

Klasificiranje mentalnog zdravlja autora analizom teksta

Sonja Grdan, Veljko Srdarević

Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva
Unska 3, 10000 Zagreb, Hrvatska
{sonja.grdan, veljko.srdarevic}@fer.hr

Sažetak

Proučavanje psihičkih poremećaja nudi raznovrsne informacije o funkcioniranju ljudskog mozga, a takve podatke moguće je analizirati i široko primjenjivim metodama strojnog učenja. U ovom radu proučavani su pisani tekstovi kao reprezentanti kognitivnog funkcioniranja ispitanika različitog mentalnog statusa (autori oboljeli od shizofrenije, bipolarnog poremećaja, depresije i kontrolne skupine bez navedenih poremećaja). Koristeći stroj s potpornim vektorima, naivni Bayesov klasifikator i naivni Bayesov klasifikator s procjenom jezgrene funkcije izgrađeni su klasifikatori tekstova autora s različitim psihičkim statusima. Dobiveni rezultati nisu potpuno očekivane točnosti, no pokazuju poboljšanje u klasifikaciji u odnosu na referentnu metodu klasificiranja u apriorno najvjerojatniju klasu, pri čemu Bayesovi klasifikatori postižu bolje rezultate. Trendovi i prijedlozi metodoloških poboljšanja prikazani u ovom radu predstavljaju poticaj daljnjem proučavanju ovog problema.

1. Uvod

Čovjekova želja za spoznavanjem samog sebe oduvijek je bila veliki motivator istraživanja ljudskih sposobnosti. Pritom vrijedne podatke pružaju kako proučavanja normalnih, prototipičnih pojava, tako i onih koje po svojim karakteristikama odskakuju po nekom od kriterija normalnosti. Mentalni poremećaji zanimljivo su područje proučavanja u kognitivnoj znanosti budući da nude mogućnost uvida u rad mozga izvan uobičajenih okvira. Oni mogu zahvaćati sve aspekte ljudskog funkcioniranja: ponašanje, emotivne doživljaje, tjelesne simptome i kogniciju. Pojedina područja različito su zahvaćena ovisno o vrsti i težini poremećaja. U ovom radu proučavane su posljedice koje određeni poremećaji mogu imati na kogniciju, pa su u nastavku navedeni neki mogući simptomi shizofrenije, bipolarnog poremećaja i depresije.

Shizofrenija se ističe sveobuhvatnošću mentalnih funkcija koje zahvaća, od čega ističemo promjene u mišljenju koje se očituju kao disocirano mišljenje: slijed misli bez smislene povezanosti, brzo skakanje misli s ideje na ideju, "ljepljivost" misli za jednu ideju i njezino teško mijenjanje i sl. (WHO, 1992). Bipolarni poremećaj primarno je poremećaj raspoloženja, no također može utjecati na kognitivne funkcije kao što su pažnja i verbalno pamćenje, rezultirajući u maničnoj fazi ubrzanim mislima i bijegom ideja: brzim promjenama teme misli, no koje u podlozi imaju logičan slijed, a nerijetko se javljaju i osjećaji velike vlastite važnosti i posebnosti. Također, dobivena je i povezanost s kreativnošću (Santosa et al., 2007). Kod depresije kognitivni simptomi su prvenstveno loša koncentracija i pesimistične, beznadne misli (Folnegović-Šmalc, 1996). Zbog jasne povezanosti procesa misli i jezika, logično je očekivati da bi se navedene karakteristike oboljelih od ovih bolesti mogle očitovati i u njihovom tekstualnom izričaju.

Simptomi bolesti vidljivi u mišljenju očituju se i na formalnoj i na sadržajnoj razini. Primjerice, sadržaj misli shizofrenih bolesnika može biti posve besmislen, no gramatička struktura pritom može biti očuvana. Također, misaoni tok depresivnih osoba često je zatvoren u krug tužnih, negativnih misli. U ovome radu osobita je pozornost po-

svećena formalnim karakteristikama teksta, budući da bi potpunija analiza semantičkog aspekta tekstova zbog svoje kompleksnosti svakako nadilazila okvire ovog istraživanja. Iznimku pritom čini uzimanje u obzir vlastite važnosti kao simptoma, koja je pomoću jednostavne operacionalizacije uzeta u obzir. U usporedbi s tekstovima mentalno zdravih osoba tekstovi oboljelih mogli bi se razlikovati po nekim karakteristikama izvedenim iz simptoma, što otvara mogućnost automatskog kategoriziranja takvih tekstova.

