

**KLASIFIKACIJA NEURODEGENERATIVNIH BOLESTI NA OSNOVU ANALIZE HODA
CLASSIFICATION OF NEURODEGENERATIVE DISEASES BASED ON GAIT ANALYSIS**Vuk Milosavljević, *Fakultet tehničkih nauka, Novi Sad***Oblast – BIOMEDICINSKO INŽENJERSTVO**

Kratak sadržaj – U ovom radu opisani su profili jednih od najpoznatijih neurodegenerativnih bolesti. Opisan je model mašinskog učenja koji bi mogao pomoći u dijagnostikovanju i praćenju progresije neurodegenerativnih bolesti. Posebno je istaknut značaj redukcije dimenzionalnosti u procesu formiranja modela mašinskog učenja. Ispitane su performanse nekoliko standardnih klasifikatora u cilju klasifikacije neurodegenerativnih bolesti na osnovu hoda, koji pokazuje potencijal da zameni redovne bolničke testove za praćenje progresije bolesti, koji su opterećenje za samog pacijenta i bolnicu.

Ključne reči: Neurodegenerativne bolesti, mašinsko učenje, klasifikacija

Abstract – This paper describes the profiles of some of the most well-known neurodegenerative diseases. A machine learning model is outlined that could aid in the diagnosis and monitoring of the progression of neurodegenerative diseases. The importance of reduction dimensionality in the process of forming the machine learning model is particularly emphasized. The performances of several standard classifiers are tested to classify neurodegenerative diseases based on gait, which shows the potential to replace regular hospital tests for monitoring disease progression, which can be burdensome for both patients and hospitals.

Keywords: Neurodegenerative disease, machine learning, classification

1. UVOD

Neurodegenerativne bolesti predstavljaju jedan od najvećih problema u savremenoj medicini. Predstavljaju teme velikog broja savremenih istraživanja jer i dalje većina njih drži nivo misterioznosti oko svog nastanka, a većina njih je progresivne prirode i neizlečivo. Potrebno je upoznati se sa karakteristikama neurodegenerativnih bolesti, kako bi se uvidelo gde leži potencijal primene modela mašinskog učenja. Kada se uspostave karakteristike problema, potrebno je upoznati se sa koracima u formiranju modela mašinskog učenja koji bi pomogao u njegovom rešenju. Formirane modele je potrebno analizirati i doneti zaključke o njihovim performansama. Nakon analize više modela, odabere se najbolji model, i donese zaključak o celokopnom postupku.

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bio dr Nikola Jorgovanović, red. prof.

2. NEURODEGENERATIVNE BOLESTI

Neurodegenerativne bolesti su skup oboljenja koji se karakteriše progresivnim gubitkom funkcija nervnog sistema [1]. Ove bolesti uzrokuju postepeno oštećenje nervnih ćelija (neurona), što dovodi do gubitka funkcija i različitih neuroloških simptoma. Kako je ova vrsta oboljenja progresivna, simptomi vremenom postaju sve izraženiji i predstavljaju najveći problem pacijentima.

Uzroci neurodegenerativnih bolesti nisu u potpunosti razjašnjeni, ali se smatra da najveći uticaj imaju genetski faktori, faktori spoljašnje sredine i starost kao jedan od najznačajnijih faktora. Većina neurodegenerativnih bolesti karakteriše nakupljanje specifičnih abnormalnih proteina u mozgu. Dijagnoza neurodegenerativnih bolesti često uključuje klinički pregled, neuropsihološke testove, genetsko testiranje itd. Većina neurodegenerativnih bolesti nema potpuno efikasno lečenje, ali postoje terapije usmerene na ublažavanje simptoma. Velikom broju današnjih istraživanja tema je upravo vezana za nove terapijske strategije kako bi se postiglo što veće ublažavanje simptoma.

S obzirom na veliku raznolikost neurodegenerativnih bolesti, važno je individualno pristupiti svakom pacijentu i pružiti adekvatnu negu i podršku. U bazi podataka obuhvaćenoj u ovom radu nalaze se Parkinsonova bolest, Huntingtonova bolest i Amiotrofična lateralna skleroza.

