaberemo pravilan vokabular koji će najbolje opisati ono što želimo prenijeti u pisani zapis.

U svrhu računanja TTR faktora na osnovu liste tokena u Matlabu je razvijena funkcija `ttr.m`. Ulaz u funkciju je lista tokena izgrađena pomoću funkcije `tokenizer.m`.

Pozivi:

```
TTR_faktor=ttr('Staza do ulazne datoteke tokena\*.lst')
```

5.6.1 Statistika TTR-a primijenjenog na KEG

Kako bismo vidjeli karakteristike KEG-a po pitanju raznovrsnosti korištenog vokabulara provedena je TTR analiza. Računali smo TTR za cijeli korpus i za pojedine klase. Svi podatci su vidljivi u dolje navedenim tablicama. Zanimljivo je povući paralelu između TTR-a po pojedinim klasama i broja tokena po klasi. Naime, TTR je najveći za klasu koja ima najmanje tokena (Strah), a najmanji za klasu koja ima najveći broj tokena (Ljutnja). Iz čega eventualno možemo zaključiti kako povećanjem broja tokena dolazi do zasićenja vokabulara u pogledu pojavljivanja sve većeg broja već korištenih riječi.

Tablica 12-Type-Token Ratio za cjelokupni KEG

	Cijeli korpus
TTR	31.34 %

Tablica 13-Type Token Ratio za pojedine emocije

	Sreća	Tuga	Strah	Ljutnja	Neutralno
TTR	57.89 %	44.15 %	62.70 %	37.69 %	50.37 %

Tablica 14-Type Token Ratio za stvarne i glumljene emocije

	„Stvarne emocije“	„Glumljene emocije“
TTR	38.43 %	40.32 %

5.7 Skup jedinstvenih tokena

Već samo po sebi, intuitivno, postavlja se pitanje koji od tokena će biti karakteristični za pojedinu klasu. Sasvim je jasno da to ne mogu biti svi tokeni. Empirijski se nameće da bi to trebali biti oni tokeni koji se u pojedinoj klasi najviše puta pojavljuju. Međutim, podatak o frekvenciji pojavljivanja u pojedinoj klasi nam i dalje ne govori ništa o tome koliko je taj token karakterističan za tu klasu, u odnosu na ostale, pošto je moguće da se isti token pojavi i u nekim od ostalih klasa. Kod gradnje primitivnog sustava ide se sa idejom da se svi tokeni koji nisu jedinstveni za pojedinu klasu izbace. Dizajnirani sustav provjerava svaki token u klasi nad kojom se vrši obrada, sa svim tokenima u svim ostalim klasama. Ako nije pronađeno pojavljivanje dotičnog tokena u nekoj od drugih klasa tada se on proglašava jedinstvenim za tu klasu, u suprotnom se briše sa liste tokena.

Funkcija u Matlabu razvijena za tu svrhu nosi naziv **unique1.m**. Ulazi u funkciju su staze do svih klasa koje se koriste u sustavu, a izlazi se nalaze na destinacijama ulaznih klasa te nose nazive prema sljedećoj sintaksi: **ime_ulazne_datoteke-unique-set.lst**

Pozivi:

```
unique1( 'Staza do 1. klase tokena\*.lst',...
         'Staza do 2. klase tokena\*.lst',...
         'Staza do N-te. klase tokena\*.lst')
```

Princip rada funkcije možemo vidjeti na primjeru prikazanom tablicom (Tablica 15) i slikom (Slika 15), gdje su vidljivi ulazni tokeni od tri različite klase i selekcija jedinstvenih tokena po svakoj klasi.

Tablica 15-Prikaz tokena prije primjene funkcije unique1

	Klasa 1	Klasa 2	Klasa 3
Tokeni	• Mercedes • BMW • Audi • Volkswagen • Opel • Honda	• Volvo • Toyota • Honda • Peugeot	• Porsche • Ferrari • Honda • Toyota

```
>> razredi.tokens1    >> razredi.tokens2    >> razredi.tokens3
ans =
    'mercedes'
    'bmw'
    'audi'
    'volkswagen'
    'opel'

ans =
    'volvo'
    'peugeot'

ans =
    'porsche'
    'ferrari'
```

Slika 15-Filtracija jedinstvenih tokena funkcijom unique1

5.7.1 Skup jedinstvenih tokena sa njihovim pripadajućim relativnim frekvencijama

Skup klasa sa jedinstvenim tokenima ne može se upotrijebiti za gradnju primitivnog sustava prepoznavanja, bez da ti jedinstveni tokeni sadrže svoje relativne frekvencije. Oni zapravo zadržavaju vjerojatnost koju su imali prije u svojoj klasi, dok nije došlo do izbacivanja tokena koji nisu jedinstveni. Takav model jedinstvenih tokena sa dodatnim zapisom relativne frekvencije je realiziran funkcijom `rfp_unique.m`

Pozivi:

```
rfp_unique(      'Staza do 1. klase tokena\*.1st',...
                 'Staza do 2. klase tokena\*.1st',...
                 'Staza do N-te. klase tokena\*.1st')
```

Primjer rada funkcije pokazati ćemo na gore navedenim razredima automobila. Nakon tog jednostavnog primjera, gdje doslovno možemo ručno izvući relativne frekvencije i usporediti rezultate, biti će dan primjer rada funkcije na stvarnim klasama s kojima radimo (sreća, tuga, strah, ljutnja, neutralno).

<code>>> razredi.tokens1</code>	<code>>> razredi.tokens2</code>	<code>>> razredi.tokens3</code>
<code>ans =</code>	<code>ans =</code>	<code>ans =</code>
mercedes: 0.1667	volvo: 0.2500	porsche: 0.2500
bmw: 0.1667	peugeot: 0.2500	ferrari: 0.2500
audi: 0.1667		
volkswagen: 0.1667		
opel: 0.1667		

Slika 16-Filtracija jedinstvenih tokena i izračun relativnih frekvencija funkcijom `rfp_unique`

sreća_cell =		tuga_cell =		strah_cell =		ljutnja_cell =		neutralno_cell =	
'haha_'	[0.0298]	'cmizdr_'	[0.0149]	'tup'	[0.0125]	'kucac'	[0.0051]	'xzenstvenost'	[0.0046]
'hehe_'	[0.0104]	'uz'	[0.0058]	'tap'	[0.0125]	'macer'	[0.0027]	'tisucxe'	[0.0034]
'hihi_'	[0.0089]	'bude'	[0.0041]	'maja'	[0.0125]	'macerinu'	[0.0024]	'neki'	[0.0034]
'hoho_'	[0.0074]	'volimo'	[0.0033]	'sss_'	[0.0063]	'jem'	[0.0024]	'recimo'	[0.0029]
'dite'	[0.0045]	'antu'	[0.0033]	'slijedecxi'	[0.0063]	'rebo'	[0.0024]	'zxene'	[0.0029]
'bok'	[0.0045]	'svojim'	[0.0025]	'neckete'	[0.0063]	'mcam'	[0.0024]	'dio'	[0.0023]
'len'	[0.0045]	'znaju'	[0.0025]	'odustanesx'	[0.0031]	'va'	[0.0024]	'dakle'	[0.0023]
'ludilo'	[0.0030]	'licu'	[0.0025]	'pomislit'	[0.0031]	'jea'	[0.0024]	'kada'	[0.0023]
'vrag'	[0.0030]	'budesx'	[0.0025]	'oteli'	[0.0031]	'ic'	[0.0020]	'godine'	[0.0023]
'teoriju'	[0.0030]	'jaka'	[0.0025]	'ubili'	[0.0031]	'pyku'	[0.0020]	'situacija'	[0.0017]
'kazxe'	[0.0030]	'losxe'	[0.0025]	'sranje'	[0.0031]	'je nom'	[0.0020]	'posto'	[0.0017]
'sisu'	[0.0030]	'odu'	[0.0017]	'stanju'	[0.0031]	'je pte'	[0.0020]	'odnosu'	[0.0017]
'kakvi'	[0.0030]	'zxivog'	[0.0017]	'napraviti'	[0.0031]	'volsx'	[0.0020]	'mozxemo'	[0.0017]
'kucxicu'	[0.0030]	'mrtvog'	[0.0017]	'miro'	[0.0031]	'zivot'	[0.0016]	'kuna'	[0.0017]
'udxite'	[0.0030]	'trebala'	[0.0017]	'minirano'	[0.0031]	'realsx'	[0.0016]	'pitanje'	[0.0017]
'pogostiti'	[0.0030]	'oprosti'	[0.0017]	'niste'	[0.0031]	'd'	[0.0016]	'period'	[0.0017]
'dovidxenja'	[0.0030]	'gazda'	[0.0017]	'trebali'	[0.0031]	'slexaj'	[0.0016]	'strane'	[0.0017]
'trkom'	[0.0015]	'ovome'	[0.0017]	'dolaziti'	[0.0031]	'sje'	[0.0016]	'trgovacykih'	[0.0017]
'ispricyati'	[0.0015]	'ponosne'	[0.0017]	'ppp_'	[0.0031]	'oc'	[0.0016]	'centara'	[0.0017]
'brezguljka'	[0.0015]	'sestra'	[0.0017]	'poslao'	[0.0031]	'ob'	[0.0016]	'koliko'	[0.0017]
'pocxi'	[0.0015]	'tvom'	[0.0017]	'dogovorili'	[0.0031]	'baco'	[0.0016]	'obzirom'	[0.0017]
'glasno'	[0.0015]	'zxelimo'	[0.0017]	'neam'	[0.0031]	'pyka'	[0.0012]	'godini'	[0.0017]
'dozivati'	[0.0015]	'misli'	[0.0017]	'broza'	[0.0031]	'macerina'	[0.0012]	've'	[0.0017]
'boljega'	[0.0015]	'stranicu'	[0.0017]	'tita'	[0.0031]	't'	[0.0012]	'danas'	[0.0017]
'komunikacije'	[0.0015]	'lovricxa'	[0.0017]	'izlazu'	[0.0031]	'gorim'	[0.0012]	'moraju'	[0.0017]
'vanjskim'	[0.0015]	'najvisxe'	[0.0017]	'groblja'	[0.0031]	'nomalan'	[0.0012]	'bile'	[0.0017]
'svijetom'	[0.0015]	'vazda'	[0.0017]	'izronio'	[0.0031]	'jano'	[0.0012]	'njena'	[0.0017]
'crvene'	[0.0015]	'dijete'	[0.0017]	'magle'	[0.0031]	'rim'	[0.0012]	'njegovom'	[0.0011]
'daljnim'	[0.0015]	'pobijediti'	[0.0017]	'cyuli'	[0.0031]	'u'	[0.0012]	'mjestu'	[0.0011]
'big'	[0.0015]	'nicyega'	[0.0017]	'korake'	[0.0031]	'je'	[0.0012]	'izvlacyi'	[0.0011]

Cenzurirano

Slika 17-Prikaz dijela filtracije jedinstvenih tokena i izračun relativnih frekvencija za svih 5 emocija

Tablica 16-Broj tokena po emociji nakon primjene funkcije `rfp_unique`

	Sreća	Tuga	Strah	Ljutnja	Neutralno	UKUPNO
Inicijalni broj tokena	672	1205	319	2552	1745	6493
Broj tokena nakon izbacivanja	193	260	73	619	614	1759

U tablici (Tablica 16) je vidljiva raspodjela tokena i jedinstvenih tokena po pojedinoj emociji. Primjećujemo da je broj jedinstvenih tokena značajno manji od inicijalnog broja tokena, što je posljedica izbacivanja svih onih tokena koji nisu jedinstveni za pojedinu klasu već su se pojavljivali i u drugim klasama.

5.8 Princip rada primitivnog sustava temeljenog na unigram modelu

Princip rada primitivnog sustava zasniva se na izračunu pogreške odstupanja od pojedinog modela. U sustav se ubacuje nepoznata rečenica te se računa pogreška odstupanja te rečenice za naših pet klasa odnosno modela (sreća, tuga, strah, ljutnja i neutralno). Onaj model koji ima najmanje odstupanje odnosno pogrešku od dotične rečenice proglašava se pobjednikom, tj. dotična rečenica pripada toj klasi. Sami izračun pogreške se odvija na način da se na početku proglasi maksimalna moguća pogreška za sve modele, tj. pogreška je onolika koliko rečenica ima tokena. Nakon toga slijedi umanjivanje pogreške za svaki pojedini model. Pogreška se umanjuje za onoliko koliko je relativna frekvencija pojavljivanja pojedinog tokena u određenoj klasi. Kada sustav izračuna pogrešku za svaku pojedinu klasu, pronade se najmanja i proglasi se da rečenica pripada tom modelu.

U svrhu realizacije gore opisanog postupka razvijena je funkcija `compereWithModels.m`. Funkcija ispisuje pogrešku za svaki pojedini model kao i ime modela, te sve tokene koji su pronađeni u pojedinome modelu.

Pozivi:

```
Errors=compereWithModels('Staza do modela\*unique-fset.lst',...
                          'Staza do unknown.txt datoteke',...
                          zastavica_za_prikaz)
```


5.8.1 Sučelje za demonstraciju rada primitivnog unigram sustava

Kako bismo mogli provjeriti rad sustava i bez nespretnog upisivanja rečenica po točno definiranim datotekama i poziva funkcija sa točno specifičnim parametrima, razvijeno je sučelje pomoću kojeg se može na intuitivan način unijeti željena rečenica u sustav i provjeriti estimirano emocionalno stanje. U sučelje se unese željena rečenica nad kojom se želi napraviti prepoznavanje i pritisne se tipka „PREPOZNAJ“. Kao povratnu informaciju dobit ćemo estimirano emocionalno stanje.



Slika 18-Sučelje za demonstraciju rada primitivnog unigram sustava

5.9 Principi redukcije dimenzionalnosti

Implementacijom primitivnog unigram sustava mogli smo uočiti kako broj tokena po pojedinoj klasi rapidno raste sa porastom broja analiziranih rečenica. U primjerima gdje primjena sustava obuhvaća veliki broj klasa i veliki broj dokumenta za treniranje sustava, takva metoda nije pogodna upravo zbog dimenzionalnosti podataka koje je potrebno analizirati. Cilj bi bio reducirati dimenzionalnost modela i maksimizirati performanse. Umjesto bilježenja svakog pojedinog tokena kod kompleksnijih sustava treba pristupiti stematizaciji, lematizaciji i mapiranju na principu leksikona sinonima tzv.

mreža riječi (engl. *WordNet*). Bilo koji od dolje spomenutih postupaka ima zadatak identificirati one izraze koji ne pridonose informaciji značajnoj za klasifikaciju uzorka u pojedinu klasu. Sukladno tome definiraju se značajni izrazi i eliminira se informacijski šum.

5.9.1 Frekvencija dokumenata

U standardnim primjenama treniranje sustava se odvija sa velikim brojem dokumenata koji pripadaju određenim klasama. Značajka koju bismo mogli iskoristiti pri redukciji dimenzionalnosti je frekvencija dokumenata DF (engl. *Document Frequency*). DF predstavlja broj dokumenta u kojima se pojavljuje izraz X . Iskoristivost ove značajke se zasniva na postupku izračuna DF svojstva za svaki izraz X u skupu podataka za trening DC_{tr} (engl. *Document training*). Pretpostavka je da rijetki izrazi ne nose informaciju pri predikciji kategorije u koju bi neklasificirani dokument bio smješten. Sukladno tome postavlja se odgovarajući prag na osnovu kojeg se odlučuje da li je izraz ima informativno značenje u predikciji ili ne.

5.9.2 Hi kvadrat (χ^2) test

Provođenje χ^2 testa u slučaju smanjenja dimenzionalnosti leksičkog modela najbolje možemo shvatiti kroz primjer (Ćavar, 2010).

Promotrimo sljedeći primjer:

Neka je $|DC|$ broj dokumenata, a c_i i $\neg c_i$ broj dokumenata označenih da pripadaju odnosno ne pripadaju klasi c_i . Oznake t_i i $\neg t_i$ predstavljaju dokumente koji sadrže odnosno ne sadrže izraz t_i .

$$|DC| = 801948 \quad (5.3.)$$

Tablica 17-Prikaz odnosa klase c_i i izraza t_i kao priprema za izračun hi kvadrat testa

	c_i	$\neg c_i$	Ukupno
t_i	49	27652	27701
$\neg t_i$	141	774106	774247
Ukupno	190	801758	801948

Hi kvadrat testom želimo istražiti hipotezu da su izraz t_i i klasa c_i međusobno zavisni.

$$P(t_j c_i) \neq P(t_j)P(c_i) \quad (5.4.)$$

Nulta hipoteza nam kaže da su izraz t_i i klasa c_i međusobno nezavisni.

$$P(t_j c_i) = P(t_j)P(c_i) \quad (5.5.)$$

Očekivane vrijednosti (nulta hipoteza) odnosa klase c_i i izraza t_i dane su u tablici (Tablica 18), te su prikazane podebljanjima fontom ispod promatranih vrijednosti. Vrijednosti su izračunate po sljedećem izrazu.

$$P(c) \cdot P(t) = \frac{Redak_ukupno \times Stupac_ukupno}{Ukupno} \quad (5.6.)$$

Tablica 18-Prikaz rezultata očekivanja odnosno nulte hipoteze

	c_i	$\neg c_i$	Ukupno
t_i	49 6.56	27652 27694.44	27701
$\neg t_i$	141 183.44	774106 774063.56	774247
Ukupno	190	801758	801948

Sami hi kvadrat se računa po dobro poznatoj formuli koja u odnos stavlja promatrane i očekivane rezultate.

$$\chi^2 = \sum \frac{(promatrano - očekivano)^2}{očekivano} \quad (5.7.)$$

$$\chi^2 = \frac{(49 - 6.56)^2}{6.56} + \frac{(27652 - 27694.44)^2}{27694.44} + \frac{(141 - 183.44)^2}{183.44} + \frac{(774106 - 774063.56)^2}{774063.56} = 284.45 \quad (5.8.)$$

Stupnjeve slobode odredimo prema izrazu:

$$stupnjeva_slobode = (redaka - 1) \times (stupaca - 1) \quad (5.9.)$$

Slijedi očitavanje iz tablice te određivanje da li je testirani izraz (token) ovisan o klasi u odnosu na koju se testirao. Ovakav postupak, odnosno χ^2 test potrebno je primijeniti na sve tokene i sve klase. Ovime se želi eliminirati sve tokene za koje se na testu utvrdi da nisu ovisni o pojedinoj klasi.

Problematični dio χ^2 testa je povećanje pogreške sustava. Za primjer možemo uzeti 1000 izbacivanja nejedinstvenih riječi za pojedinu klasu, koje recimo imaju pogrešku od $e = 0.05$. Realizacija takvog testiranja bi dovela do prosječno 50 krivih odluka. Stoga je χ^2 test potrebno promatrati kao relativnu značajku pri izbacivanju tokena iz sustava klase.

5.9.3 Značajka zajedničke informacije

Pored gore navedenih metoda izbacivanja nerelevantnih tokena za pojedinu klasu, zanimljiva značajka je i značajka zajedničke informacije. Značajka zajedničke informacije nam govori koliko informacija token t daje o klasi c . Uspoređujući izraz za značajku zajedničke informacije naspram χ^2 testa, možemo reći kako se radi o logaritmu odnosa između promatranih i očekivanih veličina.

$$I(t, c) = \log \frac{P(tc)}{P(t)P(c)} \quad (5.10.)$$

Tablica 19-Prikaz aproksimacije zajedničke informacije

	c_i	$\neg c_i$	Ukupno
t_i	A 49	B 27652	27701
$\neg t_i$	C 141	D 774106	774247
Ukupno	190	801758	N 801948

Aproksimacija značajke zajedničke informacije može se napraviti prema sljedećem izrazu sukladno (Yang, 1997).

$$I(t, c) \approx \frac{A \times N}{(A + C) \times (A + B)} \quad (5.11.)$$

$$I(U; C) = \sum_{e_t \in \{1,0\}} \sum_{e_c \in \{1,0\}} P(U = e_t, C = e_c) \log_2 \frac{P(U = e_t, C = e_c)}{P(U = e_t)P(C = e_c)} \quad (5.12.)$$

Primjere izračuna polja možemo pogledati u nastavku.

$$P(1,1) = \frac{49}{801948} \quad (5.13.)$$

$$P(U = 1) = \frac{27701}{801948} \quad (5.14.)$$

$$P(C = 1) = \frac{190}{801948} \quad (5.15.)$$

5.10 Naivni Bayesov tekstualni klasifikator

Uvjetna Bayesova vjerojatnost primijenjena na sustav tekstualnog klasifikatora prikazana je izrazom (5.16.), pri čemu je d oznaka dokumenta, c oznaka klase, a $P(t_k|c)$ uvjetna vjerojatnost pojave izraza t_k u klasi c . Oznaka n_d predstavlja broj tokena po dokumentu.

$$P(c|d) \propto P(c) \cdot \prod_{1 \leq k \leq n_d} P(t_k|c) \quad (5.16.)$$

Cilj je pronaći najbolju moguću klasu, odnosno maksimum aposteriori klasu c_{map} , pri čemu je $\hat{P}$ estimirana vjerojatnost pri korištenju korpusa za treniranje. Iz izraza je vidljivo kako možemo imati već broj decimalnih množenja što u konačnici može dovesti do preljeva. Kako bismo spriječili tu pojavu u izraz (5.17.) uvodimo logaritamske vrijednosti prijašnjeg izraza. Ovime imamo logaritamsku vjerojatnost, a pošto je logaritamska funkcija monotona, tada je najveća vjerojatnost ona koja ima najveći logaritam.

$$c_{map} = \arg \max_{c \in C} \hat{P}(c|d) = \arg \max_{c \in C} \hat{P}(c) \prod_{1 \leq k \leq n_d} \hat{P}(t_k|c) \quad (5.17.)$$

Iz analitičkog izraza (5.18.) je vidljivo kako je logaritmiranjem riješen problem decimalnih mjesta i preljeva, pošto je množenje zamijenjeno zbrajanjem.

$$c_{map} = \arg \max_{c \in C} \left[\log \hat{P}(c) + \sum_{1 \leq k \leq n_d} \log \hat{P}(t_k|c) \right] \quad (5.18.)$$

Kako ne raspolažemo stvarnim vrijednostima vjerojatnosti, tada se radi estimacija pojedinih vrijednosti. Estimacija vjerojatnosti $\hat{P}(c)$ se procjenjuje prema izrazu (5.19.), gdje je N_c broj dokumenata koji pripadaju klasi C u trening korpusu, a N broj svih dokumenata.

$$\hat{P}(c) = \frac{N_c}{N} \quad (5.19.)$$

Procjena uvjetne vjerojatnosti $\hat{P}(t|c)$ je definirana izrazom (5.20.). T_{ct} je frekvencija tokena t u skupu dokumenata koji pripadaju klasi C (T'_{ct} —svi tokeni t).

$$\hat{P}(t|c) = \frac{T_{ct}}{\sum_{t' \in V} T'_{ct}} \quad (5.20.)$$

Kako bi se izbjegle operacije sa nultim vrijednostima vrši se prilagodba izraza te se dodaje jedinica u brojniku i nazivniku.

$$\hat{P}(t|c) = \frac{T_{ct} + 1}{\sum_{t' \in V} (T'_{ct} + 1)} \quad (5.21.)$$

5.10.1 Implementacija naivnog Bayesovog tekstualnog klasifikatora

Prema gore opisanom modelu izvršena je implementacija Bayesovog tekstualnog klasifikatora u matlabu. U svrhu njegove implementacije razvijene su sljedeće funkcije:

- **KEGto_all_TXT.m**
- **TXTtoModels.m**
- **Naive_bayes_text_classifier.m**

Funkcije **KEGto_all_TXT.m** i **TXTtoModels.m** služe za prilagodbu transkripta KEG-a traženom modelu i strukturi podataka. Tako recimo, funkcija **KEGto_all_TXT.m** razbija transkript iz pojedinih klasa u pojedinačne tekstualne datotke sa nazivom klase, npr. „*ljutnjaX.txt*“, gdje je $X \in \{1, 2, 3, \dots, n\}$, a n predstavlja broj snimki za pojedinu klasu. Tako predstavljeni podatci se zatim učitavaju pomoću funkcije **TXTtoModels.m** i formiraju se modeli klasa sa ekstenzijom „*.dat“.

Dobiveni modeli su spremni za upotrebu u sustavu prepoznavanja koji se pokreće funkcijom **Naive_bayes_text_classifier.m**, a čiji poziv i paramtri poziva su prikazani u nastavku.

Pozivi:

```
[rezultati, imena_Modela]=naive_bayes_text_classifier  
  
('Staza do *.dat modela',...  
'Staza do unknown.txt datoteke',...  
zastavica_za_prikaz)
```

6 Povezivanje sa HTK alatom

Ideja razvoja sustava je bila da se na osnovu lingvističkih parametara glasa radi automatska klasifikacija diskretnih emocionalnih stanja. Lingvističku analizu smo opisali u dva sustava razvijena u prethodnim poglavljima. Ono što je preostalo je da iz govora dobijemo tekstualni zapis, nad kojim ćemo moći napraviti daljnju, gore spomenutu, lingvističku obradu. Već iz same ideje nazire se složenost takvog sustava. U okviru ovog diplomskog rada nemoguće je razviti jedan takav sustav od nule, stoga smo pribjegli već korištenom alatu HTK (engl. *Hidden Markov Model Toolkit*). Ovdje se neće ulaziti u detalje rada HTK alata, niti će se opisivati cjelokupna gradnja sustava, već će se opisati proces prilagodbe i pripreme KEG-a i lingvističkih modela. Za sve dodatne informacije od radu i podešavanju HTK alata može se pogledati u (Dropuljić, 2008.).

Osnovno što moramo znati o HTK alatu je da radi na principu skrivenih Markovljevih modela, a sami proces se može podijeliti na trening i prepoznavanje. Koraci kroz koje moramo proći kako bismo imali funkcionalni sustav prepoznavanja su:

- **Priprema podataka**—se sastoji od snimanja materijala za trening i test te pravljenja njihovih transkripcija. Definira se vokabular koji se koristi u testu te se izrađuje rječnik transkripcija. Kako mi već raspoložemo korpusom emocionalnog govora, tada je u pripremi potrebno izvršiti prilagodbu podataka KEG-a, sustavu HTK.
- **Trening**—definiraju se modeli odnosno način treniranja te se vrši višestruka estimacija parametara. Pošto je konfiguracija ove faze veoma složena koristimo se poluautomatiziranom verzijom HTK sučelja razvijenom u diplomskom radu (Dropuljić, 2008).
- **Prepoznavanje**—pomoću razvijenog ASR-a (engl. *automatic speech recognition*) i izgrađenih i estimiranih modela vrši se transkripcija nepoznatog govora i analiza rezultata.

stvra se 10 direktorija test snimki te 10 direktorija trening snimki. Unutar svakog od direktorija nalaze se nasumično odabrane snimke za trening i test, pri čemu je svaki direktorij test snimki disjunktan u odnosu na sve druge direktorije. Direktorij test snimki je razlika KEG-a i odgovarajućeg test direktorija. Opisana struktura je priprema za kros-validaciju, pošto je potrebno testirati cjelokupni korpus, odnosno sve snimke KEG-a moraju biti ujedno testni modeli u nekoj od faza. Kako bi rezultati bili realniji, prije sortiranja u direktorije, sve snimke su nasumično permutirane, tako da se dobije nasumična pokrivenost emocijama u materijalu za trening i test. Ovime su generirane sve potrebne datoteke za razvijeni ASR te ih je potrebno samo postaviti na odgovarajuće lokacije i pokreniti treniranje i testiranje korpusa.

Pozivi:

```
HTK_podjela_snimaka('Staza do direktorija SNIMKE',...  
                    'Staza do *.xlsx datoteke sa,...  
                    podacima anotiranih snimki')
```

Iako korisnik nema direktan doticaj sa funkcijom za korekciju hrvatskih grafema (poziva se iz funkcije **HTK_podjela_snimaka**), možemo je spomenuti u vidu poznavanja napravljenih korekcija. Funkcija **bezkvacice3.m** radi sljedeće pretvorbe:

- Ć=>CY
- Č=>CX
- Đ=>DY
- DŽ=>DZX
- Š=>SX
- Ž=>ZX

6.1.1 Priprema podataka za trening

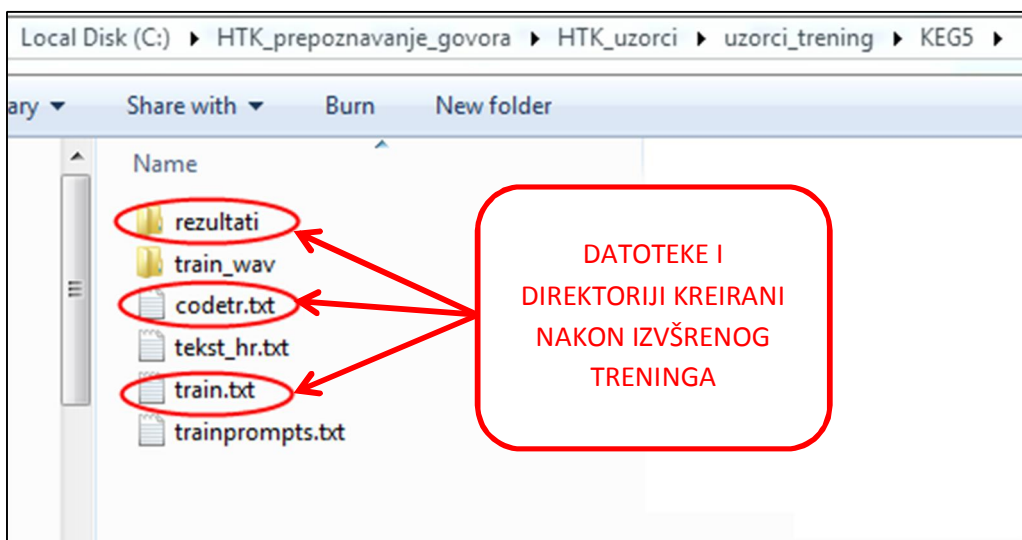
Nakon poziva i izvršenja funkcije **HTK_podjela_snimaka.m** stvoreno je 10 direktorija pod nazivom „*TreningX*“, gdje je $X \in \{1,2,3,\dots,10\}$. Unutar svakoga od direktorija nalaze se datoteke pod nazivom: „*teksts_hr.txt*“ i „*trainprompts.txt*“ koje odgovaraju transkriptu snimaka iz toga direktorija. Unutar datoteke „*tekst_hr.txt*“ nalaze

se transkripti svake pojedine snimke nanizani jedan iza drugog, dok se unutar „trainprompts.txt“ nalazi odgovarajući broj snimke i njezin transkript prema strukturi zapisa kao što je prikazno slikom (Slika 20).

```
S0070 JA PUCAM JER TI JEDNOM NESXTO KAZXEM
S0071 MA KAKVE SLIKE CXOVJECXE BOZXJI NE MOGU SAZNATI NI SXTO MI JE S
S0072 NE NE NE NEGO ONA TU PRICXA JEDNO A ONA JE VANI DRUGO
S0073 A BOLI ME [REDACTED] I ZA JASNU I ZA PERU I ZA MAGAZIN KAJ CYU TAM
S0074 NI MENE IVICE BRATE JER SIGURNO NA SVIJETU NEMA BOLJEGA TATE
S0075 SRAM TE BILO VIDIO SAM TE
```

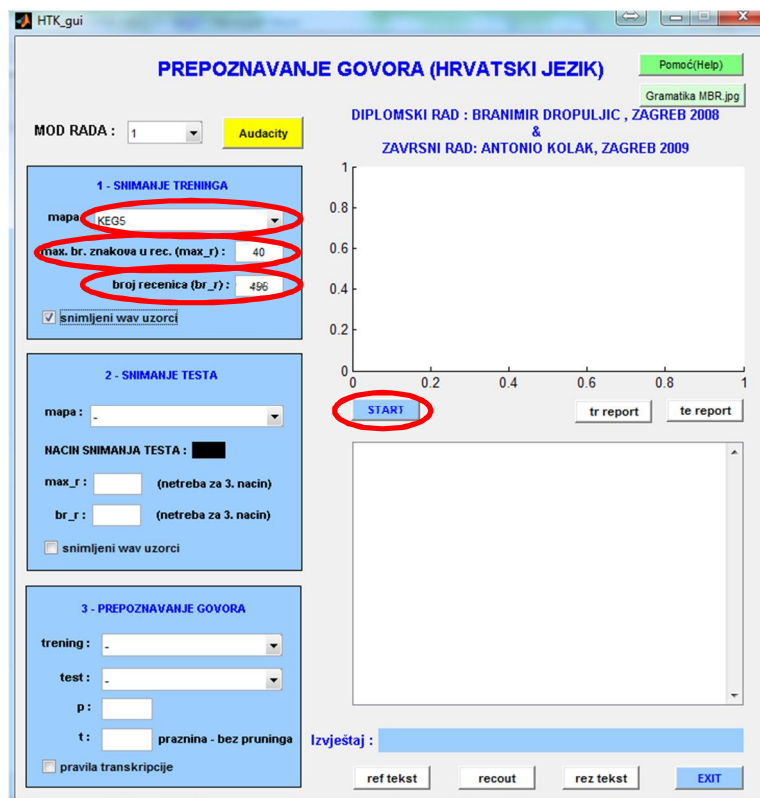
Slika 20-Primjer dijela trainprompts.txt

U pripadajućoj strukturi direktorija ASR-a „HTK_prepoznavanje_govora“ potrebno je napraviti prilagodbe i kopirati gore navedene datoteke. Unutar direktorija „HTK_uzorci\uzorci_trening“ kreira se proizvoljan direktorij u kojeg će se postaviti „trainprompts.txt“ i „tekst_hr.txt“. Kada su *.txt datoteke kopirane u proizvoljan novostvoreni direktorij, unutar njega je potrebno kreirati novi prazni direktorij pod nazivom „train_wav“. Kada je „train_wav“ kreiran u njega se kopiraju snimke iz pripadajućeg odgovarajućeg direktorija „TreningX“. Ovime bi trebala biti kreirana struktura koja je predočena na slici (Slika 21). Preostalo je dodati ime proizvoljno dodanog direktorija u tekstualnu datoteku pod nazivom „mape_sn_tr.txt“ kako bi razvijeni ASR mogao registrirati da su novi materijali za trening spremni.



Slika 21-Prikaz kreirane strukture prije i poslije treniranja

Nakon pravilno postavljene strukture i kopiranja materijala za trening pokreće se razvijeno sučelje za treniranje i testiranje, naredbom **HTK_gui**. Izgled sučelja je prikazan na slici (Slika 22). Pokretanje treninga se vrši odabirom željenog seta za trening i pritiskom na tipku start.



Slika 22-Prikaz opcija treniranja u grafičkom sučelju HTK_gui

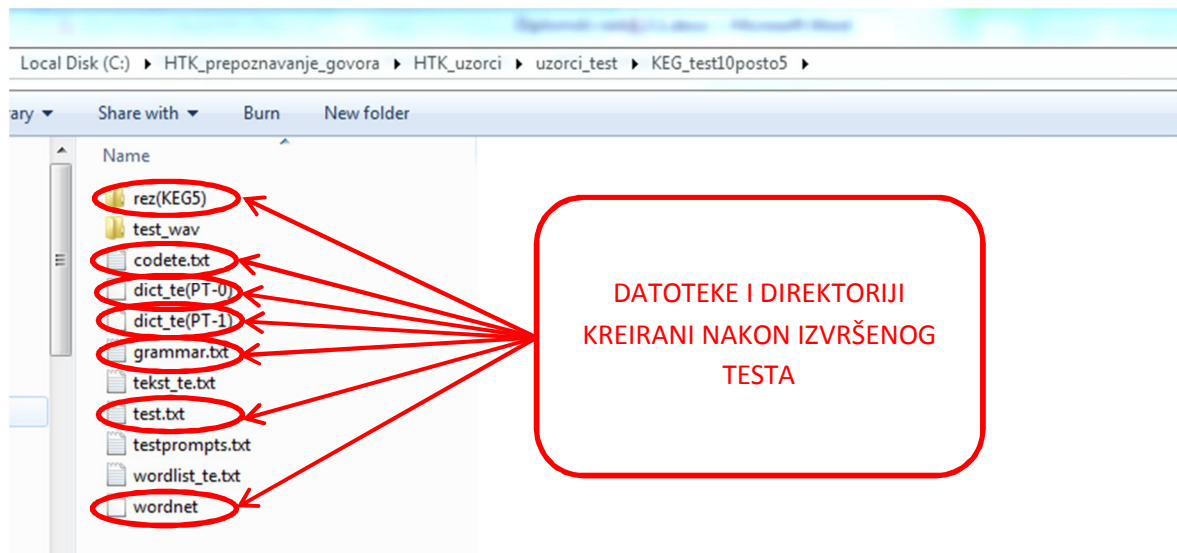
Prije generiranja podataka za trening potrebno je odabrati maksimalan broj znakova u rečenici i maksimalan broj rečenica. Sustav je zamišljen da funkcionira drugačije od onoga kako ga mi želimo iskoristiti. Odnosno, sustav će iz tekstualne datoteke „tekst_hr.txt“ kreirati datoteku „trainprompts.txt“ čija struktura rečenica će odgovarati gore spomenutim podacima (maksimalan broj znakova u rečenici i maksimalan broj rečenica). Nakon što je datoteka „trainprompts.txt“ generirana, prema njoj bi se trebale snimiti rečenice za trening. Kod nas je situacija različita pošto mi već imamo unaprijed definirane rečenice, kao i njihov transkript, te smo funkcijom **HTK_podjela_snimaka** generirali nama potrebnu datoteku „trainprompts.txt“ koja odgovara našim snimkama. Ono što nam nedostaje su datoteke „codetr.txt“ i „train.txt“ koje sustav također generira, a u kojima su zapisane staze do *.wav snimki i *.mfc parametarskih datoteka. Dakle, sustav nam je potreban samo kako bismo generirali

„*codetr.txt*“ i „*train.txt*“ datoteke, datoteke „*trainprompts.txt*“ i „*tekst_hr.txt*“ ćemo kopirati iz direktorija „*TreningX*“. Kako „*codetr.txt*“ barata sa brojem snimka, tada pod „maksimalan broj rečenica“ treba staviti točno onaj broj koji odgovara broju snimki, a pod maksimalan broj znakova u rečenici treba staviti onaj broj s kojim će se moći iz „*tekst_hr.txt*“ generirati dovoljan broj rečenica. Kada su datoteke kreirane, još jednom je potrebno kopirati datoteku „*trainprompts.txt*“ iz „*TreningX*“, pošto novonastala datoteka ne odgovara našim snimkama. Ovime je proces pripreme treninga gotov.

Napomena: U procesu treniranja sustava dogoditi će se pogreška ukoliko se iz datoteka „*trainprompts.txt*“ i „*tekst_hr.txt*“ ne ukloni specijalni znak povlake („_“) te se ne zamijeni s nekom drugom unikatnom kombinacijom slova koja ne tvori niti jednu riječ. Također unutar KEG-a ne postoji monofon „ĐŽ“ te je u slučaju korištenja gotovog, gore opisanog, ASR-a potrebno taj monofon dodati u sklopu „*trainprompts.txt*“ i na taj način namjerno izazvati pogrešku u točnosti, ili snimiti odgovarajuću rečenicu koja bi sadržavala dotični monofon.

