

Prepoznavanje emocija iz govora

Jovana Nedeljković IN32-2019, jovananedeljkovic30@gmail.com

I. UVOD

Izveštaj se bavi analizom podataka baze koja sadrži po nekoliko rečenica izgovorenih od strane različitih osoba na srpskom jeziku. Rečenice su izgovarane različitim emocijama i cilj projekta jeste odabir algoritma koji će uspešno klasifikovati rečenice u odgovarajuće emocije.

II. BAZA PODATAKA

Baza sadrži 1776 uzoraka, koji predstavljaju izgovorene rečenice od strane 6 govornika i 387 obeležja, od kojih su 3 kategorička (naziv signala, inicijali govornika i klase emocija). Klase obeležja sadrže naziv 5 osnovnih emocija i to je predstavljeno na način: L-ljutnja, N-neutralno, R-radost, S-strah i T-tuga. Ostala obeležja su numerička i predstavljaju karakteristike audio signala izdvojenih u frejmove dužine 25 ms sa korakom 10 ms. Postoji 16 deskriptora niskog nivoa izdvojenih po frejmovima i to su: srednja kvadratna energija signala, 12 mel-frekvencijskih kepstalnih koeficijenata, broj preseka sa nulom, verovatnoća zvučnosti i osnovna frekvencija. Za navedena obeležja izračunati su njihovi prvi izvodi i zatim je nad konturama obeležja primenjeno 12 statističkih funkcionala.

Baza je modifikovana tako, da je kategoričko obeležje koje predstavlja naziv signala uklonjeno, inicijali govornika su svedeni na numeričko obeležje prema polu, dok su oznake emocije izdvojene. Uklonjene su nedostajuće vrednosti, kao i 12 uzoraka sa najizrazitijim vrednostima *outlier*-a. Klase su svedene na isti broj uzoraka.

III. METODE

A. Multi-layer Perceptron klasifikator – MLP

MLP pripada klasi veštačkih neuronskih mreža. Sastoji se od više slojeva perceptrona. Ulazne karakteristike se prihvataju ulaznim slojem, prenose se preko nekoliko skrivenih slojeva i konačno se dobija klasifikacioni izlaz na izlaznom sloju [1]. Za treniranje koristi nadgledano učenje tehnikom propagacije unazad.

Smatra se dobrim za prepoznavanje emocija iz govora zato što je pogodan za skupove podataka sa složenim strukturama [1].

Unakrsnom validacijom sa 3 podskupa, utvrđeni su najpovoljniji parametri za MLP klasifikator i oni su korišćeni za treniranje. Model se sastoji od dva skrivena sloja od kojih jedan sadrži 128 neurona, a drugi 64. Vrednost tog parametra još je testirana nad modelom koji bi sadržao 3 sloja, svaki sa po 6 neurona i 4 sloja od po 32 neurona. Aktivaciona funkcija je *relu* (*Rectified Linear Unit*) koja je preferirana u odnosu na ostale zato što može

ubrzati konvergenciju procesa odlučivanja i unakrsnom validacijom je utvrđeno da, za konkretan problem, daje bolje rezultate od *logistics* i *tanh* funkcija. Za algoritam optimizacije odabran je *adam* (*Adaptive Moment Estimation*) koji pomaže u ubrzanju procesa treniranja neuronske mreže prilagođavajući stopu učenja veličini gradijenta i smanjujući verovatnoću zaglavljivanja u lokalnim minimumima. Moguće vrednosti još algoritma optimizacije još su bile: *lbfgs* i *sgd* algoritmi. Ostali parametri: *batch_size* – 50, *learning_rate* – *constant*, *max_iter* – 50, *random_state* – 42. Algoritam radi po principu ranog zaustavljanja.

B. Decision Tree Classifier - DT

Modeli zasnovani na DT-u dele skup podataka prema određenim graničnim vrednostima više puta i na taj način stvaraju podskupove koji se koriste za grananje stabla. Sastoje se od čvorova, listova i grana, gde su čvorovi obeležja koja formiraju pitanja na osnovu kojih se stablo grana, listovi su odluke, a same grane vrednosti obeležja. Nov uzorak se klasifikuje većinskim principom prisutnosti klase u čvoru u kom se našao [2].

Unakrsnom validacijom sa 3 podskupa su definisani parametri modela i to: maksimalna dubina do koje se stablo grana jeste 50 čvorova, dok se za procenu greške koristi entropija. Dodatno je testirano nad maksimalnom dubinom od 2, 5, 10 i 100 čvorova i procenjivanjem greške Ginijevim indeksom diverziteta.

