

SVEUČILIŠTE U ZAGREBU  
**FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA**

DIPLOMSKI RAD br. 401

**Metoda za automatsku klasifikaciju EEG  
signala za detekciju PTSP-a**

Iva Harbaš

Zagreb, lipanj 2012.

# Sadržaj

|                                                    |    |
|----------------------------------------------------|----|
| 1. Uvod.....                                       | 3  |
| 2. Srodni radovi.....                              | 4  |
| 3. EEG.....                                        | 6  |
| 3.1. Izvor EEG aktivnosti .....                    | 6  |
| 3.2. Snimanje EEG-a.....                           | 7  |
| 3.3. Valni oblici.....                             | 9  |
| 3.4. Korištenje EEG-a.....                         | 12 |
| 3.5. PTSP i EEG.....                               | 12 |
| 3.6. Korišteni EEG signali.....                    | 15 |
| 4. Klasifikacija EEG signala .....                 | 16 |
| 4.1 ANOVA .....                                    | 16 |
| Implementacija.....                                | 18 |
| 4.2. Značajke .....                                | 20 |
| 4.2.1. PSD.....                                    | 21 |
| 4.2.2. DWT.....                                    | 24 |
| 4.2.3. <i>Higuchijeva</i> fraktalna dimenzija..... | 30 |
| 4.2.4. <i>Skewness</i> .....                       | 35 |
| 4.2.5. <i>Kurtosis</i> .....                       | 36 |
| 4.2.6. <i>Hjorthovi</i> parametri.....             | 38 |
| 4.3 SVM.....                                       | 41 |
| Implementacija.....                                | 49 |
| 5. Rezultati i implementacija .....                | 50 |
| 5.1. Implementacija.....                           | 54 |
| 6. Zaključak.....                                  | 59 |
| 7. Literatura.....                                 | 60 |
| 8. Sažetak .....                                   | 64 |

## 1. Uvod

Ljudski mozak je kompleksan sustav koji je jako dinamičan i čija aktivnost se može mjeriti na različite načine. Onaj koji je ovdje interesantan je EEG (elektroencefalogram) pomoću kojeg se mjere električni potencijali koje stvaraju neuroni i neuronske veze u mozgu. EEG signali se koriste u svrhu istraživanja različitih oboljenja među kojima je i PTSP (posttraumatski stresni poremećaj). Dijagnosticiranje PTSP-a pomoću EEG signala nije toliko razvijeno kao npr. dijagnosticiranje i detekcija epilepsije, ali se u zadnje vrijeme sve više pažnje posvećuje tom problemu. Proučavanje električne aktivnosti mozga putem EEG-a je jedan od važnijih alata u dijagnozi. Prilikom snimanja EEG-a ne dobije se jedan signal, nego broj snimljenih signala ovisi o broju kanala koji se snima (može biti i do 128 kanala) i kao takav, jedan EEG sadži veliki broj informacija koje ponekad nije moguće obraditi vizualno. Abnormalnosti u EEG-u kod nekih ozbiljnih poremećaja su ponekad presuptilne da bi se uočile koristeći konvencionalne tehnike i upravo zbog toga se posvećuje velika pažnja korištenju računala za rješavanje problema i razvoja automatskih sustava koji bi za zadatak imali prepoznavanje promjena u EEG-u, donošenja zaključaka te postavljanje dijagnoza.

Ovaj rad je istraživački s ciljem proučavanja postojećih sustava za klasifikaciju EEG signala te primjena znanja i vještina u razvijanju nove metode za konkretnu klasifikaciju EEG signala oboljelih od PTSP-a jer još uvijek nije poznato gdje se dešavaju i kakve su to promjene EEG signala pacijenata kojima je dijagnosticiran PTSP. To je upravo jedan od ciljeva ovog rada, pronaći određene značajke po kojima bi se mogla ustanoviti razlika između bolesnih i zdravih EEG-ova.

Pronalaženje tih značajki predstavlja najvažniji, a ujedno i prvi korak u procesu klasifikacije signala. U nastavku će biti opisane pronađene značajke, njihovo predprocesiranje i odabir najboljih koje će se koristiti za treniranje i testiranje odabranog klasifikatora.

## 2. Srodni radovi

Do EEG signala se relativno lako može doći, tj. za njihovo snimanje je potrebna oprema koja u današnje vrijeme nije ni skupa ni komplicirana. Upravo zbog toga se EEG signali koriste u raznim područjima obrade signala. Najčešće se koriste za detekciju epileptičnih napadaja (Paivinen, 2005. [1], Subasi, 2005. [8]), njihovo predviđanje (Mirowski, 2009. [2]) ili za detekciju aktivnosti i promjena signala specifičnih za tu bolest (Mohamed , 2006. [3]). Jedno od novijih područja primjene EEG signala su BCI (eng. *brain-computer interface*) sustavi. BCI sustavi su dizajnirani da prevode misli korisnika u neki kontrolni signal kojeg bi računalo razumjelo. BCI predstavlja novi komunikacijski kanal između čovjeka i računala. Postoje različite vrste BCI sustava od kojih su najčešći oni koji na osnovu promjena u EEG-u klasificiraju različite pokrete ruku i/ili prstiju (Štastny [4] [5] [6] [7]). Također postoje BCI sustavi kojima se vrši detekcija različitih emocija pacijenta (Ansari-Asl, 2007. [9]) ili čak klasifikacija mentalnih zadataka (npr. korisnik zamišlja da pokreće ruku i sl.) (Chiappa , 2005. [12], Anderson [18]).

EEG snimke se sve više koriste i u svrhu procjenjivanja podložnosti Alzheimerovoj bolesti (Wu [10], Latchoumane [11]), ali konkretno, radova na temu detekcije i klasifikacije PTSP-a pomoću EEG-a nema. Upravo je zadatak ovog rada istraživanje na tu temu, pronalazak značajki koje najbolje opisuju dvije klase signala (zdrave i bolesne) te primjena odgovarajućeg klasifikatora koji će na osnovu odabranih značajki moći provesti uspješnu klasifikaciju.

Jedna od češće korištenih značajki za opisivanje EEG signala je PSD (eng. *Power Spectral Density*) (Paivinen, 2005. [1], Wu [10], Adlakha, 2002. [13], Babiloni, 2000. [14], Kerkeni [15]) koji se u suštini zasniva na *Fourierovoj* transformaciji koja se koristi od početka automatske obrade EEG signala. *Fourierova* transformacija (FT) je pogodna za analizu EEG signala jer se on sastoji od nekoliko karakterističnih valnih oblika koji se razlikuju po frekvencijama, a upravo pomoću FT-a se signal razlaže na spektralne komponente. Pored PSD-a koriste se i koeficijenti diskretne valićne transformacije, tj. DWT-a (eng. *Discrete Wavelet Transform*) (Mohamed , 2006. [3],

Subasi, 2005. [8], Guler, 2005. [16], Prochazka [17] i Yong [19]) koji su postali, recimo, popularni krajem 1980-ih godina i od tad ih se sve više koristi. Za prethodne dvije značajke se može reći da su najčešće korištene, ali pored njih se koristi i veliki broj manje poznatih značajki koje su se pokazale korisnima kao što su: entropija (Wu [10]), koeficijenti auto-regresivne analize (AR koeficijenti) (Wu [10], Adlakha, 2002. [13] i Anderson [18]), *Hjortovi* parametri (Paivinen, 2005. [1], Ansari-Asl, 2007. [9], Obermaier, 2001. [24]), *Hilbertov* spektar (Deng , 2010. [20]), asimetričnost signala (eng. *skewness*) (Paivinen, 2005. [1]), spljoštenost signala (eng. *kurtosis*) (Paivinen, 2005. [1]), fraktalna dimenzija (Paivinen, 2005. [1], Ansari-Asl, 2007. [9]) i mnoge druge.

Kad se odaberu značajke potrebno je odabrati i klasifikator. Za klasifikaciju EEG signala najčešće se koriste klasifikatori poput neuronskih mreža (NM) (Subasi, 2005. [8], Latchoumane [11], Guler, 2005. [16]) ili konkretno samoorganizirajuće NM (Prochazka [17]), unaprijedne NM (eng. *Feedforward Neural Networks*) (Anderson [18]), te višeslojni perceptron (eng. *MultiLayer Perceptron Network*) (Mohamed , 2006. [3], Kerkeni [15]). Pored neuronskih mreža koriste se još HMM (eng. *Hidden Markov Models*) (Štastny [4] [5] [6] [7] i Obermaier, 2001. [24]), LDA (eng. *Linear Discriminant Analysis*) (Ansari-Asl, 2007. [9]) i SVM (eng. *Support Vector Machines*) (Adlakha, 2002. [13], Yong [19], Li, 2009. [21], Lee [22], Hill [23]).