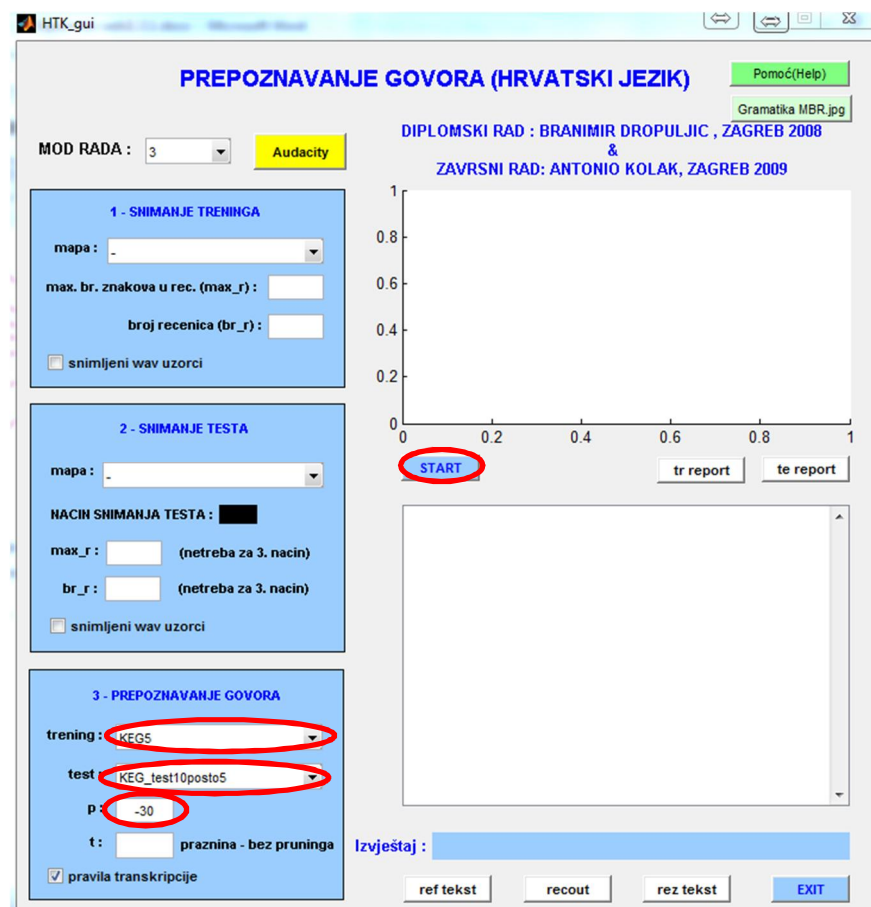
6.1.2 Priprema podataka za test

Priprema podataka za test se odvija po istim principima kao i priprema podatka za trening. U pripremi za test također koristimo datoteke generirane pomoću funkcije **HTK_podjela_snimaka**, samo što se sada generirane datoteke nalaze u direktoriju „*TestX*“, gdje je $X \in \{1,2,3,\dots,10\}$. Unutar svakog direktorija „*TestX*“ generirane su dvije datoteke, a to su: „*test_te.txt*“ i „*testprompts.txt*“. Generirane datoteke je potrebno kopirati na lokaciju „*HTK_uzorci\uzorci_test*“ u proizvoljno kreirani direktorij. U tom istom direktoriju je potrebno stvoriti novi direktorij pod nazivom „*test_wav*“ i u njega kopirati snimke iz odgovarajućeg direktorija „*TestX*“. Nakon što smo obavili sva kopiranja potrebno je u direktoriju, gdje se nalaze datoteke „*test_te.txt*“ i „*testprompts.txt*“, kreirati praznu datoteku naziva „*wordlist_te.txt*“ (čemu služi i kako postaviti rječnik od 15000 najčešćih hrvatskih izraza može se pogledati u (Dropuljić, 2008)). Preostalo je javiti sustavu da postoje novi materijali za testiranje, što obavljamo tako da u datoteku „*mape_sn_te.txt*“ upišemo naziv proizvoljno odabranog direktorija u kojem se nalaze dotični materijali (inače pokreće se tzv. 3 mod rada testa, za detalje pogledati (Dropuljić, 2008)). Slika (Slika 23) prikazuje stanje prije i poslije provođenja testiranja.



Slika 23-Prikaz kreirane strukture prije i poslije testa

Po završetku pripreme test materijala, pokreće se proces prepoznavanja s odabranim trening i test materijalima te se iščitavaju rezultati kao što je prikazano na slici (Slika 24).



Slika 24-Prikaz opcija prepoznavanja u grafičkom sučelju HTK_gui

7 Analiza rezultata

Analiza je provedena na tri odvojena sustava, od čega dva obavljaju lingvističku analizu, a treći vrši prepoznavanje govora. Sustavi lingvističke analize obavljaju jednaku zadaću samo što koriste različite metode, dok sustav prepoznavanja govora bi trebao osigurati podatke za ulaz u sustave lingvističke analize. Sustavi su testirani pojedinačno i zasebno bez zajedničke sprege. Ukupna točnost sustava automatske klasifikacije diskretnih emocionalnih stanja na osnovu lingvističkih parametara glasa kao kombinacija sustava prepoznavanja i sustava lingvističke analize nije provedena zbog dobivenih rezultata, ali se na osnovu pojedinačnih rezultata može predvidjeti maksimalna moguća točnost.

7.1 Analiza točnosti primitivnog unigram sustava na cjelokupnom korpusu

Svrha analize je provjera ispravnosti i karakteristika primitivnog unigram sustava. Testiranje sustava se provodilo na modelu prikazanom tablicom (Tablica 20). Iz tablice je vidljivo kako se sustav trenirao sa 90 % materijala KEG-a sa kojim smo raspolagali, odnosno sa 447 rečenica. Ostatak od 49 rečenica (10 %) upotrijebljen je za testiranje tako utreniranog sustava. U tablici (Tablica 21) su vidljivi rezultati provedene kros-validacije. Pri postupku kros-validacije testira se cjelokupni korpus 10×10 %.

Rezultati u tablici (Tablica 21) ne prikazuju samo točnost sustava već i cjelokupnu konfuzijsku matricu koja nam govori o rezultatima pogrešne klasifikacije za pojedinu emociju. Točno klasificirane rečenice su zabilježene na glavnoj dijagonali matrice. Iz dotične matrice vidljiv je točan broj snimaka po klasificiranim kategorijama. Međutim, s obzirom da klase korpusa nisu dimenzijski jednake tj. pojedine klase su imale više ulaznih značajki za treniranje, tada je pogodnije rezultate promatrati u postotno izraženoj matrici prikazanoj u tablici (Tablica 22).

Tablica 20-Dimenzije trening i test korpusa za testiranje primitivnog unigram sustava

Broj rečenica za trening [90 %]	Broj rečenica za test [10 %]
447	49

Tablica 21-Konfuzijska matrica primitivnog unigram sustava

		ULAZI U SUSTAV					Točno prepoznatih
		SREĆA	TUGA	STRAH	LJUTNJA	NEUTRALNO	
IZLAZ IZ SUSTAVA	SREĆA	35	9	4	9	3	35
	TUGA	5	34	1	18	5	34
	STRAH	3	3	2	5	2	2
	LJUTNJA	9	16	7	125	14	125
	NEUTRALNO	6	7	3	17	36	36
	NIJE PREPOZNATO	18	12	16	55	11	232
$\sum$ ulaznih_emocija		76	81	33	229	71	
UKUPNO		490					47.34 %

Tablica 22-Postotno izražena konfuzijska matrica primitivnog unigram sustava

		ULAZI U SUSTAV				
		SREĆA	TUGA	STRAH	LJUTNJA	NEUTRALNO
IZLAZ IZ SUSTAVA	SREĆA	46.05 %	11.11 %	12.12 %	3.93 %	4.22 %
	TUGA	6.57 %	41.97 %	3.03 %	7.86 %	7.04 %
	STRAH	3.94 %	3.70 %	6.06 %	2.18 %	2.81 %
	LJUTNJA	11.84 %	19.75 %	21.21 %	54.58 %	19.71 %
	NEUTRALNO	7.89 %	8.64 %	9.09 %	7.42 %	50.70 %
	NIJE PREPOZNATO	23.68 %	14.81 %	48.48 %	24.01 %	15.49 %

Iz postotno izražene konfuzijske matrice vidimo kako daleko najlošije rezultate prepoznavanja imamo za emociju straha, dok najbolje rezultate imamo za emociju ljutnje. Ovo je očekivani rezultat s obzirom na utreniranost sustava. Naime, najmanje trening materijala imamo upravo za emociju straha, a najviše za emociju ljutnje. Ovo je nažalost posljedica ograničenog korpusa kojeg u ovoj fazi imamo na raspolaganju.

7.2 Analiza točnosti naivnog Bayesovog tekstualnog klasifikatora na cjelokupnom korpusu

Princip testiranja Bayesovog tekstualnog klasifikatora je ostao isti kao i u prethodnom slučaju testiranja primitivnog unigram sustava. Materijali za trening i test su podijeljeni u omjeru 90 % naprama 10 % kao što to prikazuje tablica (Tablica 23). Konačni zaključak tablice (Tablica 24) je nešto veća točnost sustava (61 %) u odnosu na primitivni unigram sustav. Međutim, valjda biti veoma oprezan kod donošenja zaključaka o tome koji sustav je uistinu bolji. Pomnije promotrivši matricu konfuzije Bayesovog klasifikatora vidimo kako on svoju veću točnost ostvaruje na klasi ljutnje. Ovo je posljedica načina

računanja uvjetne vjerojatnosti za pojedini token. Naime, pri izračunu uvjetne vjerojatnosti uzima se u obzir i vjerojatnost same klase. Procjena se ostvaruje po principu broja dokumenata, odnosno u našem slučaju broja rečenica koje pripadaju pojedinoj klasi. Kako je broj rečenica ljutnje gotovo trostruko veći od sljedeće manje klase, tada je jasno da će klasa ljutnje biti dominantna u odnosu na sve ostale klase. Dakle, u onome slučaju kada niti jedan token nije prepoznat od strane sustava, on će dotičnu rečenicu proglasiti rečenicom ljutnje pošto je klasa ljutnje najvjerojatnija za pojavu. Ovo je dobra karakteristika za zatvorene sustave sa točno definiranim leksičkim korpusom, ali za općenitu primjenu nije pogodna.

Tablica 23- Dimenzije trening i test korpusa za testiranje Bayesovog tekstualnog klasifikatora

Broj rečenica za trening [90 %]	Broj rečenica za test [10 %]
447	49

Tablica 24-Konfuzijska matrica Bayesovog klasifikatora

		ULAZI U SUSTAV					Točno prepoznatih
		SREĆA	TUGA	STRAH	LJUTNJA	NEUTRALNO	
IZLAZ IZ SUSTAVA	SREĆA	30	3	0	0	3	30
	TUGA	4	25	0	8	3	25
	STRAH	3	0	2	1	1	2
	LJUTNJA	39	51	30	213	37	213
	NEUTRALNO	2	3	1	6	31	31
	NIJE PREPOZNATO	0	0	0	0	0	301
$\sum$ ulaznih_emocija		75	82	33	228	72	
UKUPNO		490					61.42 %

Tablica 25-Postotno izražena konfuzijska matrica Bayesovog sustava

		ULAZI U SUSTAV				
		SREĆA	TUGA	STRAH	LJUTNJA	NEUTRALNO
IZLAZ IZ SUSTAVA	SREĆA	40.00 %	3.66 %	0 %	0 %	0 %
	TUGA	5.33 %	30.49 %	0 %	3.51 %	4.17 %
	STRAH	0 %	0 %	6.06 %	0.44 %	1.39 %
	LJUTNJA	52.00 %	62.20 %	90.91 %	93.42 %	51.39 %
	NEUTRALNO	2.67 %	3.66 %	3.03 %	2.63 %	43.06 %
	NIJE PREPOZNATO	0 %	0 %	0 %	0 %	0 %

7.3 Analiza točnosti prepoznavanja govora HTK alatom

Nakon što su svi podatci bili spremni za HTK alat, izvršeno je testiranje po principu kros-validacije. Testiranje se vršilo na cijelome korpusu od 496 snimaka uz uključena pravila transkripcije (PT=1) i koeficijent $p=-30$. S obzirom da je korpus podijeljen nasumično na omjere od 10 % za test materijale i 90 % za trening materijale, broj govornika je varijabilan. Maksimalni broj govornika je naravno ograničen brojem govornika u KEG-u. Iz tablice (Tablica 26) je vidljiva maksimalna točnost od oko 13 %, a uprosječni rezultat je nekih 9.7 %. Ovako niski rezultati točnosti su bili i za očekivati s obzirom na svojstva KEG-a. Naime, u korpusu postoji više od 140 različitih govornika gledajući distribuciju govornika po emocijama kao što je priloženo u tablici (Tablica 3). Ovako veliki broj govornika, te loša kvaliteta snimki „stvarnih“ emocija, doprinijela je ovakvom rezultatu.

Tablica 26-Rezultati analize točnosti prepoznavanja govora HTK alatom

PREPOZNAVANJE NA RAZINI RIJEČI (nulta gramatika) PT=1; p=-30						
KORAK KROS- VALIDACIJE	TOČNO PREPOZNATIH RIJEČI (H)	OBRISANIH RIJEČI (D)	ZAMIJENJENIH RIJEČI (S)	UMETNUTIH RIJEČI (I)	UKUPNO RIJEČI (N)	TOČNOST PREPOZNAVANJA [%]
1	78	136	394	51	608	12.83
2	84	144	473	42	701	11.98
3	71	158	411	39	640	11.09
4	55	182	413	39	650	8.46
5	50	168	334	26	552	9.06
6	66	151	479	36	696	9.48
7	60	210	398	39	668	8.98
8	39	159	325	23	523	7.46
9	71	159	448	54	678	10.47
10	52	218	398	44	668	7.78
PROSJEČNO:	62.6	168.5	407.3	39.3	638.4	9.759 %

Nešto povoljniji rezultati se dobivaju pri iznosu koeficijenta $p=0$. Sve ostale izmjene utječu jako malo na promjenu točnosti sustava te možemo reći da je općenito postignuta točnost od nekih 10 do 12 posto.

Tablica 27- Rezultati analize točnosti prepoznavanja govora HTK alatom

PREPOZNAVANJE NA RAZINI RIJEČI (nulta gramatika) PT=1; p=0						
KORAK KROS- VALIDACIJE	TOČNO PREPOZNATIH RIJEČI (H)	OBRISANIH RIJEČI (D)	ZAMIJENJENIH RIJEČI (S)	UMETNUTIH RIJEČI (I)	UKUPNO RIJEČI (N)	TOČNOST PREPOZNAVANJA [%]
1	78	136	394	51	608	12.83
2	111	27	563	607	701	15.83
3	92	45	503	553	640	14.38
4	91	90	469	641	650	14.00
5	68	61	423	510	552	12.32
6	78	34	584	634	696	11.21
7	75	105	488	558	668	11.23
8	70	67	386	361	523	13.38
9	102	65	511	691	678	15.04
10	71	94	503	596	668	10.63
PROSJEČNO:	83.6.	72.4	482.4	520.2	638.4	13.085 %

8 Zaključak

Ovim diplomskim radom po prvi puta na Zavodu za elektroničke sustave i obradu informacija se proučila mogućnost automatske klasifikacije diskretnih emocionalnih stanja na osnovu lingvističkih parametara govora. U svrhu tome razvio se i prvi korpus Hrvatskog emocionalnog govora, koji može poslužiti kao osnova za daljnji razvoj i istraživanje. Iz rezultata je vidljivo kako je čista lingvistička analiza moguća uz relativno zadovoljavajuću točnost. Isto ne možemo reći i za sustav prepoznavanja govora, što je posljedica samih karakteristika KEG-a i trenutačne nesavršenosti funkcioniranja sustava prepoznavanja govora. Kako bi se istraživanje na ovom području mogla nastaviti, za primijenjene metode je dan kompletan postupak modeliranja sustava lingvističke analize, te su predložene metode koje bi se mogle istražiti i primijeniti u daljnjem razvoju.

Razviti sustav čija kompleksnost u jednom segmentu komparira sa ljudskim mozgom nije ni najmanje lak zadatak. Emocija je informacija u glasu koja je teško uočljiva u usporedbi s nekim drugim parametrima glasa, poput poruke koju nam govornik prenosi, spola govornika, dobi, ili boje glasa. Ovim diplomskim radom i diplomskim radom kolege Miłosza Chmure¹, koji se paralelno bavio razvojem sustava za automatsku klasifikaciju diskretnih emocionalnih stanja na osnovu akustičkih značajki govora, pokušalo se barem malo približiti toj kompleksnosti i otvoriti put daljnjim istraživanjima.

¹ Chmura, M. Automatska klasifikacija diskretnih emocionalnih stanja na osnovu akustičkih značajki govora, diplomski rad, ZESOI, FER, Zagreb, 2011.

9 Literatura

- [1.] Chmura, M., Dropuljić, B., Kolak A. Croatian Emotional Corpus, University of Zagreb, Faculty of electrical engineering and computing, Zagreb, Croatia, 2011.
- [2.] Ćavar, D. Quantitative and qualitative computational analysis of language and text similarities, clustering and classification slides (1,2,3,4), CLS2010 - Computational Linguistics Summer Events, Sveučilište u Zadru, Odjel za lingvistiku, Zadar, 2010.
- [3.] Ćavar, D. Quantitative and qualitative computational analysis of language and text similarities, clustering and classification,
<http://ling.unizd.hr/~cls2010/program/clusteringclassification/index.en.html>
- [4.] Young, S. i ostali, The HTK Book, Cambridge University Engineering Department, Cambridge, 2005.
- [5.] Dropuljić, B. Postupci razvoja akustičkog i leksičkog modela sustava za automatsko prepoznavanje govora za hrvatski jezik, diplomski rad, ZESOI, FER, Zagreb, 2008.
- [6.] Young, S. ATK An Application Toolkit for HTK version 1.6, Machine Intelligence Laboratory, Cambridge University Engineering Dept, Cambridge, 2007.
- [7.] Python v2.7.1 documentation, <http://docs.python.org/>, 2011.
- [8.] Matlab—R2011a Documentation,
<http://www.mathworks.com/help/techdoc/index.html>
- [9.] Banse, R., & Scherer, K. R., Acoustic profiles in vocal emotion expression. Journal of Personality and Social Psychology, Humboldt University, University of Geneva, 1996.
- [10.] R. W. Piccard, "Affective Computing", Technical Report No. 321, Media Laboratory Perceptual Computing Section, M.I.T, 1995.
- [11.] R. Fernandez and R.W. Picard, "Modeling Drivers' Speech under Stress," Speech Comm., vol. 40, 2003, pp. 145-159.
- [12.] C.M. Lee and S.S. Narayanan, "Toward Detecting Emotions in Spoken Dialogs," IEEE Trans. Speech and Audio Processing, vol. 13, no. 2, Mar. 2005, pp. 293-303.

- [13.] C. Lisetti and F. Nasoz, "Affective Intelligent Car Interfaces with Emotion Recognition", In Proceedings of 11th International Conference on Human Computer Interaction, Las Vegas, NV, USA, July 2005.
- [14.] L. Devillers and L. Vidrascu, "Real-Life Emotions Detection with Lexical and Paralinguistic Cues on Human-Human Call Center Dialogs," Proc. Ninth Int'l Conf. Spoken Language Processing, 2006.
- [15.] F. Nagel, R. Kopiez, O. Grewe and E. Altenmüller, "EMuJoy: Software for continuous measurement of perceived emotions in music", Behavior Research Methods, 39 (2), 2007, pp. 283-290.
- [16.] D. Datcu and L.J.M. Rothkrantz, "Multimodal recognition of emotions in car environments", D|C|I&I Prague, 2009.
- [17.] M. Cowan, "Pitch and intensity characteristics of stage speech", Arch. Speech, 1936.
- [18.] G. Fairbanks and W. Pronovost, "An experimental study of the pitch characteristics of the voice during the expression of emotion", Speech monograph, vol 6, 1939, pp. 87-104.
- [19.] R. Frick, "Communicating emotion: The role of prosodic features", Psychol. Bull., vol. 97, 1985, pp. 412-429.
- [20.] K. Scherer, "Vocal affect expression: A review and a model for future research", Psychological Bulletin, vol. 99, 1986, pp. 143-165.
- [21.] R. Banse and K. Scherer, "Acoustic profiles in vocal emotion expression", J. Personality Social Psych., vol. 70, 1996, no 3, pp. 614-636.
- [22.] P.N. Juslin and K.R. Scherer, "Vocal Expression of Affect," The New Handbook of Methods in Nonverbal Behavior Research, Oxford Univ. Press, 2005.
- [23.] I.R. Murray and J.L. Arnott, "Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion", J. Acoust. Soc. Amer., vol. 93, 1993, pp. 1097-1108.
- [24.] J.A. Russell, J.A. Bachorowski, and J.M. Fernandez-Dols, "Facial and Vocal Expressions of Emotion," Ann. Rev. of Psychology, vol. 54, 2003, pp. 329-349.
- [25.] B. Schuller, R.J. Villar, G. Rigoll, and M. Lang, "Meta-Classifiers in Acoustic and Linguistic Feature Fusion-Based Affect Recognition", Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing, 2005.

- [26.] C.-H. Wu, Z.-J. Chuang and Y.-C. Lin, "Emotion Recognition from Text Using Semantic Labels and Separable Mixture Models", *ACM Transactions on Asian Language Information Processing*, vol. 5, no. 2, Jun. 2006, pp. 165-182.
- [27.] J. Rong, G. Li, Y.P. Chen, "Acoustic feature selection for automatic emotion recognition from speech", *Information processing and management*, vol. 45, 2009, pp. 315-328.
- [28.] T.S. Polzin, "Verbal and non-verbal cues in the communication of emotions," in *Proc. ICASSP '00*, 2000, pp. 2429-2432.
- [29.] B. Schuller, G. Rigoll and M. Lang, "Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture", in *Proc. ICASSP '04*, 2004, pp. 577-580.
- [30.] S. Martinčić-Ipšić, "Croatian Speech Recognition and Synthesis Based on Context-dependent Hidden Markov Model," PhD Thesis, in Croatian, University of Zagreb, Croatia, Nov. 2007.
- [31.] S. Martinčić-Ipšić, S. Ribarić and I. Ipšić, "Acoustic Modelling for Croatian Speech Recognition and Synthesis," *Informatica*, vol. 19, 2008, pp. 227-254.
- [32.] S. Burger, V. Maclaren, and H. Yu, "The ISL Meeting Corpus: The Impact of Meeting Type on Speech Style", *Proc. Eighth Int'l Conf. Spoken Language Processing (ICSLP)*, 2002.
- [33.] E. Douglas-Cowie, N. Campbell, R. Cowie, and P. Roach, "Emotional Speech: Towards a New Generation of Database", *Speech Comm.*, vol. 40, nos. 1/2, 2003, pp. 33-60.
- [34.] S.T. Jovičić, Z. Kašić, M. Đorđević and M. Rajković, "Serbian emotional speech database: design, processing and evaluation", in *9th Speech and Computer Conference*, Sep. 2004.
- [35.] D. Kukolja, S. Popović, B. Dropuljić, M. Horvat and K. Ćosić, "Real-time emotional state estimator for adaptive virtual reality stimulation", *Lecture Notes in Computer Science*, vol. 5638, Jul. 2009, pp. 175-184.
- [36.] K. Ćosić, S. Popović, D. Kukolja, M. Horvat and B. Dropuljić, "Physiology-driven adaptive virtual reality stimulation for prevention and treatment of stress related disorders", *CyberPsychology, Behavior, and Social Networking*, 13, 2010, pp. 1; 73-78.

- [37.] K. Ćosić et al., "Virtual reality adaptive stimulation in stress resistance training", Proceedings RTO-MP-HFM-205 on "Mental Health and Well-Being across the Military Spectrum", 2011.
- [38.] J.A. Russell, "A circumplex model of affect", *Journal of Personality and Social Psychology*, Vol. 39, Dec. 1980, pp. 1161-1178.
- [39.] B.J. Vaughan, C. Cullen, S. Kousidis and J. McAuley, "Emotional Speech Corpus Construction, Annotation and Distribution", The 6th edition of the Language Resources and Evaluation Conference Marrakech (Morocco), 2008.
- [40.] P.J. Lang, M.M. Bradley and B.N. Cuthbert, "International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual", Technical Report A-6. University of Florida, Gainesville, FL, 2005.
- [41.] M.M. Bradley and P.J. Lang, "The International Affective Digitized Sounds, 2nd edn. (IADS-2): Affective Ratings of Sounds and Instruction Manual", Technical report B-3. University of Florida, Gainesville, FL, 2007.
- [42.] L.E. Baum and T. Petrie, "Statistical interface for probabilistic functions of finite state Markov chains", *The Annals of Mathematical Statistics*, vol. 37, 1966, pp. 1554-1563.
- [43.] S.B. Davis, and P. Mermelstein, "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 28, Aug. 1980, pp. 357-366.
- [44.] B. Dropuljić, "Development of Acoustic and Lexical model for Automatic speech recognition for Croatian Language," Diploma Thesis, in Croatian, University of Zagreb, Croatia, Jan. 2008.
- [45.] B. Dropuljić and D. Petrinović, "Development of Acoustic Model for Croatian Language Using HTK", *Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije*, vol. 51, 2010, pp. 79-88.

10 Sažetak

AUTOMATSKA KLASIFIKACIJA DISKRETNIH EMOCIONALNIH STANJA NA OSNOVU LINGVISTIČKIH ZNAČAJKI GOVORA

U sklopu ovog diplomskog rada izgradio se sustav za automatsku klasifikaciju diskretnih emocionalnih stanja na osnovu lingvističkih značajki govora. Razvoj sustava kao i cjelokupna obrada je napravljena u Matlabu. Sustav je baziran na N-gram modelima. Za potrebe diplomskog rada prikupljen je i anotiran korpus Hrvatskog emocionalnog govora, te su razvijeni popratni sadržaji i aplikacije u tu svrhu. Pripadajući transkript korpusa je upotrijebljen za lingvističku analizu i treniranje sustava. Ovaj rad predstavlja polovicu cjelokupnog sustava prepoznavanja emocija koji je razvijan, te bi se ukupni rezultati prepoznavanja trebali dobiti fuzijom ovog i sustava zasnovanog na akustičkim značajkama.

Ključne riječi:

- Automatska klasifikacija
- Diskretne emocije
- Korpus Hrvatskog emocionalnog govora
- Lingvističke značajke
- Unigram
- Bigram
- N-gram
- Automatsko prepoznavanje govora
- HTK (eng.HMM ToolKit)

11 Summary

AUTOMATIC CLASSIFICATION OF DISCRETE EMOTIONAL STATES BASED ON LINGUISTIC SPEECH FEATURES

As a part of thesis system for automatic classification of discrete emotional states based on linguistic speech features was built. System development and complete processing were made in Matlab. The system was based on N-gram models. For the purposes of the thesis Croatian emotional speech corpus was collected and annotated, and also further services and applications were developed. Associated corpus transcript was used for linguistic analysis and system training. This thesis represents half of the entire emotions recognition system that has been developed, and complete results of recognition are to be gotten by fusion of this system with acoustic features based system.

Key words:

- Automatic classification
- Discrete emotions
- Croatian emotional speech corpus
- Linguistic features
- Unigram
- Bigram
- N-gram
- Automatic speech recognition
- HTK (HMM ToolKit)

12 Privitci

12.1 Privitak — Statistika za pojedinu snimku prema spolu, dobi, ID-u govornika, ekspresiji i procijenjenoj emociji

ID snimke	spol	dob	ID govornika	ID gov. (M, Z)	ekspresija	r - Emocija
1	M	O	1	1	S	LJ
2	M	O	1	1	S	LJ
3	M	O	2	2	S	LJ
4	M	O	3	3	S	LJ
5	M	O	4	4	S	LJ
6	M	O	4	4	S	LJ
7	M	O	4	4	S	LJ
8	M	O	5	5	S	LJ
9	M	O	5	5	S	LJ
10	M	O	6	6	G	LJ
11	M	O	7	7	G	LJ
12	M	O	7	7	G	LJ
13	M	O	7	7	G	LJ
14	M	O	8	8	G	LJ
15	M	O	9	9	G	LJ
16	M	O	9	9	G	LJ
17	M	O	9	9	G	LJ
18	M	O	10	10	G	LJ
19	M	O	11	11	G	LJ
20	M	O	12	12	G	LJ
21	M	O	13	13	G	LJ
22	M	O	14	14	G	
23	M	O	15	15	G	LJ
24	M	O	16	16	G	LJ
25	M	O	16	16	G	LJ
26	M	O	17	17	G	LJ
27	M	O	17	17	G	LJ
28	M	O	17	17	G	LJ
29	M	O	17	17	G	LJ
30	M	O	17	17	G	LJ
31	M	O	17	17	G	LJ
32	M	O	18	18	G	LJ
33	M	O	19	19	G	LJ
34	M	O	19	19	G	LJ
35	M	O	19	19	G	LJ
36	M	O	19	19	G	LJ

37	M	O	19	19	G	LJ
38	M	O	19	19	G	LJ
39	M	O	20	20	G	
40	M	O	20	20	G	LJ
41	M	O	20	20	G	LJ
42	M	O	20	20	G	LJ
43	M	O	20	20	S	LJ
44	M	O	20	20	G	LJ
45	M	O	20	20	G	LJ
46	M	O	21	21	S	LJ
47	M	O	21	21	S	LJ
48	M	O	21	21	S	LJ
49	M	O	21	21	S	LJ
50	M	O	21	21	S	LJ
51	M	O	21	21	S	LJ
52	M	O	21	21	S	LJ
53	M	O	21	21	S	LJ
54	M	O	21	21	S	LJ
55	M	O	21	21	S	LJ
56	M	O	21	21	S	LJ
57	M	O	21	21	S	
58	M	O	21	21	S	LJ
59	M	O	21	21	S	LJ
60	M	O	21	21	S	LJ
61	M	O	21	21	S	NE
62	M	O	21	21	S	LJ
63	M	O	21	21	S	LJ
64	M	O	22	22	S	LJ
65	M	O	23	23	S	LJ
66	M	O	23	23	S	LJ
67	M	O	23	23	S	LJ
68	M	O	23	23	S	LJ
69	M	O	23	23	S	LJ
70	M	O	23	23	S	LJ
71	M	O	23	23	S	
72	M	O	23	23	S	
73	M	O	23	23	S	LJ
74	M	O	23	23	S	
75	M	O	23	23	S	
76	M	O	24	24	S	LJ
77	M	O	24	24	S	LJ
78	M	O	24	24	S	LJ
79	M	O	24	24	S	LJ

80	M	O	24	24	S	LJ
81	M	O	25	25	S	LJ
82	M	O	26	26	S	LJ
83	M	O	26	26	S	LJ
84	M	O	26	26	S	LJ
85	M	O	26	26	S	LJ
86	M	O	26	26	S	LJ
87	M	O	26	26	S	LJ
88	M	O	26	26	S	LJ
89	M	O	26	26	S	LJ
90	M	O	26	26	S	NE
91	M	O	26	26	S	LJ
92	M	O	26	26	S	
93	M	O	27	27	S	
94	M	O	27	27	S	NE
95	M	O	27	27	S	LJ
96	M	O	27	27	S	LJ
97	M	O	27	27	S	LJ
98	M	O	28	28	S	
99	M	O	28	28	S	
100	M	O	28	28	S	LJ
101	M	O	28	28	S	
102	M	O	28	28	S	LJ
103	M	O	28	28	S	LJ
104	M	O	29	29	G	LJ
105	M	O	29	29	G	LJ
106	M	O	29	29	G	LJ
107	M	O	29	29	G	LJ
108	M	O	29	29	G	LJ
109	M	O	29	29	G	LJ
110	M	O	29	29	G	
111	M	O	29	29	G	LJ
112	M	O	29	29	G	LJ
113	M	O	29	29	G	LJ
114	M	O	29	29	G	LJ
115	M	O	29	29	G	LJ
116	M	O	29	29	G	LJ
117	M	D	30	30	G	
118	M	D	30	30	G	LJ
119	M	D	30	30	G	LJ
120	M	O	31	31	G	LJ
121	M	O	31	31	G	LJ
122	M	O	31	31	G	LJ

123	M	O	31	31	G	LJ
124	M	O	31	31	G	LJ
125	M	O	31	31	G	LJ
126	M	O	31	31	G	LJ
127	M	O	31	31	G	LJ
128	M	O	31	31	G	LJ
129	M	O	31	31	G	LJ
130	M	O	31	31	G	LJ
131	M	O	31	31	G	LJ
132	M	O	31	31	G	LJ
133	M	O	31	31	G	LJ
134	M	O	31	31	G	LJ
135	M	O	31	31	G	
136	M	O	31	31	G	LJ
137	M	O	31	31	G	
138	M	O	31	31	G	LJ
139	M	O	31	31	G	TU
140	M	O	31	31	G	LJ
141	M	O	31	31	G	LJ
142	M	O	31	31	G	LJ
143	M	O	31	31	G	LJ
144	M	O	31	31	G	LJ
145	M	O	31	31	G	LJ
146	M	O	31	31	G	
147	M	O	31	31	G	
148	M	O	31	31	G	
149	M	O	31	31	G	LJ
150	M	O	31	31	G	LJ
151	M	O	31	31	G	LJ
152	M	O	31	31	G	LJ
153	M	O	31	31	G	LJ
154	M	O	31	31	G	
155	M	O	31	31	G	
156	M	O	31	31	G	LJ
157	M	O	31	31	G	LJ
158	Z	O	32	1	S	LJ
159	Z	O	32	1	S	LJ
160	Z	O	32	1	S	LJ
161	Z	O	33	2	S	LJ
162	Z	O	34	3	S	LJ
163	Z	O	35	4	S	LJ
164	Z	O	35	4	S	LJ
165	Z	O	35	4	S	LJ

166	Z	O	36	5	S	LJ
167	Z	O	37	6	S	LJ
168	Z	O	37	6	S	LJ
169	Z	O	37	6	S	LJ
170	Z	O	37	6	S	LJ
171	Z	O	37	6	S	LJ
172	Z	O	37	6	S	LJ
173	Z	O	38	7	S	LJ
174	Z	O	38	7	S	LJ
175	Z	O	39	8	S	
176	Z	O	40	9	G	LJ
177	Z	O	41	10	G	LJ
178	Z	O	41	10	G	LJ
179	Z	O	41	10	G	LJ
180	Z	O	41	10	G	LJ
181	Z	O	41	10	G	LJ
182	Z	O	42	11	G	LJ
183	Z	O	42	11	G	
184	Z	O	42	11	G	LJ
185	Z	O	42	11	G	LJ
186	Z	O	43	12	G	LJ
187	Z	O	44	13	G	LJ
188	Z	O	45	14	G	LJ
189	Z	O	45	14	G	
190	Z	O	46	15	G	LJ
191	Z	O	46	15	G	LJ
192	Z	O	46	15	G	LJ
193	Z	O	47	16	G	LJ
194	Z	O	48	17	S	LJ
195	Z	O	48	17	S	NE
196	Z	O	48	17	S	NE
197	Z	O	48	17	S	
198	Z	O	48	17	S	NE
199	Z	O	48	17	S	
200	Z	O	49	18	S	LJ
201	Z	O	49	18	S	
202	Z	O	49	18	S	
203	Z	O	49	18	S	
204	Z	O	50	19	S	LJ
205	Z	O	50	19	S	
206	Z	O	50	19	S	NE
207	Z	O	50	19	S	
208	Z	O	50	19	S	LJ

209	Z	O	50	19	S	
210	Z	O	50	19	S	LJ
211	Z	O	50	19	S	LJ
212	Z	O	50	19	S	
213	Z	O	50	19	S	NE
214	Z	O	50	19	S	LJ
215	Z	O	50	19	S	LJ
216	Z	O	51	20	S	LJ
217	Z	O	51	20	S	LJ
218	Z	O	51	20	S	SR
219	Z	O	52	21	S	LJ
220	Z	O	52	21	S	
221	Z	O	53	22	S	LJ
222	Z	O	53	22	S	LJ
223	Z	O	53	22	S	LJ
224	Z	O	53	22	S	LJ
225	Z	O	54	23	S	LJ
226	Z	O	54	23	S	LJ
227	Z	O	55	24	S	LJ
228	Z	O	55	24	S	
229	Z	O	55	24	S	LJ
230	Z	O	55	24	S	NE
231	Z	O	55	24	S	LJ
232	Z	O	55	24	S	LJ
233	Z	O	55	24	S	LJ
234	Z	O	55	24	S	LJ
235	Z	O	55	24	S	LJ
236	Z	O	55	24	S	LJ
237	Z	O	55	24	S	LJ
238	Z	O	55	24	S	LJ
239	Z	O	55	24	S	LJ
240	Z	O	55	24	S	NE
241	Z	O	55	24	S	
242	Z	O	55	24	S	LJ
243	Z	O	56	25	S	LJ
244	Z	O	56	25	S	LJ
245	Z	O	57	26	S	LJ
246	Z	O	58	27	S	NE
247	Z	O	58	27	S	
248	Z	O	58	27	S	
249	Z	O	58	27	S	
250	Z	O	58	27	S	
251	Z	O	59	28	S	

252	Z	O	59	28	S	LJ
253	Z	O	60	29	G	LJ
254	Z	O	60	29	G	LJ
255	Z	O	60	29	G	LJ
256	Z	O	60	29	G	LJ
257	Z	O	60	29	G	LJ
258	Z	O	60	29	G	LJ
259	Z	O	60	29	G	LJ
260	Z	O	60	29	G	LJ
261	Z	O	60	29	G	LJ
262	Z	O	60	29	G	LJ
263	Z	O	60	29	G	LJ
264	Z	O	60	29	G	
265	Z	O	60	29	G	
266	Z	O	60	29	G	LJ
267	Z	O	60	29	G	LJ
268	Z	O	60	29	G	LJ
269	Z	O	60	29	G	LJ
270	Z	O	60	29	G	LJ
271	Z	O	60	29	G	LJ
272	Z	O	60	29	G	LJ
273	Z	O	60	29	G	LJ
274	Z	O	60	29	G	LJ
275	Z	O	60	29	G	LJ
276	Z	O	60	29	G	LJ
277	Z	O	60	29	G	LJ
278	Z	O	60	29	G	LJ
279	Z	O	60	29	G	LJ
280	Z	O	60	29	G	
281	Z	O	60	29	G	LJ
282	Z	O	60	29	G	LJ
283	Z	O	60	29	G	LJ
284	Z	O	60	29	G	
285	Z	O	60	29	G	LJ
286	Z	A	61	30	G	LJ
287	Z	A	61	30	G	LJ
288	M	O	62	32	S	NE
289	M	O	63	33	S	NE
290	M	O	64	34	S	NE
291	M	O	65	35	S	NE
292	M	O	66	36	S	NE
293	M	O	66	36	S	NE
294	M	O	67	37	S	NE

295	M	O	68	38	S	NE
296	M	O	69	39	S	NE
297	M	O	70	40	S	NE
298	M	O	71	41	S	NE
299	M	O	71	41	S	NE
300	M	O	72	42	S	NE
301	M	O	73	43	S	NE
302	M	O	74	44	S	NE
303	M	O	75	45	S	NE
304	M	O	76	46	S	NE
305	M	O	77	47	S	NE
306	M	O	77	47	S	NE
307	M	O	77	47	S	NE
308	M	O	77	47	S	NE
309	M	O	77	47	S	NE
310	M	O	77	47	S	NE
311	M	O	77	47	S	NE
312	M	O	77	47	S	NE
313	M	O	78	48	S	NE
314	M	O	79	49	S	NE
315	M	O	79	49	S	NE
316	M	O	79	49	S	NE
317	M	O	80	50	S	NE
318	M	O	81	51	G	
319	M	O	81	51	G	NE
320	M	O	81	51	G	NE
321	M	O	81	51	G	NE
322	M	D	82	52	G	NE
323	M	O	83	53	G	
324	M	O	83	53	G	
325	M	O	83	53	G	NE
326	Z	O	84	31	S	NE
327	Z	O	85	32	S	NE
328	Z	O	86	33	S	NE
329	Z	O	87	34	S	NE
330	Z	O	88	35	S	NE
331	Z	O	89	36	S	NE
332	Z	O	89	36	S	NE
333	Z	O	89	36	S	NE
334	Z	O	90	37	S	
335	Z	O	91	38	S	NE
336	Z	O	92	39	S	NE
337	Z	O	92	39	S	NE

338	Z	O	93	40	S	NE
339	Z	O	94	41	S	NE
340	Z	O	94	41	S	NE
341	Z	O	95	42	G	SR
342	Z	O	95	42	S	TU
343	Z	O	95	42	G	
344	Z	O	95	42	G	NE
345	Z	O	95	42	G	
346	Z	O	95	42	G	NE
347	Z	D	96	43	G	SR
348	Z	D	96	43	G	SR
349	Z	O	97	44	G	
350	Z	O	97	44	G	
351	Z	O	97	44	G	
352	Z	O	97	44	G	
353	M	O	98	54	S	SR
354	M	O	98	54	S	SR
355	M	O	98	54	S	SR
356	M	O	99	55	G	SR
357	M	O	100	56	G	
358	M	O	101	57	G	
359	M	O	101	57	G	NE
360	M	O	102	58	G	
361	M	O	103	59	G	
362	M	O	104	60	G	
363	M	O	105	61	G	SR
364	M	O	106	62	G	
365	M	O	107	63	G	
366	M	O	108	64	G	SR
367	M	O	109	65	G	
368	M	O	110	66	G	SR
369	M	O	111	67	G	SR
370	M	O	112	68	G	
371	M	O	113	69	G	
372	M	O	113	69	G	
373	M	O	113	69	G	
374	M	O	114	70	G	SR
375	M	O	115	71	G	
376	M	O	115	71	S	NE
377	M	O	115	71	S	SR
378	M	O	116	72	S	SR
379	M	O	116	72	S	SR
380	M	O	116	72	S	