2.1 Parkinsonova bolest

Parkinsonova bolest je progresivno neurodegenerativno stanje uzrokovano gubitkom dopaminergičkih neurona u bazalnim ganglijama. Region u bazalnim ganglijama koji je pogođen gubitkom dopaminergičkih neurona je *substantia nigra*. Neuron u ovom regionu proizvode neu-rottransmiter koji se naziva *dopamin*. Parkinsonova bolest je neizlečiva bolest, ali nije sama po sebi i fatalna. Oboleli umiru od komplikacija i pridruženih bolesti. Najčešći uzrok smrti jeste pneumonija.

Opasnost od povreda usled simptoma Parkinsonove bolesti takođe doprinosi opasnosti ove bolesti, jer lomovi u starijem dobu mogu biti fatalni. Tretman Parkinsonove bolesti ima za cilj ublažavanje simptoma koje bolest prouzrokuje. Tretman Parkinsonove bolesti uključuje rehabilitaciju, primenu odgovarajućih lekova i u nekim slučajevima hiruršku intervenciju.

2.2 Huntingtonova bolest

Huntingtonova bolest, je nasledna neurodegenerativna bolest koja dovodi do progresivnog gubitka kontroliranih

pokreta, poremećaja ponašanja i kognitivnih funkcija. Uzrokovana je mutacijom gena poznatog kao *HTT gen*. Ova genetska mutacija nasleđuje se autozomno dominantno, što znači da samo jedan primerak mutiranog gena (od roditelja) može dovesti do pojave bolesti. Ako jedan roditelj ima mutirani gen, postoji 50% šanse da će ga preneti na svoje potomke. Simptomi Huntingtonove bolesti uključuju nevoljne pokrete, kao što su trzaji i grčevi mišića koji postaju sve izraženiji kako bolest napreduje.

Osim motoričkih simptoma, pacijenti često iskuse promene u ponašanju i raspoloženju, uključujući depresiju, agresivnost i impulsivnost. Prilikom dijagnostikovanja bolesti najčešće se stavlja genetski test koji može identifikovati prisutnost mutacije *HTT* gena. Nema načina da se bolest potpuno zaustavi ili izleči.

2.3 Amiotrofična lateralna skleroza

Amitrofična lateralna skleroza (ALS) je neurološka bolest koja utiče na centralni nervni sistem i izaziva slabljenje funkcije mišića čime remeti fizičku funkcionalnost. ALS uzrokuje degeneraciju i smrt motoričkih neurona u centralnom nervnom sistemu. Kako motorički neuroni prenose signal iz mozga i kičmene moždine do mišića, njihovim uništavanjem dolazi do gubitka sposobnosti kontrole mišića.

Dijagnoza ALS-a zahteva temeljno neurološko ispitivanje. Ne postoji lek za potpuno lečenje bolesti, već se takođe koriste terapije koje mogu pomoći u ublažavanju simptoma i poboljšanju kvaliteta života.

3. KLASIFIKACIJA NEURODEGENERATIVNIH BOLESTI NA OSNOVU HODA

Analiza različitih parametara hoda može odigrati značajnu ulogu u dijagnostici i klasifikaciji neurodegenerativnih bolesti [2]. Hod je složena funkcija koja uključuje niz neuro-loških, mišićnih i biomehaničkih procesa. Promene u obrascu hoda, kao što su promene u brzini, ritmu, stabilnosti, često su prisutne u različitim neurodegenerativnim bolestima.

Različite neurodegenerativne bolesti imaju različite karakteristične obrasce hoda. Na primer, pacijenti sa Parkinsonovom bolesti često imaju smanjen ritam koraka, poteškoće sa inicijacijom koraka i karakterističan nagnuti položaj tela. Prema tome analiza hoda može pomoći u razlikovanju Parkinsonove bolesti od drugih poremećaja koji imaju slične simptome.