Kao što je vidljivo iz priloženog postoji puno predloženih i realiziranih rješenja za klasifikaciju EEG signala u različite svrhe. Ovaj rad je nastao kao rezultat proučavanja svih spomenutih radova i metoda opisanih u njima te odabiranjem najboljih karakteristika koje odgovaraju ovoj temi. U nastavku će detaljno biti obrađene odabrane značajke, metode i dobiveni rezultati.

### 3. EEG

Elektroencefalografija (EEG) je proces snimanja električne aktivnosti duž skalpa. EEG mjeri promjene napona koje su posljedica protoka ionske struje među neuronima u mozgu. U kliničkom smislu, EEG se odnosi na snimanje sponatane moždane aktivnosti kroz neki vremenski period, najčešće od 20 do 40 minuta.

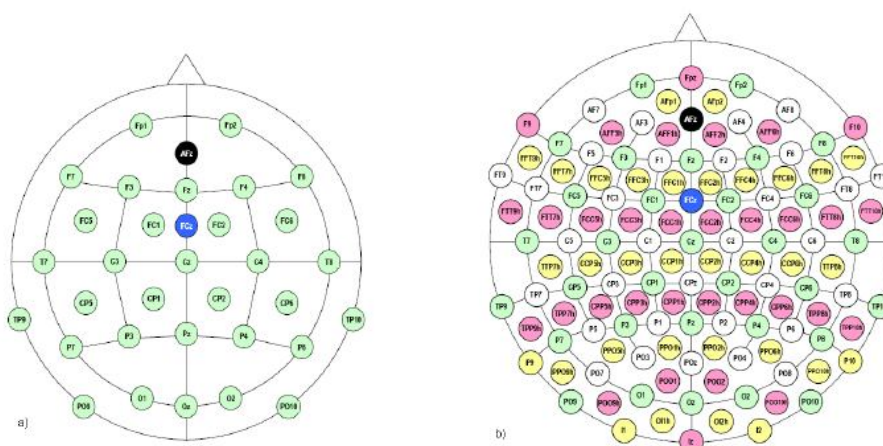
#### 3.1. Izvor EEG aktivnosti

Električni naboj mozga održavaju milijuni neurona. Neuron ili živčana stanica se smatra osnovnom jedinicom živčanog sustava i najsloženija je u ljudskom organizmu. Neuroni su električki nabijeni ili polarizirani. Svaki podražaj neurona dovodi do promjena na njihovoj membrani tako što se otvaraju pore za propust iona. Ako je stimulus dovoljno jak da prijeđe prag podražljivosti onda se ionski kanali potpuno otvore. Naglo pozitivno nabijeni  $\text{Na}^+$  ioni kroz pore ulaze u neuron, mijenja se električni naboj i neuron je depolariziran. Rezultat toga je da diferencijalni potencijal sa  $-70 \text{ mV}$  na trenutak prelazi u  $+40 \text{ mV}$ . U tom trenutku  $\text{K}^+$  ioni izlaze vani pa se potencijal uskoro vraća na  $-70 \text{ mV}$ . Ta pojava se naziva akcijski potencijal. Proces razmjene iona sa okolinom traje stalno, bilo da se održava potencijal mirovanja ili dolazi do akcijskog potencijala. Tim konstantnim kretanjem iona stvaraju se tzv. "valovi" iona koji kad dođu do elektroda na skalpu mogu odgurati ili privući elektrone metala koji se nalazi na elektrodama te dolazi do pojave napona. Napon između bilo koje dvije elektrode se može mjeriti voltmetrom i upravo snimanje tih napona određeno vrijeme kao rezultat daje EEG.

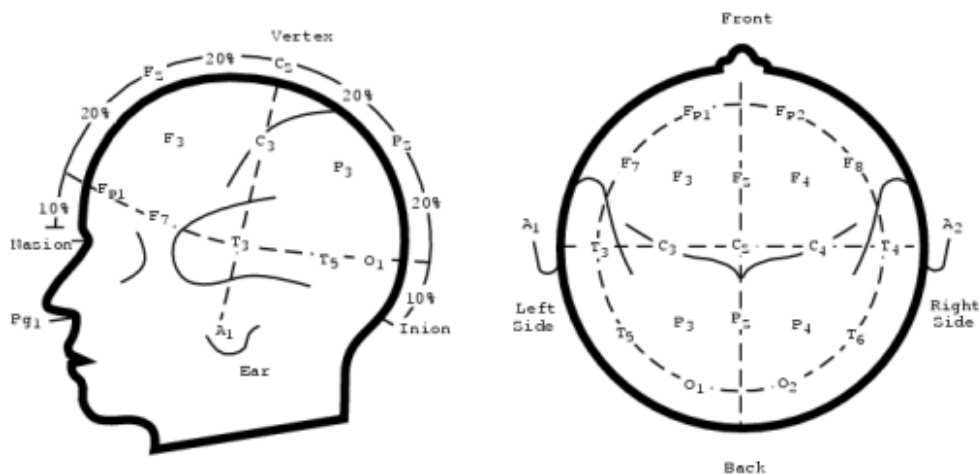
Električni potencijal kojeg stvara jedan neuron je previše malen da bi ga EEG snimio. EEG aktivnost je zapravo sumacija sinkrone aktivnosti tisuća ili milijuna neurona koji imaju sličnu orijentaciju u prostoru. Ukoliko neuroni nemaju sličnu orijentaciju u prostoru njihovi ioni se neće posložiti i kreirati valove koji se mogu detektirati. Najviše EEG signala proizvode dobro razmješteni neuroni koji djeluju sinkronizirano.

### 3.2. Snimanje EEG-a

EEG signali se snimaju preko elektroda koje se s provodljivim gelom ili pastom postavljaju na skalp. Broj elektroda koje se postavljaju na skalp nije fiksni, tj. EEG se može mjeriti pomoću manje ili više elektroda (slika 1.). Položaj elektroda na skalpu te njihovi nazivi su specificirani internacionalnim 10-20 sustavom. Raspored elektroda se može vidjeti na slici 2.



Slika 1.: Kapica sa rijede i sa gušće postavljenim elektrodama



Slika 2.: Raspored elektroda po sustavu 10-20

Svaka od elektroda je spojena na diferencijalno pojačalo (po jedno pojačalo za svaki par elektroda). Ta pojačala pojačavaju napon između aktivne i referentne elektrode. U analognim sustavima signal se filtrira i pomicanjem igle na papiru se dobije izlaz, tj. EEG signal.

Današnji EEG sustavi su većinom digitalni. Signal na izlazu iz diferencijalnog pojačala se digitalizira pomoću analogno-digitalnog pretvarača nakon što se signal propusti kroz *anti-aliasing* filter.

Prilikom snimanja EEG signala pacijent se podvrgava različitim procedurama koje mogu izazvati normalne i abnormalne EEG aktivnosti koje se nebi drugačije mogle primjetiti. Neke od procedura su hiperventilacija, fotostimulacija, zatvorene/otvorene oči i sl.

Također se razlikuju različiti načini mjerenja napona, a to su: bipolarni i monopolarni način. Kod bipolarnog načina mjeri se razlika potencijala između dvije elektrode koje su obje na potencijalima mozga, a kod monopolarnog načina samo jedna elektroda mjeri napon mozga, dok je druga na nultom, odnosno srednjem referentnom potencijalu. Vrlo često se kao elektroda nultog potencijala odabire ušna resica.

### 3.3. Valni oblici

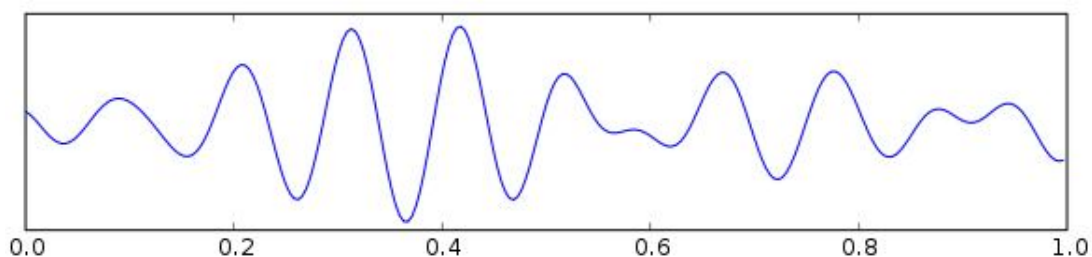
U EEG-u se mogu uočiti četiri karakteristična signala po valnom obliku, veličini amplitude i frekvenciji. Oni se nazivaju  $\alpha$ -valovi (alfa valovi),  $\beta$ -valovi (beta valovi),  $\theta$ -valovi (theta valovi) i  $\delta$ -valovi (delta valovi).