381	M	O	116	72	S	
382	M	O	117	73	S	SR
383	M	O	117	73	S	
384	M	O	117	73	S	
385	M	O	117	73	S	
386	M	O	117	73	S	
387	M	O	117	73	S	
388	M	O	117	73	S	
389	M	O	117	73	S	SR
390	M	O	117	73	S	NE
391	M	O	118	74	S	SR
392	M	O	119	75	S	SR
393	M	O	119	75	S	
394	M	O	120	76	S	
395	M	O	120	76	S	
396	M	A	121	77	G	SR
397	M	O	122	78	G	SR
398	M	D	123	79	G	
399	M	O	124	80	G	SR
400	M	O	124	80	G	SR
401	M	O	124	80	G	SR
402	M	O	124	80	G	SR
403	M	O	124	80	G	SR
404	M	O	124	80	G	SR
405	M	O	124	80	G	SR
406	Z	O	125	45	G	SR
407	Z	O	125	45	G	SR
408	Z	O	126	46	G	SR
409	Z	O	127	47	G	SR
410	Z	O	128	48	G	
411	Z	O	129	49	G	SR
412	Z	O	129	49	G	SR
413	Z	O	129	49	G	NE
414	Z	O	129	49	G	
415	Z	O	130	50	G	SR
416	Z	O	130	50	G	SR
417	Z	O	131	51	G	
418	Z	O	131	51	G	
419	Z	O	131	51	G	TU
420	Z	O	132	52	G	
421	Z	O	133	53	S	SR
422	Z	O	133	53	S	SR
423	Z	O	133	53	S	

424	Z	O	133	53	S	NE
425	Z	O	133	53	S	SR
426	Z	O	133	53	S	SR
427	Z	O	134	54	S	
428	Z	O	134	54	S	SR
429	Z	O	134	54	S	
430	Z	O	134	54	S	
431	Z	O	135	55	S	
432	Z	O	135	55	S	SR
433	Z	O	135	55	S	
434	Z	O	135	55	S	
435	Z	O	135	55	S	
436	Z	O	136	56	S	
437	Z	O	137	57	S	
438	Z	O	137	57	S	NE
439	Z	O	137	57	S	SR
440	Z	O	137	57	S	
441	Z	O	137	57	S	
442	Z	O	137	57	S	SR
443	Z	O	137	57	S	SR
444	Z	O	137	57	S	SR
445	Z	O	137	57	S	TU
446	Z	O	137	57	S	
447	Z	O	137	57	S	SR
448	Z	O	138	58	S	SR
449	Z	O	139	59	S	SR
450	Z	O	139	59	S	SR
451	Z	O	139	59	S	
452	Z	O	139	59	S	
453	Z	O	139	59	S	SR
454	Z	O	139	59	S	SR
455	Z	O	139	59	S	SR
456	Z	O	139	59	S	SR
457	Z	O	140	60	S	SR
458	Z	O	140	60	S	
459	Z	O	140	60	S	SR
460	Z	O	140	60	S	
461	Z	O	140	60	S	
462	Z	O	140	60	S	
463	Z	O	140	60	S	
464	Z	O	140	60	S	
465	Z	O	140	60	S	SR
466	Z	O	140	60	S	SR

467	Z	O	140	60	S	SR
468	Z	O	140	60	S	
469	Z	O	141	61	S	
470	Z	O	141	61	S	
471	Z	O	141	61	S	
472	Z	O	141	61	S	
473	Z	O	142	62	S	
474	Z	O	142	62	S	
475	Z	O	143	63	G	TU
476	Z	O	143	63	G	
477	Z	O	143	63	G	SR
478	Z	O	143	63	G	SR
479	Z	O	143	63	G	SR
480	Z	D	144	64	G	SR
481	Z	D	144	64	G	SR
482	Z	D	144	64	G	SR
483	Z	D	144	64	G	SR
484	Z	O	145	65	G	SR
485	Z	O	145	65	G	SR
486	Z	O	145	65	G	SR
487	Z	O	145	65	G	SR
488	Z	O	145	65	G	SR
489	Z	O	145	65	G	SR
490	Z	O	145	65	G	TU
491	Z	O	145	65	G	
492	Z	O	145	65	G	
493	Z	O	145	65	G	
494	Z	D	146	66	G	SR
495	Z	D	146	66	G	SR
496	Z	D	146	66	G	SR
497	Z	D	146	66	G	SR
498	M	O	147	81	G	ST
499	M	O	148	82	G	
500	M	O	148	82	G	LJ
501	M	O	148	82	G	ST
502	M	O	149	83	G	
503	M	O	149	83	G	ST
504	M	O	150	84	G	
505	M	O	151	85	G	ST
506	M	O	151	85	G	ST
507	M	O	151	85	G	
508	M	O	152	86	G	ST
509	M	O	152	86	G	

510	M	O	153	87	G	
511	M	O	154	88	G	
512	M	O	155	89	G	
513	M	D	156	90	G	ST
514	M	O	157	91	G	
515	M	D	158	92	G	
516	M	O	159	93	G	
517	Z	O	160	67	G	
518	Z	O	160	67	G	
519	Z	O	161	68	G	ST
520	Z	O	161	68	G	ST
521	Z	O	162	69	G	ST
522	Z	O	163	70	G	
523	Z	O	164	71	G	ST
524	Z	O	165	72	G	ST
525	Z	O	166	73	G	
526	Z	O	167	74	S	ST
527	Z	O	167	74	S	
528	Z	O	167	74	S	
529	Z	O	167	74	S	
530	Z	O	167	74	S	NE
531	Z	O	167	74	S	
532	Z	O	167	74	S	
533	Z	O	167	74	S	ST
534	Z	O	168	75	S	
535	Z	O	168	75	S	
536	Z	O	169	76	S	ST
537	Z	O	170	77	S	
538	Z	O	170	77	S	
539	Z	O	171	78	G	ST
540	Z	O	171	78	G	TU
541	Z	O	171	78	G	
542	Z	O	171	78	G	ST
543	Z	O	171	78	G	ST
544	Z	O	171	78	G	
545	Z	O	171	78	G	
546	Z	O	171	78	G	ST
547	Z	D	172	79	G	ST
548	Z	D	172	79	G	ST
549	Z	D	172	79	G	ST
550	Z	D	172	79	G	
551	Z	D	172	79	G	ST
552	Z	D	172	79	G	

553	Z	D	172	79	G	ST
554	Z	D	172	79	G	
555	Z	D	172	79	G	
556	Z	D	172	79	G	
557	Z	D	172	79	G	ST
558	Z	D	172	79	G	ST
559	Z	D	172	79	G	ST
560	Z	D	172	79	G	ST
561	Z	D	172	79	G	ST
562	Z	O	173	80	G	ST
563	Z	O	173	80	G	
564	Z	O	173	80	G	
565	Z	D	174	81	G	ST
566	Z	D	174	81	G	ST
567	Z	D	174	81	G	
568	Z	D	174	81	G	ST
569	Z	D	174	81	G	
570	M	O	175	94	S	TU
571	M	O	175	94	G	TU
572	M	O	176	95	G	
573	M	O	177	96	G	TU
574	M	O	178	97	G	
575	M	O	178	97	G	
576	M	O	179	98	G	
577	M	O	180	99	G	TU
578	M	O	181	100	G	NE
579	M	O	182	101	G	TU
580	M	O	182	101	G	NE
581	M	O	182	101	G	
582	M	O	183	102	S	
583	M	O	183	102	S	
584	M	O	183	102	S	
585	M	O	183	102	S	TU
586	M	O	183	102	S	TU
587	M	O	184	103	S	
588	M	O	185	104	G	
589	M	O	185	104	G	TU
590	M	O	185	104	G	TU
591	M	O	185	104	G	TU
592	M	O	185	104	G	TU
593	M	O	185	104	G	TU
594	M	O	185	104	G	
595	M	O	185	104	G	TU

596	M	O	185	104	G	TU
597	M	O	185	104	G	TU
598	Z	O	186	82	S	TU
599	Z	O	186	82	S	TU
600	Z	O	186	82	S	TU
601	Z	O	186	82	S	TU
602	Z	O	186	82	S	TU
603	Z	O	186	82	S	TU
604	Z	O	186	82	S	TU
605	Z	O	186	82	S	TU
606	Z	O	186	82	S	TU
607	Z	O	186	82	S	TU
608	Z	O	186	82	S	TU
609	Z	O	186	82	S	TU
610	Z	O	186	82	S	TU
611	Z	O	186	82	S	TU
612	Z	O	186	82	S	TU
613	Z	O	186	82	S	TU
614	Z	O	186	82	S	TU
615	Z	O	186	82	S	TU
616	Z	O	186	82	S	TU
617	Z	O	186	82	S	TU
618	Z	O	186	82	S	TU
619	Z	O	186	82	S	TU
620	Z	O	186	82	S	TU
621	Z	O	186	82	S	TU
622	Z	O	186	82	S	TU
623	Z	O	186	82	S	TU
624	Z	O	187	83	G	TU
625	Z	O	188	84	G	TU
626	Z	O	188	84	G	TU
627	Z	O	189	85	G	TU
628	Z	O	190	86	G	TU
629	Z	O	191	87	G	
630	Z	O	191	87	G	TU
631	Z	O	191	87	G	LJ
632	Z	O	191	87	G	TU
633	Z	O	192	88	G	TU
634	Z	O	192	88	G	TU
635	Z	O	193	89	G	
636	Z	O	193	89	G	
637	Z	O	193	89	G	
638	Z	O	194	90	G	TU

639	Z	O	194	90	G	
640	Z	O	195	91	G	TU
641	Z	O	195	91	G	
642	Z	O	195	91	G	TU
643	Z	O	196	92	G	TU
644	Z	O	197	93	G	
645	Z	O	197	93	G	TU
646	Z	O	197	93	G	
647	Z	O	197	93	G	TU
648	Z	O	198	94	S	TU
649	Z	O	198	94	S	TU
650	Z	O	198	94	S	NE
651	Z	O	198	94	S	
652	Z	O	198	94	S	TU
653	Z	O	198	94	S	TU
654	Z	O	199	95	S	TU
655	Z	O	199	95	S	TU
656	Z	O	199	95	S	TU
657	Z	O	199	95	S	TU
658	Z	O	199	95	S	TU
659	Z	O	199	95	S	TU
660	Z	O	199	95	S	TU
661	Z	O	200	96	S	TU
662	Z	O	200	96	S	TU
663	Z	O	200	96	S	
664	Z	O	201	97	S	
665	Z	O	202	98	G	TU
666	Z	O	202	98	G	
667	Z	O	202	98	G	TU
668	Z	O	202	98	G	TU
669	Z	D	203	99	G	
670	Z	D	203	99	G	TU
671	Z	O	204	100	G	TU
672	Z	O	204	100	G	TU
673	Z	O	204	100	G	TU
674	Z	O	204	100	G	

12.2 Privitak — Transkript za pojedinu snimku

ID snimke	Transkripcija
1	PA JE LI JE LI MOJA SESTRICNA MOJ ROĐAK VRBANOVIĆ PA JESMO LI MI KORLEONEOVI
2	SAD ZAMISLITE JA SE PETLJAM U PRIVATAN ŽIVOT A TI SE NE PETLJAŠ
3	JA TAK PRIČAM KAK SAD DA JA SJEDIM TU I DA JA
4	PA ČEKAJ MALO PRVO NISAM NIŠTA NIKOM NAPRAVIJA
5	ZAŠTO SI TI DOŠLA JA TI RIČ NISAM REKA SAMO ME TO ZANIMA
6	JA SE NE PONAŠAN PREMA TEBI AGRESIVNO
7	OSTAVI ME NA MIRU
8	MA IDI U PIČKU MATERINU JEBEM TEBE I TVOJE HRKANJE SVINJO
9	PUŠI KURAC
10	VAS NIŠTA NISMO PITALI ZAŠTO STE NAS ZVALI
11	MA KO SE BOJI MAMICE PIČKA TI MATERINA SMRDLJIVA AJD DOLAZI AJDE
12	A U PIZDU MATERINU
13	TI SE NE JAVLJAJ TI SI SVOJ ŽIVOT LIJEPO ZAPIO
14	NEMA MRDANJA DOK MAJOR NE DOĐE
15	ŠTA JE ROKO ŠTA OPET OČEŠ NA IZDAJU DA ME NAGOVARAŠ
16	ŠTA TI OVDE RADIŠ OSTAVI TO I NA POLOŽAJ
17	DAJ NE IGRAJ SE GORE TIM PIŠTOLJEM NEGO ČISTI
18	REKO SAM TI DA OSTAVIŠ
19	DAJTE ČOVIČE VEĆ STE MI SE POPELI III_ DA NE KAŽEM DI
20	MA IVEK KAJ JE TEBI KOLKO SAM TI PUTA REKAO DA NE UPADAŠ DOK DELAM
21	MA KAKVE SLIKE ČOVJEČE BOŽJI NE MOGU SAZNATI NI ŠTO MI JE S
22	AKO MI ŽENA TO VIDI DOŽIVJET ĆE SPONTANI I ONDA SMO NAJEBALI I TI I JA JEL TI JASNO
23	OVO TI JE ZADNJI PUT DA SI ME OVAKO PREŠO
24	PUN MI JE KURAC OVIH SELJAČKIH SRANJA
25	MA KO TI TO VIŠE SLUŠA ŠTO ZVUČI KO BLEJANJE OVACA
26	SMIRIT ĆU JA TEBE DVA METRA POD ZEMLJOM
27	DRUGI PUT ĆU PUŠKU UZET
28	SLUŠAJ JOŠ JEDNU ZUCNI UBIT ĆU BOGA U CEBI
29	NIJESAM JA SVE OVO NAPRAVIJA DA BI SAD PODVIJA REPA
30	E NEBI BILO DOBRO PIČKA MU MATERINA TODORE NE BI BILO DOBRO SLUŠAŠ LI ŠTA TI JA GOVORIM
31	MA DA MI JE OTAC SIN TETKA IDH_ NE DAM PARE
32	DA I NEMOJTE ME MOLIM VAS VIŠE ZAJEBAVAT S TIM STVARIMA UMJESTO DA SMO SRETNI ŠTO GA IMAMO ZA GOSTA VI TU STVARATE NERVOZU
33	A ŠTA SAMO TEBE POSLALO TAKVA PRIJAVA A TI SAM PA DOBRO JESI TI NORMALAN ŠTA SI TI NJIMA NAPISAO U TOJ PRIJAVI
34	JESAM TI REKA DA MORA BIT OZBILNO
35	NE MOGU JA KA GRADONAČELNIK SVE SAM
36	ŠTA JE MARINKO DOŠA SI U MOJU KONOBU DRŽAT SVOJE PARTIJSKE GOVORE JE LI
37	AJDE IZLAZI VANKA I DA TE ODE VIŠE NISAM NI VIDIJA NI ČUJA
38	MA JEBI SE I TI I TITO I NJEGOV DUH I CILI CENTRALNI KOMITET PEDESET GODINA SI NAS TLAČIJA I SAD BI TIJA OPET GOVNO JEDNO CRVENO MARŠ
39	JESAM TE PIČKA TI MATERINA A U KURAC KRASAN SAD SAM NOVINE ZJEBAL
40	A BOLI ME KURAC I ZA JASNU I ZA PERU I ZA MAGAZIN KAJ ĆU TAM
41	MA NABIJEM TE NA KURAC I TEBE I MAGAZIN NIKAM TI NE IDEŠ JEL TI JASNO
42	KAJ JE TO FORA NEKA TI KAO PROTESTIRAŠ HA_ SAMO MI NEMOJ ZVAT ONU KUĆU ZA PIČKE I PEDERE JER ĆU IM SVIMA JEBAT MATER JEL TI JASNO
43	KAJ SE VAS DVOJICA SMIJETE JEBEM VAM MATER NARKOMANSKU E PA TREBA TU JEDNOM RIJEŠIT STVARI VIŠE DA POKAŽEMO OVOJ BAGRI ČETNIČKOJ DA SU MI SE POPELI NA VRH KURCA A
44	NJEMU JE STARI SRBIN NE ZNAM KOJI JE TVOJ PROBLEM ALI STE MI OBA DVOJICA SUMNJIVI KAJ JE KURAC DI STE BILI U RATU KAJ JE
45	OVO TI JE OLTAR DOMOVINE PIZDO KLEKNI I LJUBI
46	DVINA JA SE UOPĆE NE SLAŽEM S TOBOM
47	SADA OTIŠLA DO SAD SI JU LJUBILA UDH_ KAD SU LJUDI GOVORILI U NJENU KORIST
48	SAD JA NE ŽELIM S TOBOM BITI UDH_
49	TOLIKO PODLOSTI KAO OD TEBE JA U ŽIVOTU DOŽIVIO NISAM
50	OPET KREĆEŠ U SVOJE LOBIRANJE UDH_ I TO CIJELO VREME RADIŠ
51	KAKO JE ONA LAGALA HANA DA TI TO ZNAŠ
52	IMA GLAS ONO UDH_ JA ŠTA DA RADIM
53	JA JU GLEDAM BRATE MILI
54	NE VJERUJEM ŠTA DA RADI
55	SHVAČAŠ ME ŠTA DA ONA RADI

56	DA ONA FALSIFICIRA UDH_ JA SAM SE OSUPNUO
57	A JA SAM SE ONO ZABEZEKNO JA VELIM
58	ŠTA DA UČINIM
59	DA TE RASKRINKAM ŠTA RADIŠ
60	ŠTA JE OVO ŠTA JE OVO SLADO
61	PA NE TREBA MI TOLKA MAST
62	PA TO JE ZA PET PUTA DAJ SMANJI MAST BRZO
63	PA DOBRO TKO ĆE MI ODBIT MAST
64	SAMO JEDINI POSAO KOJI MI RADIMO OVDJE DA VIŠE NEMA NITI JEDAN DRUGI POSAO NA OVOJ KRVAVOJ FARM
65	MA IDI U PIČKU MATERINU JEBEM TEBE I TVOJE HRKANJE
66	MA IDI U KURAC
67	MA IDI U KURAC
68	TI NISI NORMALAN
69	PA NE MOGU JOJ PUKO SAM OPROSTITUTE SVI
70	PA DAJ NEMOJ MOLIM TE PRDIŠ BRAJO NA USTA BOG TE MAZO
71	UPALI SI KAMERU USTVRDIT ĆEŠ ŠTA RADIŠ
72	MA TRAŽI SUTRA SNIMKE MOLIM TE
73	MA JEBE MI SE ZA TEBE MAJKE MI
74	MA BJEŽI BRATE AJDE
75	MA JA ĆU BJEŽAT NARAVNO OD TEBE
76	BRANKA NEMOJ MENI GOVORIT DA BUDEM OBJEKTIVAN AJ MOLIM TE
77	JA TE MOLIM DA SE NE RAZGOVARAŠ TAKO SA MNOM JEL TI JASNO
78	NEMOJ MENI STAVIJAT RIJEČI U USTA
79	MENE APSOLUTNO BOLI NEKA STVAR VIŠE TKO ĆE ŠTA JEST
80	NE GOVORIM TEBI
81	SLUŠAJ PRVO NISMO ISTO GODIŠTE DRUGO NISMO ISTE ZRELOSTI UDH_ I TREĆE NE MISLIM RAZGOVARAT S TOBOM NA TAKAV NAČIN ZNAŠ JER KAD BI POŠTIVALA MOJ RAD VJEROVATNO SE NE BI TAKO PONAŠALA I MOJE UDH_ ODNOS PREMA LJUDIMA
82	ZNAŠ KO JE TREBO DOŽIVJET ŠOK ONAJ KO JE STVARNO POGRJEŠIO
83	IMENOVA GA DESET PUTA
84	DA MU U DRUGIM IGRAMA KOJE NISI GLEDA
85	ONDA DOĐEŠ I PITAŠ
86	PA ZAR TI NIJE ONDA MALO ČUDNO
87	ALI KAD TI JE NEŠTO ČUDNO ONDA ODEŠ I
88	SAD ZAMISLI KAKO JE MENI JA NISAN REKO NIŠTA AL DOŽIVIN OVO
89	AL ZAMISLI KAKO JE MENI STARA MOJA IDH_ AJD ZAMISLI KAKO JE MENI
90	SVEGA JA BI NA NJEGOVOM MJESTU NENAD NAPRAVIO ISTO DA JE OVA SITUACIJA ISTO BI NAPRAVIO NA NJEGOVOM MJESTU
91	AKO ĆE TE NETKO ZBOG TOGA GLEDAT KRIVO
92	MA KO O ČEMU PRIČAŠ TI
93	ONDA SI STARIO STAVIO KRIVU OSOBU A TO JE PTO JE TVOJE VIĐENJE JEBEM MU VRAGA
94	ALI TO NIJE POŠTENO PREMA I TEBI I MENI I NJEMU I ODREĐENIM OSOBAMA JER SE TA UPORNO OSOBA IZVLAČI NIJE POŠTENO A TI SVE GLEDAŠ NA POŠTEN NAČIN TO NIJE POŠTENO
95	ZAŠTO BI ON BIO SLUGA A TO NIJE ZASLUŽIO ZAŠTO NIJE POŠTENO JEBEM MU VRAGA AKO NIJE POŠTENO NIJE
96	PA ČEKAJ PA MOGU LI ZAVRŠIT
97	ZASTAJK_ BILO BI DVOLIČNO OD MENE DA SAM JA TO TU REKAO A ONDA SE NJOJ SMIJAO JA SAM DOŠAO DO HANE I REKO SAM HANA JA SAM REKAO UDH_ KAD SI IZAŠLA VAN UDH_ DA SI IZAŠLA VAN UDH_ JER ONDA NISI IMALA ŠTO ZA REĆ
98	PA ZATO JER TI JE TO SUVIŠNO KUIŠ
99	JER JA SAM REKAO PA ČOVJEČE KAD OD PRVOG DANA MI SMO USPJELI RJEŠIT SVE I STALNO STALNO VIDIŠ PRIMJERICE U IGRI JUČER AJ RAFO TI NACRTAJ TO E
100	PA ISMIJAVA LJUDE BEZVEZE KUŽIŠ AL NEZNAM ZAŠTO JOJ TO TREBA MOŽE BIT SUPER SVE PET
101	RUŽNO MI JE TO JAKO MI JE TO RUŽNO
102	NE JA SE NE LJUTIM STALNO KUŽIŠ
103	OVO JE SAD BILO BEZOBRAZNO
104	DIŽI SE JEBEN TI MATER DROGERAŠKU DIŽI SE JEBEN TI JA MATER AJDE TREBA ROBU ISPORUČIT
105	JEBEN TI MATER ŠTAJE POGLEDAJ SI KAKAV SI JEBEŠ ME JA ĆU NJIH POBIT KO ZEČEVE
106	ŠTA SE TI DEREŠ
107	NEMOJ OPET JEBOTE
108	E JESI BLESAV CIGO
109	A VALJDA ĆEMO DOBIT JEBEŠ IN MATER
110	NAŠI NEZNAJU DA IGRAJU GLAVOM A GLAVOM SE IGRA
111	A SVAMA TREBA KO SA STOKOM
112	AJDE
113	OBIČNA SAM LUDA BIO KAD SAM GLAS KAD SAM GLAS ZABORAVIO GRRR_

114	TI ŠAPO CRNA IZDALA SI ME
115	A ONDA ŠAPA CRNA I PROPADE SVE TRAS
116	ZA SVE SU OVO KOZLIĆI KRIVI I ZATO NEĆE DUGO OSTATI ŽIVI
117	KAD ĆU ONDA NARASTI KAO VI PROTIV MENE STE SVI SVI
118	ŠTO GOD HOĆEŠ RADI ALI VIŠE NE GNJAVI
119	NEĆU BAŠ NEĆU
120	ZAVRŠIT ĆEŠ U ŽUTU KUĆU JEBA MI BOG MATER KAD MORAN TAKO REĆ
121	JEBA TE DOTUR DA TE JEBA OŠ TI TO PLATIT JELI ILI TVOJ ČAČA JEBA TE ON DA TE JEBA DA GA TRUJU
122	A I ZVONIK MI SAZIDAJTE DA SE LIPO MOGU BACIT S NJEGA
123	SE KURVAT ĐAVA TE ODNIIJA DA TI OD SVOJE PROTEZAT ZVONIK
124	ISPOVID TI NE VRIDI NAŠ NEGO AJDE U KUPINE ŽUPNIKU PA SE NJEMU ISPOVIDI S OVIM NEMAN NIKAKVE VEZE
125	U PIČKU MATERINU SVI SERU KAK NEMAJU LOVE A SVAKI DAN SVE VIŠE I VIŠE AUTI U GRADU KAK JE TO MOGUĆE
126	A JEBO GA TI JURICA SVE JA KUŽIM KAJ TI HOĆEŠ REĆ
127	DAJTE SE NOSITE U KURAC NARUČITE BILO ŠTA GLADAN SAM KO PAS
128	MAJKU TI JEBEM JOŠ ĆU TI SE JA KRVI NAPIT
129	JA GROŠA VOLIM KO ROĐENOGA SINA A ON BEZOBRAZAN KO PAS I STALNO ZAPIŠAVA U KREKET I ONDA PLAČE
130	I PO NOĆU KAD MU RADIM NA MAJKU TULI KO BIK JEBEM GA LUDA
131	TI MI NISI MAMA
132	BAŠ ME BRIGA
133	KOGA JEBE
134	DAJ ME PUSTITE NA MIRU JA JESAM GLUP ALI ZNAM DA MI JE TATA MRTAV
135	JE MRTAV JE JA IDEM GORE KNJEMU A NE ON MENI
136	NE IDEM DOMA I GOTOVO PUSTIM ME S TIM DOMA
137	PUSTI ME NA MIRU BAŠ ME BRIGA
138	MENE NIKO NE VOLI NI TI ME NE VOLIŠ VIDIO SAM TE I SRAM TE BILO
139	SRAM TE BILO VIDIO SAM TE
140	KOGA JEBE
141	VAN
142	ČITAV ŽIVOT SLUŠAM BOLI ME ZUB BOLI ME ZUB
143	BOLI ME KURAC
144	JE LI OVAJ UZMI SI GA ZA PO DOMA
145	PIČKA TI MATERINA
146	ANAMARIJA JEBEM MU MATER IZGLEDA DA MORAM NA OPERACIJU
147	MA ŠTA MOJE ŠTA SEREŠ TI GLE SLUŠAJ BURAZ ZASTAJK_ NE IDE TO VIŠE TAKO DAJ SLIKE DAJ KURTONE
148	KKK_ KOJE MALE I O KAKVOJ MALOJ TI SAD PRIČAŠ
149	ALI STALNO MI TO RADIŠ JEBOTE SVE UVIJEK ZA TEBE SVE SAMO TVOJE SAMO TI TI I TI
150	A JEBALE SE I TEBI I MOJOJ ŽENI I STAROM SVIMA VAM SE JABALO ZA MENE JA SAM BIO NITKO JEL DA A ZNAŠ ZAŠTO ZATO ŠTO SAM BIO DOBAR BRACO JE BIO DOBAR A VI STE TO SVI ISKORIŠTAVALI I ZATO JE BRACO TU DI NIJE A TI SI TU DI JESI ALI NIJE MI ŽAO NE NIJE MI ŽAO GLE UZMI NOSI AJMO
151	ZNAŠ ŠTO AJDE UZMI UZMI UZMI SVE NOSI I IDI JEBO TE KAPITALIZAM DA TE JEBO
152	NE NE I TREBAŠ MI NIŠTA OSTAVLJAT MENI NE TREBAJU TVOJI NOVCI JESI ME ČUO
153	TIČE ME SE JEBIGA I STVARNO NISI FER STALNO SEREŠ DA SAM JA ZA SVE KRIV UMJESTO DA MI KAŽEŠ BRACO ODJEBI U SKOKOVIMA KUPILA SAM SI MLAĐEG KOMADA
154	A OPROSTI A ČEGA SE ON OZBILNO UHVATIO
155	A I IMA VELIKI KURAC JEL DA
156	GRRR_ GLAD SE JAVLJA JESTI JESTI MOGAO BIH JEDNU PRIJE ALI BOLJE OBA DVIJE OBA DVIJE
157	ZAŠTO OVAKO ZAŠTO ONAKO NA SVE ČEŠ ODMAH ODGOVOR ZNATI U VREĆU ĆU TE SAD STRPATI A ONDA ĆU TE PROGUTATI
158	DA OPERE TO ŠTO TI SVIMA NAMA SVAKODNEVNO GOVORIŠ
159	NE UZIMAJ MOJE IME U USTA JER MI JE MUKA OD TEBE GADI MI SE VIŠE TVOJ GLAS
160	TO JE ČOVJEK S KIM MOŽEŠ KOM ENEŠTO MOŽEŠ OBJASNIT A TEBI NIKO NIŠTA NE MOŽE OBJASNIT
161	DOK JE HANA MELA I ĐERALA SE ZA HIGIJENU I REKLA DA DAJE TAMO IZJAVE PITALA SAM ŠTA DAJE
162	BOSANKA KOJA IZGLEDA PRLJAVO
163	MAKNI SE I TI ONDA
164	JA TEBI PRIČAM NEŠTO TI PRIČAŠ NEŠTO DRUGO
165	KAD SI REKO SAMO SMIRI SE
166	KO DA JA NE ZNAM ONO DA TO TREBA SPREMITI
167	ČINJENICA JE DA SAM SE JA ZAKAČILA VEČERAS SAMO S TEBON A TI SI SE S TRI OSOBE USPIJA POKAČIT
168	JA NE BI SAD UPLITALA VIOLETU I SVE OSTALE ODE MOLIM TE ALI SVI IMAJU OČI
169	TI SI SE U PAR NAVRATA ZADERA ONO TRI PUTA ČETRI
170	PROKLET BAZEN MENE SI POSLA U NEKU STVAR TO MENI NE TRIBA U ŽIVOTU
171	NEEĆU JER JA NIKOME NIŠTA NISAM ODE NAPRAVILA
172	JER JA NISAM ODE SMEĆE EJ_
173	POGLEDAJ SVOJE PONAŠANJE A MENI GOVORIŠ DA SAM BALAVA
174	KUŽIŠ

175	RECIMO ZAŠTO MINISTAR KULTURE BIŠKUPIĆ NE UZME GLAZBENU PROIZVODNJU
176	MA DAJ ODJEBI TEK SAM SE PROBUDILA JOŠ NISAM NI DORUČKOVALA
177	NA DANIJELA NE MISLIŠ
178	NIŠTA MAMA VAKO UPADATI MALOJ U SOBU KO JE TO VIDIJA
179	JOJ NEMOJ MI SE NA DJECU IZVLAČIT BUDI SRETNDA DA TI NISAM LUPILA KAMATE
180	MARŠ KURVE TREBA IH NAUČITI PAMETI
181	GUBITE SE U PIČKU MATERINU NAJEBAT ČETE OBA DVIJE KURVETINE
182	PUSTI ME SAMA SI KRIVA ON TE DRŽI ZA RUKU TI S NJIM RAZGOVARAŠ ZAŠTO GA NISI POSLALA U KURAC
183	NA KOG BI JA TO TREBALA BIT LJUBOMORNA NA LIDIJU MOŽDA
184	ŠTA SI JE TAJ MALI UMISLIO DA MOŽE SMRC_ SVAKU ŽENSKU KOJU VIDI PIPAT
185	ŠTA JE ŠTA VAS DVIJE KURAC BULIJE ŠTA VAS MUČI
186	A SLUŠAJ TI ROKO NEKA JE JA SAMO JOŠ JEDNOM VIDIM NA OVOJ POZORNICI IMAT ČEŠ POSLA S MENON
187	NEĆU NEĆU I NEĆU NE PRIZNAJEM JA NIKAKAV DOGOVOR
188	DA NISI MOŽDA MALO URANIO KOJIH ČETRDESET MINUTA
189	DA JE INZISTIRAO DA GA BACIŠ S NEBODERA TI BI GA BACIO JEL
190	JESI TI NORMALAN ČOVJEČE ŠTA SI MISLIO HM ŠTA SI MISLIO DA ĆU MU POPUŠIT
191	E PA NE TREBA MI TVOJA ZAŠTITA SKORO SAM DOBILA OTKAZ ZBOG TVOJE GLUPOSTI
192	JA TE NEĆU MASIRAT
193	NEČEŠ ONDA ČEMO S TOBOM DRUKČIJE NEZAHVALNICE
194	DAL SAM PITALA DAL JE SPORTAŠ ILI NIJE PA NISAM GLUPA ČOVJEČE
195	PA DOBRO JA SAM SAD SJELA I SAD SE NE MISLIM ODMAH DIGNUT
196	PA OČEMO PA JA NISAM REKLA DA NEĆE
197	PA NEĆU SADA A NEGO A ODI
198	PA NISAM REKLA DA NEĆU TEK SMO SAD SJELA I PUSTI
199	HANA TI JE PREDLAGALA ZASTAJK_ U STVARI DOGOVARALE STE SE KAKO ČETE KAKO BI BILO NAJBOLJE DA SE IZBACI ALEKS
200	PA ZATO ŠTO NORMALNO DA MI SMETA JEL NAS ŽIVI TU ČETRNAEST RAZLIČITIH
201	TI FINO TO JE ISPALO UŽASNO I TO SMO SVI SKUŽILI VANI DA JE ONA ZAKUHALA SVI
202	ZASTAJK_ O ČEM PRIČAMO ALI OD KUD JE PRAVO DA TAK NEŠ IZJAVI
203	I DA SU TO ZA NJU LJUDI DRUGOG REDA ILI TAKO NEŠTO KAKO TO MOŽEŠ REĆI
204	NAPISAO PAPIRIĆ KOJI JA DRŽIM OVDJE KOD SEBE
205	JEBOTE JA MISLIM DA SU DVINA SEFIR I ZASTAJK_ STIPE PRIČALI UDH_ ISPALO TI DA JA DA KOJI JE MENI KURAC DA UDH_ JA PODRŽAVAM ZASTAJK_ JOSIPU OVO ONO I ONDA I ONDA SAM SJELA SA STIPOM I REKLA DOBRO STIPE KOJI JE TEBI KURAC
206	ISPALO TI DA ZNAŠ ZAŠTO JER SAM JA
207	A ZNAŠ ZAŠTO UDH_ ZATO ŠTO SAM JA BILA DVA TJEDNA JEBENIH U KUHINJI MENE ONI KOJI SU BILI GAZDE S NJIMA SAM U KUHINJI JA PRALA SUĐE I ŠTA DA NE PRIČAM SA ŽENOM MA JA NISAM NIŠTA NAPRAVILA
208	I ONDA TI JE ISPALO DA KUŽIŠ SAM JA U KUHINJI DA ŠTA SE DA KOJI SE KURAC MENI DEŠAVA MENI STIPE JA KAŽEM STIPE
209	ČUJEM TE
210	ŠTA JE KOJI KURAC
211	AL MI SE NEDA MI SE BORIT BAŠ NEMA ŠANSE
212	NIJE TU SREĆU ZNAŠ TI NEMA ŠANSE
213	ZASTAJK_ DO DO IVANE DA VIDIM ONO ŠTA JE ZA JESTI I KAD JE RUČAK NE IVANA JE REKLA
214	NE NE NE NEGO ONA TU PRIČA JEDNO A ONA JE VANI DRUGO
215	MOLIM
216	DVOLIČNA SI JEDNA OBIČNA PRASICA DVOLIČNA AE
217	NAJNEGATIVNIJA OSOBA U FARMU
218	MA JOJ LJUBAVI EJ_
219	SAD SMO DOŠLI DO GRANICE BEZOBRAZLUKA
220	BIJELO VINO NISAM
221	NEMOJ ME LAGAT
222	JA TVOJU ŠALU NA TO KAD JE BILO ŠTA SPOMENEM TI ODMA TI JE U GLAVI
223	JUHA PITA AJME MENI GLAD I NEZNAN ŠTA
224	NISAN REKLA NEČEŠ IH KIKI NEGO SAM REKLA MOŽDA NEBI TRIBALO BILO BI PREVIŠE TEKUĆINE NA LIP NAČIN A TI UDH_ DOBRO TTT_ TEBI TTT_ TI NE DOZVOLJAVŠ DA TI NIKAD NNN_ NIKO NIŠTA NE SUGERIRA U VEZI
225	STAJALA SI KRAJ ĐEZVE NE SERI KVAKE STAJALA SI
226	ON TI JE NAPRAVIO MJERU
227	ŠTA SE MIŠAŠ AKO SAM JA STAVILA
228	SE ODLUČI ŠTA ŽELIŠ JEL ŽELIŠ IĆ DOMA IL ŽELIŠ OSTAT
229	JEL VOLIŠ SEBE IL NE VOLIŠ SEBE JEL VOLIŠ ANTONIJA IL NE VOLIŠ ANTONIJA
230	NEMOJ SE LJUTIT
231	ISUSE ZNALO SE DA KAD IDEMO U ŠTALU DA SE POČISTI JEBOTE UVIJEK MORAM U GOVNIMA VEĆ DVA DANA MUZEM
232	ZATO ŠTO RASPLAČEM JER TI JE OTIŠAO S DVINOM SORI NIJE UŠAO U ŠTALU
233	JA VIŠE NEĆU BIT NA KRAVAMA