Automatski klasifikator je računalni sustav koji na temelju značajki danih uzoraka određuje njihovu pripadnost određenim kategorijama. Izrada klasifikatora uključuje odabir i izvlačenje značajki iz uzorka, određivanje skupa podataka za treniranje i testiranje, odabir algoritma klasifikacije, treniranje samog klasifikatora na podacima za treniranje, te potom testiranje njegove uspješnosti na podacima za testiranje.

Dijagnosticiranje mentalnih bolesti zahtjevan je problem, pa bi izrada sustava koji procjenjuje pripadnost nekog teksta kategorijama bolesti i zdravlja mogla pružiti korisne dodatne informacije za dijagnostičare. Cilj prikazanog istraživanja bio je stoga izraditi klasifikator tekstova dobivenih od sudionika različitog mentalnog stanja (psihički zdravi, te ispitanici sa shizofrenim, depresivnim i bipolarnim poremećajem). Svrha rada je provjeriti jesu li karakteristike navedenih mentalnih poremećaja dovoljno izražene na tekstovima psihički bolesnih sudionika da bi se oni razlikovali od tekstova psihički zdravih sudionika, te, ukoliko jesu, pridonijeti postupku dijagnostike.

U idućem odjeljku dan je pregled dosadašnjih istraživanja ove teme. U trećem odjeljku opisan je metodološki pristup klasifikaciji tekstova s obzirom na mentalni status autora korišten u ovom radu. U četvrtom odjeljku prikazani su alati i algoritmi korišteni za klasifikaciju te rezultati vrednovanja. Potom je u petom odjeljku iznesen zaključak, uz prijedloge za buduća istraživanja.

2. Pregled dosadašnjih istraživanja

Problemom klasifikacije mentalnog statusa autora teksta bavilo se nekoliko istraživanja autora koji su proučavali

tekstove pacijenata psihijatrijskih ustanova. Strous et al. (2009) radili su automatsku klasifikaciju tekstova na temu *Važna osoba u mojem životu*, koje je napisalo 36 shizofrenih ispitanika nasuprot 36 ne-shizofrenih ispitanika, pomoću stroja s potpornim vektorima (engl. *support vector machine*, SVM) i Bayesove regresije. Korištene su značajke najčešćih riječi, tri-grama slova (zbog specifičnosti hebrejskog jezika u kojima su gramatički značajni) i ponavljanje riječi, te je dobivena uspješnost klasifikacije tekstova od 83,3%.

Diederich et al. (2007) su kao skup za učenje koristili tekstove 31 shizofrenih i 16 maničnih pacijenata psihijatrijske ustanove i 9 kontrolnih sudionika, koji su bilo verbalizirali tekst koji je kasnije pretipkan, bilo sami napisali tekst. Tekstovi su bili pisani na jednu od dvije zadane teme (*Piknik u parku* i *Lopov*) uz zadatak da se u tekst uklopi unaprijed određena lista riječi. Tekstovi su pretvoreni u oblik hrpe riječi (engl. *bag of words*) i korištena je frekvencija riječi, a učenje je izvedeno pomoću SVM i stabla odluke. Uspješnost klasifikacije shizofrene nasuprot kontrolne skupine bila je 77%, dok u klasifikaciji koja je uključivala manične sudionike uspješnost nije bila bolja od slučajne.

Rješavanje ovog problema donekle je slično rješavanju problema određivanja i/ili razlikovanja autora teksta, budući da na neki način pretpostavljamo postojanje određenih stilskih sličnosti u tekstovima autora koji pate od bipolarnog poremećaja, shizofrenije i depresije. Coyotl-Morales et al. (2006) navode tri glavna pristupa problemu raspoznavanja autora i njima pripadne metode. Prvi pristup je usmjeren na uspoređivanje tekstova po mjerama stila, što se odnosi na značajke poput duljine riječi i rečenica i bogatstvo vokabulara, koje ovise o žanru teksta i duljini teksta. Drugi pristup koristi sintaksna obilježja kao značajke, analizirajući sintaksnu strukturu u pozadini teksta, što je najčešće računalno zahtjevno. Naposljetku, treći pristup kao značajke koristi osobine riječi korištenih u tekstu. Tako se mogu koristiti funkcijske riječi nezavisne od teme teksta, ili pristup hrpe riječi koji je dobar ukoliko postoji povezanost između autora i tema o kojima piše.