- **$\alpha$ -valovi**

Frekvencije ovih valova se kreću od 8 Hz do 13 Hz, a amplituda im je približno 50 $\mu$ V. Hans Berger je prvu uočenu ritmičnu EEG aktivnost nazvao alfa valovima. Ti valovi dominiraju u okcipitalnoj regiji, ali se katkad mogu naći i u parijetalnom i frontalnom području. Oni se javljaju kad je pacijent opušten i ima zatvorene oči, a atenuiraju se pri otvaranju očiju. Amplituda  $\alpha$ -valova nije konstantna, nego se mijenja u obliku vretena, pa potpuno nestane da bi se nakon nekoliko sekundi opet pojavila.  $\alpha$ -valovi se gube za vrijeme spavanja kao i prilikom rješavanja nekog problema (kad se pacijent koncentrira na nešto).

$\alpha$ -valovi mogu biti i abnormalni, kao npr. EEG u kojem se oni pojavljuju i ne odgovaraju, odnosno, ne mijenjaju se prilikom nekih vanjskih podražaja. To se najčešće javlja kod pacijenata koji su u komi i tad se ti valovi nazivaju "alfa koma".

Primjer  $\alpha$ -valova je prikazan na slici 3.

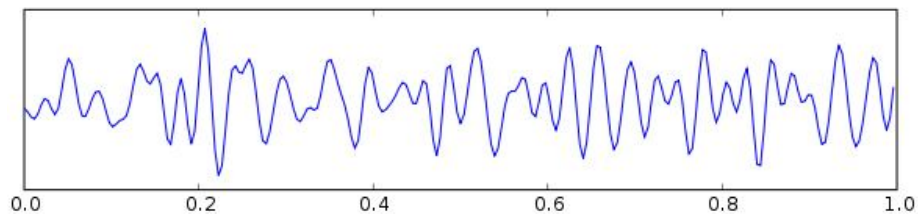


Slika 3.:  $\alpha$ -valovi

- **$\beta$ -valovi**

Frekvencije  $\beta$ -valova se kreću od 14 Hz do 30 Hz, a katkad dosežu i 50Hz dok im je amplituda manja od amplitude  $\alpha$ -valova i iznosi približno 20 $\mu$ V. Najčešće se registriraju na frontalnim i parijetalnim regijama.  $\beta$ -valovi niske amplitude se najčešće javljaju pri aktivnom stanju pacijenta, odnosno, prilikom neke duševne aktivnosti (aktivna koncentracija, razmišljanje i sl.). Ovi valovi su dominantni kod pacijenata koji su anksiozni ili koji imaju otvorene oči. Upravo otvaranjem očiju se testira blokiranje  $\alpha$ -ritma, tad se koncentracija pacijenta poveća i to prouzrokuje pojavu  $\beta$ -valova, a nestanak  $\alpha$ -valova.

Primjer  $\beta$ -valova je prikazan na slici 4.

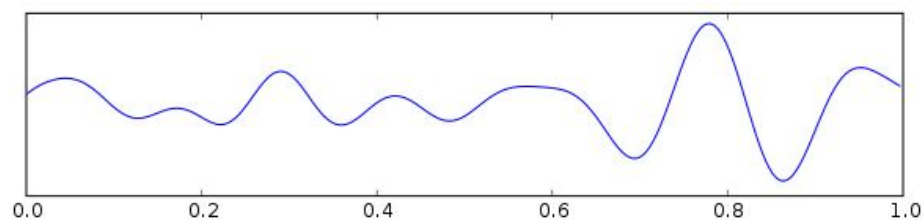


Slika 4.:  $\beta$ -valovi

- **$\theta$ -valovi**

$\theta$ -valovi imaju frekvencije između 4 i 7 Hz i amplitude oko 70 $\mu$ V. Pojavljuju se pretežno u parijetalnim i temporalnim regijama. Oni su normalno vidljivi kod mlađe djece, te prilikom napetosti i emocionalnih stresova (npr. razočarenje, frustracija) kod starije djece i kod odraslih. Ovi valovi mogu biti i patološki te se pojavljuju kod mnogih poremećaja mozga. U normalnom EEG-u može biti prisutna manja količina  $\theta$ -ritma, ali obično ne veća od 10%.

Primjer  $\theta$ -valova je prikazan na slici 5.

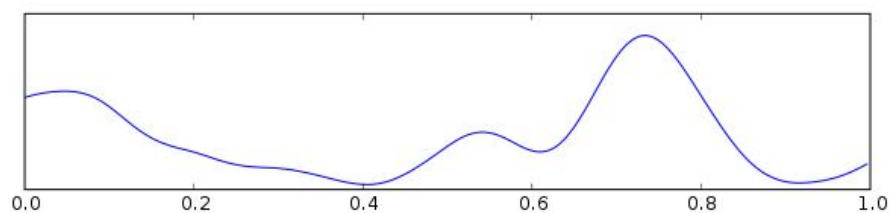


Slika 5.:  $\theta$ -valovi

- **$\delta$ -valovi**

$\delta$ -valovi zauzimaju frekvencije od 0,5 Hz do 3,5 Hz, relativno su najveće amplitude od 60 $\mu$ V do 100 $\mu$ V i ujedno su najsporiji valovi. Pojavljuju se u kori velikog mozga neovisno o aktivnosti u nižim slojevima mozga. Normalno se javljaju kod odraslih u dubokom snu i kod male djece. Ovi valovi se također mogu javiti pri težim bolestima mozga.

Primjer  $\delta$ -valova je prikazan na slici 6.



Slika 6.:  $\delta$  -valovi

U tablici 1. prikazani su svi valovi sa osnovnim karakteristikama radi lakše usporedbe ([26]).

Tablica 1.: Valni oblici EEG signala i njihove osobine

| Vrsta vala       | Frekvencije    | Amplituda           | Lokacija                        | Normalno                                                                                                | Patološki                                                                                  |
|------------------|----------------|---------------------|---------------------------------|---------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| $\alpha$ -valovi | 8 – 13 [Hz]    | 50 [ $\mu$ V]       | Okcipitalna regija              | -opušteno stanje<br>-zatvorene oči<br>-povezano s kontrolom inhibicije                                  | -koma                                                                                      |
| $\beta$ -valovi  | 14 – 30 [Hz]   | 20 [ $\mu$ V]       | Frontalna i parijetalne regije  | -aktivno stanje<br>-zaposleno ili anksiozno razmišljanje<br>-aktivna koncentracija<br>-otvorene oči     | -uzimanje nekih lijekova kao što je benzodiazepin                                          |
| $\theta$ -valovi | 4 – 8 [Hz]     | 70 [ $\mu$ V]       | Parijetalne i temporalne regije | -kod mlade djece<br>-napetost kod starije djece i odraslih<br>-kod odraslih pri emocionalnim stresovima | -subkortikalne lezije<br>-metabolička encefalopatija<br>-neke vrste <i>hydrocephalus-a</i> |
| $\delta$ -valovi | 0,5 – 3,5 [Hz] | 60 – 100 [ $\mu$ V] | Kora velikog mozga              | -duboki san kod odraslih<br>-kod beba                                                                   | -subkortikalne lezije<br>-difuzne lezije<br>-metabolička encefalopatija                    |

### 3.4. Korištenje EEG-a

EEG omogućuje dijagnostiku mnogih organskih bolesti mozga kao što su intrakranijalni tumori, različite traume, hematomi, meningitis, encefalitis, a posebno različite vrste epilepsije. Baš zahvaljujući EEG-u uspjelo se utvrditi više vrsta epilepsija.

Za tumore su karakteristični spori valovi velike amplitude u okolišu tumora, dok je samo tumorozno tkivo električki neaktivno. Spore aktivnosti su uglavnom u području  $\theta$ -valova i  $\delta$ -valova. Neposredno nakon traume pojavljuje se kratkotrajna brza aktivnost (valovi visoke frekvencije), nakon toga pojavljuje se EEG malih amplituda spore aktivnosti dok se ne uspostavi normalan ritam. Kod meningitisa i encefalitisa EEG nije jače promjenjen. Prisutna je najviše difuzna spora aktivnost u području frekvencije  $\theta$ -valova. Ako se bolest pogoršava, osnovna aktivnost je sporija, a pojavljuje se sve više  $\theta$ -valova i  $\delta$ -valova.

### 3.5. PTSP i EEG

Laici su i prije stručnjaka prihvatili činjenicu da jaka (psiho) trauma može izazvati određene psihičke tegobe koje imaju kvalitetu psihičkog poremećaja. Takav je psihički poremećaj i PTSP, što se javlja u osoba koje su proživjele izuzetan emocionalni ili fizički traumatski doživljaj koji je po karakteru traumatičan za svaku normalnu osobu. Poremećaj se može javiti u bilo kojoj dobi, uključujući i dječju.