C. Multinomial logistics regression - MLR

Algoritam predstavlja uopštenje logističke regresije koja se koristi za klasifikaciju korišćenjem logističke (sigmoidne) funkcije i logit analize [3].

Unimodalna logistička regresija opisuje vezu između jedne zavisne (varijable odziva) i jedne nezavisne varijable (prediktora), dok je zavisna varijabla dihotomna (poprima vrijednosti 0 i 1). Problem se svodi na proučavanje logit funkcije, čija je transformacija linearna i čije je parametre potrebno odrediti metodom maksimalne verodostojnosti. Kod multivarijantnog regresionog modela postoji više nezavisnih varijabli, odnosno prediktora [4].

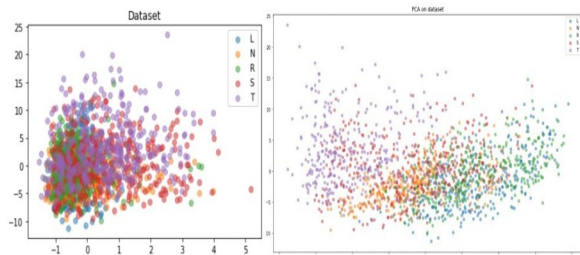
Unakrsnom validacijom sa 3 podskupa, odabrani su sledeći parametri: za maksimalan broj iteracija – 1000 (kao moguć broj iteracija testirano je: 100, 200, 500 i 1000 iteracija) i kao *solver* (algoritam za rešavanje optimizacionog problema) - *newton-cg* metod (kao mogući algoritmi testirani su: *newton-cg*, *lbfgs*, *sag* i *saga* metode).

D. Redukcija dimenzionalnosti - PCA

Odabir relevantnih karakteristika i uklanjanje irelevantnih je važan problem mašinskog učenja, odnosno važan korak u izgradnji prediktivnih modela. Stoga je potrebno primeniti postupak uklanjanja ovakvih karakteristika kako bi se povećala efikasnost, poboljšala preciznost i sprečilo

natprilagođavanje (*overfitting*) [5].

U projektu je korišćen algoritam selekcije obeležja, korišćenjem tehnike nadgledanog učenja – analize glavnih komponenti (PCA), odnosno akcenat je na tome da se uklone redundantne informacije, a da se uzorak predstavi podjednako tačno. Pre primene PCA metode izvršena je standardizacija obeležja i za broj komponentata je postavljena vrednost takva da one obuhvataju 82% ukupne varijanse. Na slici 1, prikazan je skup za obuku pre i nakon upotrebe PCA.



Slika 1: Redukcija dimenzionalnosti - PCA

IV. REZULTATI

A. Rezultati korišćenjem svih obeležja

Na slici 2 prikazani su rezultati dobijeni primenom pomenutih algoritama nad podacima.

Model	Klasa	Tačnost	Osetljivost
MLP	L	0.7011	0.2777
	N	0.7126	0.5263
	R	0.7816	0.3571
	S	0.7586	0.3125
	T	0.7356	0.2500
	Prosečno	0.7379	0.3447
DT	L	0.7126	0.2222
	N	0.7816	0.5263
	R	0.7931	0.4285
	S	0.8390	0.6250
	T	0.8620	0.6500
	Prosečno	0.7977	0.4904
MLR	L	0.6436	0.1666
	N	0.6666	0.2631
	R	0.8045	0.5000
	S	0.6781	0.0000
	T	0.7356	0.4000
	Prosečno	0.7057	0.2659

Slika 2: Rezultati dobijeni korišćenjem svih obeležja iz baze

B. Analiza

Na osnovu rezultata, može se videti da je model stabla odluke dao najbolje rezultate u pogledu tačnosti i osetljivosti. Prosečna tačnost za DT je najviša, što znači da

je ovaj algoritam najprecizniji u klasifikovanju instanci u skupu podataka. Prosečna osetljivost za DT je takođe visoka, što ukazuje na to da ovaj algoritam ima sposobnost prepoznavanja većeg broja pozitivnih instanci u skupu podataka u odnosu na MLP i MLR.

Takođe, može se primetiti značajna razlika kod sva tri algoritma između ova dva parametra, gde je tačnost veća od preciznosti. DT algoritam doseže najveći stepen osetljivosti za T – tugu. Razlog tome mogu biti varijacije u govoru ili izraženije vrednosti obeležja ove klase koje model lakše prepoznaje.

C. Rezultati nakon redukcije dimenzionalnosti

Na slici 3 prikazani su rezultati dobijeni primenom pomenutih algoritama nad redukovanim prostorom.