234	PA NISAM IŠLA RANIJE IŠLA SAM U ISTO VRIJEME NEGO SVE SAMA I POPODNE I UJUTRO TI ŠTO SI UDH_ UJUTRO ČISTIO ŠTALU ON NIJE UŠO
235	ONDA MI KAŽI BIT ĆU NA SVINJAMA PA DA JA ZNAM DA ONDA RADIM SVE
236	A SUPER ZATO ŠTO SAM JA TO REKLA DA JE NE REKO NETKO DRUGI NEBIH NIŠ
237	JA PUCAM JER TI JEDNOM NEŠTO KAŽEM
238	TI MENI GOVORIŠ JA PUCAM A JEDNOM SAM TI SAD OVO REKLA FAKAT MI SMETA
239	TI JOŠ JEDNOM TO RECI I NEĆU PRIČAT S TOBOM KUNEM
240	JA PET GODINA GODINA S NJIM NISAM NEGO ČETIRI MJESECA
241	OVAJ ZOVE PLAČE VOLIM TE OVAJ ISTO TO OVAJ SERE PO OVOM UDH_ JA KAŽEM DAJ OKANITE ME SE OBADVOJICA
242	MENI SAD NIJE ISTO
243	GLASAT ĆEMO TKO HOĆE DA IZAĐE VAN JA ILI BRANKA JA TU NEĆU S NJOM ŽIVJET
244	NEŠ TI NIKOG IZBACIT PAZI KOME SE PRIJETIŠ KIKI
245	MA JA BI JA SAM NJOJ REKLA LUPEŽICE
246	NEĆU NIŠTA NIJE NIKOM IŠTA GOVORILA KUŽIŠ ONO
247	NA NIJE ISTINA NEMA ŠANSE KUŽIŠ
248	DERALA SE RUČAK I TAKO SMO JA I NENO IZLETILI VAN
249	PODRAZUMJEVALA SAM DA ME JE TAK ČUO
250	POVIRIT KAD JE RUČAK IL DOČ VIDIT ŠTA SE RADI JA SAM TO KUHALA BAR POL SATA
251	NEDA MI SE GLEDAT NEKE LJUDE NEDA MI SE VIŠE I TO JE TO
252	JA IDEM VAN
253	MILJENKO JE BIO GOVNO TAJ JE JEBE ŠTA JE STIGA I MENE JE TIJO SAMO MU NISAM DALA ETO TI NA PA ZNAJ A U NJEMAČKOJ TI JE IMA DRUGU ŽENSKU ZA STALNO
254	POPE KUKAVICO DER_
255	PA KAKO ĆU SE ISPOVIDIT AKO MI VI NE DATE DA GOVORIN PA ČOVIČE BOŽJI PA NEMOGU SE JA NA MOTE ISPOVIDAT
256	MA ŠTA SI TI MENI DA TI SE JA IMAN OBAJŠNJAVAT ČAČA
257	DAJ RUTI KOJI TI JE JEBOTE KAJ OČEŠ OD MENE
258	KAK SI NE DAŠ DOKAZAT DA TI VELIN DA NEĆU VIŠE S TOBOM
259	TOTALNI SI DEGENERIK
260	PIŠTOLJ JE OPALIO MISLIM VAMA JEBOTE TREBA POL SATA DA SVATITE
261	BJEŽI OD MENE DALJE BAČIT ĆU SE MAJKEMI NE PRILAZI
262	NE ZANIMAM ME NIŠTA ONO ĐUBRE ME VARA SA TKO ZNA S KIM A JA PO CIJELE DANE RINTAM I BACIT ĆU SE
263	DOSTA MI JE
264	ŠTA TI GLEDAŠ DAJ JOŠ JEDNU
265	HEJ_ EJ_ EJ_ EJ_ DAJE OSTAVITE TU LETVU VANI
266	JA NE IDEM NIKAM CIJELI ŽIVOT TU SLUŽIM
267	JE LI TI TO MENI VIRIŠ KROZ PROZOR KAJ JESI TO REKO
268	TA DOSTA NE LJUTI ME
269	DONESI VODICU NALOŽI VATRICU
270	LJENJIVICE
271	NEK GORI SVE
272	UMUKNI
273	HOĆU DA VIDIM ŽAR ŽAR ŽAR
274	MAKNI SE JA ĆU
275	IDIOTI POLA ČETRI JE
276	MRŠ U PIČKU MATERINU
277	ĐAVA TE ODNJIA I KO ME NA TE NAMIRIO I URU I MINUT KAD SAN SE ZATE UDALA
278	DUŠE MI MOJE UBIT ĆU GA JESI TI ZNAO TU FUFU
279	VIDI ŠTA SI NAPRAVIO IDIOTE
280	DER_
281	IZAĐITE VAN REKLA SAM
282	NE TIČE TE SE
283	EVO VIDIŠ O TOME TI GOVORIM STALNO JE NEKO DRUGI KRIV NIJE BRACO HOĆU NEKOG TKO ĆE SE NEČEG OZBILJNO UHVATIT
284	IMA I VELIKI KURAC IMA KURČINU
285	JE LI TI MENE ČUJEŠ ŠTO JA TEBE PITAM HA_ NARAVNO DA NE ČUJEŠ KAD IMAŠ TA GOVNA NA UŠIMA
286	KRAVO GLUPA SAMO NA SEBE MISLIŠ DAJ ZOV HITNU BRZO
287	ŠTA GLEDATE DEBILI JEDNI
288	PRODAJA SKIJAŠKIH ARANMŽANA JE PORASLA ČAK DVADESET POSTO U ODNOSU NA PROŠLU GODINU ŠTO MOŽEMO ZAHVALITI PRVENSTVENO NAŠOJ BOLJOJ PRIPREMI ZA SEZONU
289	JEDNO SEDAM DANA VAM JE NEKIH TRISTO DO ČETRISTO KUNA TRISTOČETRDESET TRISTOPEDESET KUNA
290	TO VAM UKLJUČUJE PANCERICE I ŠTAPOVE NE UDH_ I TO VAM JE JOŠ I KACIGA
291	U HRVATSKOJ SE VRLO ČESTO GOVORI KAO EVO GA OPET SE SPAJAJU BLAGDANI OPET SE NE RADI OPET SE U TOM SUSTAVU NE RADI ONDA SE IZVLAČI TKO TO NE RADI ZAPRAVO A ONDA SE KAD SE UĐE DUBLJE PRVO I OSNOVNO TREBA VIDJETI DA RADNICI U SPAJANJU BLAGDANA NIKAD NE DOBIVAJU NEKAKVE DANE KOJI IM NE PRIPADAJU
292	TO JE PITANJE ZAPRAVO SVAKOG POSLODAVCA ILI AJMO REĆ RAVNATELJA U JAVNOJ VLASTI DA ODLUČI UDH_ DA LI JE

	ISPLATIVO I DA LI JE POSLOVNA SITUACIJA TAKVA DA OMOGUĆAVA RADNICIMA DA BUDU NA KOLEKTIVNOM GODIŠNJEM ODMORU
293	JEL PONEKAD JE MOŽDA BOLJE DA LJUDI BUDU KOD KUĆE I DA SE NE TROŠI NA GRIJANJE NA ČITAVU ONU OPERATIVU UDH_ FIRME AKO SE U OVOM RA VREMENU NIŠTA NE DEŠAVA I NEMA NIKAKVE POSLOVNE AKTIVNOSTI
294	NAJBITNIJI PROJEKTI A TO SU ZASTAJK_ PROJEKTI FINANCIRANI OD DRŽAVE KAO ŠTO SU AUTOCESTE STALI SU STO POSTO UDH_ VELIKE TVRTKE PRIV NEKOLKO DANA SU ZATVORILE VRATA POSLAJE LJUDE NA ČEKANJE NA GODIŠNJI UDH_
295	SVI SU SE BOJALI TO NIJE BILA PITANJE UDH_ SAMO PRVIH LJUDI NEGO KOMPLETNE ADMINISTRACIJE U POJEDINIM TVRTKAMA UDH_ JE BLOKIRANA NIKO NE ŽELI DONIJETI ODLUKU
296	ALI PRAVA INDUSTRIJA JE OSTALA I OSIM TOGA POLJSKA JE NAŠLA MOGUĆNOSTI DA PRIVUČE DOBAR DIO ZASTAJK_ ZASTAJK_ STRANIH INVESTITORA KOJI SU INVESTIRALI UPRAVO UUU_ PROIZVODNE POGONE
297	MOGU SE ODLUČITI ZA KASNIJI POČETAK U RUJNU MJESECU ILI MOGU PRODULJITI ZIMSKI RASPUST ZASTAJK_ ZA JEDAN TJEDAN U MJESECU SIJEČNJU
298	PROMATRAMO PERIOD OD DVIJE TISUĆE ZASTAJK_ DVIJE TISUĆE REC OD DVIJE TISUĆE SEDME UDH_ DAKLE TO JE PERIOD KADA JE NA NEKI NAČIN ZASTAJK_ HRVATSKA DOŽIVJELA JEDAN RECIMO TO TAKO BUM POTROŠAČKOG DRUŠTVA NE UDH_ DAKLE TO JE PERIOD KAD SE S JEDNE STRANE IZGRADIO VELIK BROJ ZAPRAVO TRGOVAČKIH CENTARA UDH_ S DRUGE STRANE KADA ZASTAJK_ SU ULAGANJA RECIMO U MARKETING OGLAŠAVANJE NAGLO RASLA I S TREĆE STRANE KADA SU ZAPRAVO ZASTAJK_ BANKE POSREDSTVOM RAZLIČITIH POTROŠAČKIH KREDITA NA NEKI NAČIN DEMOKRATIZIRALE PROCES ZADUŽIVANJA
299	JEDNOSTAVNO DA SMO NA NEKI NAČIN ZASTAJK_ SENZIBILIZIRANI NA TU NOVU POTROŠAČKU KULTURU DA ZASTAJK_ SMO ZAPRAVO POTROŠAČI U PUNOM SMISLU RIJEČI DAKLE DA SE NE RAZLIKUJEMO OD POTROŠAČA U AMERICI U ENGLJSKOJ NJEMAČKOJ ČEŠKOJ POLJSKOJ UDH_ ILI ZASTAJK_ SVIH ONIH KOJI NA NEKI NAČIN PARTICIPIRAJU U ONOME ŠTO ZOVEMO POTROŠAČKI KAPITALIZAM
300	HRVACKA NIJE IZUZETAK HRVACKA JE U BITI UDH_ DIO POSLOVNOG MODELA DIO STRATEGIJE ŠIRENJA U REGIJI SVE ZEMLJE U REGIJI DOBIVAJU ZNAČI JEDAN POSLOVNI MODEL UDH_ MODERNIH TRGOVAČKIH CENTARA SAMO JE PITANJE U TOME KOLIKO ĆE KOJA ZEMLJA DOBITI UDH_ TRGOVAČKIH CENTARA S OBZIROM NA KUPOVNU MOĆ STANOVNIŠTVA
301	TO JE JOŠ DOSTA NEDOREČENO I NEMAMO DOVOLJNO INFORMACIJA KAKVO ĆE SITUACIJA BIT KAD UĐEMO U EUROPSKU UNIJU UDH_ ZASTAJK_ ALI ONO ŠTO MOŽEMO REČI DA ĆE SIGURNO PASTI CIJENA VINA TAKO DA PROIZVOĐAČI KOJI UDH_ IMAJU SAMU PROIZVODNJU JE DOSTA SKUPA ĆE VJEROJATNO SE NAĆ U JEDNOJ SITUACIJI KOJA NEĆE BITI DOBRA
302	MANJE ĆE BITI VEZE ZA KONKRETNU PROIZVODNJU U TOJ GODINI ALI UDH_ REFERENTNA PROIZVODNJA U IDUĆOJ GODINI KOJA JE NEOBIČNO VAŽNA DVIJE TISUĆE JEDANESTA KAO REFERENTNA GODINA JE NEOBIČNO VAŽNA UDH_ ZA UTVRĐIVANJA PRAVA NA POTPORU U ONOJ GODINI DVIJE TISUĆE DVANAJSTOJ
303	MINISTARSTVO JE TU USLUGU ORGANIZIRALO NA NAČIN DA JE BESPLATNA KORISNICIMA
304	VE VE VE TOČKA ENTER JUROPE TOČKA HA ER JE INTERNETSKI SERVIS KOJI JE POKRENUO I OVAJ DOKUMENTACIJSKI CENTAR PRI INSTITUTU ZA MEĐUNARODNE ODNOSU UDH_ KRAJEM DVIJE TISUĆE ČETVIRTE GODINE UDH_ RADI DAVANJA LAKŠEG PRISTUPA ZASTAJK_ INFORMACIJA O INFORMACIJAMA VEZANIM ZA EUROPSKU UNIJU I LAKŠEG PRISTUPA DOKUMENTIMA EUROPSKE UNIJE
305	GDJE SE DANAS PRESELILA ŽENSTVENOST ISTRAŽUJE NAŠ IVAN JABUKA U PRIČI S NASLOVOM UDH_ ŽENSTVENOST POD SVAKU CIJENU
306	ZAHVALJUJUĆI FEMINIZMU ŽENE SU STEKLE RAVNOPRAVNOST PRED ZAKONOM PRAVO GLASA UDH_ PRAVO NA SUDJELOVANJE U JAVNOM ŽIVOTU UDH_ A POSTIGNUT JE I ODREĐEN NAPREDAK PREMA JEDNAKOSTI PLAĆA
307	NO I DALJE LJEPO TIJELO I ŽENSTVENOST OSTAJU ŽENI NA ODGOVORNOST
308	ONA ĆE BITI VREDNOVANA I NAGRAĐENA NA TEMELJU TOGA KOLIKO DOBRO UTJELOVI UDH_ TAJ IDEAL SAVŠRENOG ŽENSKOG TIJELA
309	ŽENE SU DANAS IZLOŽENE JEDNOJ VRSTI REPRESIJE
310	ZNA SE KAKVE PROPORCIJE MORAJU IMATI MORAJU IMATI FRIZURU I BITI NAŠMINKANE LJEPE UKRATKO SAVRŠENE
311	MEĐUTIM S OBZIROM DA ŽIVIMO U DRŽAVI KOJA NEMA SOCIJALNE RESURSE UDH_ KOJI BI ŽENI OMOGUĆILI DA POKRIJE PRIVATNU I JAVNU SFERU ŽENSTVENOST JE U NOVIJE VRIJEME NEŠTO ŠTO ŽENU UGROŽAVA
312	GDJE JE NESTALA ŽENSTVENOST TO JE TEMA NAŠE VEČERAŠNJE EMISIJE
313	UPRAVO RADI IDEJE POKAZIVANJA TIJELA I TOGA ŠTO ONI SVE MOGU U ODNOSU NA SVE NAS KOJI TO NE MOŽEMO MISLIM DA U TOM SMISLU UDH_ ZASTAJK_ NE TREBA SE NEČEMU PRETJERANO ČUDITI UDH_
314	ZNAMO FALA BOGU DA ONA POTJEČE IZ ZASTAJK_ OVAJ JEDNOG DIJELA SVIJETA GDJE SU BILI ŽIDOV I KOJI SU PO SVOJOJ RASI ILI PO SVOM NE ZNAM OKOLIŠU BILI DRUKČIJE FORMIRANI ŽENE SU BILE NIŽE UDH_ IMALI SU CRNE BILE SU ZASTAJK_ TAMNOKOSE BILE SU SA ZASTAJK_ VEĆIM OBLINAMA
315	KAD KAŽEM ŽENE MISLIM PRIJE SVEGA NA ONAJ DIO UDH_ KOJI SE ODNOSI NA ŽENSTVENOST
316	KAD KAŽEM IDEALIZIRANO UDH_ NARAVNO MORAMO VODIT RAČUNA DA TO ŠTO MI DANAS VIDIMO ŠTA JE MADONA I I NJENA RECIMO ŽENSTVENOST ILI NJENA LJEPOTA ILI NJENA ŽENSKOST NEMA VEZE S ONIM ŠTO JE BILO
317	A I ZA VATIKAN VELIKO POSTIGNUĆE NAKON PROBLEMA IZAZVANIH RAŠČIŠĆAVANJEM PEDOFILSKIH SKANDALA NEPOKRIVENIM RAČUNIMA VATIKANSKE BANKE ILI ZAVODA ZA VJERSKE POSLOVE ILI POTEZA ZBOG KOJIH SU SE VRIJEĐALI MUSLIMANI I ŽIDOV I ŠTO JE S PRAVOSLAVCIMA S DRUGIM PLUĆNIM KRILOM ISTOGA KRŠĆANSTVA KAKO JE PRAVOSLAVCE NAZIVAO POKOJNI PAPA IVAN PAVAO DRUGI VOJTLA NIJE POHODIO MOSKVI IAKO JE HTIO HOĆE LI RACINGER NOVINARKA PARALELA INKA ETEROVIĆ JEDINA JE OD STRANIH TEVE NOVINARA KOJA JE IKAD SNIMALA U RUSKOM VATIKANU U LAVRI TE RAZGOVARALA I S DRUGIM ČOVJEKOM RUSKE PRAVOSLAVNE CRKVE MITROPOLITOM ILARIJONOM
318	JAO_ DOK USTA OTVARAM DOK ZBORIM NJUŠKA KAO DA GORI
319	SAD SE ISTOM SVI PO REDU ČUDIT STALI ČINI IM SE DA SE MATI SAMO ŠALI DOK PRED SVIMA POSKAKUJE KOZLIĆ MALI KAO DA IH I NE ČUJE DO SKORA ĆE SVE DOZNATI ČIM STIGNU DO KUĆE PRIČATI I RAZMIŠLJATI I MNOGO ŠTA BOLJE

	ZNATI
320	ZAČUVŠI PJESMU POĐOŠE U SUSRET OD SREĆE SUZE BRIŠUĆI JER SAD SU ONI U BRIZI BILI KAD SU SE KUĆI PRAZNOJ VRATILI A NAKON ČVRSTOG ZAGRLJAJA PRIČANJU JEDNIH I DRUGIH DUGO JOŠ DUGO U NOĆ NE BIJEŠE KRAJA
321	TEŠKO JE DJECO DRAGA SIROMAŠAN BITI ALI JA SE ČVRSTO NADAM DA ĆEMO ZIMU OVU U NAŠEM TOPLOM DOMU U MIRU PREZIMITI
322	BEZ BRIGE MAMA SVE ĆE BITI U REDU KAO DA SI I TI S NAMA
323	MOŽDA STE I VI SRELI DJEČAKA JEDNOG OD VAŠIH VRŠNJAKA KOJI RADO IZMIŠLJA I SMIŠLJA KAO ŠTO RADO JEDE MEĐENJAKA ON SE SAMO OKRENE DESNO I LIJEVO I VEĆ PRIJEDLOGA EVO
324	A ZAŠTO SE NEBI OBLACI TAKO UREDILI DA IZ NJIH KIŠA PADA ILI SNIEG SIPA SAMO KAD SE ODVRNE PIPA
325	NASTA ŠUTNJA ŠTO ĆE BITI TKO ĆE PRVI PROZBORITI
326	ŠKOLSKI PRAZNICI POKLAPAJU SE S NAJSKUPLJIM SKIJAŠKIM TJEDNOM NAKON NOVE GODINE NO UNATOČ TOMU PRODAJA ARANŽMANA NIJE PALA ŠTOVIŠE BOLJA JE ZA DADESETAK POSTO U ODNOSU NA LANI NA GRANIČNOM PRELAZU BREGANA JE MARINA JOB
327	KOD APARTMANA JE TO OKO RECIMO DVIJE TISUĆE DO DVIJE I OSAMSTO KUNA A ŠTO SE TIČE OVIH KLASIČNIH DAKLE NAJPRIHVATLJIVIJIH UDH_ ZASTAJK_ SMJEŠTAJNIH KAPACITETA U HOTELU JE TO NEŠTO VIŠA CIJENA PA JE RECIMO OKO ČETRI I POL TISUĆE
328	ODNOSNO NE SPAJA OVE DANE NEGO SE RADI ŠTO JE NA NEKI NAČIN I DOBRO S OBZIROM DA TO POKAZUJE DA IMA POSLA
329	KAKO SU OVE GODINE BOŽIĆ I NOVA GODINA U DANE VIKENDA HRVATI INAČE MAJSTORI U SPAJANJU BLAGDANA S RADNIM DANIMA I VIKENDIMA UDH_ IMALI SU MANJE MANEVARSKOG PROSTORA
330	TVRTKE KOJE SU SE IPAK ODLUČILE NA KOLEKTIVNI GODIŠNJI ODMOR UVELE SU GA NA SAMO TJEDAN DANA
331	I MISLIM DA MOGU BEZ PRETJERIVANJA REĆI DA UDH_ STUPANJEM NA SNAGU OVOGA ZAKONA NIŠTA VIŠE NE MOŽE BITI ISTO
332	U SLUČAJU ZASTAJK_ KAD SE PROBIJE DA TO TAKO SLIKOVITO KAŽEM UDH_ ZASTAJK_ TAJ LIMIT DA IDH_ KAD GOVORIMO O VLADI O MINISTRIMA MINISTRI MORAJU DATI MANDAT NA RASPOLAGANJE
333	MEĐUTIM TRAŽIT ĆU I OSOBNO TAKVU VRSTU ODGOVORNOSTI NA SVIM RAZINAMA
334	NEKAD SE POSVAĐA NEKAD SE POMIRI NEKAD DOBRO NEKAD LOŠO AL SVE U SVEMU KAD SASTAVIMO DOBRO JE
335	TO NIJE PRIHVATLJIVO RJEŠENJE OVAJ PRAVILNIK UDH_ ZASTAJK_ KOLIKO SAM JA RAZUMIJELA NITI TO NE DOZVOLJAVJA NEGO ZASTAJK_ ODLUKA JE NA RAZINI ŽUPANIJE A NE U PREMA AUTONOMIJI SVAKE ŠKOLE
336	ODBROJAVAMO POSLJEDNJE SATE DVIJETISUĆE I DESETE GODINE
337	NEKI ĆE JE PAMTITI PO DOBRU DRUGI BI JE NAJRADIJE ZABORAVILI
338	DRUGIM RIJEČIMA KADA NAM SE OMOGUĆILO DA UĐEMO U DUŽNIČKI ODNOS BURIĆ TVRDI AKO NEŠTO ZNAMO ONDA JE TO TROŠITI
339	ZASTAJK_ ZNATE ŠTA KOD ZASTAJK_ PO MENI ŽENSTVENOST SE OČITUJE ZASTAJK_ U POGLEDU U NAČINU OPHOĐENJA U SAMOPOUZDANJU U TOJ NEKOJ SENZUALNOSTI I SLOŽILA BI SE SA GOSPODINOM OVAJ DA SEKSUALNOST ZASTAJK_ NEMA MNOGO VEZE SA ŽENSTVENOŠĆU MHM_ TO JE TO SU DVIJE RAZLIČITE ZASTAJK_ STVARI EROTIZIRANA ŽENSKA JE OVAJ SEKSUALNO KOJA JE PRIVLAČNA MISLIM DA ONA NE MORA BITI I ŽENSTVENA
340	JA BIH REKLA BEZ OBZIRA NA EMANCIPACIJU KOJA JE ŽENAMA DONELA JA KAŽEM POD NAVODNICIMA OVAJ RAVNOPRAVNOST SAMOPOUZDANJE ZASTAJK_ SAMOSVJESNOST ZASTAJK_ NEKAKO SE ZASTAJK_ OPASNO Približila TOJ GRANICI IZMEĐU MUŠKARCA I ŽENE
341	A SAD TRKOM NAŠOJ KUĆI JER VUK MOŽE DA SE VRATI TAMO ĆEMO SVE PO REDU ISPRIČATI I DOZNATI
342	SAD NAKON TOLIKO GODINA KAD SE VRATIM I KAD MI SE SVE TE SLIKE VRAĆAJU ONDA ZAPRAVO VIDIM ŠTA JE RAT I ŠTA JE UČINIO OVOJ DRŽAVI
343	A KOD NAS GLEDAJ U KUĆI DRVA NEMA DOK IZA BRDA BIJELA VEĆ SE I ZIMA SPREMA
344	ZNAM DA U SRCU NJENOM I STRAH JE I TUGA BAŠ KAO I TVOME
345	TI ĆEŠ IVICE BRIŽNO ZATVORITI VRATA I NE OTVARAJ NIKOM
346	A TI MARICE SOBU UREDI MALO A BUDETE LI GLADNI KRUHA JE PREOSTALO NA POLICI TAMO I ČEKAJTE ME ČEKAJTE ME
347	ZNAM ŠTO ĆU JA ĆU DO BREŽULJKA POĆI I GLASNO DOZIVATI
348	NI MENE IVICE BRATE JER SIGURNO NA SVIJETU NEMA BOLJEGA TATE
349	NAŠA MAJA TE MRZI IZ DNA DUŠE ALI ŠTA ĆE SE SKIRAT ZBOG TOGA NISI NI TI VOLIO SVOG TATU PUSTI BIT ĆE CURA DOBRA
350	MOGLI BI I SATOVI DA KUČAJU A NJHOVA STAKALCA KAO OGLEDALCA DA SVJETLUČAJU SVJETLUČAJU
351	ZAŠTO NEBI TAKVE OLOVKE BILE KOJE BI SE SAME ŠILJILE KOJE BI SAME RISALE KOJE BI SAME ZADAĆE PISALE
352	ONA SE LIJEPO UKOPČA NOĆU I PRENOSI U GLAVU ZNANJE IZ KOJE GOD ŽELIŠ KNJIGE A TI SPAVAŠ I LEŽIŠ BEZ BRIGE
353	VRIJEME JE KOMUNIKACIJE S VANJSKIM SVIJETOM KAD IZAĐEŠ ISPOVJEDAONICE KROZ TAJNA VRATA UĐI U HODNIK POZVONI NA ZVONCE CRVENE SOBE UĐI U SOBU U SOBI TE ČEKA PISMO S DALJNIM UPUTAMA BIG BRADER
354	TAJNA VRATA A TRESU MI SE RUKE DA SE VIDIM NA OGLEDALO HAHA_ UDH_ IZNENAĐENĐE
355	EJ_ DUŠO KAKO SI HAHA_ UDH_
356	KODA MI JE NETKO REKA KATE IMA BOGA DITE ČAČINO
357	A JEBATE DA SE I NAMA NEŠ DOGODILO DA NIJE DOSADNO
358	JOŠ JEDNA TVOJA PENZIJICA HEHE_ NAŠ HEROJ
359	CACORE PA ŠTA MI TO PRIJE NIJESI REKA JA MISLIO DA ME ZAJEBAVAŠ
360	LJEPOTICE JA SE TEBI SVIDAM A TI MENI
361	PA NARAVNO DA JE SVE U REDU KAD SAM SA ČOVJEKOM IŠAO U ŠKOLU
362	TO JE PORUKA BRAVO
363	HOMO_ ŠTO JE PARIPOVIĆI MUDA LADITE HA_

364	O RIBARSKA TRINAEST ŠTA KAŽETE NA OVO HA_ OVAKVE SU VAM SVE KOD NAS VIDI JE
365	MENI JE OVO SAN SNOVA ŠIME SAN SNOVA
366	VEČERAS ĆE BIT LUDILO KAD POČNU ROKETE KONFETE BALUNI RAZUMIŠ LI ME TO JE LUDILO
367	OVCE SMO ČUVALI KAO DJEČACI JA S OVE ON S ONE STRANE HEHE_ BOŽE KAKO SMO SE VOLJELI I SLAGALI
368	TRGOVINU ĆE PREUSMJERITI NA DUBROVNIK HEHE_ DUBROVČANINE
369	E KAK SE ZOVE VEPAR S TRI NOGE
370	BOK BOK BOK PA DA EVO DANAS SAM DOŠO ZASTAJK_ JEL DEJO DOMA
371	O PA KAJ JE KAM SI TI KRENUL HA_ KAJ ČUL SI DA TE ZOVEM SI DOŠEL BRAVO MAČAK PRAVI SI
372	NAŠ KAJ AK TE NETKO PIPNE NE TI SAM MENI RECI JA BUM GA ODMAH STRELIL NAŠ
373	K SMRC_ PA DOLAZI KRPA I EKIPA ONA PA S NJIMA IDEM NE NAŠ KAJ E BILO BU JEBENO MAJKE MI ONO NAŠ KAJ KRPA JE TAK NABRIJAN
374	MA GLAVNO JE DA SMO MI ŽIVI I ZDRAVI JEL TAKO MOMCI
375	AJDE
376	NEŠEŠ SE UDH_ PAMETNA CURA GLE KAK SE LIJEPO POČEŠAL
377	EVO GA BLATOBLA
378	PAZI TI JELENKA
379	TI SI ROĐENA GAZDARICA IDH_
380	AL DA DA DOBRO JE
381	MASTI EJ_ UDH_ SAV SAM MASAN
382	KAO JURA I HIHI_ NENO
383	GLE MOLIM TE NJIH DVA E
384	SAD SAMO SAMO SE DRŽI
385	AL MAJKE MI UDH_
386	DOĐI VIDJET PREKO PEDESET SAM IH
387	OTVORI OČI
388	OČI OTVORI
389	MISLIM DA JE INFARKT DOBILO HAHA_
390	EVA I JOŽA IZVANRETNI DOLAZAK
391	JAO_ AJME GUŠTA ČOVJEČE
392	EVO STIPE I KOZA
393	VOLE SE
394	OVI JEMA MU JE BOLJE
395	A TO TI PRED SELA MALO
396	O POČELI SMO SE I ŠMINKAT
397	NEŠTO TI OČI SVIJETLE DA NISI OPET A HAHA_
398	EVO EVO SAMO ČAS
399	TO JE VUČIJA SREĆA HAHA_
400	O VREĆA LAKA NIJE HEHE_ NESTAT VALJA S NJOM
401	MA DAJ HAHA_ UDH_ HAHA_ UDH_ HAHA_
402	MALENA BOK
403	A ZNAŠ TI KAK JE FINO ŠEVIT ŽENSKU HIHI_ KAJ MIRIŠI PO TOPLIM PICEKIMA HIHI_
404	PEPO RINGLO PA NISMO SE DUGO VIDLI
405	ALI GELDAJ GLEDAJ NA KROVU KOLAČIĆI
406	O LJEPOTICE NAŠA IDE KAKO SMO
407	TAKO I TREBA TAKO I TREBA SRCE MOJE LIJEPO
408	BOK JA SAM LIDIJA ŽIVIM TU PA SAM VAS DOŠLA KRATKO UPOZNATI BOK
409	SUPER ZNAŠ MOŽDA ĆU I JA UPISATI NOVINARSTVO IZVRANREDNO
410	ODI LUKICE LJUBAVI DOĐI
411	BAR ĆEŠ IMATI MATERIJALA ZA JEDNU SOČNU RECENZIJU
412	SUPER JE SUPER JE RECENZIJA
413	NE TO JE TO TO JE TO RAZVALIO SI IH BAŠ KAKO TREBA
414	MA DAJ PA DA NEMAŠ MUDA NE BI NIKAD NAPISAO ONAKAV TEKST O ONOM RESTORANU
415	E VRAG SI TI DITE MOJE VRAG NE MOREŠ BIT NEG CRNE KOSE HEHE_
416	AJDE AJDE LIPI MOJ ŠIME ČUVAJ MI SE NISI VIŠE DITE
417	AJME KAKO SI POGODIA BAŠ TOGA SAM TILA
418	NIJE TO BILO KOJI ANĐEO OVO TI JE BARAN ANĐEO LJUBAVI NADE
419	OČE ŠIME ZNAN JA DA OČE
420	DOK TE NIJE BILO DA ZNAŠ KAKAV SAM KRASAN GOBLEN NAPRAVILA
421	UDH_ MALO VIŠE KENJA I PAPA
422	ŠTEFICA BI TE TUKLA
423	ALEN ALEN ALEN NENO
424	TI ISTO LJUBI JAKO
425	UDH_ ON IMA TEORIJU ČIŠĆE
426	HAHA_ NA SVOJU PARTIJU
427	A KAD ĆEŠ MENE RJEŠIT

428	JE SE I JA PITAM ALENE TE JAKO BUBA KAŽE ON A A HAHA_
429	ON MENE VUČE ZA RUKU JA TU PEREM ZUBE I VIČE UDH_ IDEŠ U SAVU
430	PA NE ZNAM DAL SI MU NAPETA MISLIM JA DA ON VOLI
431	JE JE TO JE TO JE MOJ AVIONČIĆ
432	KAK TI VISIŠ HAHA_
433	ČETVRTI PUT U RINGU JE SVIZAC
434	VIDIŠ TI ŠTA JE JOGA A MICKO
435	ZASTAJK_ DRAGI NENO SAD SI ČIST
436	KOG SAM TI LJUBIJU
437	JE BLENDALA JE
438	OČE SE VRATIT IDH_ AJMO NA BUS
439	NE NO A LEN NE NO A LEN A LEN
440	ALENE PREGAZIT ĆE TE KONJ PREGAZIT ĆE TE HEHE_ UDH_ ČUJEŠ
441	NISAM TAKA NE NE NE
442	KAKO ĆEMO GA NAZAD VRATIT HAHA_
443	PROPAŠT ĆE UDH_ KROZ TOALETE IDH_
444	ALEN MORAMO IĆ PO KONJA
445	I SVI SU IMALI ROGOVE
446	A DOĐI VAMO PIČKA TI MATERINA
447	JA SAM JELA TI SI HANA NEMOJ SRAT HAHA_
448	OVAJ JE IMO NA HAUBI OD TRKAČEG AUTA ROGOVE NEMA SVE SE ZABILJEŽI HIHI_ UDH_ SVE OBILJEŽI
449	HAHA_ UDH_ ONO
450	DEČKI TKO ĆE SE KUPAT JOŠ HI UDH_ ZAGRIO JE DAVORIN VODU HAHA_
451	JA ONAKO VELIKIM I ONDA ISPOD SAM IVANA MARIĆ
452	UOPĆE NE RAZMIŠLJAM O TOME SAMO NEKAD TAKO KAD SE SJETIM NEMOGU VJEROVAT DA ONO DA NE DA NE MISLIM NA TO KUŽIŠ
453	I MISLIM SE ISUSE BOŽE ZAR JE MENI TO TOLKO NEBITNO HAHA_
454	SISA TI JE ISPALA SISA HAHA_
455	ZAGRLI GA NOGAMA OKO STRUKA
456	DA DA DA DA I OBJESI SE ZA NJEGA HAHA_
457	HA_ JDE SAD SE OKRENI HA_
458	A JELENA DOBRO KAŽE PUSTI JE DA VIDIMO ŠTA ĆE
459	HAHA_ KAŽE JELENA PA PUSTI JE DA VIDIMO
460	ŠTA ĆE UĆ ĆE U KUĆU ČUJ ŠTA ĆEŠ
461	DA NEMOJ ZEZAT
462	PA UOPŠTE VIDI JE
463	TRUDNA JEBOTE ČUJ ŠTA
464	LJUBI MI SE U USTA
465	PAZI DA NE UPATIŠ TUĐI
466	UDH_ AL PAZI ČUJ TI TEORIJU
467	STIPE POGLE OVO HOHO_
468	JAO_ KAMERE
469	TOMI ĆE NA ROG POGLE GA
470	VIDI VIDI VIDI ĐAVLIJA JE
471	POGLE TI ŠTA JE ĐAVLIJA
472	OVO JE ZA NE POGLEDATI KAKO
473	ALENE A KAKO ĆEMO GA VRATIT
474	KAKO ĆEMO GA AJME ON JE MOKAR
475	DJECO MOJA DJEČICE
476	NE JEBOTE PRIČAJ
477	DAJ NEMOJ ME JEBAT PRIČAJ
478	AH HAHA_ NEMOJ ME ZEZAT STVARNO MI RECI
479	TI NISI NORAMALAN DAJ EJ_ AJ
480	TAKVOG TE IVICE VOLIM KAD SI DOBAR
481	IVICE BRATE MILI PA TO JE KAO U PRIČI
482	PROZOR SVE SAMI ŠEĆERIĆI
483	O VI STE DOBRA BAKICE A MI SMO MI SMO OSTALI BEZ TATE I MAMICE
484	O HOHO_ HOHO_ TKO JE TO HIHI_ KAKVI TO GRICKIĆI KAKVI TO PTIČIĆI GRICKAJU MOJU SLADANU KUĆICU HA_
485	O NE BOJTE SE DJEČICE HEHE_ IMATE SREĆICE
486	HEHE_ DOĐITE U KUĆICU UĐITE NIGDJE VAM LJEPŠE HEHE_ BITI NEĆE
487	DOBRO DOBRO HOHO_ SAMO VI HEHE_ UĐITE
488	HIHI_ POGOSTITI POGOSTITI
489	PEĆICA U PEĆI VATRICA SVAŠTA ĆE VAM DATI VAŠA BAKICA
490	HM A SUNCE MOJE
491	KURAC JE DOBAR

492	PA NEMA ŠANSE JA TO NEBIH HTJELA NI ČUT
493	SUPER MI JE TATA
494	LIJEPO PJEVAŠ STRIČE LUKA SVIĐA MI SE KAKO PJESMOM TJERAŠ VUKA
495	ZOVEM SE CRVENKAPICA A MOJA JE I KOŠARICA NOŠIM NEŠTO SVOJOJ BAKI ŠTO STANUJE ČAK NA DRUGOM KRAJU ŠUME
496	O DA VIDIŠ KAKO KRASNA MALA KUĆA A OKO NJE OGRADICA SVA O PRUĆA A NAD KROVOM TRI STARA BORA HLAD JOJ ČINE ODOZGORA
497	TASICA ME MOJA ČEKA VEĆ JE VIDIM IZ DALEKA DOVIĐENJA DOVIĐENJA
498	AKO TI SAD ODUSTANEŠ ONI BI TAMO MOGAO POMISLIT DA SMO TE OTELI DA SMO TE UBILI PITAJ BOGA ŠTA TI ZNAŠ KAKVO JE ON SRANJE U STANJU NAPRAVITI
499	MAJKU TI BOŽJU KUME NE MRDAJ
500	NE MRDAJ KAD TI KAŽEM SLUŠAJ ME
501	MIRO KĆERI NEMOJ TAMO IĆ MINIRANO JE
502	E E BAŠ SAM HTIO RAZGOVARATI SA VAMA ZASTAJK_ SAD KAD JE ZASTAJK_ BBB_ BABA JEL KAD BABE VIŠE NEMA KHM_ JA IDEM S LIČOM PA BI TREBAO SVOJE PARE IDEM U ZAGREB MISLIM SAD KAD VAM VIŠE NISAM POTREBAN
503	A NISTE TREBALI DOLAZIT PPP_ POSLAO BI VAM JA PARE SSS_ SLIJEDEĆI TJEDAN TAK SMO SE DOGOVORILI JEL SSS_ SLIJEDEĆI TJEDAN EVO EVO JA ČU TO POLAKO NEAM E A TTT_ TO TO JE BAŠ ZASTAJK_ ZA VAS SAM VRATIT ČU VAM SVE
504	NE MISLIM JA NIŠTA NEŠTA JE BILO ŠTA JE BILO JA NE ZNAM
505	DA TO JE TO DUH BROZA TITA JOSIPA NA IZLAZU IZ GROBLJA IZRONIO IZ MAGLE PRVO SMO ČULI KORAKE TUP TAP TUP TAP
506	III_ IMAO JE SAMO JEDNU NOGU
507	BIO JE KAO I OVAJ OVDJE IMAO JE ISTI ŠINJAL
508	IZH UDH_ IZH UDH_ IZH UDH_ DOBRO JE IZH UDH_ IZH UDH_ SRP I ČEKIĆ
509	PRIČEKAJTE ME VANKA SAD ČU JA RAZVALIT LOKOT
510	JA JA SAM DDD_ DEIN PRVI SUSJED
511	NEMOJ PETRE
512	ZASTAJK_ REUF NE NIJE NITKO BIO NNN_ NE NIJE
513	KROZ AKO SKOČI
514	TAKVA SVIRKA ZNAČI ZLO
515	MI TO SMO MI
516	KKK_ KAŽEŠ SRCE MI TAKO LUPA DA SAMO NJEGA ČUJEM
517	PALA SAM ZASTAJK_ U KUPAONICI SAM SE POKLIZNULA SAM SE NA KADU I LUPILA
518	NISAM JA ZVALA ZASTAJK_ CURICE SU HTJELE ZVATI HITNU PA SU VALJDA UMJESTO DEVET ČETIRI NAZVALE DEVET DVA
519	JA MU NISAM NIŠ NAPRAVILA ON SE SAMO ODJEDNOM PRESTAO MICATI
520	PUSTI ME ZVAT ČU POLICIJU
521	NEMOJ TAMO TAMO SU TI MINE
522	MIRO STANI NEMOJ
523	UDH_ TUP TAP TUP TAP I ONDA JE STALO STRAŠNO
524	ŠTA ČEMO SAD HA_
525	TO SI TI PA NISAM TE PRIPOZNALA UDH_ ŠTO SE DOGODILO DITE MOJE
526	DOBRO DA SAM VIDJELA
527	TREBA TREBA UDH_
528	MLADO TREBA POSUŠIT
529	POSUŠIT GA TREBA IDH_
530	POSUŠIT UDH_ IDH_ I SKRATIT PUPČANU VRPCU
531	ČIM ČEMO IDH_
532	NA DVANAEST CENTIMETRA IDH_
533	TREBA SKRATIT PUPČANU VRPCU AL TREBA IDH_
534	ALEN
535	ISUSE O BOŽE
536	DER_
537	IZVADI GA
538	ČEKAJ NE NEMOJ GA NEMOJ GA DOK GA MAJKA NE POLIŽE
539	I ZATO SE DJECO BOJIM DA SE NEGDJE BLIZU KRIJE
540	KOZLIĆI MOJI MILI ZAŠTO STE ME OSTAVILI ŠTO LI SE JE DOGODILO DOK ME OVDE NIJE BILO UDH_ DA VAS NIJE JAO_ JAO_ VUK OPAKI PROGUTAO
541	AJME MENI JEDNICI
542	O BOŽE MOJ
543	STANISLAVE JESI DOBRO
544	DAJ PUSTI TE NAS NA MISU GROŠ IDEMO
545	DJECO VEĆ SE I NOĆ SVE TAMNIJA HVATA A MI GLEDAMO ZAUDNO NAŠA VRATA UDH_ NO NADAJMO SE JOŠ UVIJEK MOŽE STIĆI
546	U ŠUMI PUSOJ ZVIJER GA MOGLA NAPASTI NEKA I MOŽDA RANJEN TAMO POMOĆ UZALUD ČEKA
547	JAO_ IVICE BRATE ZAMISLI ZAMISLI SAMO DA SE NI OTAC NI MAJKA VIŠE NE VRATE
548	A MENE ZAR ČEŠ MENE SIROTCU SAMU TU OSTAVIT