Traumu osoba može doživjeti pojedinačno (prepad, silovanja) ili u skupini (ratno događanje). Stresovi mogu biti prirodnog podrijetla (potresi, poplave), mogu se javiti kao slučajna nesreća (prometna nesreća, veliki požari, rušenje zgrada), a mogu biti i uzrokovani ljudskom rukom (bombardiranje, granatiranje, mučenje, ubijanje). Neki stresovi gotovo uvijek izazivaju PTSP, npr. mučenje, a neki rijetko, primjerice prirodne katastrofe.

Taj je pojam našoj stručnoj javnosti postao blizak tek početkom Domovinskog rata, u jesen 1991. PTSP se u takvim situacijama ne javlja samo u osoba koje su bile izravno pogođene navedenim zbivanjima nego i u onih koji su im svjedočili.

U središtu traume je nametljivo sjećanje na središnji stresni (nasilni) događaj, poput slike i zvuka pušćane vatre, vriskova ili trenutačnih tišina, uboda nožem, praska puške, eksplozije mine, ranjavanja, zvukova zrakoplova, sirena i sl.

Osnovni simptomi PTSP-a su ([29]):

- Ponovno proživljavanje traumatskog događaja, npr. kroz sjećanja na događaj, snove i sl.
- Izbjegavanje svih podsjetnika na traumu (npr. izbjegavanja odlaska na mjesto gdje se trauma dogodila)
- Pojačana podražljivost koja se manifestira kao nesanica, razdražljivost, snižen prag na frustracije i sl.

Dijagnosticiranje PTSP-a nije jednostavno jer je ta bolest, u usporedbi sa drugim bolestima, malo istražena i relativno je nepoznata stručnjacima. Za dijagnozu se najčešće koriste psihoterapeutske razgovori i subjektivne ocjene liječnika i zbog toga se javila želja za nekim jačim pokazateljima za uspostavljanje te dijagnoze kao što je npr. EEG. Želi se pronaći način da se pomoću EEG-a može uspostaviti dijagnoza PTSP-a, kao što se pomoću EEG-a može dijagnosticirati epilepsija ili pomoću EKG-a srčane aritmije. Pored korištenja EEG-a za dijagnosticiranje PTSP-a u SAD-u se sve više razvija i koristi nova metoda za liječenje PTSP-a, tzv. *neurofeedback* ([30]) koji se bazira na EEG signalima.

PTSP se najčešće liječi lijekovima i psihoterapijom. Od lijekova se koriste antidepresivi i anksiolitici, te stabilizatori raspoloženja. Od psihoterapije koriste se različite tehnike prorađivanja trauma, ali i tehnike relaksacije. *Neurofeedback* je tek u fazi razvoja i testiranja pa se i ne spominje kao konkretna metoda liječenja. To je kompjuterizirana metoda, bazirana na praćenju električne aktivnosti mozga (EEG-a) i

davanju povratne informacije (eng. *feedback*). To je metoda kojom se mozak “uči” kako da poboljša vlastito funkcioniranje.

U Americi se razvija i koristi već tridesetak godina, te se pokazala vrlo uspješnom u tretmanu deficita pažnje/hiperaktivnog poremećaja (ADHD i ADD), poremećaja raspoloženja (depresije), poremećaja spavanja, tjeskobe, stresa, itd. Nakon postignutih dobrih rezultata u liječenju navedenih bolesti došlo se na ideju da se *neurofeedback* testira na ratnim veteranima kojima je dijagnosticiran PTSP (Othmer [25]).

*Neurofeedback* je metoda snimanja moždanih valova bez ikakvog djelovanja na njih osim kroz povratnu informaciju (nema nikakvog djelovanja elektroda ili struje na mozak).

Ova metoda se sastoji od snimanja moždanih valova koje terapeut prati na svom, posebno programiranom monitoru, dok pacijent gleda u svoj monitor. Pacijent također prati svoju moždanu aktivnost, ali u obliku video igre i zvučnih signala. Ova video igra se ne igra na klasičan način, odnosno, ne upravlja se njome na klasičan način nego se koristi snaga uma.

Tako npr. ako želimo da mozak proizvodi više brzih valova, kada se poveća aktivnost mozga upravo u tim frekvencijama, bolje se napreduje u igri (dobije se više bodova). U suprotnom, ako se poveća aktivnost u frekvencijama sporih valova (onih koje želimo smanjiti), tada će i uspjeh u igrici opasti. Postupno, mozak će sve više i više reagirati na ove “upute”, i naučit će taj novi, željeni, obrazac aktivnosti.

Na osnovu pretraga i/ili testiranja, npr. snimanje kompletnog EEG-a, kvantitativnog EEG-a, ili pomoću psihotestova koji se obave prije početka treniranja skupljaju se podaci koji služe za što preciznije određivanje najbolje vrste treninga, te na koji spektar moždane aktivnosti i na koje područje mozga se treba najviše usmjeriti.

### 3.6. Korišteni EEG signali

EEG signali koji su korišteni u ovom radu su spremljeni u *.edf* formatu koji je u Matlab učitavan pomoću *EEGLab toolboxa* jer se pomoću njega EEG signal učitava u jednu strukturu u kojoj su sadržane sve potrebne informacije kao npr. frekvencija uzorkovanja, broj kanala, nazivi kanala, signali spremljeni u matrici, itd (signal se učitava pomoću funkcije implementirane u *toolboxu pop\_biosig()*) i zbog toga je pogodniji za daljnju analizu i korištenje.

Korišteni *toolbox* je tipa *open source* i služi za analizu EEG signala. Da bi ga mogli skinuti sa interneta najprije je potrebno otići na stranicu: <http://sccn.ucsd.edu/~scott/ica.html>, tamo popuniti obrazac u kojem objasnite za što vam treba *toolbox* i onda dobijete dozvolu da ga skinete na svoje računalo. Nakon što se skinuta *.rar* arhiva otpakira na računalu potrebno se unutar Matlaba pozicionirati u taj folder te u komandnom prozoru Matlaba ukucati „eeglab“.

Svi korišteni EEG signali su snimani ujutro, u ležećem položaju, u mirovanju, sa zatvorenim očima nakon što bi pacijent doručkovao. Birani su isječki bez artefakta, odnosno s ciljanim artefaktima. Svi pacijenti su bili muškarci, ne stariji od 55 godina. Većina je bila pod terapijom, ali isječki na kojima je bio evidentan utjecaj lijekova s posljedično promijenjenim EEG-om nisu korišteni u ovom istraživanju. Također se nisu koristili isječki pacijenata koji su imali neku neurološku bolest jer tada ne bismo mogli promatrati povezanost EEG-a i psihijatrijskog poremećaja.

Korištene elektrode su kositrene i pozicionirane su po klasičnom 10-20 sustavu. Mjerenje napona je bipolarno s ukupno 20 kanala, a aparat korišten za snimanje signala je Medialov, TG Valor T40 T64 T128, Nervus v3.x.

## 4. Klasifikacija EEG signala

Postupak klasifikacije se sastoji iz tri osnovna koraka:

1. predprocesiranje i odabir značajki
2. treniranje klasifikatora koristeći odabrane značajke
3. testiranje klasifikatora.

Za predprocesiranje značajki se koristi ANOVA (eng. *Analysis of Variance*) koja će biti opisana u idućem poglavlju. Nakon razumijevanja ANOVA-e opisati će se korištene značajke, rezultati ANOVA-e za svaku od njih i konačan odabir značajki koje se prenose u fazu treniranja. Klasifikator koji je odabran za obavljanje klasifikacije u ovom radu je SVM koji će također biti opisan u nastavku rada zajedno sa rezultatima testiranja.

### 4.1 ANOVA

Struke koje svoje nove spoznaje pretežno izvode iz podataka u pravilu polaze od uzorka ispitanika. Na uzorku se izvode mjerenja i dobivaju se informacije u obliku aritmetičkih sredina, varijanci, proporcija i sličnog, a onda se dobiveno želi poopćiti na populaciju iz koje je uzorak uzet.

Testiranje hipoteze je statistički postupak kojim se određuje da li i koliko pouzdano raspoloživi podaci podupiru postavljenu pretpostavku. Testiranje hipoteza, odnosno testiranje značajnosti u osnovi je postupak kvantifikacije impresija o specifičnoj hipotezi.

ANOVA je računski postupak pomoću kojega se ispituju podaci određenoga pokusa, kroz procjenu otklona pojedinih srednjih vrijednosti od prosječne vrijednosti uzoraka uzetih iz nekog osnovnog skupa.