Belša et al. (2010) uspješno su rješavali problem automatskog prepoznavanja autora teksta, pri čemu su koristili velik broj mogućih značajki koje određuju tekst, a koje nisu ovisne o temi. Najbolje rezultate producirala je sljedeća kombinacija značajki: broj interpunkcijskih znakova, broj funkcijskih riječi, broj samoglasnika i duljine riječi: njome je dobivena uspješnost klasifikacije od 88%.

Govoreći o duljini tekstova potrebnoj da bi raznolikost rječnika bila stabilno određena i predstavljala potencijalno dobru značajku, Tweedie and Baayen (2006) navode da je minimalna potrebna duljina 1000 riječi.

U istraživanju prikazanom u ovome radu, za razliku od radova (Strous et al., 2009) i (Diederich et al., 2007), korišteni su tekstovi koje su napisali profesionalni pisci i autori blogova. Time se pokušalo utvrditi koliko su pokazatelji navedenih psihičkih bolesti u tekstu prožimajući u životu ljudi s tim poremećajima - budući da se radi o ljudima koji su u vrijeme pisanja očito dovoljno dobro funkcionirali da bi mogli objaviti knjigu, odnosno održavati blog, moguće je da im tada ni kognitivne funkcije nisu bile izmijenjene u skladu s tipičnim simptomima bolesti. Koristili

smo i veći broj značajki od navedenih radova. Korištene značajke bazirane su na onima navedenima u (Belša et al., 2010), koje su odabrane na temelju moguće povezanosti sa simptomima psihičkih poremećaja.

3. Klasifikacija mentalnog statusa autora korištenjem tekstova

Kako bi se klasificirao mentalni status autora, korištene su značajke forme teksta izvedene na temelju simptoma pojedinih psihičkih poremećaja i značajki spomenutih u srodnim radovima. Tekstovi su potom obrađeni generiranjem numeričke vrijednosti za svaku značajku, te tako dobiveni podatci o značajkama analizirani različitim algoritmima kako bi se mogli usporediti rezultati. Svaki od navedenih koraka detaljnije je opisan u nastavku.

3.1. Odabir značajki

U analizi tekstova korištene su sljedeće značajke: broj različitih riječi, broj zamjenica prvog lica jednine (*ja*, *meni*, *mene* i sl.), broj interpunkcijskih znakova, broj funkcijskih riječi i prosječna dužina riječi u tekstu. Prve dvije značajke mogu se dovesti u vezu sa simptomima istraživanih poremećaja. Korištenje broja različitih riječi ima svoju podlogu u povezanosti kreativnosti i bipolarnog poremećaja (Santosa et al., 2007). Budući da se kreativnost često operacionalizira pomoću raznolikosti produkcije, može se pretpostaviti da će tekstovi autora s bipolarnim poremećajem imati veći broj različitih riječi u tekstu nego tekstovi drugih kategorija autora. Također, sklonost ka *lijepljenju* na temu može se (iz različitih razloga) javiti kod oboljelih od shizofrenije ili depresivnog poremećaja, čime mogu biti ograničeni na manji skup riječi vezan uz tu temu, pa u tim skupinama vjerojatno možemo očekivati i manji broj različitih riječi.

Motivacija za korištenje broja zamjenica u tekstu su simptomi grandioznosti prisutni kod oboljelih od bipolarnog poremećaja. Također, u radu (Strous et al., 2009) razlikovanje između tekstova shizofrenih i ne-shizofrenih pacijenata obavljeno je prema frekvenciji pojedinih riječi, pri čemu je zamjenica *ja* bila najfrekventnija u shizofrenoj skupini. Dakako, ova značajka korištena je samo na tekstovima pisanima isključivo u prvom licu, odnosno na svim kategorijama blogova. Budući da tekstovi u uzorku pisaca nisu svi bili pisani u prvom licu, na tim skupinama nije korištena ova značajka.

Preostale tri značajke korištene su zbog prikazane uspješnosti u srodnim radovima (Belša et al., 2010) iako je povezanost s psihičkim bolestima samo spekulativna. Primjerice, moguće je da dezorganiziranost misli kod shizofrenije rezultira različitom količinom određenih sadržaja u tekstovima, što se može očitovati na broju funkcijskih riječi. Također, ekspresivnost i ubrzanost kod bipolarnih bolesnika možda bi mogla biti vidljiva i u korištenju većeg broja interpunkcija. S druge strane, sklonost kontemplaciji kod depresivnih autora može rezultirati duljim riječima i rečenicama (iako analiza duljine rečenica po zahtjevnosti nadilazi okvire ovog rada). Međutim, kako za ove pretpostavke nema jednoznačne potvrde u radovima drugih autora, uporaba ovih značajki posve je eksploratorna.