549	ALI IPAK MENE JE STRAH
550	ZAŠTO SOVA NE SPAVA
551	ŠTO KAŽETE
552	NIJE MOGUĆE UDH_ NIJE MOGUĆE
553	JAO_ MOJ BRATE
554	NEMOJTE GA ZATVORITI MOLIM VAS IVICE BRATE MOJ JAO_
555	SAD SI PROP AO
556	OČE I MAJČICE GDJE STE GDJE STE POMOZITE
557	HOĆU HOĆU ALI NE ZNAM DAL ĆU SVE ZNATI
558	HOĆU HOĆU RADIT ĆU DANJU I NOĆU NEMOJTE NEMOJTE
559	OPROSTITE GOSPODARICE DRVA SU VLAŽNA
560	NEĆETE VALJDA NEĆETE NESMIJE NITKO UČINITI TO
561	EVO VEĆ PUŠEM
562	MAJA DER_ MAJA MAJA MAJA GDJE SU MOJE TABLETE
563	EVO JE
564	KRVI MI ISUSOVE
565	I TI MENE OVAJ VUČE I SRCE MI TAKO TUČE
566	O STRAŠNO STRAŠNO PA TI SI ONAJ MI SMO SE SRELI POZNAM TI GLAS I REKAO SI DA SI PAS
567	JAO_ JAO_ I LAGAO SI SVE ŠTO SI REKAO UDH_ BAKAO MOJA UDH_ BAKICE DRAGA ŠTO JE TO
568	BAKICE JAVI SE GDJE SI JESI LI ŽIVA
569	JAO_ MENI KAKO SAM MOGLA VJEROVATI SAMO
570	UDH_ TEŠKO JE BITI BEZ BILO KOJEG RODITELJA BILO OCA BILO MAJKE UDH_ I ZATO MOJA PORUKA UDH_ SVIM LJUDIMA UŽIVAJTE U SVOJIM OČEVIMA U SVOJIM MAJKAMA UDH_ JER KAD IDH_ ODU OČEVI KAD ODU MAJKE NIKAD SE VIŠE VRATIT NEĆE A TO JE SAMO ONO ŠTO MOŽETE UŽIVATI JEDANPUT SMRC_ UDH_ MOJ OTAC JE OTIŠAO DOSTOJANSTVENO NENADANO I POLAKO
571	JADNICA PODLETILA JE POD VLAK
572	ŠTA TI SE JAKO LJUTIŠ NA MENE HA_
573	JER HRVATSKU MI MOJU OBJESIŠE KO LOPOVA UDH_ DOK NJENO IME BRIŠE
574	AL DA SI NAS ZAJEBALA ZAJEBALA SI NAS
575	AL KO ĆE SVE OVO PLATIT UDH_ OVE NJENE KREME I POMADE PUN FRIŽIDER UDH_ ZDRAVE RANE ŠTO SMO NARUČILI NOVU PATKU A SAHRANU JOJA KO ĆE SAHRANU PLATIT
576	SUTRA ĆEMO PROBAT NEŠTO DRUGO
577	OD SEBE ŠIME MOJ OD SEBE SAM POBIGA
578	POGORELCI BEZ RODBINE I OGNJIŠTA MOJ BRAT I JA OSTALI SMO SAMI NA SVIJETU SAMO JE BOG BIO S NAMA
579	NE ZAMJERAM OČE UŽASNE SU PRILIKE STRAŠNO AKO NE UHVATIMO VAŠEG BRATA REVAČA AKO GA ŽIVOG ILI MRTVOG NE IZRUČIMO BOLJE ŽIVOG NEGO MRTVOG BOSANSKI BEGLER BEG PREKINUT ĆE TRGOVINU UDH_ S MLETAČKOM REPUBLIKOM PREUSMJERITI NA DUBROVNIK
580	U IME KRŠĆANSKE VJERE I BRATSKJE LJUBAVI
581	VJERUJ MI REVAČ JE ŽIV UDH_ BIT ĆE UŽASNO AKO MI NE POVJERUJEŠ
582	NAJGORE ŠTO SE MOŽE
583	NIKOM PREPORUČIO
584	ŠTO IDH_ SADA KUDA KAKO NA KOJI NAČIN DOKLE
585	ETO IDH_
586	TO JE BILO TEŠKO
587	JEBOTE DOĐEM OVDJE I ODJEDNOM NEKAKVA NEGATIVNA ENERGIJA PA KOJI KURAC JEBOTE I SAD KAD MI TI OVO ISPRIČAŠ SRCE MI SE PARA JER NE MOGU VIROVAT
588	A VIDIŠ DA OČE S TOBOM OSTAT
589	TAK MLADA TAK LEPA NEMOJ JASNA AJDE JASNICE PUSTI TO
590	TREBALA SI SE JAČE BRANIT
591	CMIZDR_ SMRC_ MOGAO BIH SI JEDNU PAPIGU POFARBANU KUPIT CMIZDR_
592	BEBA MORA IMATI TATU JA ZNAM KAK JE TO BEZ TATE ONDA TE GAĐAJU S KAMENJEM
593	CMIZDR_ OPROSTI OPROSTI NISAM ZNAO
594	GOTOV SAM TI BRACO MOJ NEĆU TI JA VIŠE DUGO
595	RAK JE TO SMRT
596	JER ME GAZDA NEDUŽNOGA SA DVORIŠTA POTJERAO
597	GAZDA MENE GRDIT STANE I BATINOM OVOLIKOM JOŠ VEĆOM POTJERA ME OTJERA ME
598	KAD IZADEŠ IZ ISPOVJEDAONICE KROZ TAJNA VRATA UĐI U HODNIK UDH_ POZVONI NA ZVONCE ZELENE SOBE UĐI U SOBU U SOBI TE ČEKA PISMO S IDH_ DALJNIM IDH_ UPUTAMA
599	MOŽDA SE NISI NADALA OVOME PISMU IDH_ UDH_ AL ODMA NA POČETKU DA NAPIŠEM DA SMO SVI UZ TEBE I DA TE SVI VOLIMO I PODRŽAVAMO JOJ UDH_
600	JAKO SMO PONOSNI NA TEBE SVI KOJI TE ZNAJU I NE ZNAJU TE PODRŽAVAJU I VOLE CMIZDR_
601	MAMA I TVOJE SESTRE SU PONOSNE NA TEBE PONOSNE SU ŠTO SI NJEZINA KĆER A NAŠA SESTRA
602	JAKO TE VOLIMO I TUŽNI SMO KAD NA TVOM LICU VIDIMO SUZE SMRC_ ŽELIMO DA TVOM LICU STALNO BUDE OSMJEH
603	ON VIŠE NE MISLI NA TEBE
604	OKRENUO JE NOVU STRANICU U SVOM ŽIVOTU

605	IMA NOVU DJEVOJKU UDH_ I ZA TEBE VIŠE NE ŽELI ZNAT
606	ALI MOLE TE TVOJA MAMA I SESTRA DA TE TO NE RASTUŽI ŠMRC I DA TI U SVOM ŽIVOTU OKRENEŠ NOVU STRANICU JER ON TE NIJE VRIJEDAN
607	UDH_ SLATKI STE POPUT NAJSLAĐEG BOMBONA NA OVOME SVIJETU
608	TANJA NEMOJ SLUČAJNO DA VIDIM VIŠE SUZE NA TVOME LICU TO NE ŽELIM AJME MAMA VOLIM TE
609	TVOJA MAMA ŽELI DA TI BUDEŠ JAKA KAO I DO SADA I DA NE MISLIŠ UDH_ NA ANTU LOVRIČA NEGO NA ANTU SMRC_ KOJI JE SAD UZ TEBE SMRC_
610	VOLIM TE NAJVIŠE UZ TEBE SAM
611	VAZDA SI GOVORILA DA SAMO IDEŠ DO KRAJA PA TAKO CMIZDR_ ŽELIM DA I SAD BUDE
612	MAMA JE DOBRO HVALA BOGU
613	TOBOM IAKO JE NAJSRETNJE DIJETE ŠTA TE IMA ZA TETKU CMIZDR_
614	SAMO TE MOLIMO SVI DA BUDEŠ JAKA KAO I DO SADA UDH_ JER TI SI BORAC IDH_ ZNAŠ SE BORITI UDH_ I U SVAKOJ BORBI ZNAMO DA ČEŠ POBIJEDITI
615	NEMOJ NIČEGA DA TI BUDE ŽAO SAMO NASTAVI TAKO SVI SMO UZ TEBE I SVI TE VOLIMO NAJDRAŽA NAŠA
616	VOLIM I JA VAS
617	TVOJ ANTE IDH_ KAKO TI KAŽEŠ SKROZ TE ZABORAVIO VIŠE NE MISLI NA TEBE SMRC_
618	I TI NEŠTO JE NAJLJPŠE ŠTO ČOVJEK MOŽE VIDJET I POŽELITI IDH_
619	TVOJA MAMA ŽELI DA TI BUDEŠ JAKA KAO I DO SADA I DA NE MISLIŠ NA ANTU LOVRIČA CMIZDR_
620	NEGO NA ANTU KOJI JE SAD UZ TEBE IDH_
621	VOLIM TE NAJVIŠE I UZ TEBE SAM
622	VAZDA SI GOVORILA DA U SVEMU IDEŠ DO KRAJA PA TAKO ŽELIM DA SAD BUDE
623	NEMOJ NIČEGA DA TI BUDE ŽAO SAMO NASTAVI TAKO SVI SMO UZ TEBE I SVI TE VOLIMO NAJDRAŽA NAŠA
624	PA ŠTA BI TREBALA OD SVAKOG TRAŽIT PISMENU POTVRDU DA NIJE SRČANI BOLESNIK
625	KAKAV SPROVOD MAMA
626	MAMA TO NIJE ISTINA CMIZDR_ TO NIJE ZATO
627	A NAMA JE BILO PRIJE TAKO DIVNO TAKO SMO LIJEPO ŽIVJELI
628	E NEĆE ONI S OVIM ŠTO IZNOSU RUŠIT MOJU HRVACKU
629	A ŠTA VI OČETE SA OVIM JEL VVV_ VI VIDITE ŠTA VI RADITE OD MENE
630	KODA SAM JA KRIVA ZA SVE
631	O ODEM U TU NJEMAČKU VRATIM SE ISTA SI KO ŠTO SI I BILA
632	DOBRO ZNALA SAM I ŠTA SAD ISTO SAM TAKO MOGLA ŽIVIT I SA CIGANIMA
633	ČOVIK MORA BIT NA SVAŠTA PRIPRAVAN IDH_
634	KAKAV JE JEL ZADOVOLJAN ŠTA TI SE ČINI
635	FALA VAM ČASNA JA SE NADAM DA NEĆE
636	A KAKVI SU MOJI KORACI
637	A NE ZNAM BAŠ NISAM SIGURNA
638	A ZNAŠ ŠTA MI JOŠ PIŠE DA JE ONAJ NJEZIN OPET TUČE DA JE PUNO NESRETNIA BOG NEKA JOJ SNAGE DA
639	JE JE TO JE PUNO LIPO ŠIME TO ME GRIJE ONA JE MOJE DITE MI SMO OBITEU
640	KĆERI MOJA VRATI SE SAMNOM KUĆI
641	NE TRAŽIM DA SE ODREKNEŠ LJUBAVI PREMA KRŠĆANSKOM BOGU AKO JE TO RAZLOG
642	VRATI SE KUĆI NE SRAMOTI SVOJU OBITEU
643	IDH_ UDH_ BOŽE MOJ LUDOG LI DJETETA TI MISLIŠ DA IZNAD NAŠE VOLJE NE POSTOJI NIŠTA NA OVOM SVIJETU DA SVE MORA BITI ONAKO KAO ŠTO MI ŽELIMO NIJE TAKO ZNAŠ I SAMA DA NIJE TAKO
644	NE MOGU SE JA VRATIT SELMA
645	NETKO JE KRIVO VIDIO DUŠO MOJA POGRIJEŠILI STE
646	ISTINE VIŠE NEMA
647	LAŽEŠ MI JA TE RAZUMIJEM ALI LAŽEŠ DUBROVČANINE
648	KAK SE OSJEĆAM UNUTRA CMIZDR_
649	SUPER PRIKAZANA KO I JOSIPA I KO NEZNAM KO CMIZDR_ A OPĆE SVI ZNAJU UNUTRA
650	I NEKE OSOBE VANI ISPALA DA MOŽEŠ
651	NE GOVORI O NEKIM POKVARENIM LJILJ LJUDIMA I O LICOMJERSTVU A PRVA KOJA JE LICOMJERNA JE ONA
652	NE NEGO CMIZDR_ NAJGORE ONDA CMIZDR_ DA VANI ISPADNE NEŠTO A TOTALNO JE DRUGO CMIZDR_
653	I ON JUČER SA SVOJIM IZJAVAMA I NEKIM ISTO KO LUPINO POVRIJEDIO ME NA LUPINA ISTO CMIZDR_
654	NIKAD U ŽIVOTU NISAM TOLIKO SMRC_ PRIČALA O GLUPOSTIMA TOLIKO SE PROPUCAVALA RADI OVOG RADI ONOG DA SVE SLUŠAM OD JUTRA DO MRAKA
655	IMALA SAM SEDAM GODINA TOLIKU PROBLEMATIČNU I TEŠKU VEZU U ŽIVOTU NISAM PUSTILA SUZU NI RIJEŠVALA S DRUGIMA PROBLEME
656	I JADILA SE GLUPOSTIMA MENI SU OVO GLUPOSTI
657	DA JA IMAM GRČ U ŽELUCU STALNO CMIZDR_ I DA SLUŠAM OD JUTRA DO MRAKA SVADE PA ONDA UZMI RADI PA NE MOGU NEDA MI SE
658	STVARNO MI JE DOSTA
659	NE MOGU RAZMIŠLJAT EVO BUDEMO
660	NEZNAM GLUMIT NIJE BITNO NEGO ZASTAJK_ LOŠE MI JE OD SVEGA I U GLAVI MI JE LOŠE III_ ONO KUŽI SE DA MI JE LOŠE I SMRC_ I NEMOGU I TEŠKO JE NEMOJTE DA TEŠKO DIŠEM DA IMAM NEKAKAV NAPAD I TU SMRC_ I NE MOGU NI PRIČAT JEL STE SKUŽILI NORMALNO NI NIŠTA NE MOGU SMRC_ CMIZDR_

661	MENE BOLI OVA DRUGA STVAR RAZUMIJEŠ
662	JA ZNAM KAKVA SAM OSOBA JA NE PLAČEM OVDJE ZATO ŠTO JE MENE PRIKAZANO OVAKO IL ONAKO PRIKAZANO ME ONAKO KAKVA JESAM
663	AL ZA MENE UOPČE TO NE BRINE SVI PRIČAJU O NEČEM KAKO NAS
664	TREBAM JA TRČAT ZA TOBOM
665	JOŠ NE MOGU K SEBI DOĆI
666	DAN DA MISLIM DA ĆE ON OD NEKUD IPAK DOĆI
667	ALI JA SE NADAM DA ĆE IPAK PRAVDA POBIJEDITI DA SE NEBI VIŠE TA ZLA DOGAĐALA DA NEBI BILO VIŠE RATOVA
668	CMIZDR_ JA NE ZNAM KAJ JA DELAM GROŠ NE ZNAM KAJ JE MENI BILO UDH_ KAD SI TI ZAMAHNUO NOŽEM JA SAM SI REKLA KAJ UDH_ CMIZDR_ UDH_ ZAKAJ IDH_ MORA OVO DIJETE SA MONOM UMRIJETI
669	DA SU NAM BAREM ŽIVI DJEDICA ILI BAKA OVAKO STRAH ME ŠTO LI ĆE S NAMA BITI AKO
670	ALI JA SAM GLADNA I NOGE ME VEĆ BOLE ALI MISLIM
671	TU JA SIROTA PROVODIM DANE DALEKO OD LJUDI DALEKO OD PUTA I JEDVA ČEKAM DA NETKO DOĐE
672	NE PITAJ NIJE TO NIŠTA ŠTO SE MENI INAČE NE DOGAĐA JEDEM TUĐA GOVNA JER SAM GLUPAČA UDH_ OSTANEM SA MAMOM OSTANEM U BRAKU I EVO SAD NE VIDIM ŠTO MI RADI NEMAM HRABROSTI GLUPAČA
673	BILO MI JE TEŠKO KAD SI OTIŠAO U AMERIKU PLAKALA SAM MJESECIMA
674	NEMOJ DAJ DAJ DAJ

12.3 Privitak — VBA kôd aplikacije korištene za anotaciju korpusa

12.3.1 VBA kôd—Modul1

```
Sub StartUserForm()  
  
    'Pokretanje UserForma  
  
    UserForm1.Show  
  
End Sub  
  
Public Sub OpenMp3(strFile As String)  
    Dim x As Variant  
    Dim Path As String  
    Dim File As String  
  
    Path = "C:\Program Files\Windows Media Player\wmplayer.exe"  
    File = "C:\KEGprocjena\SNIMKE\  
    File = File & strFile  
  
    Call Shell(Path + " " + File, vbNormalFocus)  
  
End Sub
```

12.3.2 VBA kôd—UserForm1

```
Public brojZaPlay As Integer  
Public countClick As Integer  
  
Private Sub CloseButton_Click()  
    Unload Me  
End Sub  
  
Private Sub DrugoButton_Click()  
  
End Sub  
  
Private Sub FrameDrugo_Click()  
    Me.DrugButton.Value = False  
End Sub  
  
Private Sub PlayButton_Click()  
  
    Dim nazivSnimke As String  
    brojZaPlay = brojZaPlay + 1  
    rangeZaPlay = "A" & CStr(brojZaPlay)  
    dohvaceniBrojZaPlay = Range(rangeZaPlay).Value  
  
    nazivSnimke = CStr(dohvaceniBrojZaPlay) & ".wav"  
    Call OpenMp3(nazivSnimke)  
  
End Sub  
  
Private Sub SaveButton_Click()  
  
    ' Provjera ocjene za emociju SREB  
    If Me.HappyButton1.Value = True Then  
        ocjenaSreca = 1  
    ElseIf Me.HappyButton2.Value = True Then  
        ocjenaSreca = 2  
    ElseIf Me.HappyButton3.Value = True Then  
        ocjenaSreca = 3  
    ElseIf Me.HappyButton4.Value = True Then  
        ocjenaSreca = 4  
    ElseIf Me.HappyButton5.Value = True Then  
        ocjenaSreca = 5  
    ElseIf Me.HappyButton6.Value = True Then  
        ocjenaSreca = 6
```

```

ElseIf Me.HappyButton7.Value = True Then
    ocjenaSreca = 7
ElseIf Me.HappyButton8.Value = True Then
    ocjenaSreca = 8
ElseIf Me.HappyButton9.Value = True Then
    ocjenaSreca = 9
ElseIf Me.HappyButton10.Value = True Then
    ocjenaSreca = 10
ElseIf Me.HappyButton11.Value = True Then
    ocjenaSreca = 0
End If

' Provjera ocjene za emociju TUGA
If Me.SadButton1.Value = True Then
    ocjenaTuga = 1
ElseIf Me.SadButton2.Value = True Then
    ocjenaTuga = 2
ElseIf Me.SadButton3.Value = True Then
    ocjenaTuga = 3
ElseIf Me.SadButton4.Value = True Then
    ocjenaTuga = 4
ElseIf Me.SadButton5.Value = True Then
    ocjenaTuga = 5
ElseIf Me.SadButton6.Value = True Then
    ocjenaTuga = 6
ElseIf Me.SadButton7.Value = True Then
    ocjenaTuga = 7
ElseIf Me.SadButton8.Value = True Then
    ocjenaTuga = 8
ElseIf Me.SadButton9.Value = True Then
    ocjenaTuga = 9
ElseIf Me.SadButton10.Value = True Then
    ocjenaTuga = 10
ElseIf Me.SadButton11.Value = True Then
    ocjenaTuga = 0
End If

' Provjera ocjene za emociju STRAH
If Me.FearButton1.Value = True Then
    ocjenaStrah = 1
ElseIf Me.FearButton2.Value = True Then
    ocjenaStrah = 2
ElseIf Me.FearButton3.Value = True Then
    ocjenaStrah = 3
ElseIf Me.FearButton4.Value = True Then
    ocjenaStrah = 4
ElseIf Me.FearButton5.Value = True Then
    ocjenaStrah = 5
ElseIf Me.FearButton6.Value = True Then
    ocjenaStrah = 6
ElseIf Me.FearButton7.Value = True Then
    ocjenaStrah = 7
ElseIf Me.FearButton8.Value = True Then
    ocjenaStrah = 8
ElseIf Me.FearButton9.Value = True Then
    ocjenaStrah = 9
ElseIf Me.FearButton10.Value = True Then
    ocjenaStrah = 10
ElseIf Me.FearButton11.Value = True Then
    ocjenaStrah = 0
End If

' Provjera ocjene za emociju LJUTNJA
If Me.AngryButton1.Value = True Then
    ocjenaLjutnja = 1
ElseIf Me.AngryButton2.Value = True Then
    ocjenaLjutnja = 2
ElseIf Me.AngryButton3.Value = True Then
    ocjenaLjutnja = 3
ElseIf Me.AngryButton4.Value = True Then
    ocjenaLjutnja = 4
ElseIf Me.AngryButton5.Value = True Then
    ocjenaLjutnja = 5
ElseIf Me.AngryButton6.Value = True Then
    ocjenaLjutnja = 6
ElseIf Me.AngryButton7.Value = True Then
    ocjenaLjutnja = 7

```



```

ElseIf Me.AngryButton8.Value = True Then
    ocjenaLjutnja = 8
ElseIf Me.AngryButton9.Value = True Then
    ocjenaLjutnja = 9
ElseIf Me.AngryButton10.Value = True Then
    ocjenaLjutnja = 10
ElseIf Me.AngryButton11.Value = True Then
    ocjenaLjutnja = 0
End If

'''''''''' Provjera ocjene za emociju NEUTRALNO
'If Me.NeutralButton1.Value = True Then
'    ocjenaNeutralno = 1
'ElseIf Me.NeutralButton2.Value = True Then
'    ocjenaNeutralno = 2
'ElseIf Me.NeutralButton3.Value = True Then
'    ocjenaNeutralno = 3
'ElseIf Me.NeutralButton4.Value = True Then
'    ocjenaNeutralno = 4
'ElseIf Me.NeutralButton5.Value = True Then
'    ocjenaNeutralno = 5
'ElseIf Me.NeutralButton6.Value = True Then
'    ocjenaNeutralno = 6
'ElseIf Me.NeutralButton7.Value = True Then
'    ocjenaNeutralno = 7
'ElseIf Me.NeutralButton8.Value = True Then
'    ocjenaNeutralno = 8
'ElseIf Me.NeutralButton9.Value = True Then
'    ocjenaNeutralno = 9
'ElseIf Me.NeutralButton10.Value = True Then
'    ocjenaNeutralno = 10
'ElseIf Me.NeutralButton11.Value = True Then
'    ocjenaNeutralno = 0
'End If

' Provjera ocjene za UGODU
If Me.UgodaButton1.Value = True Then
    ocjenaUgoda = 1
ElseIf Me.UgodaButton2.Value = True Then
    ocjenaUgoda = 2
ElseIf Me.UgodaButton3.Value = True Then
    ocjenaUgoda = 3
ElseIf Me.UgodaButton4.Value = True Then
    ocjenaUgoda = 4
ElseIf Me.UgodaButton5.Value = True Then
    ocjenaUgoda = 5
ElseIf Me.UgodaButton6.Value = True Then
    ocjenaUgoda = 6
ElseIf Me.UgodaButton7.Value = True Then
    ocjenaUgoda = 7
ElseIf Me.UgodaButton8.Value = True Then
    ocjenaUgoda = 8
ElseIf Me.UgodaButton9.Value = True Then
    ocjenaUgoda = 9
End If

' Provjera ocjene za POBUDENOST
If Me.PobudenostButton1.Value = True Then
    ocjenaPobudenost = 1
ElseIf Me.PobudenostButton2.Value = True Then
    ocjenaPobudenost = 2
ElseIf Me.PobudenostButton3.Value = True Then
    ocjenaPobudenost = 3
ElseIf Me.PobudenostButton4.Value = True Then
    ocjenaPobudenost = 4
ElseIf Me.PobudenostButton5.Value = True Then
    ocjenaPobudenost = 5
ElseIf Me.PobudenostButton6.Value = True Then
    ocjenaPobudenost = 6
ElseIf Me.PobudenostButton7.Value = True Then
    ocjenaPobudenost = 7
ElseIf Me.PobudenostButton8.Value = True Then
    ocjenaPobudenost = 8
ElseIf Me.PobudenostButton9.Value = True Then
    ocjenaPobudenost = 9
End If

```

```

If Me.DrugoButton.Value = True Then
    ocjenaDrugo = 1
End If

'Definiranje pokazivačna ȋije (cells)
brojredka = brojac + 1
rangeZaStatus = "C" & CStr(brojredka)
rangeZaSrecu = "D" & CStr(brojredka)
rangeZaTugu = "E" & CStr(brojredka)
rangeZaStrah = "F" & CStr(brojredka)
rangeZaLjutnju = "G" & CStr(brojredka)
'rangeZaNeutralno = "H" & CStr(brojredka)
rangeZaKomentar = "I" & CStr(brojredka)
rangeZaDrugo = "J" & CStr(brojredka)
rangeZaUgodu = "K" & CStr(brojredka)
rangeZaPobudenost = "L" & CStr(brojredka)

'Upis vrijednosti u cell-ove
Range(rangeZaStatus).Value = "SAVE"
Range(rangeZaSrecu).Value = ocjenaSreca
Range(rangeZaTugu).Value = ocjenaTuga
Range(rangeZaStrah).Value = ocjenaStrah
Range(rangeZaLjutnju).Value = ocjenaLjutnja
'Range(rangeZaNeutralno).Value = ocjenaNeutralno
Range(rangeZaKomentar).Value = Me.KomentartBox.Value
Range(rangeZaDrugo).Value = ocjenaDrugo
Range(rangeZaUgodu).Value = ocjenaUgoda
Range(rangeZaPobudenost).Value = ocjenaPobudenost

Me.StatusLight.Caption = "SPREMLJENO"
Me.StatusLight.BackColor = &H80FF80

' MsgBox CStr(ocjenaSreca) & " " & CStr(ocjenaTuga) & " " & CStr(ocjenaStrah) & " "
& CStr(ocjenaLjutnja) & " " & CStr(ocjenaNeutralno)
' MsgBox CStr(ocjenaUgoda) & " " & CStr(ocjenaPobudenost)

'Spremanje datoteke na disk

Application.DisplayAlerts = False
ThisWorkbook.Save
Application.DisplayAlerts = True

End Sub

Private Sub PreviousButton_Click()
    brojac = brojac - 1

    If brojac < 1 Then
        brojac = 1
    End If
    Me.Label8.Caption = Str(brojac)

    brojredka = brojac + 1
    rangeZaStatus = "C" & CStr(brojredka)
    Range(rangeZaStatus).Select
    If Range(rangeZaStatus).Value = "" Then
        Me.StatusLight.Caption = "NIJE SPREMLJENO"
        Me.StatusLight.BackColor = &H8080FF

        Me.HappyButton11.Value = True
        Me.SadButton11.Value = True
        Me.FearButton11.Value = True
        Me.AngryButton11.Value = True
        'Me.NeutralButton11.Value = True
        Me.DrugoButton.Value = False
        Me.KomentartBox.Value = ""
        Me.UgodaButton5.Value = True
        Me.PobudenostButton5.Value = True

    ElseIf Range(rangeZaStatus).Value = "SAVE" Then
        Me.StatusLight.Caption = "SPREMLJENO"
        Me.StatusLight.BackColor = &H80FF80

        ' Preuzimanje informacija o ocjeni emocije iz tablice kako bi se prikazali u aplikaciji

```

```
brojredka = brojrac + 1  
rangeZaSrecu = "D" & CStr(brojredka)  
rangeZaTugu = "E" & CStr(brojredka)  
rangeZaStrah = "F" & CStr(brojredka)  
rangeZaLjutnju = "G" & CStr(brojredka)  
''''''''rangeZaNeutralno = "H" & CStr(brojredka)  
rangeZaKomentar = "I" & CStr(brojredka)  
rangeZaDrugo = "J" & CStr(brojredka)  
rangeZaUgodu = "K" & CStr(brojredka)  
rangeZaPobudenost = "L" & CStr(brojredka)  
  
ocjenaSreca = Range(rangeZaSrecu).Value  
ocjenaTuga = Range(rangeZaTugu).Value  
ocjenaStrah = Range(rangeZaStrah).Value  
ocjenaLjutnja = Range(rangeZaLjutnju).Value  
''''''''ocjenaNeutralno = Range(rangeZaNeutralno).Value  
Me.KomentartBox.Value = Range(rangeZaKomentar).Value  
ocjenaDrugo = Range(rangeZaDrugo).Value  
ocjenaUgoda = Range(rangeZaUgodu).Value  
ocjenaPobudenost = Range(rangeZaPobudenost).Value  
  
' Povratak informacija u sućenje za svaku pojedinu emociju  
' Povratak ocjene za emociju SREB  
If ocjenaSreca = 1 Then  
    Me.HappyButton1.Value = True  
ElseIf ocjenaSreca = 2 Then  
    Me.HappyButton2.Value = True  
ElseIf ocjenaSreca = 3 Then  
    Me.HappyButton3.Value = True  
ElseIf ocjenaSreca = 4 Then  
    Me.HappyButton4.Value = True  
ElseIf ocjenaSreca = 5 Then  
    Me.HappyButton5.Value = True  
ElseIf ocjenaSreca = 6 Then  
    Me.HappyButton6.Value = True  
ElseIf ocjenaSreca = 7 Then  
    Me.HappyButton7.Value = True  
ElseIf ocjenaSreca = 8 Then  
    Me.HappyButton8.Value = True  
ElseIf ocjenaSreca = 9 Then  
    Me.HappyButton9.Value = True  
ElseIf ocjenaSreca = 10 Then  
    Me.HappyButton10.Value = True  
ElseIf ocjenaSreca = 0 Then  
    Me.HappyButton11.Value = True  
End If  
  
' Povratak ocjene za emociju TUGA  
If ocjenaTuga = 1 Then  
    Me.SadButton1.Value = True  
ElseIf ocjenaTuga = 2 Then  
    Me.SadButton2.Value = True  
ElseIf ocjenaTuga = 3 Then  
    Me.SadButton3.Value = True  
ElseIf ocjenaTuga = 4 Then  
    Me.SadButton4.Value = True  
ElseIf ocjenaTuga = 5 Then  
    Me.SadButton5.Value = True  
ElseIf ocjenaTuga = 6 Then  
    Me.SadButton6.Value = True  
ElseIf ocjenaTuga = 7 Then  
    Me.SadButton7.Value = True  
ElseIf ocjenaTuga = 8 Then  
    Me.SadButton8.Value = True  
ElseIf ocjenaTuga = 9 Then  
    Me.SadButton9.Value = True  
ElseIf ocjenaTuga = 10 Then  
    Me.SadButton10.Value = True  
ElseIf ocjenaTuga = 0 Then  
    Me.SadButton11.Value = True  
End If  
  
' Povratak ocjene za emociju STRAH  
If ocjenaStrah = 1 Then  
    Me.FearButton1.Value = True  
ElseIf ocjenaStrah = 2 Then
```

```
        Me.FearButton2.Value = True
    ElseIf ocjenaStrah = 3 Then
        Me.FearButton3.Value = True
    ElseIf ocjenaStrah = 4 Then
        Me.FearButton4.Value = True
    ElseIf ocjenaStrah = 5 Then
        Me.FearButton5.Value = True
    ElseIf ocjenaStrah = 6 Then
        Me.FearButton6.Value = True
    ElseIf ocjenaStrah = 7 Then
        Me.FearButton7.Value = True
    ElseIf ocjenaStrah = 8 Then
        Me.FearButton8.Value = True
    ElseIf ocjenaStrah = 9 Then
        Me.FearButton9.Value = True
    ElseIf ocjenaStrah = 10 Then
        Me.FearButton10.Value = True
    ElseIf ocjenaStrah = 0 Then
        Me.FearButton11.Value = True
    End If

' Povratak ocjene za emociju LJUTNJA
If ocjenaLjutnja = 1 Then
    Me.AngryButton1.Value = True
ElseIf ocjenaLjutnja = 2 Then
    Me.AngryButton2.Value = True
ElseIf ocjenaLjutnja = 3 Then
    Me.AngryButton3.Value = True
ElseIf ocjenaLjutnja = 4 Then
    Me.AngryButton4.Value = True
ElseIf ocjenaLjutnja = 5 Then
    Me.AngryButton5.Value = True
ElseIf ocjenaLjutnja = 6 Then
    Me.AngryButton6.Value = True
ElseIf ocjenaLjutnja = 7 Then
    Me.AngryButton7.Value = True
ElseIf ocjenaLjutnja = 8 Then
    Me.AngryButton8.Value = True
ElseIf ocjenaLjutnja = 9 Then
    Me.AngryButton9.Value = True
ElseIf ocjenaLjutnja = 10 Then
    Me.AngryButton10.Value = True
ElseIf ocjenaLjutnja = 0 Then
    Me.AngryButton11.Value = True
End If

' Povratak ocjene za emociju NEUTRALNO
'If ocjenaNeutralno = 1 Then
'    Me.NeutralButton1.Value = True
'ElseIf ocjenaNeutralno = 2 Then
'    Me.NeutralButton2.Value = True
'ElseIf ocjenaNeutralno = 3 Then
'    Me.NeutralButton3.Value = True
'ElseIf ocjenaNeutralno = 4 Then
'    Me.NeutralButton4.Value = True
'ElseIf ocjenaNeutralno = 5 Then
'    Me.NeutralButton5.Value = True
'ElseIf ocjenaNeutralno = 6 Then
'    Me.NeutralButton6.Value = True
'ElseIf ocjenaNeutralno = 7 Then
'    Me.NeutralButton7.Value = True
'ElseIf ocjenaNeutralno = 8 Then
'    Me.NeutralButton8.Value = True
'ElseIf ocjenaNeutralno = 9 Then
'    Me.NeutralButton9.Value = True
'ElseIf ocjenaNeutralno = 10 Then
'    Me.NeutralButton10.Value = True
'ElseIf ocjenaNeutralno = 0 Then
'    Me.NeutralButton11.Value = True
'End If

' Povratak ocjene za emociju UGODU
If ocjenaUgoda = 1 Then
    Me.UgodaButton1.Value = True
ElseIf ocjenaUgoda = 2 Then
    Me.UgodaButton2.Value = True
ElseIf ocjenaUgoda = 3 Then
```

```

        Me.UgodaButton3.Value = True
    ElseIf ocjenaUgoda = 4 Then
        Me.UgodaButton4.Value = True
    ElseIf ocjenaUgoda = 5 Then
        Me.UgodaButton5.Value = True
    ElseIf ocjenaUgoda = 6 Then
        Me.UgodaButton6.Value = True
    ElseIf ocjenaUgoda = 7 Then
        Me.UgodaButton7.Value = True
    ElseIf ocjenaUgoda = 8 Then
        Me.UgodaButton8.Value = True
    ElseIf ocjenaUgoda = 9 Then
        Me.UgodaButton9.Value = True
    End If

    ' Povratak ocjene za emociju POBUĐENOST
    If ocjenaPobudenost = 1 Then
        Me.PobudenostButton1.Value = True
    ElseIf ocjenaPobudenost = 2 Then
        Me.PobudenostButton2.Value = True
    ElseIf ocjenaPobudenost = 3 Then
        Me.PobudenostButton3.Value = True
    ElseIf ocjenaPobudenost = 4 Then
        Me.PobudenostButton4.Value = True
    ElseIf ocjenaPobudenost = 5 Then
        Me.PobudenostButton5.Value = True
    ElseIf ocjenaPobudenost = 6 Then
        Me.PobudenostButton6.Value = True
    ElseIf ocjenaPobudenost = 7 Then
        Me.PobudenostButton7.Value = True
    ElseIf ocjenaPobudenost = 8 Then
        Me.PobudenostButton8.Value = True
    ElseIf ocjenaPobudenost = 9 Then
        Me.PobudenostButton9.Value = True
    End If

    'Prvo pobrisati dugme Drugo, pa onda ako je postavljeno ga vratiti
    Me.DrugoButton.Value = False
    If ocjenaDrugo = 1 Then
        Me.DrugoButton.Value = True
    End If

End If

End Sub

Private Sub NextButton_Click()
    brojac = brojac + 1

    If brojac > 674 Then
        brojac = 674
    End If

    Me.Label8.Caption = Str(brojac)
    brojredka = brojac + 1
    rangeZaStatus = "C" & CStr(brojredka)
    Range(rangeZaStatus).Select
    If Range(rangeZaStatus).Value = "" Then
        Me.StatusLight.Caption = "NIJE SPREMLJENO"
        Me.StatusLight.BackColor = &H8080FF

        Me.HappyButton11.Value = True
        Me.SadButton11.Value = True
        Me.FearButton11.Value = True
        Me.AngryButton11.Value = True
        '.....Me.NeutralButton11.Value = True
        Me.DrugoButton.Value = False
        Me.KomentartBox.Value = ""
        Me.UgodaButton5.Value = True
        Me.PobudenostButton5.Value = True

    ElseIf Range(rangeZaStatus).Value = "SAVE" Then
        Me.StatusLight.Caption = "SPREMLJENO"
        Me.StatusLight.BackColor = &H80FF80
    End If
End Sub

```

```

Preuzimanje informacija o ocjeni emocije iz tablice kako bi se prikazali u aplikaciji
brojredka = brojac + 1
rangeZaSrecu = "D" & CStr(brojredka)
rangeZaTugu = "E" & CStr(brojredka)
rangeZaStrah = "F" & CStr(brojredka)
rangeZaLjtnju = "G" & CStr(brojredka)
''''''''rangeZaNeutralno = "H" & CStr(brojredka)
rangeZaKomentar = "I" & CStr(brojredka)
rangeZaDrugo = "J" & CStr(brojredka)
rangeZaUgodu = "K" & CStr(brojredka)
rangeZaPobudenost = "L" & CStr(brojredka)

ocjenaSreca = Range(rangeZaSrecu).Value
ocjenaTuga = Range(rangeZaTugu).Value
ocjenaStrah = Range(rangeZaStrah).Value
ocjenaLjtnja = Range(rangeZaLjtnju).Value
''''''''ocjenaNeutralno = Range(rangeZaNeutralno).Value
Me.KomentarTextBox.Value = Range(rangeZaKomentar).Value
ocjenaDrugo = Range(rangeZaDrugo).Value
ocjenaUgoda = Range(rangeZaUgodu).Value
ocjenaPobudenost = Range(rangeZaPobudenost).Value

' Povratak informacija u sučelje za svaku pojedinu emociju
' Povratak ocjene za emociju SREB
If ocjenaSreca = 1 Then
    Me.HappyButton1.Value = True
ElseIf ocjenaSreca = 2 Then
    Me.HappyButton2.Value = True
ElseIf ocjenaSreca = 3 Then
    Me.HappyButton3.Value = True
ElseIf ocjenaSreca = 4 Then
    Me.HappyButton4.Value = True
ElseIf ocjenaSreca = 5 Then
    Me.HappyButton5.Value = True
ElseIf ocjenaSreca = 6 Then
    Me.HappyButton6.Value = True
ElseIf ocjenaSreca = 7 Then
    Me.HappyButton7.Value = True
ElseIf ocjenaSreca = 8 Then
    Me.HappyButton8.Value = True
ElseIf ocjenaSreca = 9 Then
    Me.HappyButton9.Value = True
ElseIf ocjenaSreca = 10 Then
    Me.HappyButton10.Value = True
ElseIf ocjenaSreca = 0 Then
    Me.HappyButton11.Value = True
End If

' Povratak ocjene za emociju TUGA
If ocjenaTuga = 1 Then
    Me.SadButton1.Value = True
ElseIf ocjenaTuga = 2 Then
    Me.SadButton2.Value = True
ElseIf ocjenaTuga = 3 Then
    Me.SadButton3.Value = True
ElseIf ocjenaTuga = 4 Then
    Me.SadButton4.Value = True
ElseIf ocjenaTuga = 5 Then
    Me.SadButton5.Value = True
ElseIf ocjenaTuga = 6 Then
    Me.SadButton6.Value = True
ElseIf ocjenaTuga = 7 Then
    Me.SadButton7.Value = True
ElseIf ocjenaTuga = 8 Then
    Me.SadButton8.Value = True
ElseIf ocjenaTuga = 9 Then
    Me.SadButton9.Value = True
ElseIf ocjenaTuga = 10 Then
    Me.SadButton10.Value = True
ElseIf ocjenaTuga = 0 Then
    Me.SadButton11.Value = True
End If

' Povratak ocjene za emociju STRAH
If ocjenaStrah = 1 Then
    Me.FearButton1.Value = True
ElseIf ocjenaStrah = 2 Then