Kroz niz relativno jednostavnih izračunavanja potrebno je dobiti  $F$  vrijednost. Sinonim za izračunavanje  $F$  vrijednosti je  $F$ -test<sup>1</sup> ili grupni test za ispitivanje hipoteze pokusa.  $F$ -testom se ispituje, pokusom postavljena, nulta hipoteza<sup>2</sup> da su aritmetičke sredine  $k$  osnovnih skupova međusobno jednake, odnosno, da u cjelini nema statistički značajne razlike. Cilj je ispitati odnos varijacija između uzoraka s varijacijama unutar uzoraka. Ako je taj odnos, tzv. empirijski  $F$ -omjer, statistički značajan zaključuje se da promatrani uzorci ne pripadaju istoj populaciji, odnosno aritmetičke sredine se značajno razlikuju.

ANOVA se u ovom radu koristi za predprocesiranje značajki, tj. za pronalaženje onih značajki po kojima se dvije grupe signala (bolesni i zdravi) najviše razlikuju. ANOVA-om se testira nul-hipoteza da obje grupe značajki (značajka izračunata za zdrave signale i značajke izračunate za bolesne signale) pripadaju istoj populaciji što znači da će u ovom slučaju biti zanimljiv rezultat gdje se dobije niska vjerojatnost sličnosti tih grupa.

Cilj predprocesiranja značajki je smanjenje dimenzionalnosti vektora značajki koji se koristi u treniranju i testiranju sustava. Smanjivanje dimenzionalnosti je od velikog značaja jer svaki EEG signal ima 20 kanala, što zapravo predstavlja 20 signala. Tako bi se za jednu značajku jednog EEG signala dobio vektor od 20 značajki. Pored smanjivanja dimenzionalnosti, rezultatima ANOVA-e se može dobiti bolji uvid u promjene EEG signala po kanalima. Tako npr. prilikom računanja snage beta pojasa (jedna značajka) primjenom ANOVA-e kao rezultat se dobije u kojem kanalu snaga beta pojasa pokazuje najveću razliku između normanih i PTSP EEG signala i taj kanal se smatra zanimljivim i njega ćemo koristiti za klasifikaciju.

U idućem poglavlju će biti detaljnije opisane korištene značajke i rezultati ANOVA-e za svaku od njih.

---

<sup>1</sup>  $F$ -test je samo jedan od testova za provjeru hipoteza. Pored  $F$ -testa najčešće se koriste još  $T$ -test i  $\chi^2$  test

<sup>2</sup> Nul-hipoteza,  $H_0$  (eng. *null hypothesis*) je pretpostavka o izostanku efekta, tj. da ne postoji razlika među uzorcima u populaciji od interesa (npr. nema razlike u aritmetičkim sredinama). To je hipoteza koja se testira, *hipoteza da nema razlike* (eng. *hypothesis of no difference*).

## Implementacija

ANOVA je implementirana u Matlabu kao dio *Statistics Toolbox*-a i u sklopu istog postoje tri različite izvedbe i to:

- *One - way ANOVA* - ANOVA s jednim promjenjivim faktorom (karakteristikom)
- *Two – way ANOVA* - ANOVA s dva promjenjiva faktora
- *N – way ANOVA* - ANOVA s N promjenjivih faktora

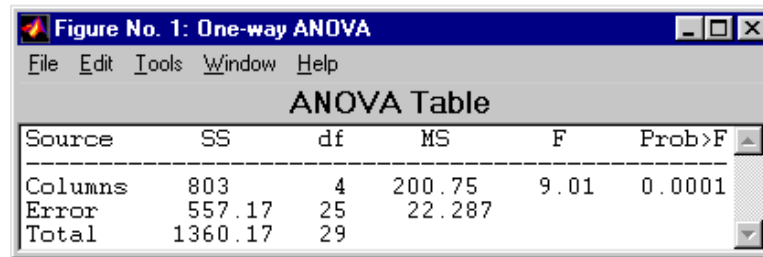
Svrha *one-way* ANOVA-e je ustanoviti da li podaci iz više skupina imaju zajedničku aritmetičku sredinu, tj. da li se skupine stvarno razlikuju u izmjerenoj karakteristici. *Two-way* ANOVA se razlikuje od *one-way* ANOVA-e u tome što su skupine u *two-way* ANOVA-i opisane sa dvije karakteristike. Analogno, skupine u *N-way* ANOVA-i se opisuju sa N karakteristika.

U svrhu predprocesiranja značajki EEG signala u ovom radu korištena je *one-way* ANOVA u kojoj se testiraju odabrane značajke (karakteristike) EEG signala za svaki kanal pojedinačno. Značajke su podjeljene u dvije grupe, značajke dobivene iz normalnih EEG-ova i značajke dobivene iz PTSP EEG-ova.

Implementirana funkcija u Matlabu koja realizira *one-way* ANOVA-u je *anova1()* kojoj se kao ulazni parametar daje matrica sa dva stupca od kojih svaki predstavlja jednu skupinu (zdravi i bolesni) dok se u redovima matrice nalaze izračunate vrijednosti značajki, a kao rezultat se dobije tzv. ANOVA tablica (tablica 2. i slika 7.) i vjerojatnost  $p$  sličnosti dvije skupine.

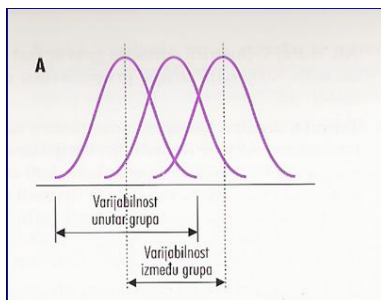
Tablica 2.: ANOVA tablica

| Izvor varijacije | Zbroj kvadrata odstupanja (SS)                     | Stupnjevi slobode (df) | Sredine kvadrata odstupanja (MS)                                         | F                     |
|------------------|----------------------------------------------------|------------------------|--------------------------------------------------------------------------|-----------------------|
| Između uzoraka   | $\sum_{j=1}^k n_j (\bar{X}_j - \bar{X})^2$         | $k - 1$                | $S_A^2 = \frac{\sum_{j=1}^k n_j (\bar{X}_j - \bar{X})^2}{k - 1}$         | $\frac{S_A^2}{S_U^2}$ |
| Unutar uzoraka   | $\sum_{j=1}^k \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2$ | $n - k$                | $S_U^2 = \frac{\sum_{j=1}^k \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2}{n - k}$ |                       |
| Ukupno           | $\sum_{j=1}^k \sum_{i=1}^n (X_{ij} - \bar{X})^2$   | $n - 1$                |                                                                          |                       |

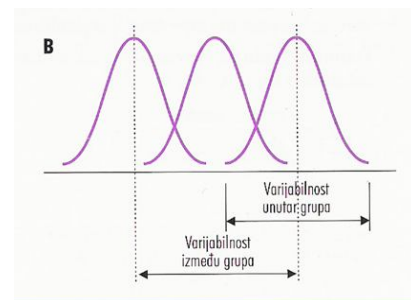


Slika 7.: ANOVA tablica dobivena Matlabom

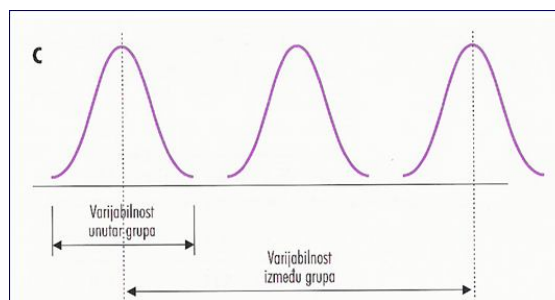
Grafički prikaz rezultata kad je varijabilnost između grupa manja od varijabilnosti unutar grupa je prikazan na slici 8. Za klasifikaciju EEG signala ovaj rezultat nije povoljan. Ono što se želi dobiti je veća varijabilnost između grupa kao što je prikazano na slici 9. To bi značilo da su grupe različite, tj. da postoji razlika između zdravih i bolesnih signala za zadanu značajku. Idealno bi bilo dobiti odnos kao na slici 10. gdje je varijabilnost između grupa puno veća od varijabilnosti unutar grupa.



Slika 8.: Varijabilnost između grupa < varijabilnost unutar grupa



Slika 9.: Varijabilnost između grupa > varijabilnost unutar grupa



Slika 10.: Varijabilnost između grupa >> varijabilnost unutar grupa

## 4.2. Značajke

Odabir značajki je najbitnija faza u procesu klasifikacije signala. Potrebno je naći one značajke koje dobro opisuju razlike između klasa koje se žele klasificirati. Tu nam u pomoć pristiže ANOVA, ali najprije se trebaju odabrati značajke na kojima će se ANOVA primijeniti.

U prvoj fazi su uzete sve značajke korištene za klasifikaciju EEG signala koje su spomenute u srodnoj literaturi. Zatim je slijedio proces eliminacije onih koje su bile računski (vremenski) zahtjevne. Nakon toga su eliminirane one značajke koje su specifične za druge bolesti (npr. epilepsija), a koje nisu pokazivale bitne razlike uspoređujući PTSP i normalne EEG-ove.