3.2. Prikupljanje uzorka tekstova

Podaci za analizu napravljeni su iz dvije skupine uzoraka: tekstova poznatih pisaca i tekstova prikupljenih s blogova. Psihički poremećaji kod pisaca (bipolarni poremećaj i depresija) identificirani su detaljnom analizom njihovih dostupnih biografija, te su u uzorak uzeti oni tekstovi koji se bave upravo autorovim osvrtnom na psihičku bolest (ukoliko takvi postoje) ili su pisani u vremenskom razdoblju najbližem onom u kojem je poremećaj identificiran. Tekstovi kontrolne skupine pisaca koji ne pate od depresije, bipolarnog poremećaja ili shizofrenije odabrani su također analizom biografija, uz preferenciju prema opširnim biografijama te onima u kojima se spominje neko drugo zdravstveno stanje autora, nevezano uz psihičku bolest. Svi tekstovi pisani su suvremenim engleskim jezikom, i pretežno se radi o prvim poglavljima romana.

Blogovi autora koji imaju neki psihički poremećaj pronađeni su preko web-stranica koje se bave tim poremećajima (npr. ^{1,2}) i na njima se autori deklariraju kao oboljeli od određenog poremećaja (bipolarni, shizofrenija). Blogovi psihički zdravih autora odabrani su s lista popularnih osobnih blogova, pod uvjetom da sadrže odjeljak u kojem autor navodi podatke o sebi, a ne navodi neku psihičku bolest, te također na kojima pretraživanjem nisu pronađene ključne riječi (npr. *mental*, *depression*, *illness* itd.) u kontekstu psihičke bolesti autora. Svi blogovi bave se osobnim doživljajima autora i pisani su u prvom licu, engleskim jezikom. U slučaju da autor bloga piše samo kratke zabilješke, spojeno je po nekoliko zabilježaka u jedan tekst, budući da su tekstovi klasificirani prema značajkama forme, a ne sadržaja (iako forma i sadržaj nisu potpuno nepovezani).

Radi odsutnosti dovoljnog broja poznatih pisaca koji boluju od shizofrenije ta skupina tekstova nije prikupljena, kao ni skupina tekstova autora blogova koji boluju od depresije, zbog nerijetko olakog korištenja riječi *depresija* u svakodnevnom govoru i time uzrokovane pretpostavljeno niske pouzdanosti procjene stvarne depresivnosti kod autora. Tako su se podaci za treniranje i testiranje sastojali od šest kategorija: poznati pisci s bipolarnim poremećajem, poznati pisci koji boluju od depresije, poznati pisci bez identificiranog psihičkog poremećaja, autori blogova s dijagnosticiranom shizofrenijom, autori blogova oboljeli od bipolarnog poremećaja i autori blogova bez identificiranog psihičkog poremećaja. Sljedeći kriterij o 3–5 puta većem broju primjera u pojedinoj kategoriji od broja značajki (Ribarić, 2010) u svakoj kategoriji skupa podataka za treniranje prikupljeno je 20 primjera, osim u kategoriji blogova autora sa shizofrenijom, gdje ih zbog slabe dostupnosti ima samo 14. U skupu podataka za testiranje svaka je kategorija imala 15 primjera, osim, ponovno, kategorije blogova autora sa shizofrenijom, u kojoj ih ima samo 12.

Veličina tekstova varirala je od 301 do 2519 riječi za blogove (prosječno 1131 riječi), te od 874 do 6749 riječi za tekstove pisaca (prosječno 1993 riječi). Gdje god je bila dostupna dovoljna količina teksta poštivana je preporuka o minimalno 1000 riječi potrebnih za ispravnost nekih značajki (Tweedie and Baayen, 2006). Sve značajke

određivane su uzimajući u obzir duljinu teksta.

3.3. Predobrada podataka

Obrada teksta je izvršena pomoću skripte pisane u programskom jeziku *Perl*³, kako bi se svakoj značajki pridružila brojčana vrijednost. Značajke su određene na sljedeći način:

1. broj različitih riječi u tekstu određen je kao postotni udio različitih riječi u broju svih riječi u tekstu;
2. broj zamjenica za prvo lice jednine određivan je kao frekvencija pojavljivanja tih zamjenica u odnosu na broj svih riječi u tekstu;
3. broj interpunkcijskih znakova određen je kao frekvencija pojavljivanja interpunkcijskih znakova u odnosu na dužinu teksta. Korištena je manja grupa interpunkcijskih znakova koja se sastoji od točke, upitnika, uskličnika, zareza, dvotočja, točke-zareza i apostrofa. Znakovi navoda ni crtica nisu prebrojavani, budući da zbog različitog načina navođenja dijaloga to može biti odrednica dizajna knjige i slova, a ne značajka samog teksta;
4. korištenje funkcijskih riječi određeno je kao frekvencija pojavljivanja funkcijskih riječi u tekstu u odnosu na broj svih riječi u tekstu. Takve riječi su u engleskom jeziku prilozi, prijedlozi, odrednice, veznici, pomoćni glagoli i čestice. Korištena je lista od 159 čestih funkcijskih riječi, prebrojano je njihovo pojavljivanje u tekstu, te podijeljeno brojem svih riječi u tekstu;
5. prosječna dužina riječi u tekstu izračunata je zbrajanjem duljina svih riječi u tekstu, te dijeljenjem sa brojem riječi.

3.4. Metode klasifikacije

Provedba postupaka strojnog učenja na obrađenim tekstovima rađena je u programskom alatu *Weka*⁴, koji podržava velik broj algoritama analize podataka i prediktivnog modeliranja. S obzirom na jednostavnost uporabe različitih algoritama nadziranog učenja u *Weka*-i korišteno ih je nekoliko te su uspoređeni rezultati.

Naivni Bayesov klasifikator pridjeljuje primjeru onu klasu za koju je vjerojatnost najveća na temelju njegovog vektora svojstava, uz pretpostavku da su svojstva nezavisna. Temelji se na Bayesovom teoremu (Bayes, 1763). Iako je prilično jednostavan, i počiva na često pogrešnoj pretpostavci o nezavisnosti svojstava, u uporabi redovito daje dobre rezultate (Hand and Yu, 2001)

Naivni Bayesov klasifikator s procjenom jezgrene funkcije temelji se na klasičnom Bayesovom klasifikatoru uz razliku da ne pretpostavlja normalnu raspodjelu neprekidnih varijabli, već radi procjenu jezgrene funkcije, tj. distribucije iz koje proizlaze kontinuirane varijable (John and Langley, 1995).

Stroj s potpornim vektorima (engl. *Support Vector Machine*, *SVM*) je metoda strojnog učenja koja se temelji na

¹www.psychologytoday.com

²www.schizophrenia.com

³www.perl.com

⁴<http://www.cs.waikato.ac.nz/ml/weka/>

konstrukciji hiperravnine ili hiperravnina koje odjeljuju primjere. Hiperravnine se pronalaze traženjem *potpornih vektora*, tj. onih primjera koji pripadaju različitim klasama, a najmanje su međusobne udaljenosti (Cortes and Vapnik, 1995).

4. Rezultati

Za obradu prikupljenih podataka korišteni su naivni Bayesovi klasifikatori implementirani u programskom alatu Weka. SVM s radijalnom baznom funkcijom korišten je pomoću biblioteke WLSVM⁵ (EL-Manzalawy and Honavar, 2005). Sve značajke su normalizacijom svedene na raspon [0, 1]. Prilikom izrade modela korištena je deseterostruka unakrsna provjera. Potom je na najboljim modelima za pojedine kombinacije skupina i značajki testirana uspješnost modela na odvojenim skupovima za testiranje.

Referentna metoda čiji su rezultati uspoređeni s rezultatima klasifikatora je klasifikacija u apriori najvjerojatniju klasu. U realnim uvjetima dijagnostike i na velikim uzorcima mogla bi se koristiti metoda koja u obzir uzima prevalenciju pojedinih psihičkih bolesti u populaciji, prevalenciju u populaciji ljudi koji dolaze na psihologijsku procjenu, ili pak prevalenciju u populaciji pisaca, međutim, takvi podaci nisu dostupni. U preliminarnim ispitivanjima korištene su i druge metode izrade klasifikatora, no njihova uspješnost nije bila bolja od referentne metode, stoga nisu dalje korišteni.

