

Detekcija Parkinsonove bolesti na osnovu glasa

Milica Tomić

Biomedicinsko inženjerstvo

Univerzitet u Novom Sadu, Fakultet tehničkih nauka

Novi Sad, Srbija

tomic.e29.2024@uns.ac.rs

Sažetak—U ovom radu rešava se klasifikacioni problem pri čemu je potrebno tačno klasifikovati ispitanike kojima je dijagnostikovana Parkinsonova bolest na osnovu audio snimaka. Upoređivane su različite vrednosti hiperparametara, a evaluacija je izvršena unakrsnom validacijom sa 5 podskupova, za Mašine na bazi vektora nosača (SVM) i AdaBoost algoritam, kao i tri modela različitih kompleksnosti za Duboke neuronske mreže. Kao mera uspešnosti modela uzeta je osetljivost. SVM i AdaBoost sa određenim kombinacijama vrednosti hiperparametra pokazali su se kao najefikasniji modeli sa osetljivošću od 100.0 %, dok je za jedan od modela Dubokih neuronskih mreža postignuta osetljivost od 75.0 %.

Index Terms—Parkinsonova bolest, klasifikacioni problem, SVM, AdaBoost, Duboke neuronske mreže.

I. UVOD

Parkinsonova bolest je drugi najčešći neurodegenerativni poremećaj, koji pogađa približno 1 od 100 ljudi starijih od 65 godina u Evropi i procenjenih 7–10 miliona ljudi širom sveta. Predstavlja progresivni poremećaj nervnog sistema koji pre svega utiče na pokrete, izazivajući tremor, ukočenost, usporenost pokreta i probleme sa ravnotežom. Rani znaci su često suptilni, poput blagog drhtanja, smanjene izražajnosti lica ili blage promene govora, a simptomi se pogoršavaju kako bolest napreduje [1].

Rano otkrivanje Parkinsonove bolesti je ključno, ali su trenutne dijagnostičke procedure često zahtevne, subjektivne i nedovoljno pouzdane, naročito kod blagih slučajeva [2]. Nedavna istraživanja pokazuju da suptilne abnormalnosti u glasu i govoru postoje i u izuzetno ranim fazama bolesti koje pogađa do 90 % pacijenata, a mogu se uočiti čak 5 godina pre kliničke dijagnoze [3]. Zato, analiza govora može biti vredan i objektivan alat za ranu detekciju i praćenje ove bolesti uz primenu modela mašinskog učenja [1].

Cilj ovog rada jeste da se ispituju različite metode za rešavanje klasifikacionog problema u mašinskom učenju i da se pronađe najefikasniji algoritam za detekciju Parkinsonove bolesti na osnovu audio zapisa.

II. OPIS BAZE

Baza [4] sadrži akustičke karakteristike izdvojene iz glasovnih snimaka kontinuirane /a/ fonacije, specifično dizajniranih za detekciju i istraživanje Parkinsonove bolesti. Sastoji se od 240 uzoraka i 48 obeležja, pri čemu uzorci predstavljaju snimke od 80 jedinstvenih ispitanika, a svaki ispitanik je ponovio snimanje tri puta.

A. Ispitanici

Od ukupnih 80 ispitanika, 40 je zdravo (22 muškaraca - 55 % i 18 žena - 45 %), dok je kod preostalih 40 dijagnostikovana Parkinsonova bolest (27 muškaraca - 67.5 % i 13 žena - 32.5 %) i svi ispitanici su stariji od 50 godina. Kod ispitanika kojima je dijagnostikovana Parkinsonova bolest prisutna su jedan od dva sledeća simptoma: tremor u mirovanju, bradikinezija (usporeni pokreti) ili rigidnost.

B. Eksperiment - snimanje govora

Vokalni zadatak je bio kontinuirana fonacija vokala /a/ na određenoj visini i jačini, što je konstantnije moguće. Ova fonacija je trebalo da se održi najmanje 5 sekundi i u jednom dahu. Zadatak je ponovljen tri puta po ispitaniku, i svi su smatrani replikacijama.

C. Obeležja

Od 48 obeležja, 44 su akustička obeležja i mogu se podeliti u 5 grupa:

Jitter_rel, *Jitter_abs*, *Jitter_RAP*, *Jitter_PPQ* - kvantifikuju nepravilnosti u fundamentalnoj frekvenciji.