Nakon svih odabira i eliminacija došlo se do konačnog popisa značajki koje će se koristiti i to:

1. PSD srednja snaga  $\alpha, \beta, \delta$  i  $\theta$ -pojasa
2. srednja snaga  $\alpha, \beta, \delta$  i  $\theta$ -pojasa izračunata pomoću DWT koeficijenata
3. *Higuchieva* fraktalna dimenzija
4. *skewness*
5. *kurtosis*
6. *Hjortovi* parametri: aktivnost, mobilnost i kompleksnost.

#### 4.2.1. PSD

PSD (eng. *Power Spectral Density*) prikazuje kako je snaga nekog vremenskog signala raspoređena po frekvencijama. Drugim riječima, pomoću PSD-a se može vidjeti kolika je snaga signala na određenoj frekvenciji. To svojstvo je jako korisno kad se radi o EEG signalima jer se iz PSD-a mogu jasno očitati snage svih karakterističnih pojaseva EEG-a. Tako npr. pošto znamo da se  $\alpha$ -valovi nalaze u frekvencijskom području od 8 Hz do 13 Hz iz PSD-a se može izračunati srednja snaga signala tog frekvencijskog pojasa i na taj način dobiti snaga  $\alpha$ -pojasa. Analogno tome se računaju snage ostalih pojaseva EEG signala i upravo tako dobivene vrijednosti snage će se koristiti kao značajke EEG signala u procesu klasifikacije.

PSD se definira kao:

$$PSD(\omega) = F_T(\omega)F_T^*(\omega) \quad (1.)$$

gdje je  $F_T(\omega)$  normalizirana *Fourierova* transformacija signala  $f(t)$  koja se računa na slijedeći način:

$$F_T(\omega) = \frac{1}{\sqrt{T}} \int_0^T f(t)e^{-j\omega t} dt \quad (2.)$$

a  $F_T^*(\omega)$  predstavlja konjugiranu *Fourierovu* transformaciju  $F_T(\omega)$ .

#### **Implementacija**

Za računanje PSD-a signala korišten je *EEGlab toolbox*, konkretno, korištena je funkcija *spectopo()* koja je implementirana unutar *toolboxa*. Funkcija *spectopo()* poziva Matlabovu implementiranu funkciju *pwelch()* pomoću koje se računa PSD

zadanog signala koristeći *Welchovu* metodu. *Welchova* metoda se sastoji iz sljedećeg:

1. signal se podijeli na preklapajuće segmente. Originalni signal se podijeli na  $L$  segmenata duljine  $M$  koji se preklapaju u  $D$  točaka (ako je  $D = M/2$  znači da je preklapanje od 50%).
2. nakon što se signal podijeli, na  $L$  individualnih segmenata se primjeni prozor (u vremenskoj domeni). Većina prozora više značaja pridaje centralnim uzorcima nego onima na rubovima prozora i zbog toga dolazi do gubitka informacija. Preklapanje segmenata služi zato da bi se umanjio taj gubitak.
3. zatim se na svaki od prozora primjeni Fourierova transformacija i računa se snaga po jednačbi (1.).

Koristeći navedene funkcije računaju se prosječne snage za sva četiri pojasa EEG-signala i te prosječne snage predstavljaju korištene značajke.

### ***ANOVA i odabir značajki***

Kao što je već spomenuto susrećemo se sa problemom velikog vektora značajki. Za svaki signal koji se sastoji od 20 kanala se izračuna prosječna snaga četiri pojasa što ukupno čini 80 izračunatih vrijednosti. Budući da će se koristiti još značajki to bi vektor značajki učinilo još većim. Tu dolazi na red predprocesiranje značajki ANOVA-om.

Dakle, konstruira se matrica  $M$  (vidi poglavlje ANOVA) tako da se u prvi stupac matrice uvrste vrijednosti dobivene za snagu  $\alpha$ -pojasa u prvom kanalu izračunate za zdrave EEG-ove, a u drugi stupac ide isto to samo što se računa za PTSP EEG-ove. Ova radnja se dalje ponavlja sa svaki kanal posebno, a kad se potroše svi kanali onda se cijeli postupak ponavlja za  $\beta$ -pojas. Ono što se dobije kao rezultat je vrijednost  $p$  za svaki kanal svake značajke. Vrijednost  $p$  predstavlja sličnost dvije skupine (zdravih i bolesnih) što znači da nas zanimaju oni kanali gdje će vrijednost  $p$  biti najniža, tj. skupine se razlikuju.

Dobiveni rezultati se nalaze u tablici 3.

Tablica 3.: Rezultati ANOVA-e

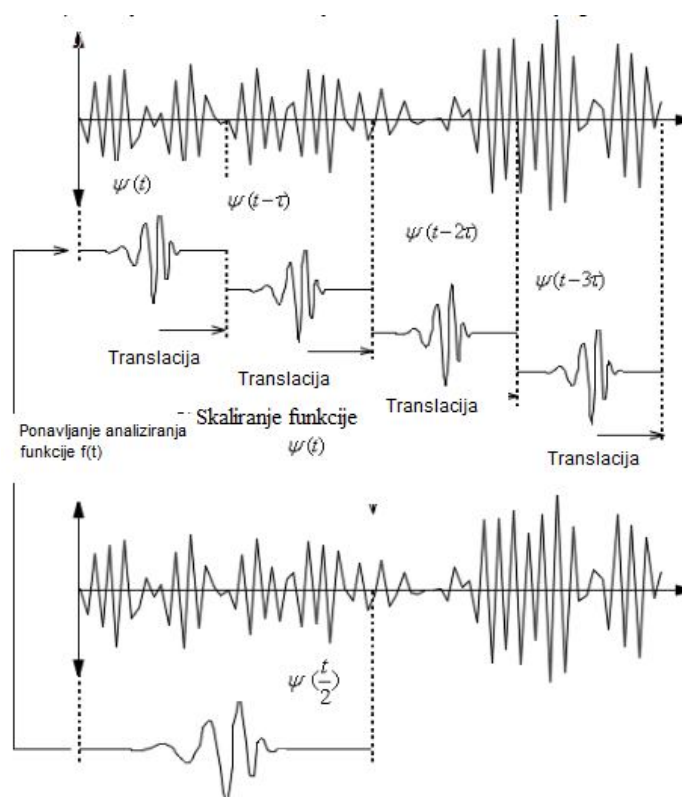
| Značajka | PSD snaga<br>$\alpha$ -pojasa | PSD snaga<br>$\beta$ -pojasa | PSD snaga<br>$\delta$ -pojasa | PSD snaga<br>$\theta$ -pojasa |
|----------|-------------------------------|------------------------------|-------------------------------|-------------------------------|
| Kanal    | Vrijednost p                  | Vrijednost p                 | Vrijednost p                  | Vrijednost p                  |
| 1        | 0,010779549                   | 0,371136116                  | 0,1868229                     | 0,14376142                    |
| 2        | 0,007604161                   | 0,48317885                   | 0,31433651                    | 0,228254023                   |
| 3        | 0,000110411                   | 0,977935527                  | 0,21123688                    | 0,320246688                   |
| 4        | 0,000469741                   | 0,212278285                  | 0,00237774                    | 0,534677476                   |
| 5        | 0,020787258                   | 0,9160973                    | 0,18758672                    | 0,521787709                   |
| 6        | 0,024423882                   | 0,925467797                  | 0,00152299                    | 0,977214571                   |
| 7        | 1,8931E-05                    | 0,742513586                  | 0,6906467                     | 0,248651438                   |
| 8        | 0,001167888                   | 0,005971705                  | 0,17370577                    | 0,108749003                   |
| 9        | 0,063115991                   | 0,799064022                  | 0,16246078                    | 0,852341592                   |
| 10       | 0,03066745                    | 0,48151352                   | 0,97587208                    | 0,745847288                   |
| 11       | 2,1813E-06                    | 0,010540659                  | 0,11481453                    | 0,041568912                   |
| 12       | 0,000673414                   | 0,050129709                  | 0,04118806                    | 0,999217796                   |
| 13       | 0,129883883                   | 0,297825394                  | 0,14067593                    | 0,422901096                   |
| 14       | 0,087179283                   | 0,020403484                  | 0,43944913                    | 0,462010643                   |
| 15       | 0,005084867                   | 0,000542046                  | 0,82026176                    | 0,121701412                   |
| 16       | 0,000102953                   | 0,02256572                   | 0,55793158                    | 0,113940177                   |
| 17       | 0,000292805                   | 0,67865018                   | 0,0108696                     | 0,325894954                   |
| 18       | 0,000470357                   | 0,28071924                   | 0,02556371                    | 0,251498487                   |
| 19       | 0,002279444                   | 0,012096576                  | 0,00082267                    | 0,206534409                   |
| 20       | 0,000641105                   | 0,032760105                  | 0,00675998                    | 0,125276225                   |