Dobiveni rezultati razvijenih klasifikatora prikazani su u tablici 1 na skupu blogova i tablici 2 na skupu pisaca. Iz tablica se vidi da naivni Bayesov klasifikator u obje svoje varijante nudi poboljšanje u odnosu na klasificiranje u apriori najvjerojatniju klasu, te da je uspješniji od SVM. Međutim, s obzirom na veličinu ovih razlika ne mogu se donositi konkretni zaključci. Naime, ukoliko pretpostavimo da je proporcija uspješnosti normalno distribuirana, možemo odrediti standardnu pogrešku proporcije prema formuli:

$$s_p = \sqrt{\frac{pq}{N}}$$

gdje je p proporcija točne klasifikacije, q proporcija pogrešne klasifikacije, a N broj primjera u uzorku. Tako se npr. za naivni Bayesov klasifikator s procjenom jezgrene funkcije na skupu pisaca dobiva $s_p = 0,074$, odnosno 7,4%. Ovako velika standardna pogreška proporcije uz dobivene rezultate onemogućuje zaključak da je razvijeni klasifikator bolji od jednostavne referentne metode, no ipak ukazuje na postojanje takvog trenda.

Uspješnost klasifikacije testirana je i na klasifikatorima izrađenim za razlikovanje skupina po parovima, čije rezultate vidimo u tablicama 3 i 4. Najuspješnija je klasifikacija na skupu blogova naivnim Bayesovim algoritmom, i to shizofrenih autora nasuprot onih bez poremećaja. Standardna pogreška proporcije za ovu procjenu uspješnosti iznosi $s_p = 0,075$, odnosno 7,5%. Najveće poboljšanje s obzirom na referentnu metodu ostvareno je u razlikovanju pisaca s bipolarnim poremećajem nasuprot onih sa depresijom, koristeći Bayesov algoritam s procjenom jezgrene funkcije. Standardna pogreška za ovu procjenu us-

pješnosti iznosi $s_p = 0,077$ odnosno 7,7%. Ponovno, u odnosu na dobivene standardne pogreške, razlike uspješnosti izgrađenih klasifikatora od referentne metode najvjerojatnije nisu dovoljne da bi potvrđivale stvarnu razliku u uspješnosti, no pokazuju trend koji bi trebalo dodatno istražiti.

Iz svih prikazanih tablica s rezultatima vidljivo je da se SVM algoritam nije pokazao osobito uspješnim u klasifikaciji ovih podataka, generirajući rezultate u rangu s onima koj pridjeljuju primjere u apriori najvjerojatniju klasu, do najviše 17% povećanja uspješnosti. Manja poboljšanja u klasifikaciji mogu se postići izbacivanjem nekih značajki. Npr. na uzorku blogova izbacivanje značajke broja interpunkcije i broja različitih riječi, u slučaju kad se uspoređuju samo autori s bipolarnim poremećajem i oni bez poremećaja, povećava točnost za nekoliko postotaka. No, i dalje su ovi rezultati vidljivo lošiji od onih dobivenih Bayesovim klasifikatorima. Postoji mogućnost da je SVM algoritmu zbog malog raspona rezultata i sličnosti primjera otežano naći dobre potporne vektore, pa su stoga rezultati lošiji.

Usporedbom rezultata za Bayesov klasifikator i Bayesov klasifikator s procjenom jezgrene funkcije uočavamo neke pojave očekivane s obzirom na njihov način rada. Naime, iz rezultata, kao i iz razdioba značajki, vidi se da korištenje procjene jezgrene funkcije ima više smisla u slučaju pisaca, budući da razdiobe broja funkcijskih riječi, interpunkcijskih znakova i postotak različitih riječi ne podliježu normalnoj razdiobi, što je osobito naglašeno između skupina pisaca s bipolarnim poremećajem i onih s depresijom. U skladu s objašnjenjem našeg odabira značajki, kod onih s bipolarnim poremećajem distribucija je pomaknuta prema različitijim riječima, dok je kod depresivnih obrnuto. Zanimljivo je da je to naglašeno kod profesionalnih pisaca, a ne kod autora blogova, iako bismo očekivali da profesionalni pisci naspram ljudi koji se ne bave pisanjem profesionalno imaju bogatiji vokabular pa će te razlike biti manje.

U analizi podataka nije primijećeno da bi klasifikatori (nakon standardizacije podataka) imali trend činiti pogreške favorizirajući jednu kategoriju. To potvrđuje i velika sličnost postotaka točnosti klasifikacije i F_1 mjera vidljivih u tablicama 1 i 2.

Neke dobivene rezultate možemo usporediti s onima srodnih radova. Strous et al. (2009) dobili su uspješnost klasifikacije u shizofrenu nasuprot neshizofrenoj skupini od 83,3%, Diederich et al. (2007) točno su klasificirali 77% u iste kategorije, dok je u ovom radu najbolji rezultat na takvim skupinama 81,5% (vidljiv u tablici 3). S obzirom na različitost korištenih značajki i drugačiji tip tekstova, nije moguće donijeti jednoznačne zaključke uspoređujući ove rezultate. Moguće je da bi uspješnost klasifikatora izvedenog u ovom radu bila bolja da su korišteni tekstovi autora u akutnoj fazi bolesti, kao što jesu u navedenim radovima. S druge strane, prilično je sigurno zaključiti da veći broj (dobro odabranih) značajki omogućuje i bolje rezultate.