Shim_loc, *Shim_dB*, *Shim_APQ3*, *Shim_APQ5*, *Shim_APQ11* - procenjuju varijacije intenziteta glasa.

HNR05, *HNR015*, *HNR025*, *HNR035*, *HNR038* - evaluiraju kvalitet glasa i disanje kroz različite frekventne opsege. *HNR* kvantifikuje količinu aditivnog šuma u glasovnim signalima i obrnuto je povezan sa dahom i promuklošću [1].

MFCC0-MFCC12 (Mel frekventni cepstralni koeficijenti) i *Delta0-Delta12* (Prvi izvod *MFCC* koeficijenata) - pružaju aproksimaciju ljudske auditivne percepcije i posebno su efikasni za hvatanje spektralnih karakteristika relevantnih za detekciju patološkog govora. *MFCC2*, posebno, može se interpretirati kao ponderisan odnos energije niskih prema visokim frekvencijama, što je povezano sa promenama glasa izazvanim bolešću [1].

RPDE, *DFA*, *PPE*, *GNE* - napredne akustičke karakteristike koje hvataju složenu dinamiku glasa.

Pored pomenutih obeležja, postoje još 4 obeležja koja nisu vezana za snimanje govora, a to su: *ID*, *Recording*, *Status*, *Gender*, a koja predstavljaju ID ispitanika, redni broj snimanja, prisustvo Parkinsonove bolesti ispitanika i pol ispitanika, respektivno.

III. OBRADA BAZE

Baza sadrži više merenja akustičkih karakteristika iz tri glasovna snimka po ispitaniku, što stvara izazov za standardne metode mašinskog učenja, zato što one pretpostavljaju da su svi uzorci nezavisni. Pre treniranja modela mašinskog učenja, potrebno je obezbediti nezavisnost podataka. Najjednostavnije rešenje je izračunavanje srednje vrednosti tri glasovna snimka za svakog ispitanika, čime se baza smanjuje sa 240 na 80 uzoraka i rešava problem zavisnosti.

Dodatno, pre treniranja, obeležja *ID* i *Recording* treba ukloniti zato što ne sadrže korisne informacije za obučavanje modela, a *Status* predstavlja ciljne izlaze modela za evaluaciju modela.

Nakon obrade baze, baza sadrži 80 uzoraka i 44 obeležja.

A. Standardizacija obeležja

U medicinskim podacima često postoje obeležja čije su vrednosti različitih opsega, što može prouzrokovati dominaciju obeležja sa znatno većim vrednostima u poređenju sa ostalim i zbog toga je potrebno standardizovati obeležja [5].

Standardizacija obeležja podrazumeva skaliranje vrednosti obeležja. Korisna je kod nekih modela mašinskog učenja kako bi se obeležja skalirala na sličan opseg vrednosti i time smanjio uticaj velikih vrednosti koje mogu dominirati, kao i kod metode opadanja gradijenta za ubrzavanje obuke [5].

IV. METODOLOGIJA

Ideja ovog rada jeste da se ispituju različiti algoritmi sa različitim vrednostima hiperparametara i odaberu one vrednosti hiperparametara sa kojima se dobijaju najbolji rezultati na osnovu zadate metrike.

Baza je podeljena na skup za trening i skup za test. Trening skup sadrži 90 % nasumičnih uzoraka, dok se 10 % nasumičnih uzoraka nalazi u test skupu.

GridSearchCV [6] funkcija u *Python* - u [7] se koristi za automatsko pretraživanje kroz sve moguće kombinacije hiperparametara koje su definisane u okviru funkcije i model evaluiira koristeći unakrsnu validaciju. Unakrsnom validacijom [5] se dostupni skup uzoraka za testiranje deli na 5 jednakih podskupova i u svakoj rundi u datom podskupu se izdvajaju uzorci za trening i validaciju. Parametri ove funkcije su:

- klasifikator koji se koristi,
- parametri i njihove vrednosti zadatog klasifikatora,
- metrika za procenu performansi modela,
- broj podskupova za unakrsnu validaciju.