U tablici su zelenom bojom označene najniže, a crvenom najviše vrijednosti parametra p. Iz priloženog se vidi da će se koristiti slijedeće značajke:

1. PSD snaga  $\alpha$ -pojasa u 11. kanalu
2. PSD snaga  $\beta$ -pojasa u 15. kanalu
3. PSD snaga  $\delta$ -pojasa u 19. kanalu
4. PSD snaga  $\theta$ -pojasa u 11. kanalu

#### 4.2.2. DWT

*Waveleti* su funkcije koje mogu imati bilo kakav oblik, ali su vremenski ograničene. Multirezolucijsko predstavljanje signala je osnovno načelo *wavelet* transformacije, koja za razliku od *Fourierove* transformacije signal prikazuje istodobno u vremenskoj i frekvencijskoj domeni. Signal se promatra u vremenskim intervalima i za svaki takav interval se računa spektar. Kod *waveleta* se ne koristi pojam spektra, već je uveden pojam skala. Skala je obrnuto proporcionalna frekvencijskom pojasu. Kad se dođe do kraja signala vremensko ograničavanje se ponavlja s dužim ili kraćim intervalima. Kao rezultat se dobije niz *vrijeme-skala* funkcija sve sa različitim rezolucijama. *Wavelet* analiza je veoma slična *Fourierovoj*. *Fourierovom* analizom signal se predstavlja pomoću kosinusnih i sinusnih funkcija dok se kod *waveleta* prikazuje tzv. *wavelet* funkcijama. Sve *wavelet* funkcije generirane su iz iste funkcije, koja se zove osnovna ili *mother wavelet* funkcija, postupkom skaliranja i translacije koji je prikazan na slici 11.



Slika 11. Postupak skaliranja i translacije

Općenito, *wavelet* transformacija može biti kontinuirana (eng. *CWT-Continuous Wavelet Transform*) i diskretna (eng. *DWT-Discrete Wavelet Transform*). CWT djeluje na bilo kakvoj skali dok DWT koristi specifični podskup skala i vrijednosti translacije.

Za dobivanje značajki u ovom radu koristiće se DWT. DWT se definira prema izrazu

$$X[m, k] = \frac{1}{\sqrt{a_0^k}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t}{a_0^k} - mT\right) dt \quad (3.)$$

gdje su:

- $X[m, k]$  - frekvencijski sastav signala  $x(t)$  u određenom vremenskom intervalu
- $x(t)$  - originalni signal
- $a_0^k$  – skala (logaritamska podjela u skali)
- $mT$  - pomak (translacija)

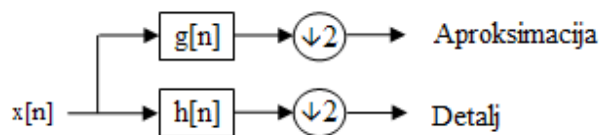
i gdje je  $\psi_{m,k}(t) = \frac{1}{\sqrt{a_0^k}} \psi\left(\frac{t}{a_0^k} - mT\right)$  familija *wavelet* funkcija, pri čemu je  $\psi(t)$  *mother wavelet*.

DWT signala se računa njegovim propuštanjem kroz niz filtera. Istovremeno se uzorci signala propuštaju kroz niskopropusni filter impulsnog odziva  $g$  i kroz visokopropusni filter impulsnog odziva  $h$ . Na izlazu iz filtera dobije se konvolucija signala i impulsnih odziva filtera, tj.:

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n - k] \quad (4.)$$

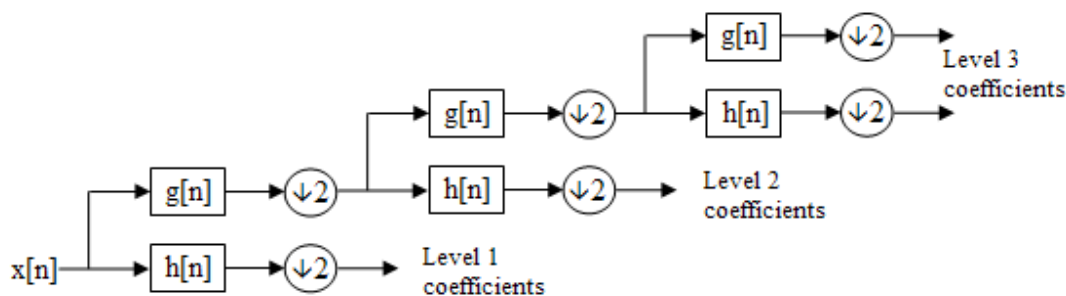
$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n - k] \quad (5.)$$

Izlazi iz niskopropusnog filtra se nazivaju aproksimacijama (eng. ca – *approximation coefficients*), a izlazi visokopropusnog filtra detaljima (eng. cd – *detail coefficients*). Ovaj postupak je prikazan na slici 12.



Slika 12.: Dobivanje koeficijenata aproksimacija i detalja

Ovakva dekompozicija se ponavlja da bi se povećala frekvencijska rezolucija. Taj postupak se prikazuje pomoću binarnog stabla koje se još naziva filtarskim slogom (slika 13.).



Slika 13.: Filtarski slog

Postoji puno familija *waveleta* i unutar svake familije postoje različite izvedbe tog *waveleta* i zbog toga je njihov odabir dugotrajan posao. Za EEG signale najboljom se pokazala familija *waveleta* pod imenom *Daubechies*<sup>3</sup> wavelet (u Matlabu skraćeno *db wavelet*). Skoro u svim srodnim radovima koji koriste DWT kao sredstvo ekstrakcije značajki koriste se upravo *Daubechijevi waveleti* [3] [8] [16].

Formulacija *Daubechies waveleta* se temelji na korištenju ponavljajućih odnosa za generiranje progresivno finijih diskretnih uzoraka implicitno nastalih od *mother waveleta*. Svaka rezolucija je dva puta veća od prethodne skale (prethodne razine dekompozicije). Postoji nekoliko varijacija *Daubechiesovog* originalnog

<sup>3</sup> Dobili naziv po belgijskoj matematičarki Ingrid Daubechies koja je fomulirala svoju familiju *waveleta* 1988. godine.

*waveleta*, a onaj koji se koristi u ovom radu je *Daubechiesov wavelet* četvrtog reda (tzv. *db4*).

Poznavanjem centralne frekvencije *waveleta* može se odrediti frekvencijsko područje i centralne frekvencije svih koeficijenata aproksimacije i detalja u svakoj razini dekompozicije. To svojstvo, zajedno sa svojstvom *Daubechiesovih waveleta* da im je rezolucija u svakoj razini dekompozicije duplo veća od prethodne, se koristi za pronalazak onih koeficijenata dekompozicije koji padaju u frekvencijska područja alfa, beta, delta i theta pojasa.

Pomoću Matlaba se dobiju centralne frekvencije *db4 waveleta* i to:

| Scale   | Frequency | Period |     |
|---------|-----------|--------|-----|
| 2.0000  | 91.4286   | 0.0109 | cd1 |
| 4.0000  | 45.7143   | 0.0219 | cd2 |
| 8.0000  | 22.8571   | 0.0437 | cd3 |
| 16.0000 | 11.4286   | 0.0875 | cd4 |
| 32.0000 | 5.7143    | 0.1750 | cd5 |
| 64.0000 | 2.8571    | 0.3500 | ca5 |

Iz dobivenih rezultata se može zaključiti da npr. detalj pete razine dekompozicije (*cd5*) ima centralnu fekvenciju 5,7143 Hz i da zauzima područje od, približno 3,7143 Hz do 7,7413 Hz što čini ukupnu širinu pojasa 4Hz, što bi značilo da *cd4* zauzima duplo veći frekvencijski pojas (osobina *db waveleta*), tj. područje od 8Hz koje otprilike odgovara intervalu od 8 do 16 Hz. Na osnovu izračunatog zaključujemo da *cd5* po frekvencijama odgovara  $\theta$ -valu, a *cd4*  $\alpha$ -valu.

Pomoću svih do sad spomenutih svojstava i dobivenih informacija dobiju se frekvencijska područja koeficijenata dekompozicije koji su prikazani u tablici 4. zajedno sa prethodno utvrđenim granicama alfa, beta, delta i theta valova. Treba napomenuti da su frekvencije zaokruživane na cijele brojeve radi preglednosti zapisa.