Promene u hodu mogu da se jave u ranoj fazi neurodegenerativne bolesti pre nego što se pojave drugi očigledni simptomi. Prema tome analiza hoda mogla bi omogućiti rano otkrivanje bolesti i pružiti mogućnost za intervenciju pre nego što se simptomi pogoršaju.

Promene u hodu, kao što su smanjenje brzine ili povećanje fluktuacija iz koraka u korak, mogu ukazivati na napredovanje bolesti. U zavisnosti od pozitivnih ili negativnih parametara eksperimenta koji uključuje hodanje, može se zaključiti pozitivan ili negativan odgovor na terapiju.

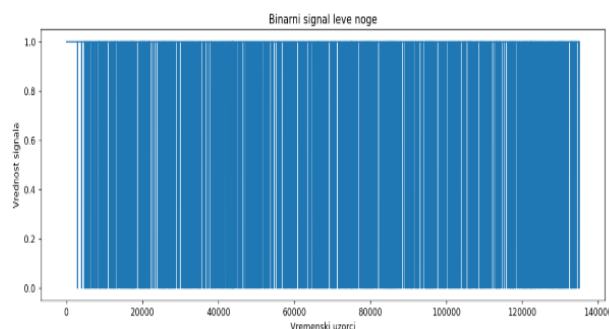
3.1 Baza podataka

Baza podataka na kojoj se zasniva analiza hoda, prikupljena je u bolnici *Massachusetts General Hospital* od strane naučnika predvođenih *Jeffrey Hausdorff-om*. Ispitanici su u toku eksperimenta hodali samostalno 5 minuta, pri čemu nisu koristili nikakva pomagala za kretanje koja bi mogla uticati na hod. Baza podataka sadrži podatke o hodu pacijenata sa oboljenjem Parkinsonove bolesti, Huntingtonove bolesti i Amiotrofične lateralne skleroze. Baza podataka sadrži uzorke 15 pacijenata sa Parkinsonovom bolesti, 20 pacijenata sa Huntingtonovom bolesti i 13 pacijenata sa ALS-om. Uključeni su i podaci iz 16 uzoraka zdravih kontrolnih ispitanika.

Baza podataka je formirana kroz eksperiment u kojem je svakom od učesnika snimljen hod. Eksperiment se sastojao iz šetanja uz hodnik dužine 77m, tokom kojeg se snimala sila koju levo i desno stopalo pacijenta vrše na pod i na osnovu te sile, za svaki korak, određene su karakteristike koraka vremenski i procentualno.

Osim toga baza podataka sadrži demografske i kliničke podatke o svakom pacijentu. Podaci su dobijeni korišćenjem otpornika osetljivih na silu, pri čemu je izlaz proporcionalan sili pod stopalom.

Za obe noge dobijaju se binarni signali kao na slici 1. Na osnovu ovih signala, za svakog subjekta su određene karakteristične vrednosti za svaki korak.



Slika 1. Binarni signal dobijen sa leve noge

3.2 Predobrada podataka

Obeležja koja su poslata klasifikatorima, formirana su tako što su uzeta srednja, minimalna, maksimalna vrednost i varijansa, svake od 5 karakterističnih vrednosti u toku čitavog testa za svakog subjekta, u procentima [3].

Time je formiran skup od 20 numeričkih obeležja. Jedan pacijent sa Huntingtonovom bolesti izbačen je iz razmatranja zbog nedostajućih podataka o određenim parametrima.

Prema tome u procesu formiranja modela učestvovali su rezultati 63 ispitanika koji će biti podeljeni na skup za trening i test. Za svaki od klasifikatora proveren je uticaj normalizacije. Prilikom pripreme podataka za binarnu klasifikaciju je urađen *oversampling*.