Djelomični izostanak očekivanog uspjeha izrade automatskog klasifikatora mentalnog statusa autora tekstova u ovom radu u konačnici ipak ima optimističnu poruku: s obzirom da su tekstovi prikupljeni od autora koji imaju dijagnozu, no tijekom pisanja tekstova najvjerojatnije nisu

⁵www.cs.iastate.edu/~yasser/wlsvm

Tablica 1: Postotak točno klasificiranih primjera i F_1 mjera na skupu blogova korištenjem različitih algoritama.

Metoda	Točno klasificiranih	F_1
Naivni Bayesov klasifikator	50,0%	0,477
Naivni Bayesov klasifikator s procjenom jezgrene funkcije	50,0%	0,500
SVM	35,7%	0,319
Apriorno najvjerojatnija klasa	36,3%	0,357

Tablica 2: Postotak točno klasificiranih primjera i F_1 mjera na skupu pisaca korištenjem različitih algoritama.

Metoda	Točno klasificiranih	F_1
Naivni Bayesov klasifikator	52,2%	0,473
Naivni Bayesov klasifikator s procjenom jezgrene funkcije	56,5%	0,565
SVM	50,0%	0,486
Apriorno najvjerojatnija klasa	33,3%	0,326

Tablica 3: Postotak točno klasificiranih primjera na skupu blogova po parovima klasa korištenjem različitih algoritama.

Grupe	Metoda	Točno klasificiranih
Normalni, shizofreni	Bayes	81,5%
	Bayes s procjenom jezgrene funkcije	70,4%
	SVM	62,0%
	Apriorno najvjerojatnija klasa	59,0%
Bipolarni, shizofreni	Bayes	70,4%
	Bayes s procjenom jezgrene funkcije	63,0%
	SVM	59,0%
	Apriorno najvjerojatnija klasa	59,0%
Bipolarni, normalni	Bayes	66,7%
	Bayes s procjenom jezgrene funkcije	66,7%
	SVM	46,6%
	Apriorno najvjerojatnija klasa	50,0%

Tablica 4: Postotak točno klasificiranih primjera na skupu pisaca po parovima klasa korištenjem različitih algoritama.

Grupe	Metoda	Točno klasificiranih
Depresivni, normalni	Bayes	60,0%
	Bayes s procjenom jezgrene funkcije	70,0%
	SVM	67,4%
	Apriorno najvjerojatnija klasa	50,0%
Bipolarni, normalni	Bayes	64,5%
	Bayes s procjenom jezgrene funkcije	67,7%
	SVM	67,7%
	Apriorno najvjerojatnija klasa	50,0%
Bipolarni, depresivni	Bayes	66,7%
	Bayes s procjenom jezgrene funkcije	76,7%
	SVM	60,0%
	Apriorno najvjerojatnija klasa	50,0%

bili u akutnoj fazi bolesti, moguće je da utjecaj bolesti izvan akutne faze nije bio dovoljno jak da bi dovoljno ugrozio kognitivno funkcioniranje bolesnika toliko da bi se ono očitovalo na njihovom pisanju.

5. Zaključak

U svrhu boljeg razumijevanja psihičkih bolesti i mogućeg poboljšanja njihove dijagnostike u ovom smo radu pokušali izraditi klasifikator tekstova prema mentalnom statusu autora. Skupine tekstova sastojale su se od radova autora sa bipolarnim poremećajem, shizofrenijom i depresijom, te one bez tih psihičkih poremećaja. Korišteni su tekstovi poznatih pisaca i tekstovi preuzeti s blogova, a pokušali smo ih klasificirati na temelju značajki broja različitih riječi, broja zamjenica prvog lica, broja interpunkcijskih znakova, broja funkcijskih riječi i prosječne dužine riječi u tekstu.

Ovakav pristup problemu nije potpuno postigao očekivani uspjeh. Na skupu blogova najuspješnija je bila klasifikacija shizofrenih autora nasuprot onih bez poremećaja, a na skupu pisaca najbolja je klasifikacija pisaca s bipolarnim poremećajem nasuprot onih sa depresijom. Bayesov algoritam i Bayesov algoritam s procjenom jezgrene funkcije bili su uspješniji od algoritma SVM, koji načelno nije postizao rezultate bolje od one dobivene jednostavnom referentnom metodom. Iako su dobiveni rezultati skromni, daju naznaku da postoje trendovi u pozadini ovog problema koje valja dalje istražiti.