Kao mera za procenu performanse modela [5] koristi se osetljivost koja predstavlja udeo ispravno klasifikovanih uzoraka iz klase bolesnih. Što je veća osetljivost, to znači da je veći broj bolesnih otkriveno, a mali broj nije otkriven. Takođe, veća je greška bolesnog ispitanika svrstati u klasu zdravih, nego zdravog ispitanika u klasu bolesnih, što predstavlja još jedan razlog za korišćenje osetljivosti.

U ovom radu korišćena su 3 algoritma za zadati klasifikacioni problem:

- Mašina na bazi vektora nosača (*SVM*)

- *Aadaptivno poboljšanje (AdaBoost)*
- *Duboke neuronske mreže*

A. SVM

SVM (engl. *Support vector machine*) je algoritam za nadgledano mašinsko učenje čija je osnovna ideja pronalaženje optimalne hiperravni koja maksimizuje marginu između različitih klasa. Margina se može definisati kao rastojanje između hiperravni i najbližih uzoraka iz svake klase, poznatih kao vektori nosači. Vektori nosači određuju poziciju i orijentaciju hiperravni [5].

Ovaj algoritam postiže dobre rezultate čak i za linearno neseeparabilne uzorke i uzorke sa nelinearnom granicom odlučivanja. Koncept meke margine se uvodi za linearno neseeparabilne uzorke pri čemu je ključan parametar *C* (regularizacioni parametar) koji definiše toleranciju prelaska uzorka na pogrešnu stranu margine (vektora nosača). Velike vrednosti parametra *C* mogu dovesti do natprilagođenja (engl. *overfitting*). Uvođenjem kernel funkcija, *SVM* omogućava efikasno rešavanje nelinearnih granica odlučivanja. Najpoznatije kernel funkcije su: polinomijalni i radijalni kernel [5].

Parametri i njihove vrednosti korišćene pri rešavanju klasifikacionog problema su:

- *C* sa vrednostima 0.01, 0.1, 1, 10, 100,
- *kernel* sa vrednostima *linear*, *rbf*, *poly*,
- *gamma* sa vrednostima *scale*, *auto*, 0.001, 0.01, 0.1, 1,
- *degree* sa vrednostima 2, 3, 4,

Pri čemu je *gamma* parametar za *rbf* (radijalni kernel), dok parametar *degree* određuje red polinoma i koristi se za *poly* (polinomijalni kernel).

B. AdaBoost

Adaptivno poboljšanje (engl. *AdaBoost*) predstavlja algoritam ansambl učenja koji kombinuje više slabih klasifikatora kako bi formirao jak klasifikator. Funkcioniše tako što se slabi klasifikatori, najčešće jednostavna stabla odlučivanja, treniraju sekvencijalno, pri čemu veću težinu dodeljuju pogrešno klasifikovanim uzorcima u svakoj iteraciji. Ovo adaptivno ponderisanje usmerava naredne klasifikatore na teže slučajeve i postepeno poboljšava tačnost modela. AdaBoost je jednostavan i često otporan na prekomerno prilagođavanje, što ga čini efikasnim u različitim klasifikacionim zadacima [8].

Parametri i njihove vrednosti korišćene pri rešavanju klasifikacionog problema su:

- *n_estimators* sa vrednostima 50, 100, 200,
- *learning_rate* sa vrednostima 0.3, 0.5, 1.0,

Pri čemu, *n_estimators* predstavlja maksimalni broj slabih klasifikatora, a *learning_rate* uticaj tih slabih klasifikatora na konačnu predikciju. Veliki broj slabih klasifikatora može dovesti do natprilagođenja. Takođe, postoji kompromis između *learning_rate* i *n_estimators*. Smanjenje *learning_rate* obično zahteva više slabih klasifikatora, dok veći *learning_rate* može ubrzati konvergenciju, ali povećava rizik od natprilagođenja.

C. Duboke neuronske mreže

Duboke neuronske mreže (engl. *Deep Neural Network - DNN*) su računarski modeli inspirisani ljudskim mozgom i sastoje se od slojeva međusobno povezanih neurona. Informacije se prenose kroz mrežu pomoću propagacije unapred (engl. *forward propagation*), pri čemu se ulazi transformišu u izlaze putem ponderisanih veza i aktivacionih funkcija. Učenje se ostvaruje propagacijom unazad (engl. *back propagation*), koja podešava težine minimizacijom greške između predviđenih i stvarnih izlaza korišćenjem metode gradijentnog spusta (engl. *gradient descent*). Ovaj mehanizam omogućava neuronskim mrežama da modeluju složene, nelinearne odnose, što ih čini veoma efikasnim za zadatke kao što su klasifikacija, regresija i prepoznavanje obrazaca [5].