```

```
Me.FearButton2.Value = True
ElseIf ocjenaStrah = 3 Then
Me.FearButton3.Value = True
ElseIf ocjenaStrah = 4 Then
Me.FearButton4.Value = True
ElseIf ocjenaStrah = 5 Then
Me.FearButton5.Value = True
ElseIf ocjenaStrah = 6 Then
Me.FearButton6.Value = True
ElseIf ocjenaStrah = 7 Then
Me.FearButton7.Value = True
ElseIf ocjenaStrah = 8 Then
Me.FearButton8.Value = True
ElseIf ocjenaStrah = 9 Then
Me.FearButton9.Value = True
ElseIf ocjenaStrah = 10 Then
Me.FearButton10.Value = True
ElseIf ocjenaStrah = 0 Then
Me.FearButton11.Value = True
End If

' Povratak ocjene za emociju LJUTNJA
If ocjenaLjutnja = 1 Then
Me.AngryButton1.Value = True
ElseIf ocjenaLjutnja = 2 Then
Me.AngryButton2.Value = True
ElseIf ocjenaLjutnja = 3 Then
Me.AngryButton3.Value = True
ElseIf ocjenaLjutnja = 4 Then
Me.AngryButton4.Value = True
ElseIf ocjenaLjutnja = 5 Then
Me.AngryButton5.Value = True
ElseIf ocjenaLjutnja = 6 Then
Me.AngryButton6.Value = True
ElseIf ocjenaLjutnja = 7 Then
Me.AngryButton7.Value = True
ElseIf ocjenaLjutnja = 8 Then
Me.AngryButton8.Value = True
ElseIf ocjenaLjutnja = 9 Then
Me.AngryButton9.Value = True
ElseIf ocjenaLjutnja = 10 Then
Me.AngryButton10.Value = True
ElseIf ocjenaLjutnja = 0 Then
Me.AngryButton11.Value = True
End If

'''''''''' Povratak ocjene za emociju NEUTRALNO
'If ocjenaNeutralno = 1 Then
' Me.NeutralButton1.Value = True
'ElseIf ocjenaNeutralno = 2 Then
' Me.NeutralButton2.Value = True
'ElseIf ocjenaNeutralno = 3 Then
' Me.NeutralButton3.Value = True
'ElseIf ocjenaNeutralno = 4 Then
' Me.NeutralButton4.Value = True
'ElseIf ocjenaNeutralno = 5 Then
' Me.NeutralButton5.Value = True
'ElseIf ocjenaNeutralno = 6 Then
' Me.NeutralButton6.Value = True
'ElseIf ocjenaNeutralno = 7 Then
' Me.NeutralButton7.Value = True
'ElseIf ocjenaNeutralno = 8 Then
' Me.NeutralButton8.Value = True
'ElseIf ocjenaNeutralno = 9 Then
' Me.NeutralButton9.Value = True
'ElseIf ocjenaNeutralno = 10 Then
' Me.NeutralButton10.Value = True
'ElseIf ocjenaNeutralno = 0 Then
' Me.NeutralButton11.Value = True
'End If

' Povratak ocjene za emociju UGODU
If ocjenaUgoda = 1 Then
Me.UgodaButton1.Value = True
ElseIf ocjenaUgoda = 2 Then
Me.UgodaButton2.Value = True
ElseIf ocjenaUgoda = 3 Then
```

```

        Me.UgodaButton3.Value = True
    ElseIf ocjenaUgoda = 4 Then
        Me.UgodaButton4.Value = True
    ElseIf ocjenaUgoda = 5 Then
        Me.UgodaButton5.Value = True
    ElseIf ocjenaUgoda = 6 Then
        Me.UgodaButton6.Value = True
    ElseIf ocjenaUgoda = 7 Then
        Me.UgodaButton7.Value = True
    ElseIf ocjenaUgoda = 8 Then
        Me.UgodaButton8.Value = True
    ElseIf ocjenaUgoda = 9 Then
        Me.UgodaButton9.Value = True
    End If

    ' Povratak ocjene za emociju POBUĐENOST
    If ocjenaPobudenost = 1 Then
        Me.PobudenostButton1.Value = True
    ElseIf ocjenaPobudenost = 2 Then
        Me.PobudenostButton2.Value = True
    ElseIf ocjenaPobudenost = 3 Then
        Me.PobudenostButton3.Value = True
    ElseIf ocjenaPobudenost = 4 Then
        Me.PobudenostButton4.Value = True
    ElseIf ocjenaPobudenost = 5 Then
        Me.PobudenostButton5.Value = True
    ElseIf ocjenaPobudenost = 6 Then
        Me.PobudenostButton6.Value = True
    ElseIf ocjenaPobudenost = 7 Then
        Me.PobudenostButton7.Value = True
    ElseIf ocjenaPobudenost = 8 Then
        Me.PobudenostButton8.Value = True
    ElseIf ocjenaPobudenost = 9 Then
        Me.PobudenostButton9.Value = True
    End If

    'Prvo pobrisati dugme Drugo, pa onda ako je postavljeno ga vratiti
    Me.Drugobutton.Value = False
    If ocjenaDrugo = 1 Then
        Me.Drugobutton.Value = True
    End If

End If

End Sub

Private Sub UserForm_Activate()

    brojac = 1
    countClick = 0

    Range("C2").Select
    Do Until Selection.Value = ""
        brojac = brojac + 1
        If brojac > 674 Then
            Exit Do
        End If
        Selection.Offset(1, 0).Select
    Loop
    Me.Label8.Caption = Str(brojac)

    If brojac > 674 Then
        brojac = 1
        Me.Label8.Caption = Str(brojac)
        brojredka = brojac + 1
        rangeZaStatus = "C" & CStr(brojredka)
        Range(rangeZaStatus).Select
        If Range(rangeZaStatus).Value = "" Then
            Me.StatusLight.Caption = "NIJE SPREMLJENO"
            Me.StatusLight.BackColor = &H8080FF

            Me.HappyButton11.Value = True
            Me.SadButton11.Value = True
        End If
    End If
End Sub

```



```

Me.FearButton11.Value = True
Me.AngryButton11.Value = True
'*****Me.NeutralButton11.Value = True
Me.DrugoButton.Value = False
Me.UgodaButton5.Value = True
Me.PobudenostButton5.Value = True

ElseIf Range(rangeZaStatus).Value = "SAVE" Then
Me.StatusLight.Caption = "SPREMLJENO"
Me.StatusLight.BackColor = &H80FF80

' Preuzimanje informacija o ocjeni emocije iz tablice kako bi se prikazali u
aplikaciji
brojredka = brojac + 1
rangeZaSrecu = "D" & CStr(brojredka)
rangeZaTugu = "E" & CStr(brojredka)
rangeZaStrah = "F" & CStr(brojredka)
rangeZaLjutnju = "G" & CStr(brojredka)
'*****rangeZaNeutralno = "H" & CStr(brojredka)
rangeZaKomentar = "I" & CStr(brojredka)
rangeZaDrugo = "J" & CStr(brojredka)
rangeZaUgodu = "K" & CStr(brojredka)
rangeZaPobudenost = "L" & CStr(brojredka)

ocjenaSreca = Range(rangeZaSrecu).Value
ocjenaTuga = Range(rangeZaTugu).Value
ocjenaStrah = Range(rangeZaStrah).Value
ocjenaLjutnja = Range(rangeZaLjutnju).Value
'*****ocjenaNeutralno = Range(rangeZaNeutralno).Value
Me.KomentartBox.Value = Range(rangeZaKomentar).Value
ocjenaDrugo = Range(rangeZaDrugo).Value
ocjenaUgoda = Range(rangeZaUgodu).Value
ocjenaPobudenost = Range(rangeZaPobudenost).Value

'Povratak informacija u sučelje za svaku pojedinu emociju
' Povratak ocjene za emociju SREB
If ocjenaSreca = 1 Then
Me.HappyButton1.Value = True
ElseIf ocjenaSreca = 2 Then
Me.HappyButton2.Value = True
ElseIf ocjenaSreca = 3 Then
Me.HappyButton3.Value = True
ElseIf ocjenaSreca = 4 Then
Me.HappyButton4.Value = True
ElseIf ocjenaSreca = 5 Then
Me.HappyButton5.Value = True
ElseIf ocjenaSreca = 6 Then
Me.HappyButton6.Value = True
ElseIf ocjenaSreca = 7 Then
Me.HappyButton7.Value = True
ElseIf ocjenaSreca = 8 Then
Me.HappyButton8.Value = True
ElseIf ocjenaSreca = 9 Then
Me.HappyButton9.Value = True
ElseIf ocjenaSreca = 10 Then
Me.HappyButton10.Value = True
ElseIf ocjenaSreca = 0 Then
Me.HappyButton11.Value = True
End If

' Povratak ocjene za emociju TUGA
If ocjenaTuga = 1 Then
Me.SadButton1.Value = True
ElseIf ocjenaTuga = 2 Then
Me.SadButton2.Value = True
ElseIf ocjenaTuga = 3 Then
Me.SadButton3.Value = True
ElseIf ocjenaTuga = 4 Then
Me.SadButton4.Value = True
ElseIf ocjenaTuga = 5 Then
Me.SadButton5.Value = True
ElseIf ocjenaTuga = 6 Then
Me.SadButton6.Value = True
ElseIf ocjenaTuga = 7 Then
Me.SadButton7.Value = True
ElseIf ocjenaTuga = 8 Then
Me.SadButton8.Value = True

```

```
ElseIf ocjenaTuga = 9 Then
    Me.SadButton9.Value = True
ElseIf ocjenaTuga = 10 Then
    Me.SadButton10.Value = True
ElseIf ocjenaTuga = 0 Then
    Me.SadButton11.Value = True
End If

' Povratak ocjene za emociju STRAH
If ocjenaStrah = 1 Then
    Me.FearButton1.Value = True
ElseIf ocjenaStrah = 2 Then
    Me.FearButton2.Value = True
ElseIf ocjenaStrah = 3 Then
    Me.FearButton3.Value = True
ElseIf ocjenaStrah = 4 Then
    Me.FearButton4.Value = True
ElseIf ocjenaStrah = 5 Then
    Me.FearButton5.Value = True
ElseIf ocjenaStrah = 6 Then
    Me.FearButton6.Value = True
ElseIf ocjenaStrah = 7 Then
    Me.FearButton7.Value = True
ElseIf ocjenaStrah = 8 Then
    Me.FearButton8.Value = True
ElseIf ocjenaStrah = 9 Then
    Me.FearButton9.Value = True
ElseIf ocjenaStrah = 10 Then
    Me.FearButton10.Value = True
ElseIf ocjenaStrah = 0 Then
    Me.FearButton11.Value = True
End If

' Povratak ocjene za emociju LJUTNJA
If ocjenaLjutnja = 1 Then
    Me.AngryButton1.Value = True
ElseIf ocjenaLjutnja = 2 Then
    Me.AngryButton2.Value = True
ElseIf ocjenaLjutnja = 3 Then
    Me.AngryButton3.Value = True
ElseIf ocjenaLjutnja = 4 Then
    Me.AngryButton4.Value = True
ElseIf ocjenaLjutnja = 5 Then
    Me.AngryButton5.Value = True
ElseIf ocjenaLjutnja = 6 Then
    Me.AngryButton6.Value = True
ElseIf ocjenaLjutnja = 7 Then
    Me.AngryButton7.Value = True
ElseIf ocjenaLjutnja = 8 Then
    Me.AngryButton8.Value = True
ElseIf ocjenaLjutnja = 9 Then
    Me.AngryButton9.Value = True
ElseIf ocjenaLjutnja = 10 Then
    Me.AngryButton10.Value = True
ElseIf ocjenaLjutnja = 0 Then
    Me.AngryButton11.Value = True
End If

'''''''''' Povratak ocjene za emociju NEUTRALNO
'If ocjenaNeutralno = 1 Then
'    Me.NeutralButton1.Value = True
'ElseIf ocjenaNeutralno = 2 Then
'    Me.NeutralButton2.Value = True
'ElseIf ocjenaNeutralno = 3 Then
'    Me.NeutralButton3.Value = True
'ElseIf ocjenaNeutralno = 4 Then
'    Me.NeutralButton4.Value = True
'ElseIf ocjenaNeutralno = 5 Then
'    Me.NeutralButton5.Value = True
'ElseIf ocjenaNeutralno = 6 Then
'    Me.NeutralButton6.Value = True
'ElseIf ocjenaNeutralno = 7 Then
'    Me.NeutralButton7.Value = True
'ElseIf ocjenaNeutralno = 8 Then
'    Me.NeutralButton8.Value = True
'ElseIf ocjenaNeutralno = 9 Then
'    Me.NeutralButton9.Value = True
```

```
'ElseIf ocjenaNeutralno = 10 Then
'    Me.NeutralButton10.Value = True
'ElseIf ocjenaNeutralno = 0 Then
'    Me.NeutralButton11.Value = True
'End If

' Povratak ocjene za emociju UGODU
If ocjenaUgoda = 1 Then
    Me.UgodaButton1.Value = True
ElseIf ocjenaUgoda = 2 Then
    Me.UgodaButton2.Value = True
ElseIf ocjenaUgoda = 3 Then
    Me.UgodaButton3.Value = True
ElseIf ocjenaUgoda = 4 Then
    Me.UgodaButton4.Value = True
ElseIf ocjenaUgoda = 5 Then
    Me.UgodaButton5.Value = True
ElseIf ocjenaUgoda = 6 Then
    Me.UgodaButton6.Value = True
ElseIf ocjenaUgoda = 7 Then
    Me.UgodaButton7.Value = True
ElseIf ocjenaUgoda = 8 Then
    Me.UgodaButton8.Value = True
ElseIf ocjenaUgoda = 9 Then
    Me.UgodaButton9.Value = True
End If

' Povratak ocjene za emociju POBUDENOST
If ocjenaPobudenost = 1 Then
    Me.PobudenostButton1.Value = True
ElseIf ocjenaPobudenost = 2 Then
    Me.PobudenostButton2.Value = True
ElseIf ocjenaPobudenost = 3 Then
    Me.PobudenostButton3.Value = True
ElseIf ocjenaPobudenost = 4 Then
    Me.PobudenostButton4.Value = True
ElseIf ocjenaPobudenost = 5 Then
    Me.PobudenostButton5.Value = True
ElseIf ocjenaPobudenost = 6 Then
    Me.PobudenostButton6.Value = True
ElseIf ocjenaPobudenost = 7 Then
    Me.PobudenostButton7.Value = True
ElseIf ocjenaPobudenost = 8 Then
    Me.PobudenostButton8.Value = True
ElseIf ocjenaPobudenost = 9 Then
    Me.PobudenostButton9.Value = True
End If

'Prvo pobrisati dugme Drugo, pa onda ako je postavljeno ga vratiti
Me.DrugoButton.Value = False
If ocjenaDrugo = 1 Then
    Me.DrugoButton.Value = True
End If

End If

Else
    Me.HappyButton11.Value = True
    Me.SadButton11.Value = True
    Me.FearButton11.Value = True
    Me.AngryButton11.Value = True
    'Me.NeutralButton11.Value = True
    Me.DrugoButton.Value = False
    Me.UgodaButton5.Value = True
    Me.PobudenostButton5.Value = True
End If

End Sub
```

12.4 Privitak—Matlab kôdovi

12.4.1 tokenizer.m

```
function [ output_args ] = tokenizer( fname_in, fname_out )
%TOKENIZER Funkcija za tokenizaciju dokumenta
% (C) 2011 by Antonio Kolak <antonio.kolak@gmail.hr>
% Simple tokenization algorithm

try
    %fname_in='D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\primjeri\ulaz3.txt'
    %fname_out='D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\primjeri\tokens3.lst'
    fp_in = fopen(fname_in,'r');
    fp_out=fopen(fname_out,'w');

    % izmjena na \n
    delimiterSet = ';.,!?\\"()':[]/+==??{}><*'""|';
    delimiterSet(end+1)=native2unicode(13);
    digits = '0123456789';
    chars = 'abcdefghijklmnopqrstuvwxyz';
    chars =strcat(chars,upper(chars));
%     spaces = ' \t\n';
%     spaces='\x20\xA\xD';
%     spaces=strcat(strcat(char(20),char(10)),char(13));
    spaces=[native2unicode(32),native2unicode(10),native2unicode(9)];
    numberdelimiters = ',.';

    count_token=1;
    temp = fread(fp_in,inf,'char=>char');
    token = '';
    ch=temp(count_token);
    count_token=count_token+1;
    lookahead=temp(count_token);
    while true
        if ~ ch
            if token
                fprintf(fp_out,'%s\n',token);
                disp(token)
            end
            break;
        end
        if strfind(delimiterSet,ch)
            if token
                ptoken=size(token,2);
                if ~isempty(strfind(digits,token(ptoken))) &&
~isempty(strfind(digits,lookahead)) && ~isempty(strfind(numberdelimiters,ch))
                    token=strcat(token,ch);
                elseif ~isempty(strfind(chars,token(ptoken))) &&
~isempty(strfind(chars,lookahead)) && ~isempty(strfind(numberdelimiters,ch))
                    token=strcat(token,ch);
                else
                    fprintf(fp_out,'%s\n',token);
                    disp(token)
                    token='';
                    if ~(ch==native2unicode(32) || ch==native2unicode(10) ||
ch==native2unicode(9))
                        fprintf(fp_out,'%c\n',ch);
                        disp(ch)
                    end
                end
            end
            elseif (ch==native2unicode(32) || ch==native2unicode(10) ||
ch==native2unicode(9)) % must use this notation instead of variable spaces
                if token
                    fprintf(fp_out,'%s\n',token);
                    disp(token)
                    token='';
                end
            else
                token=strcat(token,ch);
            end
            ch=lookahead;
        end
    end
end
```

```
        count_token=count_token+1;
        lookahead=temp(count_token);

    end
    fclose(fp_in)
    fclose(fp_out);

    catch exception
        if (strcmp(exception.identifier,'MATLAB:inputArgUndefined'))
            error('MATLAB:myFunction:Dimensionality','Progrešna staza do ulazne datoteke');
        else if (strcmp(exception.identifier,'MATLAB:badsubscript'))
            % pokazivač je zašao dalje nego ima teksta
            token=strcat(token,ch);
            fprintf(fp_out,'%s\n',token);
            disp(token)
            fclose(fp_in);
            fclose(fp_out);
        else
            rethrow(exception);
        end
    end
end
```

12.4.2 fp.m

```

function [ output_args ] = fp( fname_in, fname_out )
%FP Function count number of same tokens
%
%   output=> Frequencies of tokens in text
try
    %fname_in='C:\Users\Antonio\Desktop\tokens2.lst'
    %fname_out='C:\Users\Antonio\Desktop\tokens-fp2.lst'
    fp_in = fopen(fname_in,'r');
    fp_out=fopen(fname_out,'w');

    % izmjena na \n
    delimiterSet = '.,!?\\"(){}[]/+--=?><*'";";
    delimiterSet(end+1)=native2unicode(13);
    delimiterSet(end+1)=native2unicode(32);
    delimiterSet(end+1)=native2unicode(10);
    delimiterSet(end+1)=native2unicode(9);

    tline = fgetl(fp_in);
    counter=1;
    dictStruct=struct;
    while ischar(tline)
        if isempty(strfind(delimiterSet,tline)) && ~isempty(tline)
            if isfield(dictStruct,tline)
                counter=getfield(dictStruct,tline)+1;
                dictStruct=setfield(dictStruct,tline,counter);
                counter=1;
            else
                dictStruct=setfield(dictStruct,tline,counter);
            end

            % disp(tline)
        end
        tline = fgetl(fp_in);
    end

    namesDict = fieldnames(dictStruct);
    valuesDict = cell2mat(struct2cell(dictStruct));
    for k=1:size(namesDict)
        fprintf(fp_out,'%s\t%d\n',char(namesDict(k)), valuesDict(k));
    end

    fclose(fp_in);
    fclose(fp_out);
catch exception
    if (strcmp(exception.identifier,'MATLAB:inputArgUndefined'))
        error('MATLAB:myFunction:Dimensionality','Progrešna staza do ulazne datoteke');
    else
        rethrow(exception);
    end
end
end

```

12.4.3 rfp

```

function [ output_args ] = rfp( fname_in, fname_out )
%RFP Function for count relative number of same tokens
%
%   output Relative Frequencies of tokens in text
try
    %fname_in='C:\Users\Antonio\Desktop\tokens2.lst'
    %fname_out='C:\Users\Antonio\Desktop\tokens-rfp2.lst'
    fp_in = fopen(fname_in,'r');
    fp_out=fopen(fname_out,'w');

    % izmjena na \n
    delimiterSet = '.,!?\\"()\'':[]/+--=??{}><*'""|';
    delimiterSet(end+1)=native2unicode(13);
    delimiterSet(end+1)=native2unicode(32);
    delimiterSet(end+1)=native2unicode(10);
    delimiterSet(end+1)=native2unicode(9);

    tline = fgetl(fp_in);
    counter=1;
    dictStruct=struct;
    while ischar(tline)
        if isempty(strfind(delimiterSet,tline)) && ~isempty(tline)
            if isfield(dictStruct,tline)
                counter=getfield(dictStruct,tline)+1;
                dictStruct=setfield(dictStruct,tline,counter);
                counter=1;
            else
                dictStruct=setfield(dictStruct,tline,counter);
            end

            % disp(tline)
        end
        tline = fgetl(fp_in);
    end

    namesDict = fieldnames(dictStruct);
    valuesDict = cell2mat(struct2cell(dictStruct));

    % sum all valus in stuct to make relative freq.
    total=sum(valuesDict);

    for k=1:size(namesDict)
        fprintf(fp_out,'%s\t%s\n',char(namesDict(k)),
num2str(valuesDict(k)/total,'%.12f'));
    end

    fclose(fp_in);
    fclose(fp_out);
catch exception
    if (strcmp(exception.identifier,'MATLAB:inputArgUndefined'))
        error('MATLAB:myFunction:Dimensionality','Progrešna staza do ulazne datoteke');
    else
        rethrow(exception);
    end
end
end

```

12.4.4 ttr.m

```

function [ TTR_var ] = ttr( fname_in, lowerFlag )
%TTR Funkcija računa Type Token ratio
% fname_in je putanja do ulazne datoteke tokena (zbog spremanja u strukturu
% moraju biti uklonjeni HR slova s kvačicom što se obavlja funkcijom
% bezkvacice.m
% kao izlaz iz funkcije dobiva se TTR faktor u varijabli TTR_var i kao
% zapis u matlab sučelju.
%
    try

        % lowerFlag=1;
        fp_in = fopen(fname_in,'r');

        % izmjena na \n
        delimiterSet = ';.,!?\"'():'[]/+--=??{}><*'""|';
        delimiterSet(end+1)=native2unicode(13);
        delimiterSet(end+1)=native2unicode(32);
        delimiterSet(end+1)=native2unicode(10);
        delimiterSet(end+1)=native2unicode(9);

        tline = fgetl(fp_in);
        counter=1;
        dictStruct=struct;
        while ischar(tline)
            if lowerFlag
                tline=lower(tline);
            end
            if isempty(strfind(delimiterSet,tline)) && ~isempty(tline)
                if isfield(dictStruct,tline)
                    counter=getfield(dictStruct,tline)+1;
                    dictStruct=setfield(dictStruct,tline,counter);
                    counter=1;
                else
                    dictStruct=setfield(dictStruct,tline,counter);
                end

                % disp(tline)
            end
            tline = fgetl(fp_in);
        end

        namesDict = fieldnames(dictStruct);
        valuesDict = cell2mat(struct2cell(dictStruct));

        % sum all values in struct to make relative freq.
        total=sum(valuesDict);
        sizeofDict=size(namesDict,1);

        [ 'Type/Token ratio for ',fname_in,': ',num2str(100*sizeofDict/total,'%2f'), '%']
        TTR_var=100*sizeofDict/total;

        %         for k=1:size(namesDict)
        %             fprintf(fp_out,'%s\t%s\n',char(namesDict(k)),
        num2str(valuesDict(k)/total,'%12f'));
        %         end

        fclose(fp_in);
        %         fclose(fp_out);
    catch exception
        if (strcmp(exception.identifier,'MATLAB:inputArgUndefined'))
            error('MATLAB:myFunction:Dimensionality','Progrešna staza do ulazne datoteke');
        else
            rethrow(exception);
        end
    end
end

```


12.4.5 unique1.m

```

function [ dictStructFieldNew ] = unique1( varargin )
%UNIQUE1 Vadi jedinstvene tokene
% Funkcija prima sve liste koje predstavljaju razrede tokena, te za svaki
% razred izdvađa jedinstvene tokene u odnosu na ostale razrede i zapisuje
% ih u novu datoteku po nazivom ime_stare_datoteke-unique-set.lst
% pored ispisa funkcija vraća strukturu sa svim razredima i njihovim
% jedinstvenim tokenima

numOfArg=nargin;
%fname_out='D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\primjeri\tokens.lst';

dictStructField=struct;
for k=1:numOfArg
    fname_in=cell2mat(varargin(k));

    lowerFlag=1;
    fp_in = fopen(fname_in,'r');

    % izmjena na \n
    delimiterSet = ';.,!?\\"() '':[]/+--=??{}><*'"/\|';
    delimiterSet(end+1)=native2unicode(13);
    delimiterSet(end+1)=native2unicode(32);
    delimiterSet(end+1)=native2unicode(10);
    delimiterSet(end+1)=native2unicode(9);

    tline = fgetl(fp_in);
    counter=1;
    dictStruct=struct;
    while ischar(tline)
        if lowerFlag
            tline=lower(tline);
        end
        if isempty(strfind(delimiterSet,tline)) && ~isempty(tline)
            if isfield(dictStruct,tline)
                counter=getfield(dictStruct,tline)+1;
                dictStruct=setfield(dictStruct,tline,counter);
                counter=1;
            else
                dictStruct=setfield(dictStruct,tline,counter);
            end
            % disp(tline)
        end
        tline = fgetl(fp_in);
    end
    namesDict = fieldnames(dictStruct);
    dictStructField=setfield(dictStructField,['tokens',num2str(k)],namesDict);
    fclose(fp_in);

end

numOfTokoenLists=size(fieldnames(dictStructField),1);
dictStructFieldNew=struct;
cell1={};
cell2={};
%oldDictStructField=dictStructField;

% Find the difference between two cell fields and save it into
% dictStructFieldNew

for k=1:numOfTokoenLists
    cell1=getfield(dictStructField,['tokens',num2str(k)]);
    for j=1:numOfTokoenLists
        cell2=getfield(dictStructField,['tokens',num2str(j)]);
        if k==j
            continue
        end
        cell1=differenceBtwCell(cell1,cell2);
        cell1=cell1';
    end
    dictStructFieldNew=setfield(dictStructFieldNew,['tokens',num2str(k)],cell1);
    cell1={};
    cell2={};
end

```

```

end

% Print data into N files
for i=1:numOfArg
    fname_in=cell2mat(varargin(i));
    fname_out=[fname_in(1:(strfind(fname_in, '.')-1)), '-unique-set.lst'];
    %fname_out=['D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\primjeri\tokens', num2str(i), '-unique-set', '.lst'];
    fp_out=fopen(fname_out, 'w');

    cellNewtokens=getfield(dictStructFieldNew, ['tokens', num2str(i)]);
    for k=1:size(cellNewtokens)
        fprintf(fp_out, '%s\n', char(cellNewtokens(k)));
    end
    fclose(fp_out);
end
end

```

12.4.6 rfp_unique.m

```

function [ dictStructFieldWithRfp ] = rfp_unique( varargin )
%RPF_UNIQUE Vadi jedinstvene tokene njihove relativne frekvencije
% Funkcija prima sve liste koje predstavljaju razrede tokena, te za svaki
% razred izdvaja jedinstvene tokene u odnosu na ostale razrede i zapisuje
% ih u novu datoteku po nazivom ime_stare_datoteke-unique-set.lst
% pored ispisa funkcija vraća strukturu sa svim razredima i njihovim
% jedinstvenim tokenima
% Funkcija je slična funkciji unique1, ali pored jedinstvenih tokena za
% svaki token zapisuje i njegovu relativnu frekvenciju pojavljivanja u
% njegovoj klasi

    numOfArg=nargin;
    %fname_out='D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\primjeri\tokens.lst';

    dictStructField=struct;
    for k=1:numOfArg
        fname_in=cell2mat(varargin(k));

        lowerFlag=1;
        fp_in = fopen(fname_in,'r');

        % izmjena na \n
        delimiterSet = ';.,!?\()"':[]/+--=??{}><*"\'|';
        delimiterSet(end+1)=native2unicode(13);
        delimiterSet(end+1)=native2unicode(32);
        delimiterSet(end+1)=native2unicode(10);
        delimiterSet(end+1)=native2unicode(9);

        tline = fgetl(fp_in);
        counter=1;
        dictStruct=struct;
        while ischar(tline)
            if lowerFlag
                tline=lower(tline);
            end
            if isempty(strfind(delimiterSet,tline)) && ~isempty(tline)
                if isfield(dictStruct,tline)
                    counter=getfield(dictStruct,tline)+1;
                    dictStruct=setfield(dictStruct,tline,counter);
                    counter=1;
                else
                    dictStruct=setfield(dictStruct,tline,counter);
                end
                % disp(tline)
            end
            tline = fgetl(fp_in);
        end
        %namesDict = fieldnames(dictStruct);
        dictStructField=setfield(dictStructField,['tokens',num2str(k)],dictStruct);
        fclose(fp_in);
    end

    % Relative freq.
    numOfTokoenLists=size(fieldnames(dictStructField),1);
    for h=1:numOfTokoenLists
        structN=(getfield(dictStructField,['tokens',num2str(h)]));
        values_structN = cell2mat(struct2cell(structN));
        total=sum(values_structN);
        names_of_structN=fieldnames(structN);
        numOfNames=size(names_of_structN,1);
        for g=1:numOfNames
            nameOfOne=names_of_structN(g);
            valueOfOne=getfield(structN,char(names_of_structN(g)));
            valueOfOne=valueOfOne/total;
            structN=setfield(structN,char(names_of_structN(g)),valueOfOne);
        end
        dictStructField=setfield(dictStructField,['tokens',num2str(h)],structN);
    end

```

```
dictStructFieldNew=struct;
dictStructFieldWithRfp=struct;
NewStruct=struct;
cell1={};
cell2={};
%oldDictStructField=dictStructField;

% Find the difference between two cell fields and save it into
% dictStructFieldNew

for k=1:numOfTokoenLists
    cell1=fieldnames(getfield(dictStructField, ['tokens', num2str(k)]));
    for j=1:numOfTokoenLists
        cell2=fieldnames(getfield(dictStructField, ['tokens', num2str(j)]));
        if k==j
            continue
        end
        cell1=differenceBtwCell(cell1, cell2);
        cell1=cell1';
    end
    dictStructFieldNew=setfield(dictStructFieldNew, ['tokens', num2str(k)], cell1);

    numOfCell=size(cell1,1);
    OneTokenList=getfield(dictStructField, ['tokens', num2str(k)]);
    for g=1:numOfCell
        nameOfCell=cell1(g);
        valueOfCell=getfield(OneTokenList, char(nameOfCell));
        NewStruct=setfield(NewStruct, char(nameOfCell), valueOfCell);
    end
end

dictStructFieldWithRfp=setfield(dictStructFieldWithRfp, ['tokens', num2str(k)], NewStruct);
NewStruct={};
end

% Print date into N file-s
for i=1:numOfArg
    fname_in=cell2mat(varargin(i));
    fname_out=[fname_in(1:(strfind(fname_in, '.')-1)), '-unique-fset.lst'];
    %fname_out=['D:\FER\Cetvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\tokens', num2str(i), '-unique-fset', '.lst'];
    fp_out=fopen(fname_out, 'w');

    StructNewtokens=getfield(dictStructFieldWithRfp, ['tokens', num2str(i)]);
    namesStructNewtokens=fieldnames(StructNewtokens);
    valuesStructNewtokens=cell2mat(struct2cell(StructNewtokens));
    for k=1:size(namesStructNewtokens,1)
        fprintf(fp_out, '%s\t%s\n', char(namesStructNewtokens(k)),
num2str(valuesStructNewtokens(k)));
    end
    fclose(fp_out);
end
end
```

12.4.7 differenceBtwCell.m

```
function [ newDiffCell ] = differenceBtwCell( CellStr1,CellStr2 )
%DIFFERENCEBTWCELL compere two cells and return difference
% Detailed explanation goes here

newDiffCell={};
sizeCellStr1=size(CellStr1,1);
sizeCellStr2=size(CellStr2,1);
counter=1;
logic=0;

for k=1:sizeCellStr1
    for j=1:sizeCellStr2
        if strcmp(CellStr1(k),CellStr2(j))
            logic=1;
            break
        end
    end
    if logic==0;
        newDiffCell(counter)=CellStr1(k);
        counter=counter+1;
    else
        logic=0;
    end
end
end
```

12.4.8 loadModels.m

```
function [ outStruct, outPaths ] = loadModels( modelsPath )
%LOADMODELS funkcija učitava modele iz datoketa sa dijelom naziva
% *unique-rfp.lst i sprema ih u strukturu
% izlaz iz funkcije je lista učitanih datoteka i struktura u koju su
% pohrenji tokeni iz razreda (modela).

%   shoud load in this order
%   sreca
%   tuga
%   strah
%   ljutnja
%   neutralno

outStruct=struct;
outPaths=struct;
tokenStruct=struct;
currentFolder = pwd;
cd(modelsPath)
r=dir('*unique-fset.lst');
cd(currentFolder)

numOfModels=size(r,1);
for k=1:numOfModels
    pathOfModel=[modelsPath,r(k,1).name];
    outPaths=setfield(outPaths,['outPath',num2str(k)],r(k,1).name);
    fp_in = fopen(pathOfModel,'r');
    tline = fgetl(fp_in);
    while ischar(tline)
        tline=strtrim(tline);
        [tlineKey,tlineValue] = strread(tline,'%s%f','delimiter','\t');
        tokenStruct=setfield(tokenStruct,char(tlineKey),tlineValue);
        tline = fgetl(fp_in);
    end
    fclose(fp_in);
    outStruct=setfield(outStruct,['tokens',num2str(k)],tokenStruct);
    tokenStruct={};
end
end
```

12.4.9 compareWithModels.m

```
function [ var_errors, errPaths ] = compareWithModels( modelsPath
, unknownModelPath, dispFlag )
%COMPAREWITHMODELS Za izračunavanja greške odstupanja od pojedinih modela
% Funkcija služi za prikaz rada primitivnog sustava (model unigram)
% Princip rada:
%         Učitaju se svi modeli pomoću loadMoels te se proglašava
%         onolika pogreška koliko ima tokena u rečenici. Nakon
%         toga se za svaki razred izračuna kolik bi bila pogreška
%         uzimajući vjerojatnosti tokena u pojedinoj klasi i
%         umanjujući pogrešku. Klasa koja ima najmanji ererror
%         predstavlja klasu kojoj prirpada dotična ubačena
%         rečanica
% Nepoznata rečenica se upisuje u unknown.txt datoteku, te se funkciji
% predaje staza do te datoteke.
% Zastavica dispFlag služi za prikaz rezultata, pošto se ista funkcija
% koristi i za računanje statistike.

txt2lower(unknownModelPath);
[ ModelsStruct, outPaths ] = loadModels( modelsPath );
errPaths=outPaths;
tokenizedStruct=tokenize(unknownModelPath);
fp_Struct=fpStruct(tokenizedStruct);
rfp_Struct=rfpStruct(fp_Struct);

numOfModels=size(fieldnames(ModelsStruct),1);
var_errors=[];

for k=1:numOfModels
    compSet=fieldnames(rfp_Struct);
    modelNSet=fieldnames(getfield(ModelsStruct, ['tokens', num2str(k)]));
    ModelNStruct=getfield(ModelsStruct, ['tokens', num2str(k)]);
    compSet=intersect(compSet, modelNSet);
    if dispFlag
        disp(sprintf('\n\nMatching %d tokens in unknown and
%s', size(compSet,1), char(getfield(outPaths, ['outPath', num2str(k)]))));
    end
    numOfCompSet=size(compSet,1);
    %     for h=1:numOfCompSet
    %         disp(sprintf('%s ', char(compSet)))
    %     end
    if dispFlag
        compSet
    end
    if dispFlag
        disp(sprintf('\n'))
    end
    var_error=0.0;
    rfp_Struct_Keys=fieldnames(rfp_Struct);
    numOfUnknown=size(fieldnames(rfp_Struct),1);
    for g=1:numOfUnknown
        if ~isempty(intersect(modelNSet, rfp_Struct_Keys(g)))
            var_error=var_error+abs(getfield(rfp_Struct, char(rfp_Struct_Keys(g)))-getfield(
ModelNStruct, char(rfp_Struct_Keys(g))));
        else
            var_error=var_error+1.0;
        end
    end
    var_errors(k)= var_error;
    if dispFlag
        disp(sprintf('Error: %.6f', var_error));
    end
end
numOfErrors=size(var_errors,2);
for g=1:numOfErrors
    if dispFlag
        disp(sprintf('Summary:'));
        disp(sprintf('%s error:
%.6f', char(getfield(outPaths, ['outPath', num2str(g)])), var_errors(g)));
    end
end
end
```

12.4.10 KEGto5class.m

```
function [ output_args ] = KEGto5class( metaKEGPath,classPath,FlagForNewMetapodaci )
%KEGT05CLASS Stvara 5 disjunktnih klasa (datoteka) iz transkripta KEG-a
% (to su sreca.txt, tuga.txt, strah.txt, ljutnja.txt, neutralno.txt

% metaKEGPath - Staza do metaKEG.xlsx datoteke u kojoj se nalazi
% transkript svih snimaka

% classPath - Staza do direktorija u kojem želimo da nam budu tekstualne
% datoteke klasa (sreca.txt, tuga.txt, strah.txt, ljutnja.txt,
% neutralno.txt)

%
% FlagForNewMetapodaci- zastavica koja označava kreiranje nove *.xlsx
% datoteke sa samo anotiranim snimkama, tj. onima koje imaju dobru ocjenu
% emocije (koje su pogodne za gradnju sustava)

M={};
[num, txt] = xlsread(metaKEGPath,'Sheet1');
KEGstruct=struct;

orderNum=num(:,1);
sizeOf_orderNum=size(orderNum,1);
sentences=txt(2:sizeOf_orderNum+1,7);
emotions=txt(2:sizeOf_orderNum+1,55);

srecaPath=[classPath,'sreca.txt'];
tugaPath=[classPath,'tuga.txt'];
ljutnjaPath=[classPath,'ljutnja.txt'];
strahPath=[classPath,'strah.txt'];
neutralnoPath=[classPath,'neutralno.txt'];

fp_out_sreca=fopen(srecaPath,'w');
fp_out_tuga=fopen(tugaPath,'w');
fp_out_ljutnja=fopen(ljutnjaPath,'w');
fp_out_strah=fopen(strahPath,'w');
fp_out_neutralno=fopen(neutralnoPath,'w');

counter_sreca=1;
counter_ljutnja=1;
counter_tuga=1;
counter_strah=1;
counter_neutralno=1;

counter_sve_emocije=1;
%incijalizacija M-a, upis headera
M(counter_sve_emocije,1)={'ID snimke'};
M(counter_sve_emocije,7)={'transkripcija'};
M(counter_sve_emocije,55)={'r - Emocija'};
counter_sve_emocije=counter_sve_emocije+1;

for k=1:sizeOf_orderNum
    if strcmp(char(emotions(k)),'SR')
        sreca(counter_sreca)=sentences(k);
        fprintf(fp_out_sreca,'%s\n',char(sentences(k)));

        M(counter_sve_emocije,1)=num2cell(counter_sve_emocije);
        M(counter_sve_emocije,7)=sentences(k);
        M(counter_sve_emocije,55)=emotions(k);

        counter_sreca=counter_sreca+1;
        counter_sve_emocije=counter_sve_emocije+1;
    elseif strcmp(char(emotions(k)),'LJ')
        ljutnja(counter_ljutnja)=sentences(k);
        fprintf(fp_out_ljutnja,'%s\n',char(sentences(k)));

        M(counter_sve_emocije,1)=num2cell(counter_sve_emocije);
        M(counter_sve_emocije,7)=sentences(k);
        M(counter_sve_emocije,55)=emotions(k);

        counter_ljutnja=counter_ljutnja+1;
        counter_sve_emocije=counter_sve_emocije+1;
    elseif strcmp(char(emotions(k)),'TU')
        tuga(counter_tuga)=sentences(k);
        fprintf(fp_out_tuga,'%s\n',char(sentences(k)));
    end
end
```



```

        M(counter_sve_emocije,1)=num2cell(counter_sve_emocije);
        M(counter_sve_emocije,7)=sentences(k);
        M(counter_sve_emocije,55)=emotions(k);

        counter_tuga=counter_tuga+1;
        counter_sve_emocije=counter_sve_emocije+1;
    elseif strcmp(char(emotions(k)),'ST')
        strah(counter_strah)=sentences(k);
        fprintf(fp_out_strah,'%s\n',char(sentences(k)));

        M(counter_sve_emocije,1)=num2cell(counter_sve_emocije);
        M(counter_sve_emocije,7)=sentences(k);
        M(counter_sve_emocije,55)=emotions(k);

        counter_strah=counter_strah+1;
        counter_sve_emocije=counter_sve_emocije+1;
    elseif strcmp(char(emotions(k)),'NE')
        neutralno(counter_neutralno)=sentences(k);
        fprintf(fp_out_neutralno,'%s\n',char(sentences(k)));

        M(counter_sve_emocije,1)=num2cell(counter_sve_emocije);
        M(counter_sve_emocije,7)=sentences(k);
        M(counter_sve_emocije,55)=emotions(k);

        counter_neutralno=counter_neutralno+1;
        counter_sve_emocije=counter_sve_emocije+1;
    end
end
sreca=sreca';
tuga=tuga';
ljutnja=ljutnja';
strah=strah';
neutralno=neutralno';

if FlagForNewMetapodaci==1
    mataNewKEPath=[metaKEGPath(1:(strfind(metaKEGPath, '.')-1)),'-
anotirane_snimke.xlsx'];
    xlswrite(mataNewKEPath,M)
end

fclose(fp_out_sreca);
fclose(fp_out_tuga);
fclose(fp_out_ljutnja);
fclose(fp_out_strah);
fclose(fp_out_neutralno);
%fclose('all');

end