3.3 Redukcija dimenzionalnosti

Intuicija govori da za određeni skup uzoraka je bolje imati veći broj obeležja. Međutim ono što se dešava nakon prekoračenja nekog optimalnog broja obeležja je zapravo suprotno. Ako se premaši optimalan broj obeležja i povećava dimenzionalnost, dobija se retko popunjen uzorački prostor. Skup uzoraka tada prestaje da bude reprezentativan i tačnost klasifikatora na nekim novim podacima će opadati. Opisani problem je poznat kao prokletstvo dimenzionalnosti. Ovaj problem opisuje problem koji se javlja kada se radi sa podacima visoke dimenzionalnosti, odnosno kada postoji veliki broj atributa u odnosu na broj uzoraka, kao što je slučaj u ovom radu.

Sa povećanjem broja dimenzija, potrebno je mnogo više podataka da bi se dobio pouzdan model. U visokodimenzionalnom prostoru, potrebno je znatno više uzoraka kako bi se popunio prostor podacima i izbegla natprikladnost.

Sa više dimenzija, takođe dolazi i povećana računrska složenost. Računanje udaljenosti, konstruisanje modela i obrada podataka postaju izazovnije i zahtevaju više resursa. Problem se najčešće rešava nekim od postupaka redukcije dimenzionalnosti. Time se uklanjaju redundantna i nebitna obeležja. Smanjuje se i računrska složenost postupka, kao i zauzeće memorijskog prostora. Time se otvara i mogućnost upotrebe algoritama koji nisu pogodni za slučajeve kada je broj dimenzija veliki.

4. REZULTATI ANALIZE

Kako bi se uporedili različiti modeli, koristi se validacija. Validacija omogućava eksperimentisanje sa različitim vrednostima hiperparametara na osnovu čega se mogu odabrati najbolji parametri. U ovom radu korišćen je uobičajeni postupak validacije koji uključuje podelu podataka na tri skupa: skup za trening, validacioni skup i skup za test.

Korišćena su dva slučaja unakrsne validacije za validaciju modela.

Prvi slučaj je unakrsna validacija sa 10 podskupova (*cv-10*). On uključuje podelu skupa podataka na 10 jednakih podskupova, nakon čega se model trenira i testira 10 puta, svaki put koristeći drugi podskup za testiranje, dok se preostalih 9 podskupova koristi za obuku. Na kraju, uzima

se prosečna vrednost rezultata kako bi se dobila konačna procena performansi modela.

Drugi slučaj predstavlja specijalan slučaj unakrsne validacije poznat kao *Leave-One-Out (LOO)*. Kod LOO svaki pojedinačni uzorak koristi se kao uzorak za test, dok se svi preostali uzorci koriste za obuku. Drugi rečima, za svaki od N uzoraka, $N-1$ uzorak se koristi za obuku, a 1 uzorak za testiranje.

Ovaj proces se ponavlja N puta (za svaki uzorak) i uzima prosečan rezultat.

LOO može biti koristan u slučaju kada se radi sa malim skupom podataka, kao što je slučaj sa bazom podataka u ovom radu.

4.1 Rezultati predikcije

Klasifikacija predstavlja problem nadgledanog učenja čiji je zadatak da neki objekat pridruži određenoj klasi. U slučaju ovog rada razmatrane su dve vrste klasifikacije. U prvom postavku problema klasifikacije razmatrano kojoj klasi neurodegenerativnih bolesti ispitanik pripada ili da li je ispitanik iz kontrolne grupe.

Ispitane su performanse nekoliko standardnih klasifikatora, prilikom čega je pri obučavanju svakog klasifikatora, istrenirano više modela sa različitim parametrima, i uticajem različite validacije, normalizacije i redukcije dimenzionalnosti. Sumirani rezultati višeklasne klasifikacije dati su u tabeli 1.

Performanse klasifikatora za problem višeklasne klasifikacije su ujednačene. Pokazuje se da je potrebno izvršiti redukciju dimenzionalnosti sa jednom od metoda (u ovom radu su korišćene PCA i LDA), normalizovati podatke i validirati ih.