U ovom radu, za implementaciju *DNN* korišćena je biblioteka *keras* [9] i ispitani su duboke neuronske mreže različitih složenosti:

- model_1 - sastoji se iz dva sloja, skriveni sloj koji ima 20 neurona sa ReLU aktivacionom funkcijom i izlaznim slojem sa jednim neuronom (zato što je u pitanju binarna klasifikacija) sa sigmoidalnom aktivacionom funkcijom.
- model_2 - sastoji se iz 3 sloja (32 → 16 → 1 neurona), skriveni slojevi sadrže ReLU aktivacionu funkciju i izlazni sloj je isti kao u prethodnom modelu.
- model_3 - sastoji se iz 4 sloja (64 → 32 → 16 → 1 neurona), skriveni slojevi sadrže ReLU aktivacionu funkciju i izlazni sloj je isti kao u prethodna dva modela.

Za trening proces korišćen je batch size od 8 uzoraka i 50 epoha sa 10 % odvojenih podataka za validacioni skup. Pored toga, korišćen je *Adam optimizer* [10] i *early stopping* [10]. Što se tiče regularizacionih slojeva, u modelu 2 i 3 korišćen je *dropout* [10].

Za evaluaciju modela, korišćena je *binary crossentropy* funkcija gubitka (engl. *loss function*) [5] koja se izračunavana na sledeći način:

$$L = -\frac{1}{N} \sum_{i=1}^N (y_i \log(p_i) + (1 - y_i) \log(1 - p_i))$$

V. REZULTATI

Nakon obuke modela, vrednosti hiperparametara koje su se pokazale kao najbolje su:

- *SVM* sa vrednostima hiperparametrima : *C: 0.01, degree: 2, gamma: scale, kernel: linear*
- *AdaBoost* sa vrednostima hiperparametrima : *learning_rate: 1.0, n_estimators: 50*

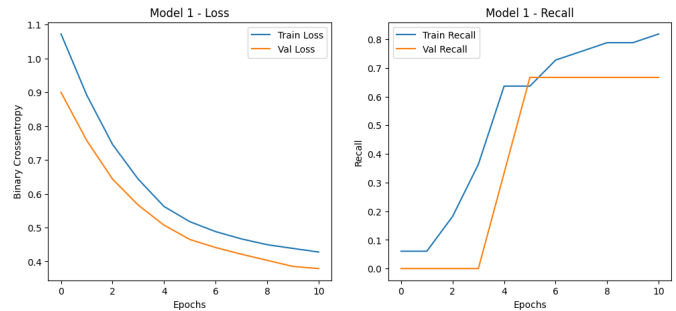
Što se tiče prvog modela *DNN - a*, može se videti na Slici 1. (levo) da *loss* za trening (*train loss*) i validacioni (*val loss*) skup konstantno opada sa povećanjem broja epoha, što ukazuje da *overfitting* nije prisutan. Takođe, može se videti da je *val loss* malo niži od *train loss* tokom epoha, što se može desiti ako su uzorci u trening skupu malo kompleksniji od uzoraka u validacionom skupu. Ovo je moguće zbog baze koja ima izuzetno malo uzoraka.

Na Slici 1. (desno), može se uočiti da se performanse modela kod trening skupa relativno konstantno povećavaju sa

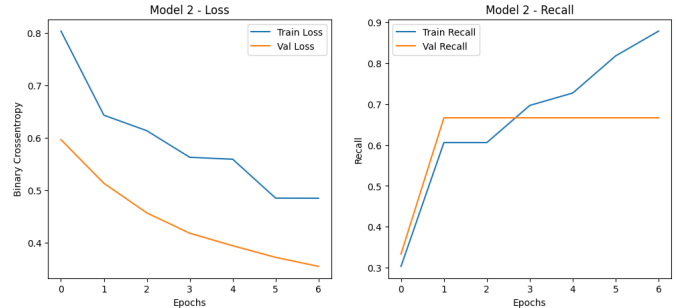
brojem epoha, međutim kod validacionog skupa vidimo da je postignut plato. Ovo ponašanje može se objasniti kao da se desio *underfitting* usled manje kompleksnosti ovog modela, odnosno model ne može da nađe dodatne, kompleksnije šablone koji mogu pripomoći u daljoj generalizaciji podataka.