Naime, postoji mnogo prostora za poboljšanje primjenjenog metodološkog pristupa. U budućim istraživanjima bilo bi dobro koristiti tekstove za koje pouzdano znamo da su ih autori pisali u akutnoj fazi svoje bolesti, kada su kognitivne funkcije nile najviše pogođene, čime bi se povećala izglednost dobivanja uspješnijeg razlikovanja među tekstovima. Također, bilo bi dobro koristiti tekstove koji su pisani na istu temu (na tragu korištenja osobnih blogova), kako bi se smanjila mogućnost da semantički kontekst utječe na rezultate, ili pak da se lakše koriste neke potencijalno korisne semantičke značajke. Osim toga, moglo bi se kontrolirati vještinu pisanja autora tekstova koristeći autore koji nisu izrazito vješti u tome kao što su profesionalni pisci, te koristiti nezavisne procjene stručnjaka za određivanje kvalitete teksta. Moguće je da bi i veći uzorak tekstova donio bolje rezultate, no problematično je njegovo prikupljanje s obzirom na malu populaciju oboljelih. Također, mogle bi se koristiti još neke značajke, a čija je implementacija bila preopširna za okvire ovog rada, primjerice frekvencije izraza (engl. *term frequency*, *TF*), inverzne frekvencije dokumenta (engl. *inverse document frequency*) i njihove međudnose sa frekvencijom funkcijskih riječi.

Zahvale

Zahvaljujemo anonimnim recenzentima na konstruktivnim kritikama kojima je poboljšana prva verzija ovog rada.

Literatura

Thomas Bayes. 1763. An essay towards solving a problem in doctrine of chances. In Richard Price, editor, *Philosophical Transactions of the Royal Society of London*, pages 370–418, London.

Igor Belša, Tomislav Reicher, and Ivan Krišto. 2010. Automatizirano prepoznavanje autora teksta. Dokumentacija za projekt iz kolegija Strojno učenje: Fakultet elektrotehnike i računarstva: Zagreb.

Corinna Cortes and Vladimir Vapnik. 1995. Support-vector networks. *Machine Learning*, pages 273–297.

R.M. Coyotl-Morales, L. Villaseñor-Pineda, M. Montes y Gómez, P. Rosso, and L.T. del Lenguaje. 2006. Authorship attribution using word sequences. *Lecture Notes in Computer Science*, 4225:844.

Joachim Diederich, Aqeel Al-Ajmi, and Peter Yellowlees. 2007. Ex-ray: Data mining and mental health. *Applied Soft Computing*, 7(3):923–927.

Yasser EL-Manzalawy and Vasant Honavar, 2005. *WL-SVM: Integrating LibSVM into Weka Environment*.

Vera Folnegović-Šmalc, editor. 1996. *Dijagnostički i statistički priručnik za duševne poremećaje : DSM-IV : međunarodna verzija s MKB-10 šiframa*. Naklada Slap, Jastrebarsko. objavila Američka psihijatrijska udruga.

David Hand and Keming Yu. 2001. Idiot's bayes – not so stupid after all? *International Statistical Review*, pages 385–398.

George H. John and Pat Langley. 1995. Estimating continuous distributions in bayesian classifiers. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann Publishers.

Slobodan Ribarić. 2010. Predavanja iz kolegija raspoznavanje uzoraka: Fakultet elektrotehnike i računarstva: Zagreb.

Claudia M. Santosa, Connie M. Strong, Cecylia Nowakowska, Po W. Wang, Courtney M. Rennie, and Terence A. Ketter. 2007. Enhanced creativity in bipolar disorder patients: A controlled study. *Journal of Affective Disorders*, 100:31–39.

Rael Strous, Moshe Koppel, Jonathan Fine, Smadar Nachliel, Ginette Shaked, and Ari Zivotofsky. 2009. Automated characterization and identification of schizophrenia in writing. *Journal of Nervous and Mental Disease*, 197(8):585–588.

F.J. Tweedie and R.H. Baayen. 2006. How variable may a constant be? measures of lexical richness in perspective. *Computers and the Humanities*, 32(5):323–352.

World Health Organization WHO. 1992. Icd-10 : The icd-10 classification of mental and behavioural disorders : Clinical descriptions and diagnostic guidelines. <http://www.who.int/classifications/icd/en/>.