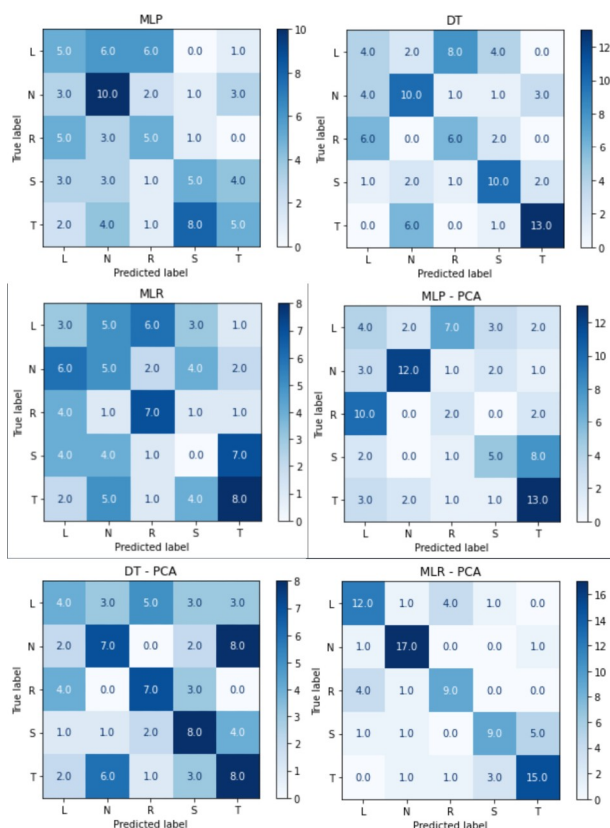
Model	Klasa	Tačnost	Osetljivost
MLP - PCA	L	0.6321	0.2222
	N	0.8735	0.6315
	R	0.7471	0.1428
	S	0.8045	0.3125
	T	0.7701	0.6500
	Prosečno	0.7655	0.3918
DT - PCA	L	0.7356	0.2222
	N	0.7471	0.3684
	R	0.8275	0.5000
	S	0.7816	0.5000
	T	0.6896	0.4000
	Prosečno	0.7563	0.3981
MLR - PCA	L	0.8620	0.6666
	N	0.9310	0.8947
	R	0.8850	0.6428
	S	0.8735	0.5625
	T	0.8735	0.7500
	Prosečno	0.8850	0.7033

Slika 3: Rezultati dobijeni nad redukovanim prostorom

D. Analiza

Nakon redukcije dimenzionalnosti, MLP i MLR klasifikator daju bolje rezultate i po meri tačnosti i po meri osetljivosti, dok algoritam DT ima pad performansi i takođe značajnu razliku između ove dve mere.

E. Uporedni prikaz matricama konfuzije



Slika 4: Matrice konfuzije

V. ZAKLJUČAK

Pre upotrebe metode PCA, metod stabla odluke je imao najbolje performanse u pogledu tačnosti i osjetljivosti i imao je najmanju razliku između ta dva parametra. Nakon redukcije dimenzionalnosti, MLP algoritam je popravio performanse, kao i MLR, dok su kod DT algoritma one opale. Uzrok tome može biti to što manji set podataka pogoduje performansama MLP klasifikatora i modelu logističke regresije koji su osjetljivi na ulaze velikih dimenzija, dok je DT algoritam robustan na besmislena obeležja, pa redukcija dimenzionalnosti može u tom slučaju predstavljati gubitak relevantnijih informacija. Takođe, razlog ostvarivanja najboljih performansi kod modela logističke regresije nakon redukcije dimenzionalnosti može ukazivati na linearnu separabilnost između klasa.

VI. LITERATURA

- [1] Singh, A., Kumar, A., & Kumar, V. (2021). *Speech Emotion Recognition using Ensemble Learning Approach*. *EAI Endorsed Transactions on Scalable Information Systems*, 8(32), e5. [2021.2314726](https://doi.org/10.18187/pjsor.v8i2.234)
- [2] Erdelić, M. (2022). *Klasifikacija modova prijevoza višerazinskom metodom zasnovanom na matricama prijelaznih stanja i stablima odluke u stvarnom vremenu*.
- [3] El-Habil, A. M. (2012) *An Application on*

Multinomial Logistic Regression Model. Pakistan Journal of Statistics and Operation research, Vol.8 No 2
<https://doi.org/10.18187/pjsor.v8i2.234>

[4] Šarlija, A. (2022) *Logistička regresija u predikciji odljeva korisnika mobilne igre*.

[5] Janković, O., Babić, Đ. (2019) *Inženjering karakteristika: Implikacije selekcije karakteristika na performanse predikcije*.