Ponašanje drugog modela (Slika 2.) slično je ponašanju prvog modela, s tim što se može videti da se performanse modela brzo poboljšavaju i kod trening i kod validacionog skupa, što ukazuje da model brzo uči. Ovo je moguće zbog korišćenja kompleksnijeg modela i *dropout* regularizacije, zbog čega je ovaj model bolji od prvog modela.

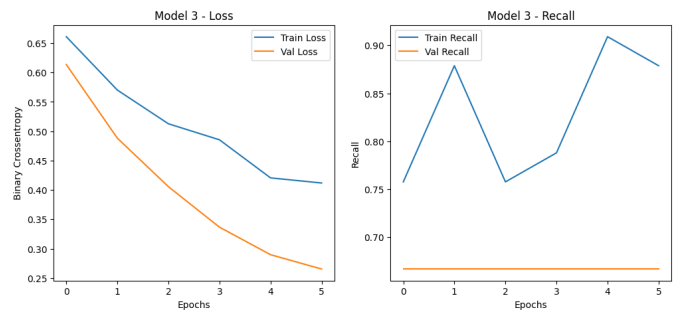
Na Slici 3. može se videti nestabilno ponašanje trećeg modela, što je znak *overfitting - a* ili prevelike varijacije u trening podacima. Model je kompleksan za količinu podataka kojom raspolaže i time se ovaj model pokazao kao najgori.



Slika 1: Ponašanje modela 1 tokom treniranja.



Slika 2: Ponašanje modela 2 tokom treniranja.

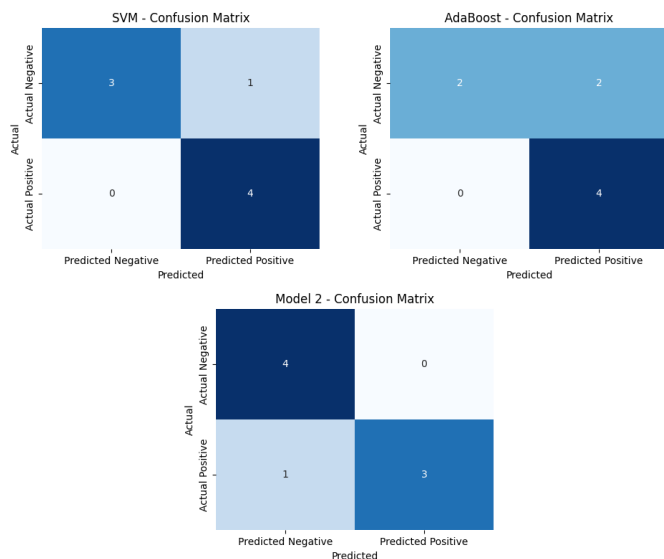


Slika 3: Ponašanje modela 3 tokom treniranja.

Nakon testiranja modela nad test skupom, može se uočiti (Slika 4.) da su *SVM* i *AdaBoost* postigli osetljivost od 100

% odnosno da modeli nisu pogrešno klasifikovali ni jednog ispitanika kojem je detektovana Parkinsonova bolest, što je i najbitnije kod ovakvih medicinskih modela.

Model 2 kod *DNN* - a, koji se pokazao kao najefikasniji model u poređenju sa modelima 1 i 3, postigao je malo niže rezultate, odnosno samo jedan ispitanik koji boluje od Parkinsonove bolesti nije detektovan.



Slika 4: Matrica konfuzije za *SVM*, *AdaBoost* i *DNN*.

Takođe, u Tabeli 1 može se videti i da su dobijeni relativno dobri rezultati i za ostale metrike za procenu performansi modela.

Tabela I: Mere uspešnosti za tri najefikasnija algoritma

	<i>SVM</i>	<i>AdaBoost</i>	<i>DNN</i>
tačnost	87.5 %	75.0 %	87.5 %
preciznost	80.0 %	66.7 %	100.0 %
osetljivost	100.0 %	100.0 %	75.0 %
f - mera	88.9 %	80.0 %	85.7 %

VI. UPOREĐIVANJE RADOVA

Ovaj rad i rad [1] bave se detekcijom Parkinsonove bolesti koristeći isti skup podataka, ali koriste različite metodologije. *Naranjo et al.* predlažu sofisticirani Bejesov statistički model, specifično dizajniran za obradu ponovljenih, zavisnih merenja. Ovaj pristup je rezultirao skromnijim performansama sa osetljivošću od 71.8 %.