Drugi problem koji je posmatran u ovom radu jeste problem binarne klasifikacije u kojem je potrebno kontrolnog ispitanika izdvojiti od pacijenta koji poseduju neku vrstu neurodegenerativne bolesti. Performanse klasifikatora su date u tabeli 2.

Izdvojili su se modeli klasifikatora logističke regresije, k -najbližih suseda i neuronske mreže, kao klasifikatori sa najboljim performansama za problem binarne klasifikacije u ovom slučaju, uz primenu normalizacije podataka, validacije i redukcije dimenzionalnosti.

Model	Tačnost	Osetljivost	Preciznost	F mera	Normalizacija	Redukcija	Validacija
SVM	70%	70%	77%	0.72	Da	PCA	cv-10
LR	70%	70%	73%	0.70	Da	LDA	cv-10
KNN	70%	70%	73%	0.70	Da	LDA	LOO
SO	70%	70%	73%	0.70	Da	LDA	cv-10
NM	70%	70%	73%	0.70	Da	LDA	cv-10

Tabela 1. Rezultati performansi standardnih klasifikatora za višeklasnu klasifikaciju

Model	Tačnost	Osetljivost	Preciznost	F mera	Normalizacija	Redukcija	Validacija
SVM	80%	80%	80%	0.80	Da	LDA	LOO
LR	84%	84%	88%	0.83	Da	PCA	cv-10
KNN	84%	84%	88%	0.83	Da	PCA	cv-10
SO	79%	79%	85%	0.78	Da	PCA	cv-10
NM	84%	84%	88%	0.84	Da	PCA	cv-10

Tabela 2. Rezultati binarne klasifikacije

5.ZAKLJUČAK

Neurodegenerativne bolesti same po sebi nisu fatalne bolesti. Pacijenti najviše problema muče sa simptomima koji mogu biti motorički i nemotorički. Cilj je dijagnostikovati bolest u ranoj fazi, kako bi se odgovarajućom terapijom znatno produžio životni vek pacijenta i poboljšao kvalitet života. Potencijal rane dijagnostike u slučaju ispitanika, ili praćenja progresije bolesti u slučaju pacijenata, leži u razvijanju brzih i jednostavnih testova koji ne bi zahtevali seriju testova u bolnici.

Analiza hoda nameće se kao jedno od potencijalnih rešenja. Analizirati parametre hoda pomoću određenih modela mašinskog učenja, ima potencijal da značajno pomogne u dijagnostici i praćenju progresije bolesti. Međutim, da bi modeli mašinskog učenja bili pouzdaniji, neophodno im je omogućiti velike količine podataka na osnovu kojih bi učili. Treba biti oprezan prilikom formiranja modela i biti upoznat sa pravilima koje je potrebno ispoštovati. Nakon formiranja modela, on ne bi služio za direktne odluke da li je pacijent oboleo od neke bolesti ili ne, nego kao savetnik lekaru, obraćajući pažnju na određene parametre, koje bi lekar mogao da protumači i time možda bar umanjiti pacijentu nekoliko kontrola u bolnici.

6.LITERATURA

- [1] Harvey Checkoway, Jessica I. Lundin, and Samir N. Kelada. Neurodegenerative diseases. Chapter 22
- [2] Rashad Hussain, Hira Zubair, Sarah Pursell. Neurodegenerative diseases: Regenerative Mechanisms and Novel Therapeutic Approches. DOI: <https://doi.org/10.3390/brainsci8090177>
- [3] Jeffrey M. Hausdorff, Apinya Lertratanakul, Merit E. Cudkowicz, Amie L. Peterson. Dynamyc markers of altered gait rhytm in amyotrophic lateral sclerosis.
- [4] Nosek Tijana, Brkljač Branko, Despotović Danica, Sečujski Milan, Lončar-Turukalo Tatjana - Praktikum iz mašinskog učenja.

Kratka biografija:



Vuk Milosavljević rođen je u Novom Sadu 2000. godine. Osnovne studije na Fakultetu tehničkih nauka na odseku biomedicinsko inženjerstvo završio je 2022.godine