```

12.4.11 bezkvacice.m

```

function [ output_args ] = bezkvacice( pathUTF )
%BEZKVACICE uklanja HR slova sa kvačicama
% Funkcija uklanja donju listu hrvatski slova i zamjenjuje ih sa
% ekvivalentima.
% Izlaz iz funkcije je *.txt file sa istim nazivom kao i ulazni uz
% dodatak -bez_kvacice.txt

% Č=>CX
% Ć=>CY
% Đ=>DX
% DŽ=>DZX
% Š=>SX
% Ž=>ZX

fp_in = fopen(pathUTF, 'r');
pathOut=[pathUTF(1:(strfind(pathUTF, '.')-1)), '-bez_kvacice.txt'];
fp_out = fopen(pathOut, 'w');
%temp = fread(fp_in,inf,'char=>char');
%temp=temp';
z=1;
tline = fgetl(fp_in);
while ischar(tline)
    varCX=strfind(tline, 'Ć');
    numOfVarCX=size(varCX,2);
    for k=1:numOfVarCX
        tline(varCX(z))='C';
        tline=strcat(tline(1:varCX(z)), 'X', tline(varCX(z)+1:end));
        varCX=strfind(tline, 'Ć');
    end

    varCY=strfind(tline, 'Č');
    numOfVarCY=size(varCY,2);
    for k=1:numOfVarCY
        tline(varCY(z))='C';
        tline=strcat(tline(1:varCY(z)), 'Y', tline(varCY(z)+1:end));
        varCY=strfind(tline, 'Č');
    end

    varDX=strfind(tline, 'Đ');
    numOfVarDX=size(varDX,2);
    for k=1:numOfVarDX
        tline(varDX(z))='D';
        tline=strcat(tline(1:varDX(z)), 'X', tline(varDX(z)+1:end));
        varDX=strfind(tline, 'Đ');
    end

    varDZX=strfind(tline, 'DŽ');
    numOfVarDZX=size(varDZX,2);
    for k=1:numOfVarDZX
        tline(varDZX(z))='D';
        tline=strcat(tline(1:varCX(z)), 'ZX', tline(varCX(z)+1:end));
        varDZX=strfind(tline, 'DŽ');
    end

    varSX=strfind(tline, 'Š');
    numOfVarSX=size(varSX,2);
    for k=1:numOfVarSX
        tline(varSX(z))='S';
        tline=strcat(tline(1:varSX(z)), 'X', tline(varSX(z)+1:end));
        varSX=strfind(tline, 'Š');
    end

    varZX=strfind(tline, 'Ž');
    numOfVarZX=size(varZX,2);
    for k=1:numOfVarZX
        tline(varZX(z))='Z';
        tline=strcat(tline(1:varZX(z)), 'X', tline(varZX(z)+1:end));
        varZX=strfind(tline, 'Ž');
    end
    fprintf(fp_out, '%s\n', tline);
    tline = fgetl(fp_in);
end
fclose(fp_in);
fclose(fp_out);end

```

12.4.12 bezkvacice2.m

```

function [ output_strings ] = bezkvacice2( string_sa_kvacicama )
%BEZKVACICE2 uklanja HR slova sa kvačicama
% Funkcija uklanja donju listu hrvetski slova i zamjenjuje ih sa
% ekvivalentima.
% Funkcija je ekvivalent funkciji bezkvacice osim što je ovdje ulaz
% string te se izlaz također zapisuje string. Funkcija zamjenjuje
% isključivo znakove nad jednom rečenicom.

% Ć=>CX
% Č=>CY
% Đ=>DX
% DŽ=>DZX
% Š=>SX
% Ž=>ZX

%fp_in = fopen(pathUTF,'r');
%pathOut=[pathUTF(1:(strfind(pathUTF,'.')-1)),'-bez_kvacice.txt'];
%fp_out = fopen(pathOut,'w');

%temp = fread(fp_in,inf,'char=>char');
%temp=temp';
z=1;
%tline = fgetl(fp_in);
tline=string_sa_kvacicama;

%while ischar(tline)
    varCX=strfind(tline,'Ć');
    numOfVarCX=size(varCX,2);
    for k=1:numOfVarCX
        tline(varCX(z))='C';
        tline=strcat(tline(1:varCX(z)),'X',tline(varCX(z)+1:end));
        varCX=strfind(tline,'Ć');
    end

    varCY=strfind(tline,'Č');
    numOfVarCY=size(varCY,2);
    for k=1:numOfVarCY
        tline(varCY(z))='C';
        tline=strcat(tline(1:varCY(z)),'Y',tline(varCY(z)+1:end));
        varCY=strfind(tline,'Č');
    end

    varDX=strfind(tline,'Đ');
    numOfVarDX=size(varDX,2);
    for k=1:numOfVarDX
        tline(varDX(z))='D';
        tline=strcat(tline(1:varDX(z)),'X',tline(varDX(z)+1:end));
        varDX=strfind(tline,'Đ');
    end

    varDZX=strfind(tline,'DŽ');
    numOfVarDZX=size(varDZX,2);
    for k=1:numOfVarDZX
        tline(varDZX(z))='D';
        tline=strcat(tline(1:varCX(z)),'ZX',tline(varCX(z)+1:end));
        varDZX=strfind(tline,'DŽ');
    end

    varSX=strfind(tline,'Š');
    numOfVarSX=size(varSX,2);
    for k=1:numOfVarSX
        tline(varSX(z))='S';
        tline=strcat(tline(1:varSX(z)),'X',tline(varSX(z)+1:end));
        varSX=strfind(tline,'Š');
    end

    varZX=strfind(tline,'Ž');
    numOfVarZX=size(varZX,2);
    for k=1:numOfVarZX
        tline(varZX(z))='Z';
        tline=strcat(tline(1:varZX(z)),'X',tline(varZX(z)+1:end));
        varZX=strfind(tline,'Ž');
    end
end

```

```
        end
        output_strings=tline;
        %fprintf(fp_out,'%s\n',tline);
        %tline = fgetl(fp_in);
    %end

    %fclose(fp_in);
    %fclose(fp_out);

end
```

12.4.13 sakvacicom.m

```

function [ output_args ] = sakvacicom( pathASCII )
%SAKVACIOM funkcija vraća znakove sa kvačicom
% Funkcija vraća donju listu hrvetski slova i zamjenjuje ih sa
% ekvivalentima.
% Izlaz iz funkcije je *.txt file sa istim nazivom kao i ulazni uz
% dodatak -sa_kvacice.txt
% funkcija rekonstruira i velika i mala slova

% CX=>Ć (ć)
% CY=>Č (č)
% DX=>Đ (đ)
% DZX=>DŽ (dž)
% SX=>Š (š)
% ZX=>Ž (ž)

fp_in = fopen(pathASCII,'r');
pathOut=[pathASCII(1:(strfind(pathASCII,'.')-1)),'-sa_kvacicom.txt'];
fp_out = fopen(pathOut,'w');
z=1;
tline = fgetl(fp_in);
while ischar(tline)
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    varCX=strfind(tline,'CX');
    numOfVarCX=size(varCX,2);
    for k=1:numOfVarCX
        tline(varCX(z))='Ć';
        tline=strcat(tline(1:varCX(z)),tline(varCX(z)+2:end));
        varCX=strfind(tline,'CX');
    end
    varCX=strfind(tline,'cx');
    numOfVarCX=size(varCX,2);
    for k=1:numOfVarCX
        tline(varCX(z))='ć';
        tline=strcat(tline(1:varCX(z)),tline(varCX(z)+2:end));
        varCX=strfind(tline,'cx');
    end
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    varCY=strfind(tline,'CY');
    numOfVarCY=size(varCY,2);
    for k=1:numOfVarCY
        tline(varCY(z))='Č';
        tline=strcat(tline(1:varCY(z)),tline(varCY(z)+2:end));
        varCY=strfind(tline,'CY');
    end
    varCY=strfind(tline,'cy');
    numOfVarCY=size(varCY,2);
    for k=1:numOfVarCY
        tline(varCY(z))='č';
        tline=strcat(tline(1:varCY(z)),tline(varCY(z)+2:end));
        varCY=strfind(tline,'cy');
    end
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    varDX=strfind(tline,'DX');
    numOfVarDX=size(varDX,2);
    for k=1:numOfVarDX
        tline(varDX(z))='Đ';
        tline=strcat(tline(1:varDX(z)),tline(varDX(z)+2:end));
        varDX=strfind(tline,'DX');
    end
    varDX=strfind(tline,'dx');
    numOfVarDX=size(varDX,2);
    for k=1:numOfVarDX
        tline(varDX(z))='đ';
        tline=strcat(tline(1:varDX(z)),tline(varDX(z)+2:end));
        varDX=strfind(tline,'dx');
    end
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    varDZX=strfind(tline,'DZX');
    numOfVarDZX=size(varDZX,2);
    for k=1:numOfVarDZX
        tline(varDZX(z))='DŽ';
        tline=strcat(tline(1:varDZX(z)),tline(varDZX(z)+2:end));
        varDZX=strfind(tline,'DZX');
    end
end

```

```
varDZX=strfind(tline,'dzx');
numOfVarDZX=size(varDX,2);
for k=1:numOfVarDZX
    tline(varDZX(z))='ďž';
    tline=strcat(tline(1:varDZX(z)),tline(varDZX(z)+2:end));
    varDZX=strfind(tline,'dzx');
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
varSX=strfind(tline,'SX');
numOfVarSX=size(varSX,2);
for k=1:numOfVarSX
    tline(varSX(z))='Š';
    tline=strcat(tline(1:varSX(z)),tline(varSX(z)+2:end));
    varSX=strfind(tline,'SX');
end
varSX=strfind(tline,'sx');
numOfVarSX=size(varSX,2);
for k=1:numOfVarSX
    tline(varSX(z))='š';
    tline=strcat(tline(1:varSX(z)),tline(varSX(z)+2:end));
    varSX=strfind(tline,'sx');
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
varZX=strfind(tline,'ZX');
numOfVarZX=size(varZX,2);
for k=1:numOfVarZX
    tline(varZX(z))='Ž';
    tline=strcat(tline(1:varZX(z)),tline(varZX(z)+2:end));
    varZX=strfind(tline,'ZX');
end
varZX=strfind(tline,'zx');
numOfVarZX=size(varZX,2);
for k=1:numOfVarZX
    tline(varZX(z))='ž';
    tline=strcat(tline(1:varZX(z)),tline(varZX(z)+2:end));
    varZX=strfind(tline,'zx');
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
fprintf(fp_out,'%s\n',tline);
tline = fgetl(fp_in);
end
fclose(fp_in);
fclose(fp_out);
end
```

12.4.14 **txt2lower.m**

```
function [ output_args ] = txt2lower(fname_in)
%TXT2LOWER otvara dokument i pretvara sva velika slova u mala
%

    fp_in = fopen(fname_in, 'r');
    temp = fread(fp_in, inf, 'char=>char');
    temp=lower(temp);
    tline = fgetl(fp_in);
    fclose(fp_in);

    fp_out=fopen(fname_in, 'w');
    fwrite(fp_out, temp, 'char');
    fclose(fp_out);

end
```

12.4.15 tokenize.m

```

function [ outStructEmpty ] = tokenize( fname_in )
%TOKENIZE Vrší tokenizaciju dokumenta i tokene pohranjuje u izlaznu
%strukturu
%   fname_in- staza do ulazne datoteke koju je potrebno tokenizirati
%   outStructEmpty- je izlazna struktura u koju su tokeni pohranjeni u
%   sljedećoj strukturi: tokenBROJ: 'token'

%   (C) 2011 by Antonio Kolak <antonio.kolak@gmail.hr>
%   Simple tokenization algorithm
outStructEmpty=struct;
try
    %fname_in='D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\primjeri\ulaz3.txt'

    fp_in = fopen(fname_in,'r');

    % izmjena na \n
    delimiterSet = ';.,!?\\"(){}[]/+--=?{><*'""|' ;
    delimiterSet(end+1)=native2unicode(13);
    digits = '0123456789';
    chars = 'abcdefghijklmnopqrstuvwxyz';
    chars =strcat(chars,upper(chars));

    %   spaces = ' \t\n';
    %   spaces='\x20\xA\xD';
    %   spaces=strcat(strcat(char(20),char(10)),char(13));
    spaces=[native2unicode(32),native2unicode(10),native2unicode(9)];
    numberdelimiters = ',.';

    count_token=1;
    count_token_struct=1;
    temp = fread(fp_in,inf,'char=>char');
    token = '';
    ch=temp(count_token);
    count_token=count_token+1;
    lookahead=temp(count_token);
    while true
        if ~ ch
            if token
                fprintf(fp_out,'%s\n',token);
                disp(token)
            end
            break;
        end
        if strfind(delimiterSet,ch)
            if token
                ptoken=size(token,2);
                if ~isempty(strfind(digits,token(ptoken))) &&
~isempty(strfind(digits,lookahead)) && ~isempty(strfind(numberdelimiters,ch))
                    token=strcat(token,ch);
                elseif ~isempty(strfind(chars,token(ptoken))) &&
~isempty(strfind(chars,lookahead)) && ~isempty(strfind(numberdelimiters,ch))
                    token=strcat(token,ch);
                else
                    fprintf(fp_out,'%s\n',token);
                    disp(token)
                end
            end
            outStructEmpty=setfield(outStructEmpty,['token',num2str(count_token_struct)],token);
            count_token_struct=count_token_struct+1;
            token='';
            if ~(ch==native2unicode(32) || ch==native2unicode(10) ||
ch==native2unicode(9))
                fprintf(fp_out,'%c\n',ch);
                disp(ch)
            end
            outStructEmpty=setfield(outStructEmpty,['token',num2str(count_token_struct)],ch);
            count_token_struct=count_token_struct+1;
        end
    end
    elseif (ch==native2unicode(32) || ch==native2unicode(10) ||
ch==native2unicode(9)) % must use this notation instead of variable spaces
        if token
            fprintf(fp_out,'%s\n',token);
            disp(token)
        end
    end
end

```



```
outStructEmpty=setfield(outStructEmpty,['token',num2str(count_token_struct)],token);
    count_token_struct=count_token_struct+1;
    token='';
end
else
    token=strcat(token,ch);
end
ch=lookahead;
count_token=count_token+1;
lookahead=temp(count_token);

end
fclose(fp_in);
%fclose(fp_out);

catch exception
    if (strcmp(exception.identifier,'MATLAB:inputArgUndefined'))
        error('MATLAB:myFunction:Dimensionality','Progrešna staza do ulazne datoteke');
    else if (strcmp(exception.identifier,'MATLAB:badsubscript'))
        % pokazivač je zašao dalje nego ima teksta
        token=strcat(token,ch);
        %fprintf(fp_out,'%s\n',token);
%        disp(token)

outStructEmpty=setfield(outStructEmpty,['token',num2str(count_token_struct)],token);
    count_token_struct=count_token_struct+1;
    fclose(fp_in);
    %fclose(fp_out);

    else
        rethrow(exception);
    end
end
end
```

12.4.16 tokenize2.m

```

function [ outStructEmpty ] = tokenize2( recenica )
%TOKENIZE2 Vrší tokenizaciju iz stringa i tokene pohranjuje u izlaznu
%strukturu- inače strukturno identična funkciji tokenize
% outStructEmpty- je izlazna struktura u koju su tokeni pohranjeni u
% sljedećoj strukturi: tokenBROJ: 'token'
% (C) 2011 by Antonio Kolak <antonio.kolak@gmail.hr>
% Simple tokenization algorithm
outStructEmpty=struct;
try
    %fname_in='D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\primjeri\ulaz3.txt'

    %fp_in = fopen(fname_in,'r');

    % izmjena na \n
    delimiterSet = '.,!?"()':[]/+--=?{}><*"\'";
    delimiterSet(end+1)=native2unicode(13);
    digits = '0123456789';
    chars = 'abcdefghijklmnopqrstuvwxyz';
    chars =strcat(chars,upper(chars));
%
    spaces = ' \t\n';
%
    spaces='\x20\xA\xD';
%
    spaces=strcat(strcat(char(20),char(10)),char(13));
spaces=[native2unicode(32),native2unicode(10),native2unicode(9)];
numberdelimiters = ',.:';

count_token=1;
count_token_struct=1;
%temp = fread(fp_in,inf,'char=>char');
temp=recenica;
token = '';
ch=temp(count_token);
count_token=count_token+1;
lookahead=temp(count_token);
while true
    if ~ ch
        if token
            fprintf(fp_out,'%s\n',token);
%
            disp(token)
        end
        break;
    end
    if strfind(delimiterSet,ch)
        if token
            ptoken=size(token,2);
            if ~isempty(strfind(digits,token(ptoken))) &&
~isempty(strfind(digits,lookahead)) && ~isempty(strfind(numberdelimiters,ch))
                token=strcat(token,ch);
            elseif ~isempty(strfind(chars,token(ptoken))) &&
~isempty(strfind(chars,lookahead)) && ~isempty(strfind(numberdelimiters,ch))
                token=strcat(token,ch);
            else
                %fprintf(fp_out,'%s\n',token);
%
                disp(token)
            end
        end
        outStructEmpty=setfield(outStructEmpty,['token',num2str(count_token_struct)],token);
        count_token_struct=count_token_struct+1;
        token='';
        if ~(ch==native2unicode(32) || ch==native2unicode(10) ||
ch==native2unicode(9))
            %fprintf(fp_out,'%c\n',ch);
%
            disp(ch)
        end
        outStructEmpty=setfield(outStructEmpty,['token',num2str(count_token_struct)],ch);
        count_token_struct=count_token_struct+1;
    end
end
elseif (ch==native2unicode(32) || ch==native2unicode(10) ||
ch==native2unicode(9)) % must use this notation instead of variable spaces
    if token
        %fprintf(fp_out,'%s\n',token);
%
        disp(token)
    end
end

```

```
outStructEmpty=setfield(outStructEmpty,['token',num2str(count_token_struct)],token);
    count_token_struct=count_token_struct+1;
    token='';
end
else
    token=strcat(token,ch);
end
ch=lookahead;
count_token=count_token+1;
lookahead=temp(count_token);

end
%fclose(fp_in);
%fclose(fp_out);

catch exception
    if (strcmp(exception.identifier,'MATLAB:inputArgUndefined'))
        error('MATLAB:myFunction:Dimensionality','Progrešna staza do ulazne datoteke');
    else if (strcmp(exception.identifier,'MATLAB:badsubscript'))
        % pokazivač je zašao dalje nego ima teksta
        token=strcat(token,ch);
        %fprintf(fp_out,'%s\n',token);
        % disp(token)
    end

outStructEmpty=setfield(outStructEmpty,['token',num2str(count_token_struct)],token);
    count_token_struct=count_token_struct+1;
    % fclose(fp_in);
    % fclose(fp_out);
end
    rethrow(exception);
end
end
```

12.4.17 fpStruct.m

```

function [ outStruct ] = fpStruct( inStruct )
%FPSTRUCT Izračunava frekvenciju pojavljivanja tokena iz ulazne strukture
%   inStruct-ulazna struktura sa listom tokena
%   outStruct-izlazna struktura sa frekvencijom tokena
outStruct=struct;

try

    % izmjena na \n
    delimiterSet = ';.,!?\()"':[]/+--=??{}><*'""\|';
    delimiterSet(end+1)=native2unicode(13);
    delimiterSet(end+1)=native2unicode(32);
    delimiterSet(end+1)=native2unicode(10);
    delimiterSet(end+1)=native2unicode(9);

    numOfTokens=size(fieldnames(inStruct),1);
    counter=1;
    for k=1:numOfTokens
        nameOfToken=char(getfield(inStruct,['token',num2str(k)]));

        if isempty(strfind(delimiterSet,nameOfToken)) && ~isempty(nameOfToken)
            if isfield(outStruct,nameOfToken)
                counter=getfield(outStruct,nameOfToken)+1;
                outStruct=setfield(outStruct,nameOfToken,counter);
                counter=1;
            else
                outStruct=setfield(outStruct,nameOfToken,counter);
            end

            % disp(tline)
        end
    end
catch exception
    if (strcmp(exception.identifier,'MATLAB:inputArgUndefined'))
        error('MATLAB:myFunction:Dimensionality','Progrešna staza do ulazne datoteke');
    else
        rethrow(exception);
    end
end
end

```

12.4.18 rfpStruct.m

```
function [ outStruct ] = rpfStruct( inStruct )
%RPFSTRUCT računa relativnu frekvenciju pojavljivanja iz početne strukture
% inStruct-ulazna struktura koja ima izračunatu frekvenciju pojavljivanja
% tokena
% outStruct-izlazna struktura sa izračunatom relativnom frekvencijom
% pojavljivanja
outStruct=struct;

% Relative freq.
values_structN = cell2mat(struct2cell(inStruct));
total=sum(values_structN);
names_of_structN=fieldnames(inStruct);
numOfTokoens=size(names_of_structN,1);
for g=1:numOfTokoens
    nameOfOne=names_of_structN(g);
    valueOfOne=getfield(inStruct,char(names_of_structN(g)));
    valueOfOne=valueOfOne/total;
    outStruct=setfield(outStruct,char(names_of_structN(g)),valueOfOne);
end
end
```

12.4.19 `upis_u_unknown.m`

```
function [ output_args ] = upis_u_unknown( fname_in, sentenceCell )
%UPIŠ_U_UNKNOWN upisuje zadanu rečenicu u datoteku definiranu stazom
%fname_in

    fp_out=fopen(fname_in,'w');
    fprintf(fp_out,'%s',char(sentenceCell));
    fclose(fp_out);

end
```

12.4.20 KEG_randperm.m

```

function [ all_errors ] = KEG_randperm( metaKEGPath,metaNewKEGPath,unknownPath,
transkripti_po_emocijama_Path )
%KEG_RANDPERM funkcija koja provodi kros-validaciju nad primitivnim unigram sustavom

% metaKEGPath-staza do ulazne *xlsx datoteke sa zapisom transkripcije
% metaNewKEGPath-staza do nove datoteke u koju se upisuje samo 90 %
% korpusa kako bismo dobili podatke za trening, te se tih 90% opet
% srotira u 5 klasa i vrši se testiranje sa preostalih 10 %; taj proces
% se ponavlja 10 puta pošto kros-validacija zahtjeva da se testira
% cijeli korpus tj. 10 x 10% trening podataka.
% unknownPath-staza do datoteke u kojoj se nalazi rečenica koja se upravo
% testira, u ovu datoteku se upisuje svaki put nova rečenica
% transkripti_po_emocijama_Path-staza do datoteka klasa (sreca.txt,
% tuga.txt, strah.txt, ljutnja.txt, neutralno.txt)
% Na kraju je izračunata matrica konfuzije u kojoj se vide korelacije
% klasifikacije među pojedinim emocijama

[num, txt] = xlsread(metaKEGPath,'Sheet1');

orderNum=num(:,1);
sizeOf_orderNum=size(orderNum,1);
sentences=txt(2:sizeOf_orderNum+1,7);
emotions=txt(2:sizeOf_orderNum+1,55);

TranskriptCell={};
TranskriptCell(:,1)=num2cell(orderNum);
TranskriptCell(:,2)=num2cell(randperm(sizeOf_orderNum)); %generira random niz po kojem
se sortiraju rečenice
TranskriptCell(:,3)=emotions;
TranskriptCell(:,4)=sentences;

TranskriptCellPerm=sortrows( TranskriptCell,2);

sizeOfTestCorpus=floor(sizeOf_orderNum*0.1);
sizeOfTreningCorpus=sizeOf_orderNum-sizeOfTestCorpus;

% 10 permutacija kako bismo napravili kros-validaciju na cijelom
% korpusu

TestVectorStart=1;
TestVectorEnd=sizeOfTestCorpus;
all_errors=[];

all_srecasreca=[];
all_srecatuga=[];
all_srecastrah=[];
all_srecaljutnja=[];
all_srecaneutralno=[];
all_srecanepoznato=[];

all_tugasreca=[];
all_tugatuga=[];
all_tugastrah=[];
all_tugaljutnja=[];
all_tuganeutralno=[];
all_tuganepoznato=[];

all_strahsreca=[];
all_strahtuga=[];
all_strahstrah=[];
all_strahljutnja=[];
all_strahneutralno=[];
all_strahnepoznato=[];

all_ljutnjasreca=[];
all_ljutnjatuga=[];
all_ljutnjastrah=[];
all_ljutnjajlutnja=[];
all_ljutnjaneutralno=[];
all_ljutnjanepoznato=[];

all_neutralnosreca=[];
all_neutralnotuga=[];

```

```

all_neutralnostrah=[];
all_neutralnoljutnja=[];
all_neutralnoneutralno=[];
all_neutralnonepoznato=[];

all_srecaCount=[];
all_tugaCount=[];
all_strahCount=[];
all_ljutnjaCount=[];
all_neutralnoCount=[];

for k=1:10
    TestVector=[TestVectorStart:TestVectorEnd];
    TestVectorStart=TestVectorEnd+1;
    TestVectorEnd=TestVectorEnd+sizeOfTestCorpus;

    TreningVector=[1:sizeOf_orderNum];
    TreningVector=setdiff(TreningVector,TestVector);

    M={};
    M(1,1)={'ID snimke'};
    M(1,7)={'transkripcija'};
    M(1,55)={'r - Emocija'};

    for h=1:size(TreningVector,2)
        numberForTrening=TreningVector(h);
        M(h+1,1)=TranskriptCellPerm(numberForTrening,2); %broj
        M(h+1,7)=TranskriptCellPerm(numberForTrening,4); %transkrip
        M(h+1,55)=TranskriptCellPerm(numberForTrening,3); %emocija
    end
    xlsxwrite(metaNewKEGPath,M)

    % Kreiranje pomoćne strukture, tokeni, skidanje kvačica,
    % definiranje unique klasa
    KEGto5class(metaNewKEGPath,transkripti_po_emocijama_Path,0);
    bezkvacice([transkripti_po_emocijama_Path,'\sreca.txt']);
    bezkvacice([transkripti_po_emocijama_Path,'\tuga.txt']);
    bezkvacice([transkripti_po_emocijama_Path,'\strah.txt']);
    bezkvacice([transkripti_po_emocijama_Path,'\ljutnja.txt']);
    bezkvacice([transkripti_po_emocijama_Path,'\neutralno.txt']);

    tokenizer([transkripti_po_emocijama_Path,'\sreca-bez_kvacice.txt'],...
        [transkripti_po_emocijama_Path,'\sreca-bez_kvacice.lst']);
    tokenizer([transkripti_po_emocijama_Path,'\tuga-bez_kvacice.txt'],...
        [transkripti_po_emocijama_Path,'\tuga-bez_kvacice.lst']);
    tokenizer([transkripti_po_emocijama_Path,'\strah-bez_kvacice.txt'],...
        [transkripti_po_emocijama_Path,'\strah-bez_kvacice.lst']);
    tokenizer([transkripti_po_emocijama_Path,'\ljutnja-bez_kvacice.txt'],...
        [transkripti_po_emocijama_Path,'\ljutnja-bez_kvacice.lst']);
    tokenizer([transkripti_po_emocijama_Path,'\neutralno-bez_kvacice.txt'],...
        [transkripti_po_emocijama_Path,'\neutralno-bez_kvacice.lst']);

    rfp_unique([transkripti_po_emocijama_Path,'\sreca-bez_kvacice.lst'],...
        [transkripti_po_emocijama_Path,'\tuga-bez_kvacice.lst'],...
        [transkripti_po_emocijama_Path,'\strah-bez_kvacice.lst'],...
        [transkripti_po_emocijama_Path,'\ljutnja-bez_kvacice.lst'],...
        [transkripti_po_emocijama_Path,'\neutralno-bez_kvacice.lst']);

    % testiranje svake pojedine rečenice iz test korpusa
    positiveCounter=0;
    in_sreca_out_sreca=0;
    in_sreca_out_tuga=0;
    in_sreca_out_strah=0;
    in_sreca_out_ljutnja=0;
    in_sreca_out_neutralno=0;
    in_sreca_out_nepoznato=0;

    in_tuga_out_sreca=0;
    in_tuga_out_tuga=0;
    in_tuga_out_strah=0;
    in_tuga_out_ljutnja=0;
    in_tuga_out_neutralno=0;
    in_tuga_out_nepoznato=0;

```



```

in_strah_out_sreca=0;
in_strah_out_tuga=0;
in_strah_out_strah=0;
in_strah_out_ljutnja=0;
in_strah_out_neutralno=0;
in_strah_out_nepoznato=0;

in_ljutnja_out_sreca=0;
in_ljutnja_out_tuga=0;
in_ljutnja_out_strah=0;
in_ljutnja_out_ljutnja=0;
in_ljutnja_out_neutralno=0;
in_ljutnja_out_nepoznato=0;

in_neutralno_out_sreca=0;
in_neutralno_out_tuga=0;
in_neutralno_out_strah=0;
in_neutralno_out_ljutnja=0;
in_neutralno_out_neutralno=0;
in_neutralno_out_nepoznato=0;

srecaCount=0;
tugaCount=0;
strahCount=0;
ljutnjaCount=0;
neutralnoCount=0;

for g=1:sizeOfTestCorpus
    %upis_u_unknown(unknownPath,TranskriptCellPerm(TestVector(g),4));
    UnknownString=char(TranskriptCellPerm(TestVector(g),4));
    UnknownString=bezkvacice2(UnknownString);
    %unknownPathBezKvacice=[unknownPath(1:(strfind(unknownPath,'.')-1)),'-
bez_kvacice.txt'];

[errorsForOneSentence,errPaths]=compareWithModels2(transkripti_po_emocijama_Path,UnknownString,0);

errPathsCells=struct2cell(errPaths);
[C,I]=min(errorsForOneSentence);
charEmocija=char(TranskriptCellPerm(TestVector(g),3));

if strcmp(charEmocija,'SR')
    srecaCount=srecaCount+1;
elseif strcmp(charEmocija,'TU')
    tugaCount=tugaCount+1;
elseif strcmp(charEmocija,'ST')
    strahCount=strahCount+1;
elseif strcmp(charEmocija,'LJ')
    ljutnjaCount=ljutnjaCount+1;
elseif strcmp(charEmocija,'NE')
    neutralnoCount=neutralnoCount+1;
end

if std(errorsForOneSentence)~=0
    if ~isempty(strfind(char(errPathsCells(I)),'sreca'))&&
strcmp(charEmocija,'SR')
        in_sreca_out_sreca=in_sreca_out_sreca+1;
    elseif ~isempty(strfind(char(errPathsCells(I)),'tuga'))&&
strcmp(charEmocija,'SR')
        in_sreca_out_tuga= in_sreca_out_tuga+1;
    elseif ~isempty(strfind(char(errPathsCells(I)),'strah'))&&
strcmp(charEmocija,'SR')
        in_sreca_out_strah=in_sreca_out_strah+1;
    elseif ~isempty(strfind(char(errPathsCells(I)),'ljutnja'))&&
strcmp(charEmocija,'SR')
        in_sreca_out_ljutnja=in_sreca_out_ljutnja+1;
    elseif ~isempty(strfind(char(errPathsCells(I)),'neutralno'))&&
strcmp(charEmocija,'SR')
        in_sreca_out_neutralno=in_sreca_out_neutralno+1;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    elseif ~isempty(strfind(char(errPathsCells(I)),'sreca'))&&
strcmp(charEmocija,'TU')
        in_tuga_out_sreca=in_tuga_out_sreca+1;
    elseif ~isempty(strfind(char(errPathsCells(I)),'tuga'))&&
strcmp(charEmocija,'TU')
        in_tuga_out_tuga=in_tuga_out_tuga+1;

```

```

        elseif ~isempty(strfind(char(errPathsCells(I)), 'strah')) &&
strcmp(charEmocija, 'TU')
            in_tuga_out_strah=in_tuga_out_strah+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'ljutnja')) &&
strcmp(charEmocija, 'TU')
            in_tuga_out_ljutnja=in_tuga_out_ljutnja+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'neutralno')) &&
strcmp(charEmocija, 'TU')
            in_tuga_out_neutralno=in_tuga_out_neutralno+1;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        elseif ~isempty(strfind(char(errPathsCells(I)), 'sreca')) &&
strcmp(charEmocija, 'ST')
            in_strah_out_sreca=in_strah_out_sreca+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'tuga')) &&
strcmp(charEmocija, 'ST')
            in_strah_out_tuga=in_strah_out_tuga+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'strah')) &&
strcmp(charEmocija, 'ST')
            in_strah_out_strah=in_strah_out_strah+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'ljutnja')) &&
strcmp(charEmocija, 'ST')
            in_strah_out_ljutnja=in_strah_out_ljutnja+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'neutralno')) &&
strcmp(charEmocija, 'ST')
            in_strah_out_neutralno=in_strah_out_neutralno+1;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        elseif ~isempty(strfind(char(errPathsCells(I)), 'sreca')) &&
strcmp(charEmocija, 'LJ')
            in_ljutnja_out_sreca=in_ljutnja_out_sreca+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'tuga')) &&
strcmp(charEmocija, 'LJ')
            in_ljutnja_out_tuga=in_ljutnja_out_tuga+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'strah')) &&
strcmp(charEmocija, 'LJ')
            in_ljutnja_out_strah=in_ljutnja_out_strah+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'ljutnja')) &&
strcmp(charEmocija, 'LJ')
            in_ljutnja_out_ljutnja=in_ljutnja_out_ljutnja+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'neutralno')) &&
strcmp(charEmocija, 'LJ')
            in_ljutnja_out_neutralno=in_ljutnja_out_neutralno+1;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        elseif ~isempty(strfind(char(errPathsCells(I)), 'sreca')) &&
strcmp(charEmocija, 'NE')
            in_neutralno_out_sreca=in_neutralno_out_sreca+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'tuga')) &&
strcmp(charEmocija, 'NE')
            in_neutralno_out_tuga=in_neutralno_out_tuga+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'strah')) &&
strcmp(charEmocija, 'NE')
            in_neutralno_out_strah=in_neutralno_out_strah+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'ljutnja')) &&
strcmp(charEmocija, 'NE')
            in_neutralno_out_ljutnja=in_neutralno_out_ljutnja+1;
        elseif ~isempty(strfind(char(errPathsCells(I)), 'neutralno')) &&
strcmp(charEmocija, 'NE')
            in_neutralno_out_neutralno=in_neutralno_out_neutralno+1;
    end
    elseif strcmp(charEmocija, 'SR')
        in_sreca_out_nepoznato=in_sreca_out_nepoznato+1;
    elseif strcmp(charEmocija, 'TU')
        in_tuga_out_nepoznato=in_tuga_out_nepoznato+1;
    elseif strcmp(charEmocija, 'ST')
        in_strah_out_nepoznato=in_strah_out_nepoznato+1;
    elseif strcmp(charEmocija, 'LJ')
        in_ljutnja_out_nepoznato=in_ljutnja_out_nepoznato+1;
    elseif strcmp(charEmocija, 'NE')
        in_neutralno_out_nepoznato=in_neutralno_out_nepoznato+1;
    end
end
all_errors(k)=positiveCounter/sizeOfTestCorpus;
all_srecasreca(k)=in_sreca_out_sreca;
all_srecatuga(k)=in_sreca_out_tuga;
all_srecastrah(k)=in_sreca_out_strah;

```

```

all_srecaljutnja(k)=in_sreca_out_ljutnja;
all_srecaneutralno(k)=in_sreca_out_neutralno;
all_srecanepoznato(k)=in_sreca_out_nepoznato;

all_tugasreca(k)=in_tuga_out_sreca;
all_tugatuga(k)=in_tuga_out_tuga;
all_tugastrah(k)=in_tuga_out_strah;
all_tugaljutnja(k)=in_tuga_out_ljutnja;
all_tuganeutralno(k)=in_tuga_out_neutralno;
all_tuganepoznato(k)=in_tuga_out_nepoznato;

all_strahsreca(k)=in_strah_out_sreca;
all_strahtuga(k)=in_strah_out_tuga;
all_strahstrah(k)=in_strah_out_strah;
all_strahljutnja(k)=in_strah_out_ljutnja;
all_strahneutralno(k)=in_strah_out_neutralno;
all_strahnepoznato(k)=in_strah_out_nepoznato;

all_ljutnjasreca(k)=in_ljutnja_out_sreca;
all_ljutnjatuga(k)=in_ljutnja_out_tuga;
all_ljutnjastrah(k)=in_ljutnja_out_strah;
all_ljutnjajlutnja(k)=in_ljutnja_out_ljutnja;
all_ljutnjaneutralno(k)=in_ljutnja_out_neutralno;
all_ljutjanepoznato(k)=in_ljutnja_out_nepoznato;

all_neutralnosreca(k)=in_neutralno_out_sreca;
all_neutralnotuga(k)=in_neutralno_out_tuga;
all_neutralnostrah(k)=in_neutralno_out_strah;
all_neutralnoljutnja(k)=in_neutralno_out_ljutnja;
all_neutralnoneutralno(k)=in_neutralno_out_neutralno;
all_neutralnonepoznato(k)=in_neutralno_out_nepoznato;

all_srecaCount(k)=srecaCount;
all_tugaCount(k)=tugaCount;
all_strahCount(k)=strahCount;
all_ljutnjaCount(k)=ljutnjaCount;
all_neutralnoCount(k)=neutralnoCount;
TestVector=[];
k
end
disp(sprintf('SUMA: sreca-sreca %d',sum(all_srecasreca)));
disp(sprintf('SUMA: sreca-tuga %d',sum(all_srecatuga)));
disp(sprintf('SUMA: sreca-strah %d',sum(all_srecastrah)));
disp(sprintf('SUMA: sreca-ljutnja %d',sum(all_srecaljutnja)));
disp(sprintf('SUMA: sreca-neutralno %d',sum(all_srecaneutralno)));
disp(sprintf('SUMA: sreca-nepoznato %d',sum(all_srecanepoznato)));
disp(sprintf('\n'));

disp(sprintf('SUMA: tuga-sreca %d',sum(all_tugasreca)));
disp(sprintf('SUMA: tuga-tuga %d',sum(all_tugatuga)));
disp(sprintf('SUMA: tuga-strah %d',sum(all_tugastrah)));
disp(sprintf('SUMA: tuga-ljutnja %d',sum(all_tugaljutnja)));
disp(sprintf('SUMA: tuga-neutralno %d',sum(all_tuganeutralno)));
disp(sprintf('SUMA: tuga-nepoznato %d',sum(all_tuganepoznato)));
disp(sprintf('\n'));

disp(sprintf('SUMA: strah-sreca %d',sum(all_strahsreca)));
disp(sprintf('SUMA: strah-tuga %d',sum(all_strahtuga)));
disp(sprintf('SUMA: strah-strah %d',sum(all_strahstrah)));
disp(sprintf('SUMA: strah-ljutnja %d',sum(all_strahljutnja)));
disp(sprintf('SUMA: strah-neutralno %d',sum(all_strahneutralno)));
disp(sprintf('SUMA: strah-nepoznato %d',sum(all_strahnepoznato)));
disp(sprintf('\n'));

disp(sprintf('SUMA: ljutnja-sreca %d',sum(all_ljutnjasreca)));
disp(sprintf('SUMA: ljutnja-tuga %d',sum(all_ljutnjatuga)));
disp(sprintf('SUMA: ljutnja-strah %d',sum(all_ljutnjastrah)));
disp(sprintf('SUMA: ljutnja-ljutnja %d',sum(all_ljutnjajlutnja)));
disp(sprintf('SUMA: ljutnja-neutralno %d',sum(all_ljutnjaneutralno)));
disp(sprintf('SUMA: ljutnja-nepoznato %d',sum(all_ljutjanepoznato)));
disp(sprintf('\n'));

disp(sprintf('SUMA: neutralno-sreca %d',sum(all_neutralnosreca)));
disp(sprintf('SUMA: neutralno-tuga %d',sum(all_neutralnotuga)));
disp(sprintf('SUMA: neutralno-strah %d',sum(all_neutralnostrah)));
disp(sprintf('SUMA: neutralno-ljutnja %d',sum(all_neutralnoljutnja)));
disp(sprintf('SUMA: neutralno-neutralno %d',sum(all_neutralnoneutralno)));