U ovom radu, kao što je pomenuto, korišćeni su drugi algoritmi mašinskog učenja kao i drugačija obrada baze. Drugi rad (*Naranjo et al.*) tvrdi da se usrednjavanjem odbacuju ključne informacije o varijabilnosti unutar glasa svakog pojedinca i da su njegovi rezultati statistički robusniji i verovatnije je da će se generalizovati u realnim uslovima.

VII. ZAKLJUČAK

U ovom radu ispitivana su tri algoritma sa različitim vrednostima hiperparametara i kompleksnošću modela, pri čemu su dobijene osetljivosti od 100 % za *SVM* i *AdaBoost* i osetljivost od 75 % za *DNN*. *SVM* i *AdaBoost* su postigli savršene rezultate, pokazavši veliku sposobnost prepoznavanja obrazaca za ovaj specifičan set akustičkih karakteristika.

Jedan od razloga zašto *DNN* nije pokazao podjednako dobre rezultate kao *SVM* i *AdaBoost* može biti da pomenute akustičke karakteristike ne pružaju dovoljno informacija za ovaj vid dubokog učenja, pri čemu se performanse ovog modela mogu poboljšati proširenjem pomenute baze. Takođe, drugačija arhitektura neuronske mreže može poboljšati performanse modela, kao i dodavanje drugih regularizacionih parametara. Dodatno, *fine tuning DNN* -a može biti sledeća opcija za efikasniju detekciju Parkinsonove bolesti.

Takođe, uzimajući u obzir tvrdnje u radu [1], koji kaže da se usrednjavanjem odbacuju ključne informacije o varijabilnosti podataka, jedna od opcija za dalji rad može biti pronalaženje neke druge metode za sređivanje baze ili nekog pristupa koji će dovesti do robusnijih rezultata i bolje generalizacije u realnim medicinskim problemima.

LITERATURA

- [1] Naranjo, L., Pérez, C.J., Campos-Roca, Y., Martín, J. (2016). "Addressing voice recording replications for Parkinson's disease detection." *Expert Systems With Applications*, 46, 286-292.
- [2] Mayo Clinic Staff. (2024). "Parkinson's disease — Symptoms and causes." Mayo Clinic. Retrieved September 27, 2024, from <https://www.mayoclinic.org/diseases-conditions/parkinsons-disease/symptoms-causes/syc-20376055>.
- [3] Xu, H., Xie, W., Pang, M., Li, Y., Jin, L., Huang, F., & Shao, X. (2025). Non-invasive detection of Parkinson's disease based on speech analysis and interpretable machine learning. *Frontiers in Aging Neuroscience*, 17, Article 1586273. <https://doi.org/10.3389/fnagi.2025.1586273>
- [4] Souza, H. (2019). Parkinson Replicated Acoustic Features Dataset [Data set]. Kaggle. <https://www.kaggle.com/datasets/hugosouza/parkinson-replicated-acoustic-features-dataset>
- [5] Nosek T., Brkljac B., Despotovic D., Secujski M., Loncar-Turukalo T. (2020). "Praktikum iz mašinskog učenja" Univerzitet u Novom sadu, Fakultet tehničkih nauka.
- [6] Python, <https://www.python.org/>
- [7] Scikit-Learn, <https://scikit-learn.org/stable/>
- [8] Baladram, S. (2024, November 10). AdaBoost classifier, explained: A visual guide with code examples. *Towards Data Science* (Medium). Retrieved September 21, 2025, from <https://medium.com/data-science/adaboost-classifier-explained-a-visual-guide-with-code-examples-fc0f25326d7b>
- [9] Keras, <https://www.tensorflow.org/guide/keras>
- [10] ZEMIM. (2024, April 14). Deep Neural Network (DNN) Explained. *Medium*. Retrieved September 28, 2025, from <https://medium.com/@zemim/deep-neural-network-dnn-explained-0f7311a0e869>