```

```

disp(sprintf('SUMA: neutralno-nepoznato %d',sum(all_neutralnonepoznato)));
disp(sprintf('\n'));

disp(sprintf('SUMA: CountSreca %d',sum(all_srecaCount)));
disp(sprintf('SUMA: CountTuga %d',sum(all_tugaCount)));
disp(sprintf('SUMA: CountStrah %d',sum(all_strahCount)));
disp(sprintf('SUMA: CountLjutnja %d',sum(all_ljutnjaCount)));
disp(sprintf('SUMA: CountNeutralno %d',sum(all_neutralnoCount)));
disp(sprintf('\n\n'));

disp('CONFUSION MATRIX');
confusionM(1,1)=sum(all_srecasreca);
confusionM(2,1)=sum(all_srecatuga);
confusionM(3,1)=sum(all_srecastrah);
confusionM(4,1)=sum(all_srecaljutnja);
confusionM(5,1)=sum(all_srecaneutralno);
confusionM(6,1)=sum(all_srecanepoznato);

confusionM(1,2)=sum(all_tugasreca);
confusionM(2,2)=sum(all_tugatuga);
confusionM(3,2)=sum(all_tugastrah);
confusionM(4,2)=sum(all_tugaljutnja);
confusionM(5,2)=sum(all_tuganeutralno);
confusionM(6,2)=sum(all_tuganepoznato);

confusionM(1,3)=sum(all_strahsreca);
confusionM(2,3)=sum(all_strahtuga);
confusionM(3,3)=sum(all_strahstrah);
confusionM(4,3)=sum(all_strahljutnja);
confusionM(5,3)=sum(all_strahneutralno);
confusionM(6,3)=sum(all_strahnepoznato);

confusionM(1,4)=sum(all_ljutnjasreca);
confusionM(2,4)=sum(all_ljutnjatuga);
confusionM(3,4)=sum(all_ljutnjastrah);
confusionM(4,4)=sum(all_ljutnjajlutnja);
confusionM(5,4)=sum(all_ljutnjaneutralno);
confusionM(6,4)=sum(all_ljutnjanepoznato);

confusionM(1,5)=sum(all_neutralnosreca);
confusionM(2,5)=sum(all_neutralnotuga);
confusionM(3,5)=sum(all_neutralnostrah);
confusionM(4,5)=sum(all_neutralnoljutnja);
confusionM(5,5)=sum(all_neutralnoneutralno);
confusionM(6,5)=sum(all_neutralnonepoznato);

confusionM(7,1)=sum(all_srecaCount);
confusionM(7,2)=sum(all_tugaCount);
confusionM(7,3)=sum(all_strahCount);
confusionM(7,4)=sum(all_ljutnjaCount);
confusionM(7,5)=sum(all_neutralnoCount);

confusionM

disp(sprintf('\n\n'));
disp('CONFUSION MATRIX-Percentage');

confusionMposto(:,1)=confusionM(:,1)/confusionM(7,1);
confusionMposto(:,2)=confusionM(:,2)/confusionM(7,2);
confusionMposto(:,3)=confusionM(:,3)/confusionM(7,3);
confusionMposto(:,4)=confusionM(:,4)/confusionM(7,4);
confusionMposto(:,5)=confusionM(:,5)/confusionM(7,5);
confusionMposto

```

end

12.4.21 KEGto2class_G_S.m

```
function [ glumljeneEmocije,stvarneEmocije ] = KEGto2class_G_S( metaKEGPath,classPath )
%KEGT02CLASS_G_S Razdvaja cijelokupni transkript na dva dijela
% Razdvaja cjelokupni transkript na dva dijela, na glumljene emocije i na
% stvarne emocije kako bi se mogao poslije pronaći TTR
%
% metaKEGPath-putanja do *.xls datoteke
% classPath-putanja do direktorija u kojeg želimo pohraniti nove dije
% klase glumljene.txt i stvarne.txt
% izlaz su cell-ovi rečenica glumljenihEmocija i stvarnihEmocija
% kao i *.txt datoteke naziva glumljeneEmocije.txt i stvarneEmocije.txt

[num, txt] = xlsread(metaKEGPath, 'Sheet1');
orderNum=num(:,1);
sizeof_orderNum=size(orderNum,1);
sentences=txt(2:sizeof_orderNum+1,7);
emotions=txt(2:sizeof_orderNum+1,55);
vrsta_emocije=txt(2:sizeof_orderNum+1,6);

glumljeneEmocijePath=[classPath, 'glumljeneEmocije.txt'];
stvarneEmocijePath=[classPath, 'stvarneEmocije.txt'];

fp_out_glumljene=fopen(glumljeneEmocijePath, 'w');
fp_out_stvarne=fopen(stvarneEmocijePath, 'w');

counter_glumljene=1;
counter_stvarne=1;

for k=1:sizeof_orderNum
    if strcmp(char(emotions(k)), 'SR') || ...
        strcmp(char(emotions(k)), 'TU') || ...
        strcmp(char(emotions(k)), 'ST') || ...
        strcmp(char(emotions(k)), 'LJ') || ...
        strcmp(char(emotions(k)), 'NE')
        if strcmp(char(vrsta_emocije(k)), 'G')
            glumljeneEmocije(counter_glumljene)=sentences(k);
            fprintf(fp_out_glumljene, '%s\n', char(sentences(k)));
            counter_glumljene=counter_glumljene+1;
        elseif strcmp(char(vrsta_emocije(k)), 'S')
            stvarneEmocije(counter_stvarne)=sentences(k);
            fprintf(fp_out_stvarne, '%s\n', char(sentences(k)));
            counter_stvarne=counter_stvarne+1;
        end
    end
end

glumljeneEmocije=glumljeneEmocije';
stvarneEmocije=stvarneEmocije';

fclose(fp_out_glumljene);
fclose(fp_out_stvarne);

end
```

12.4.22 KEGto_all_TXT.m

```
function [ output_args ] = KEGto_all_TXT( metaKEGPath,TXTPath )
%KEGTO_ALL_TXT Pretvara zadani transkrip iz xls tablice u niz txt datoteka
%
% metaKEGPath - Staza do metaKEG.xlsx datoteke u kojoj se nalazi
% transkript snimka koji želimo prebaciti u pojedinačne račenica zapisane
% u txt datoteke
% VAŽNO: unutar metaKEGPath- se moraju nalaziti samo anotirane snimke za
% koje postoji ocjena emocionalnog stanja (SR,TU,ST,LJ,NE)
%
% classPath - Staza do direktorija u kojem želimo da nam budu tekstualne
% datoteke (transkript svake snimke se pretvara u uzasebnu txt datoteku)
%

[num, txt] = xlsread(metaKEGPath, 'Sheet1');

orderNum=num(:,1);
sizeOf_orderNum=size(orderNum,1);
sentences=txt(2:sizeOf_orderNum+1,7);
emotions=txt(2:sizeOf_orderNum+1,55);
cout_sreca=1;
cout_tuga=1;
cout_strah=1;
cout_ljutnja=1;
cout_neutralno=1;

for k=1:(sizeOf_orderNum-1)
    if strcmp(char(emotions(k)), 'SR')
        TXTPath_with_number=[TXTPath, 'sreca', num2str(cout_sreca), '.txt'];
        cout_sreca=cout_sreca+1;
    elseif strcmp(char(emotions(k)), 'TU')
        TXTPath_with_number=[TXTPath, 'tuga', num2str(cout_tuga), '.txt'];
        cout_tuga=cout_tuga+1;
    elseif strcmp(char(emotions(k)), 'ST')
        TXTPath_with_number=[TXTPath, 'strah', num2str(cout_strah), '.txt'];
        cout_strah=cout_strah+1;
    elseif strcmp(char(emotions(k)), 'LJ')
        TXTPath_with_number=[TXTPath, 'ljutnja', num2str(cout_ljutnja), '.txt'];
        cout_ljutnja=cout_ljutnja+1;
    elseif strcmp(char(emotions(k)), 'NE')
        TXTPath_with_number=[TXTPath, 'neutralno', num2str(cout_neutralno), '.txt'];
        cout_neutralno=cout_neutralno+1;
    else
        continue
    end
    fp_out=fopen(TXTPath_with_number, 'w');
    fprintf(fp_out, '%s\n', bezkvacice2(char(sentences(k))));
    fclose(fp_out);
end

end
```

12.4.23 TXTtoModels.m

```

function [ output_args ] = TXTtoModels( TXTsPath, relName_polje, imeModela_polje )
%TXTTOMODELS uzima niz txt datoteka i od njih tvori jedinstveni model za tu
%klasu ekstenzije *.dat
%   TXTsPath sadrži stazu do direktorija u kojem se nalaze txt datoteke za
%   pojedinu emociju (1 datoteka -1 snimka(transkript)
%   Potrebno je navesti stazu i relativno ime za sve datoteke oblika *.txt
%   odnosno npr. (sreca*.txt,tuga*.txt, strah*.txt, ljutnja*.txt,
%   neutralno*.txt)
%   relName=relativno ime txt datoteka
%   imeModela=ime koje želimo za izlaznu datoteku

    outStruct=struct;
    outPaths=struct;
    tokenStruct=struct;
    brojModela=size(relName_polje,2);
    for g=1:brojModela
        currentFolder = pwd;
        cd(TXTsPath)
        dir_list=dir([TXTsPath,char(relName_polje(g))]);
        cd(currentFolder)
        sizeOfTXTList=size(dir_list,1);
        sizeOfTXTList_polje(g)=sizeOfTXTList;

        PathForModel=[TXTsPath,'MODEL',char(imeModela_polje(g)),'.txt'];
        fp_out = fopen( PathForModel,'w');

        for k=1: sizeOfTXTList
            imena(k)={dir_list(k).name};
            %fp_in=fopen([TXTsPath,dir_list(k).name],'r');
            fid = fopen([TXTsPath,dir_list(k).name]);
            temp = fgetl(fid);
            %temp = fread(fp_in,inf,'char=>char');
            %temp=temp';
            temp=bezkvacice2(temp);
            fprintf(fp_out,'%s\n',temp);
            %fclose(fp_in);
            fclose(fid);
        end
        fclose(fp_out);
        %tokenizacija

        tokenizer([TXTsPath,'MODEL',char(imeModela_polje(g)),'.txt'],[TXTsPath,'MODEL',char(imeModela_polje(g)),'.lst']);
        end
        %bezkvacice([TXTsPath,imeModela,'MODEL.txt'])
        % kreiranje stringa koji će sadržavati putanje do svih datoteka za funkciju unique
        combinedStr={};
        for g1=1:brojModela
            %combinedStr =
            strcat(combinedStr,','', [TXTsPath,'MODEL',char(imeModela_polje(g1)),'.lst'],'','');
            combinedStr(g1)=mat2cell([TXTsPath,'MODEL',char(imeModela_polje(g1)),'.lst']);
        end
        %combinedStr=combinedStr(1:end-1);

        %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% HARDCODE 5 emocija
        dictStructFieldNew =
        unique1(char(combinedStr(1)),char(combinedStr(2)),char(combinedStr(3)),char(combinedStr(4)),char(combinedStr(5)))

        % prikaz u obliku tokeniziranih modela (napravljena je struktura sa
        % dekeniziranim modelima)
        tokenize_polje=struct;
        sva_tokenize_polja=struct;
        numOfTokensModels=size(fieldnames(dictStructFieldNew),1);
        for g2=1:numOfTokensModels
            cell_tokena=getfield(dictStructFieldNew,['tokens',num2str(g2)]);
            numOfTokensInCell=size(cell_tokena,1);
            for g3=1:numOfTokensInCell

                tokenize_polje=setfield(tokenize_polje,['token',num2str(g3)],char(cell_tokena(g3)));
            end

            sva_tokenize_polja=setfield(sva_tokenize_polja,['tokens',num2str(g2)],tokenize_polje);
        end
    end
end

```

```
        tokenize_polje=[];
        cell_tokena={};
    end

    for g4=1:numOfTokensModels
        %tokensStruct=tokenize([TXTsPath, 'MODEL', imeModela, '.txt']);
        tokensModelN=getfield(sva_tokenize_polja, ['tokens', num2str(g4)]);
        fp_tokensStruct=fpStruct(tokensModelN);

        % Zamjena frekvencije 1(zato što su sad liste jedinstvene) sa prijašnjim
        frekvencijama
        txt2lower([TXTsPath, 'MODEL', char(imeModela_polje(g4)), '.txt']);

        tokenize_freq_org=fpStruct(tokenize([TXTsPath, 'MODEL', char(imeModela_polje(g4)), '.txt']));
        numOfUniqueTokens=size(fieldnames(fp_tokensStruct),1);
        fp_token_names=fieldnames(fp_tokensStruct);
        for g5=1:numOfUniqueTokens
            stara_vrijednost=getfield(tokenize_freq_org, char(fp_token_names(g5)));

            fp_tokensStruct=setfield(fp_tokensStruct, char(fp_token_names(g5)), stara_vrijednost);

        end

        % sum(cell2mat(struct2cell(fp_tokensStruct)))

        % zapis u *.dat
        PathForModel=[TXTsPath, 'MODEL', char(imeModela_polje(g4)), '.dat'];
        fp_cell={};
        fp_cell(:,1)=fieldnames(fp_tokensStruct);
        fp_cell(:,2)=struct2cell(fp_tokensStruct);
        fp_cell=sortrows(fp_cell,-2);

        sizeOfCell=size(fp_cell,1);
        fp_out = fopen( PathForModel, 'w');

        fprintf(fp_out, '%s\t%d\n', 'count__', sizeOfTXTList_polje(g4));
        for k=1:sizeOfCell
            fprintf(fp_out, '%s\t%d\n', char(fp_cell(k,1)), cell2mat(fp_cell(k,2)));
        end
        fclose(fp_out);
    end
end

end
```


12.4.24 loadDAT.m

```
function [ outStruct,nameOfModels ] = loadDAT( modelsPath )
%LOADDAT učitava podatke iz modela *.DAT za Bayes TC
% modelsPath-staza do direktorija u kojem se nalaze zapisi modela u
% obliku *.dat te ih pohranjuje u strukturu
% outStruct-sadrži tokene i frekvencije njihova pojavljivanja za svaki od
% modela
% nameOfModels-sadrži imena svakog modela, tako da bi se mogla povezati
% pripadnost tokena pojedinom modelu.

outStruct=struct;
nameOfModels=struct;
tokenStruct=struct;
currentFolder = pwd;
cd(modelsPath)
dir_list=dir('*.dat');
cd(currentFolder)

sizeOfModelList=size(dir_list,1);

for k=1: sizeOfModelList
    imena(k)={dir_list(k).name};
    pathOfModel=[modelsPath,dir_list(k,1).name];
    imeModela=dir_list(k,1).name;
    imeModela=imeModela(1:(strfind(imeModela,'.')-1));
    nameOfModels=setfield(nameOfModels,['modelName',num2str(k)],dir_list(k,1).name);
    fp_in = fopen(pathOfModel,'r');
    tline = fgetl(fp_in);
    while ischar(tline)
        tline=strtrim(tline);
        [tlineKey,tlineValue] = strread(tline,'%s%f','delimiter','\t');
        tokenStruct=setfield(tokenStruct,char(tlineKey),tlineValue);
        tline = fgetl(fp_in);
    end
    fclose(fp_in);
    outStruct=setfield(outStruct,['tokens',num2str(k)],tokenStruct);
    tokenStruct={};
end
end
```

12.4.25 trainMNNB.m

```

function [ allNBModels ,priorsC] = trainMNNB( modelsStruct )
%TRAINMNNB funkcija za treniranje modela po Bayesovim vjerojatnostima
% modelsStruct-Ulazna struktura sa tokenima iz modela *.dat
% allNBModels-bayesove vjerojatnosti za svaki pojedini token i model
% priorsC-priori vjerojatnosti pojedine klase

%[modelsStruct,nameOfModels]=loadDAT('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski
rad\Python kodovi\KEGprimjeri\Emocije u txt\test_primjer\');

countDocs=0;
numOfModels=size(fieldnames(modelsStruct),1);
tAllTerms=struct;
for k=1:numOfModels
    tempModel=getfield(modelsStruct,['tokens',num2str(k)]);
    countDocs=countDocs+tempModel.count__;
    numOfDiffTokens=size(fieldnames(tempModel),1);
    ModelNames = fieldnames(tempModel);
    ModelValues = cell2mat(struct2cell(tempModel));
    for j=1:numOfDiffTokens
        if strcmp(char(ModelNames(j)),'count__')
            continue
        end
        if isfield(tAllTerms,ModelNames(j))
            postojeca_vrijednost=getfield(tAllTerms,char(ModelNames(j)));
            nova_vrijednost=postojeca_vrijednost+ModelValues(j);
            tAllTerms=setfield(tAllTerms,char(ModelNames(j)),nova_vrijednost);
        else
            tAllTerms=setfield(tAllTerms,char(ModelNames(j)),ModelValues(j));
        end
    end
end
totalTerms=sum(cell2mat(struct2cell(tAllTerms)));

numOfTotalTerms=size(fieldnames(tAllTerms),1);
allTerms=struct;
allNBModels=struct;

struct_polje=[];
ModelNames = fieldnames(tAllTerms);
%ModelValues = cell2mat(struct2cell(tAllTerms));
for h=1:numOfTotalTerms
    postojeca_vrijednost=getfield(tAllTerms,char(ModelNames(h)));
    nova_vrijednost=postojeca_vrijednost/totalTerms;
    allTerms=setfield(allTerms,char(ModelNames(h)),nova_vrijednost);
    allNBModels=setfield(allNBModels,char(ModelNames(h)),struct_polje);
end
tAllTerms=[];

countCurrentDocs = 0;
for k=1:numOfModels
    tempModel=getfield(modelsStruct,['tokens',num2str(k)]);
    countDocsInClass=tempModel.count__;
    priorC = countDocsInClass/countDocs;
    numOfallTerms=size(fieldnames(allTerms),1);
    allTermsNames = fieldnames(allTerms);
    allTermsValues = cell2mat(struct2cell(allTerms));
    for g=1:numOfallTerms
        if isfield(tempModel,allTermsNames(g))
            condProb=(getfield(tempModel,char(allTermsNames(g)))+1)/(allTermsValues(g)+1);
            struct_polje=getfield(allNBModels,char(ModelNames(g)));
            struct_polje(k)=condProb;
            allNBModels=setfield(allNBModels,char(ModelNames(g)),struct_polje);
        else
            condProb=1/(allTermsValues(g)+1);
            struct_polje=getfield(allNBModels,char(ModelNames(g)));
            struct_polje(k)=condProb;
            allNBModels=setfield(allNBModels,char(ModelNames(g)),struct_polje);
        end
    end
    priorsC(k)=priorC;
end
end

```

12.4.26 `compare_with_NBM.m`

```
function [ result, imenaModela ] = compare_with_NBM( allNBModels, priorsC,
modelsStruct, nameOfModels, UnknownPath, dispFlag )
%COMPARE_WITH_NBM fukcija za izračun najvjerojatnijeg bayesovog modela
%
    %[modelsStruct, nameOfModels]=loadDAT(modelsStruct);
    %[ allNBModels ,priorsC] = trainMNNB( modelsStruct );

    tokeniUnknown=tokenize(UnknownPath);
    result=[];
    numOfModels=size(fieldnames(modelsStruct),1);
    numOfUnknownTokens=size(fieldnames(tokeniUnknown),1);
    for k=1:numOfModels
        score=log(priorsC(k));
        for h=1:numOfUnknownTokens
            token=getfield(tokeniUnknown, ['token', num2str(h)]);
            if isfield(allNBModels, token)
                polje=getfield(allNBModels, token);
                add_score=polje(k);
                score=score+log(add_score);
            end
        end
        result(k)=score;
    end
    imenaModela=struct2cell(nameOfModels);
    if dispFlag==1
        for h=1:size(result,2)
            disp(sprintf('Model %s: score %.10f', char(imenaModela(h)), result(h)));
        end
        [C,I] = max(result);
        disp(sprintf('The winner is: %s\n', char(imenaModela(I))));
    end

end
```

12.4.27 naive_bayes_text_classifier.m

```
function [ result, imenaModela ] = naive_bayes_text_classifier( modelsPath, UnknownPath,
dispFlag )
%NAIVE_BAYES_TEXT_CLASSIFIER koja radi klasifikaciju na principu Naive
%Bayes teorije
%  modelsPath-Staza do *dat modela
%  UnknownPath-Staza do datoteke sa nepoznatim(neklasificiranim) tekstom
%  dispFlag-zastavica čijim postavljanjem se omogućuje ispis rezultata

[modelsStruct, nameOfModels]=loadDAT(modelsPath);
[ allNBModels, priorsC] = trainMNNB( modelsStruct );
[ result, imenaModela ] = compare_with_NBM( allNBModels, priorsC,
modelsStruct,nameOfModels, UnknownPath,dispFlag );

end
```

12.4.28 auto_prim_model.m

```
% Niz naredbi koji služi za prikaz primjera rada primitivnog unigram
% sustava
KEGto5class('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\metapodaciKEG_stat_final.xlsx','D:\FER\Četvrti semestar
DIPLOMSKI\Diplomski rad\Pyton kodovi\KEGprimjeri\','0)
bezkvacice('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\ljutnja.txt')
bezkvacice('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\nutralno.txt')
bezkvacice('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\sreca.txt')
bezkvacice('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\strah.txt')
bezkvacice('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\tuga.txt')

tokenizer('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton kodovi\KEGprimjeri\tuga-
bez_kvacice.txt',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton kodovi\KEGprimjeri\tuga-
bez_kvacice.lst')

tokenizer('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton kodovi\KEGprimjeri\sreca-
bez_kvacice.txt',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton kodovi\KEGprimjeri\sreca-
bez_kvacice.lst')

tokenizer('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\ljutnja-bez_kvacice.txt',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\ljutnja-bez_kvacice.lst')

tokenizer('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\nutralno-bez_kvacice.txt',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\nutralno-bez_kvacice.lst')

tokenizer('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton kodovi\KEGprimjeri\strah-
bez_kvacice.txt',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton kodovi\KEGprimjeri\strah-
bez_kvacice.lst')

rfp_unique('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton kodovi\KEGprimjeri\sreca-
bez_kvacice.lst',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton kodovi\KEGprimjeri\tuga-
bez_kvacice.lst',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton kodovi\KEGprimjeri\strah-
bez_kvacice.lst',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\ljutnja-bez_kvacice.lst',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\nutralno-bez_kvacice.lst')

compareWithModels('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\','D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\unknown.txt',1)
```

12.4.29 auto_prim_model2.m

```

% Niz naredbi koje služe izvlačenju statistike frekvencije pojavljivanja i
% relativnje frekvencije za pet razreda tokena (emocija: sreca, tuga,
% strah, ljutnja, neutralno)

sreca=tokenize('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\sreca-bez_kvacice.txt')
tuga=tokenize('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\tuga-bez_kvacice.txt')
strah=tokenize('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\strah-bez_kvacice.txt')
ljutnja=tokenize('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\ljutnja-bez_kvacice.txt')
neutralno=tokenize('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\nneutralno-bez_kvacice.txt')

fp_sreca=fpStruct(sreca);
fp_tuga=fpStruct(tuga);
fp_strah=fpStruct(strah);
fp_ljutnja=fpStruct(ljutnja);
fp_neutralno=fpStruct(neutralno);

fp_sreca_cell={};
fp_tuga_cell={};
fp_strah_cell={};
fp_ljutnja_cell={};
fp_neutralno_cell={};

fp_sreca_cell(:,1)=fieldnames(fp_sreca);
fp_sreca_cell(:,2)=struct2cell(fp_sreca);

fp_tuga_cell(:,1)=fieldnames(fp_tuga);
fp_tuga_cell(:,2)=struct2cell(fp_tuga);

fp_strah_cell(:,1)=fieldnames(fp_strah);
fp_strah_cell(:,2)=struct2cell(fp_strah);

fp_ljutnja_cell(:,1)=fieldnames(fp_ljutnja);
fp_ljutnja_cell(:,2)=struct2cell(fp_ljutnja);

fp_neutralno_cell(:,1)=fieldnames(fp_neutralno);
fp_neutralno_cell(:,2)=struct2cell(fp_neutralno);

fp_sreca_cell=sortrows(fp_sreca_cell,-2);
fp_tuga_cell=sortrows(fp_tuga_cell,-2);
fp_strah_cell=sortrows(fp_strah_cell,-2);
fp_ljutnja_cell=sortrows(fp_ljutnja_cell,-2);
fp_neutralno_cell=sortrows(fp_neutralno_cell,-2);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
fp_sreca_struct=cell2struct(fp_sreca_cell(:,2),fp_sreca_cell(:,1),1)
fp_tuga_struct=cell2struct(fp_tuga_cell(:,2),fp_tuga_cell(:,1),1)
fp_strah_struct=cell2struct(fp_strah_cell(:,2),fp_strah_cell(:,1),1)
fp_ljutnja_struct=cell2struct(fp_ljutnja_cell(:,2),fp_ljutnja_cell(:,1),1)
fp_neutralno_struct=cell2struct(fp_neutralno_cell(:,2),fp_neutralno_cell(:,1),1)

rfp_sreca=rfpStruct(fp_sreca_struct);
rfp_tuga=rfpStruct(fp_tuga_struct);
rfp_strah=rfpStruct(fp_strah_struct);
rfp_ljutnja=rfpStruct(fp_ljutnja_struct);
rfp_neutralno=rfpStruct(fp_neutralno_struct);

```

12.4.30 auto_prim_model3.m

```
% Niz naredbi koje služe izvlačenju statistike relativne frekvencije
% pojavljivanja i relativnje frekvencije za pet razreda tokena (emocija:
% sreca, tuga, strah, ljutnja, neutralno)

razredi=rfp_unique('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Unique proba\sreca-bez_kvacice.lst',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Unique proba\tuga-bez_kvacice.lst',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Unique proba\strah-bez_kvacice.lst',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Unique proba\ljutnja-bez_kvacice.lst',...
'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Unique proba\nneutralno-bez_kvacice.lst');

sreca=razredi.tokens1;
tuga=razredi.tokens2;
strah=razredi.tokens3;
ljutnja=razredi.tokens4;
neutralno=razredi.tokens5;

sreca_cell={};
tuga_cell={};
strah_cell={};
ljutnja_cell={};
neutralno_cell={};

sreca_cell(:,1)=fieldnames(sreca);
sreca_cell(:,2)=struct2cell(sreca);

tuga_cell(:,1)=fieldnames(tuga);
tuga_cell(:,2)=struct2cell(tuga);

strah_cell(:,1)=fieldnames(strah);
strah_cell(:,2)=struct2cell(strah);

ljutnja_cell(:,1)=fieldnames(ljutnja);
ljutnja_cell(:,2)=struct2cell(ljutnja);

neutralno_cell(:,1)=fieldnames(neutralno);
neutralno_cell(:,2)=struct2cell(neutralno);

sreca_cell=sortrows(sreca_cell,-2);
tuga_cell=sortrows(tuga_cell,-2);
strah_cell=sortrows(strah_cell,-2);
ljutnja_cell=sortrows(ljutnja_cell,-2);
neutralno_cell=sortrows(neutralno_cell,-2);
```

12.4.31 auto_prim_model4.m

```
KEGto_all_TXT('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Bayes TC\metapodaciKEG_stat_final-anotirane_snimke.xlsx',...
            'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Bayes TC\');

TXTtoModels('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python kodovi\KEGprimjeri\Bayes
TC\',...
            {'sreca*.txt','tuga*.txt','strah*.txt','ljutnja*.txt','neutralno*.txt'},...
            {'sreca','tuga','strah','ljutnja','neutralno'})

naive_bayes_text_classifier('D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Bayes TC\',...
            'D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Bayes TC\unknown.txt',...
            1 );

%modelsPath='D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Emocije u txt\test_primjer\';
%UnknownPath='D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Python
kodovi\KEGprimjeri\Emocije u txt\test_primjer\unknown.txt';
```


12.4.32 HTK_podjela_snimaka.m

```
function [ output_args ] = HTK_podjela_snimaka( inputPathToSnimke,metaKEGPath )
%HTK_PODJELA_SNIMAKA razdvaja snimke u datoteke tako kao priprema za kros
%validaciju. Razdvaja se korpus na 10 Trening i 10 Test direktorija kako
%bi se pokrilo 100% KEG-a pri testiranju i obavila kros-validacija.
% inputPathToSnimke-Staza do snimaka (u našem slučaju direktorij SNIMKE)
% metaKEGPath-Staza do *.xlsx datoteke u kojoj se nalazi popis snimaka i
% pripadajući transkript

%*****
% Dio koji kreira direktorij "Anotirane snimke" i u njega kopira isključivo
% one snimke iz korpusa koje su anotirane (a nalaze se u direktoriju
% SNIMKE čiju stazu smo prosljedili u inputPathSnimke) u isti direktorij
% se ujedno kreira i nova *.xlsx datoteka sa transkriptom anotiranih
% snimaka. Po završetku kopiranja sva daljnja obrada se vrši samo na
% anotiranim snimka, pošto su samo one pogodne za daljnu obradu.
%*****
inputPathToSnimke='D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\HTK\SNIMKE\';
metaKEGPath='D:\FER\Četvrti semestar DIPLOMSKI\Diplomski rad\Pyton
kodovi\KEGprimjeri\metapodaciKEG_stat_final.xlsx';

mkdir(inputPathToSnimke,'Anotirane snimke');
inputPathToSnimkeAnotirane=[inputPathToSnimke,'Anotirane snimke\'];
delete([inputPathToSnimkeAnotirane,'*.wav']);
delete([inputPathToSnimkeAnotirane,'*.xlsx']);
M={};
[num, txt] = xlsread(metaKEGPath,'Sheet1');

orderNum=num(:,1);
sizeOf_orderNum=size(orderNum,1);
sentences=txt(2:sizeOf_orderNum+1,7);
emotions=txt(2:sizeOf_orderNum+1,55);

counter_sreca=1;
counter_ljutnja=1;
counter_tuga=1;
counter_strah=1;
counter_neutralno=1;
counter_sve_emocije=1;

%incijalizacija M-a, upis headera
M(counter_sve_emocije,1)={'ID snimke'};
M(counter_sve_emocije,7)={'transkripcija'};
M(counter_sve_emocije,55)={'r - Emocija'};
counter_sve_emocije=counter_sve_emocije+1;

for k=1:sizeOf_orderNum
    if strcmp(char(emotions(k)),'SR')
        sreca(counter_sreca)=sentences(k);

        M(counter_sve_emocije,1)=num2cell(counter_sve_emocije-1);
        M(counter_sve_emocije,7)=sentences(k);
        M(counter_sve_emocije,55)=emotions(k);

        copyfile([inputPathToSnimke,num2str(k),'.wav'],[inputPathToSnimkeAnotirane,num2str(counter_sve_emocije-1),'.wav']);

        counter_sreca=counter_sreca+1;
        counter_sve_emocije=counter_sve_emocije+1;

    elseif strcmp(char(emotions(k)),'LJ')
        ljutnja(counter_ljutnja)=sentences(k);

        M(counter_sve_emocije,1)=num2cell(counter_sve_emocije-1);
        M(counter_sve_emocije,7)=sentences(k);
        M(counter_sve_emocije,55)=emotions(k);

        copyfile([inputPathToSnimke,num2str(k),'.wav'],[inputPathToSnimkeAnotirane,num2str(counter_sve_emocije-1),'.wav']);

        counter_ljutnja=counter_ljutnja+1;
        counter_sve_emocije=counter_sve_emocije+1;
```

```

elseif strcmp(char(emotions(k)),'TU')
    tuga(counter_tuga)=sentences(k);

    M(counter_sve_emocije,1)=num2cell(counter_sve_emocije-1);
    M(counter_sve_emocije,7)=sentences(k);
    M(counter_sve_emocije,55)=emotions(k);

copyfile([inputPathToSnimke,num2str(k),'.wav'],[inputPathToSnimkeAnotirane,num2str(counter_sve_emocije-1),'.wav']);

    counter_tuga=counter_tuga+1;
    counter_sve_emocije=counter_sve_emocije+1;
elseif strcmp(char(emotions(k)),'ST')
    strah(counter_strah)=sentences(k);

    M(counter_sve_emocije,1)=num2cell(counter_sve_emocije-1);
    M(counter_sve_emocije,7)=sentences(k);
    M(counter_sve_emocije,55)=emotions(k);

copyfile([inputPathToSnimke,num2str(k),'.wav'],[inputPathToSnimkeAnotirane,num2str(counter_sve_emocije-1),'.wav']);

    counter_strah=counter_strah+1;
    counter_sve_emocije=counter_sve_emocije+1;
elseif strcmp(char(emotions(k)),'NE')
    neutralno(counter_neutralno)=sentences(k);

    M(counter_sve_emocije,1)=num2cell(counter_sve_emocije-1);
    M(counter_sve_emocije,7)=sentences(k);
    M(counter_sve_emocije,55)=emotions(k);

copyfile([inputPathToSnimke,num2str(k),'.wav'],[inputPathToSnimkeAnotirane,num2str(counter_sve_emocije-1),'.wav']);

    counter_neutralno=counter_neutralno+1;
    counter_sve_emocije=counter_sve_emocije+1;
end
end
sreca=sreca';
tuga=tuga';
ljutnja=ljutnja';
strah=strah';
neutralno=neutralno';
mataNewKEPath=[inputPathToSnimkeAnotirane,'anotirane_snimke.xlsx'];
xlswrite(mataNewKEPath,M);
%*****
%*****

for k=1:10
    mkdir(inputPathToSnimke,['Test',num2str(k)]);
    mkdir(inputPathToSnimke,['Trening',num2str(k)]);
    delete([inputPathToSnimke,'Test',num2str(k),'\','*.wav']);
    delete([inputPathToSnimke,'Test',num2str(k),'\','*.xlsx']);
    delete([inputPathToSnimke,'Test',num2str(k),'\','*.txt']);
    delete([inputPathToSnimke,'Trening',num2str(k),'\','*.wav']);
    delete([inputPathToSnimke,'Trening',num2str(k),'\','*.xlsx']);
    delete([inputPathToSnimke,'Trening',num2str(k),'\','*.txt']);
end

[num, txt] = xlsread([inputPathToSnimkeAnotirane,'anotirane_snimke.xlsx'],'Sheet1');

orderNum=num(:,1);
sizeOf_orderNum=size(orderNum,1);
sentences=txt(2:sizeOf_orderNum+1,7);
emotions=txt(2:sizeOf_orderNum+1,55);

TranskriptCell={};
TranskriptCell(:,1)=num2cell(orderNum);
TranskriptCell(:,2)=num2cell(randperm(sizeOf_orderNum)); %generira random niz po kojem
se sortiraju rečenice
TranskriptCell(:,3)=emotions;
TranskriptCell(:,4)=sentences;

TranskriptCellPerm=sortrows(TranskriptCell,2);

sizeOfTestCorpus=floor(sizeOf_orderNum*0.1);

```

```

sizeOfTreningCorpus=sizeOf_orderNum-sizeOfTestCorpus;

TestVectorStart=1;
TestVectorEnd=sizeOfTestCorpus;
for k=1:10
    TreningPath=[inputPathToSnimke, 'Trening', num2str(k), '\'];
    TestPath=[inputPathToSnimke, 'Test', num2str(k), '\'];

    TestVector=[TestVectorStart:TestVectorEnd];
    TestVectorStart=TestVectorEnd+1;
    TestVectorEnd=TestVectorEnd+sizeOfTestCorpus;

    TreningVector=[1:sizeOf_orderNum];
    TreningVector=setdiff(TreningVector,TestVector);

    M1={};
    M1(1,1)={'ID snimke'};
    M1(1,7)={'transkripcija'};
    M1(1,55)={'r - Emocija'};

    TreningTXTPath=[inputPathToSnimke, 'Trening', num2str(k), '\', 'trainprompts', '.txt'];
    TreningTXTPath2=[inputPathToSnimke, 'Trening', num2str(k), '\', 'tekst_hr', '.txt'];
    fp_out_txt=fopen(TreningTXTPath, 'w');
    fp_out_txt2=fopen(TreningTXTPath2, 'w');
    for h=1:size(TreningVector,2)
        numberForTrening=TreningVector(h);
        %M1(h+1,1)=TranskriptCellPerm(numberForTrening,2); %broj
        M1(h+1,1)={h}; %broj
        M1(h+1,7)=TranskriptCellPerm(numberForTrening,4); %transkrip
        M1(h+1,55)=TranskriptCellPerm(numberForTrening,3); %emocija

        %***** KOPIRANJE SNIMKE *****%

copyfile([inputPathToSnimkeAnotirane, num2str(cell2mat(TranskriptCellPerm(numberForTrening,1)))], '.wav'), [inputPathToSnimke, 'Trening', num2str(k), '\', sprintf('S%04.0f', h), '.wav']);
        %*****KREIRANJE trainprompts.txt*****%
        string_bez_kvaciceHTK=bezkvacice3(char(M1(h+1,7)));
        fprintf(fp_out_txt, 'S%04.0f %s\n', h, string_bez_kvaciceHTK);
        fprintf(fp_out_txt2, '%s\n', char(M1(h+1,7)));
    end
    fclose(fp_out_txt);
    fclose(fp_out_txt2);
    xlswrite([TreningPath, 'Trening', num2str(k), '.xlsx'], M1)

    M2={};
    M2(1,1)={'ID snimke'};
    M2(1,7)={'transkripcija'};
    M2(1,55)={'r - Emocija'};

    TestTXTPath=[inputPathToSnimke, 'Test', num2str(k), '\', 'testprompts', '.txt'];
    TestTXTPath2=[inputPathToSnimke, 'Test', num2str(k), '\', 'tekst_te', '.txt'];
    fp_out_txt=fopen(TestTXTPath, 'w');
    fp_out_txt2=fopen(TestTXTPath2, 'w');
    for h=1:size(TestVector,2)
        numberForTest=TestVector(h);
        %M2(h+1,1)=TranskriptCellPerm(numberForTest,2); %broj
        M2(h+1,1)={h}; %broj
        M2(h+1,7)=TranskriptCellPerm(numberForTest,4); %transkrip
        M2(h+1,55)=TranskriptCellPerm(numberForTest,3); %emocija

        %***** KOPIRANJE SNIMKE *****%

copyfile([inputPathToSnimkeAnotirane, num2str(cell2mat(TranskriptCellPerm(numberForTest,1)))], '.wav'), [inputPathToSnimke, 'Test', num2str(k), '\', sprintf('T%04.0f', h), '.wav']);
        %*****KREIRANJE testprompts.txt*****%
        string_bez_kvaciceHTK=bezkvacice3(char(M2(h+1,7)));
        fprintf(fp_out_txt, 'T%04.0f %s\n', h, string_bez_kvaciceHTK);
        fprintf(fp_out_txt2, '%s\n', char(M2(h+1,7)));
    end
    fclose(fp_out_txt);
    fclose(fp_out_txt2);
    xlswrite([TestPath, 'Test', num2str(k), '.xlsx'], M2)
end
end

```