Tablica 4.: Odgovarajući koeficijenti i valovi sa frekvencijskim područjima

| Karakteristični val | Frekvencije    | Odgovarajući koeficijent aproksimacije/detalja | Frekvencije koeficijenata dekompozicije |
|---------------------|----------------|------------------------------------------------|-----------------------------------------|
| $\alpha$ -val       | 8 – 13 [Hz]    | <i>cd4</i>                                     | 8 – 16 [Hz]                             |
| $\beta$ -val        | 14 – 30 [Hz]   | <i>cd3</i>                                     | 16 – 32 [Hz]                            |
| $\theta$ -val       | 4 – 8 [Hz]     | <i>cd5</i>                                     | 4 – 8 [Hz]                              |
| $\delta$ -val       | 0,5 – 3,5 [Hz] | <i>ca5</i>                                     | do 4 Hz                                 |

Kao što se vidi iz priloženog, frekvencijska područja se ne poklapaju 100%. To je proizašlo iz same prirode *wavelet* transformacije koja ovisi o centralnoj frekvenciji *waveleta* i njegovoj efektivnoj širini. Ovaj problem nismo imali kod PSD-a jer se pomoću Fourierove transformacije dobije kompletan spektar i točno se može uzeti dio spektra od-do one frekvencije koja nas zanima, dok kod *wavelet* transformacije sve ovisi o korištenom valiću i njegovim osobinama. Pretpostavlja se da pogreška u frekvencijskim područjima nije prevelika i da neće utjecati na rezultate.

### **Implementacija**

Nakon što se ustanove odnosi između pojedinih EEG karakterističnih valova i koeficijenata detalja i aproksimacija moguće je izračunati snage  $\alpha, \beta, \delta$  i  $\theta$ -područja. Te snage su značajke koje će se koristiti.

Kad se pomoću Matlabove funkcije *dwt()* izračunaju odgovarajući koeficijenti *ca* i *cd* onda se pomoću njih izračuna snaga pomoću funkcije *mean()* kao srednja kvadratna vrijednost aproksimacije, odnosno, detalja.

### ANOVA i odabir značajki

Dolazi do istog problema kao i kod računanja značajki PSD-om koji se rješava na isti način. Rezultati ANOVA-e se nalaze u tablici 5.

Tablica 5.: Rezultati ANOVA-e

| Značajka | DWT snaga $\alpha$ -pojasa | DWT snaga $\beta$ -pojasa | DWT snaga $\delta$ -pojasa | DWT snaga $\theta$ -pojasa |
|----------|----------------------------|---------------------------|----------------------------|----------------------------|
| Kanal    | Vrijednost p               | Vrijednost p              | Vrijednost p               | Vrijednost p               |
| 1        | 0,47900587                 | 0,47707528                | 0,0060633                  | 0,43439567                 |
| 2        | 0,02265506                 | 0,87351286                | 0,00481014                 | 0,04012505                 |
| 3        | 0,00022857                 | 0,25765672                | 0,00618835                 | 0,00644564                 |
| 4        | 0,00071654                 | 0,03593928                | 0,04835356                 | 0,23585472                 |
| 5        | 0,08493003                 | 0,44696957                | 0,02160262                 | 0,65756351                 |
| 6        | 0,02709487                 | 0,38109341                | 3,7947E-06                 | 0,0028364                  |
| 7        | 0,00011587                 | 0,05883079                | 5,2571E-05                 | 0,0002881                  |
| 8        | 0,00250104                 | 0,00040912                | 0,00034648                 | 0,04202799                 |
| 9        | 0,08451197                 | 0,63647354                | 0,00032306                 | 0,12598324                 |
| 10       | 0,02538949                 | 0,13311484                | 0,00120958                 | 0,09436288                 |
| 11       | 1,3856E-06                 | 0,00057499                | 0,00231972                 | 5,0308E-05                 |
| 12       | 0,00110623                 | 0,0038997                 | 0,03395096                 | 0,06728001                 |
| 13       | 0,08448865                 | 0,08476468                | 0,00025704                 | 0,17128372                 |
| 14       | 0,04462739                 | 0,01252272                | 0,00690521                 | 0,09313063                 |
| 15       | 0,00449565                 | 0,00116695                | 0,01338974                 | 0,01803944                 |
| 16       | 0,0005499                  | 0,00089709                | 0,02389061                 | 0,00578772                 |
| 17       | 0,00485481                 | 0,52431566                | 0,00372831                 | 0,29376748                 |
| 18       | 0,00276053                 | 0,82560986                | 0,00295346                 | 0,20411006                 |
| 19       | 0,00090498                 | 0,00324311                | 0,0126841                  | 0,02344261                 |
| 20       | 0,00129574                 | 0,0037546                 | 0,01266762                 | 0,01522986                 |

Koristiti će se slijedeće značajke:

1. DWT snaga  $\alpha$ -pojasa u 11. kanalu
2. DWT snaga  $\beta$ -pojasa u 8. kanalu
3. DWT snaga  $\delta$ -pojasa u 6. kanalu
4. DWT snaga  $\theta$ -pojasa u 11. kanalu

### 4.2.3. Higuchijeva fraktalna dimenzija

#### *Fraktalne strukture*

Za opisivanje objekata i fenomena u prirodi koriste se različite mjere koje se zovu dimenzije. Govoreći o dimenzijama u klasičnom smislu može se reći da su u Euklidovom prostoru poznate tri dimenzije. Linija ima samo jednu dimenziju (dužina), ravnina ima dvije dimenzije (dužina i širina), a prostor ima tri dimenzije (dužina, visina i širina).

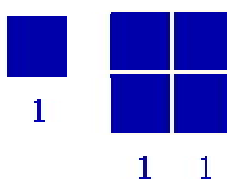
Pokušajmo sada definirati dimenziju na malo drugačiji način promatrajući samoslične objekte. Poznato je iz matematike da su dva objekta slična ukoliko su im kutevi isti, a stranice proporcionalno umanjene ili uvećane. Samosličan objekt je onaj objekt koji unutar sebe sadrži dijelove koji su slični velikom izvornom dijelu.

Uzmimo npr. jedan linijski segment. Kako bi načinili samosličan objekt uduplajmo ga na način kao na slici 14.



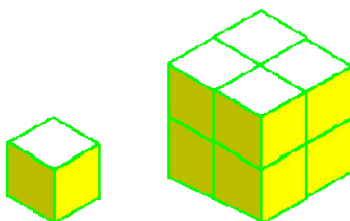
Slika 14.: Kreiranje samosličnog objekta - linija

Međutim, da bi se od jednog kvadrata načinio samosličan objekt potrebno je postaviti četiri takva ista kvadrata kao na slici 15.



Slika 15.: Kreiranje samosličnog objekta - kvadrat

Za dobivanje samoslične kocke potrebno je osam istih kockica kao na slici 16.



Slika 16.: Kreiranje samosličnog objekta - kocka

Sad se već može uočiti da je za svaku dimenziju potrebno  $n = 2^d$  istih manjih dijelova kako bi načinili samosličan geometrijski objekt, gdje je  $d$  broj dimenzija. Umjesto da se multiplicira mali geometrijski objekt samo 2 puta to se može učiniti proizvoljan broj  $s$  puta pa će općenita jednačina biti  $n = s^d$ . Dakle,  $n$  je potreban broj manjih dijelova za formiranje objekta,  $d$  je broj dimenzija, a  $s$  je faktor multiplikacije. Iz te jednačine proizlazi da će broj dimenzija  $d$  iznositi:

$$d = \frac{\log(n)}{\log(s)} \quad (6.)$$

Dimenzija ne mora uvijek biti cijeli broj. Dimenzija fraktala nije cijeli broj, oni imaju razlomljenu, fraktalnu dimenziju (npr. *Sierpinski-jev Trokut* [41]).

U prirodi, primjer fraktala je morska obala kojoj je gotovo nemoguće izmjeriti duljinu jer ovisi o skali kojom se mjeri. Npr. ako se mjeri na jako uvećanoj karti dobije se puno veća dužina nego na umanjenoj karti jer neće prikazivati manje detalje.

### **Fraktalna dimenzija (FD)**

Izraz "fraktalna dimenzija" odnosi se na necjelobrojnu ili fraktalnu dimenziju bilo kojeg objekta. FD analiza se često koristi u procesiranju biomedicinskih signala, kao što su EEG, HRV (eng. *heart rate variability*) i koračni intervali.

U fraktalnoj geometriji, fraktalna dimenzija  $D$ , je statistička veličina koja pokazuje koliko fraktalni objekt popunjava prostor i to na različitim skalama uvećanja. Postoji puno specifičnih definicija fraktalne dimenzije i nijedna od njih se ne tretira kao univerzalna.

Postoje različite metode i algoritmi za računanje fraktalne dimenzije, a za EEG signale koji se mogu promatrati kao vremenske serije koristi se *Higuchiev* algoritam izračunavanja fraktalne dimenzije. Fraktalna dimenzija, u ovom slučaju, predstavlja jednu moguću mjeru kompleksnosti signala.

*Higuchijev* algoritam za računanje fraktalne dimenzije  $D$  vremenske serije:

Za zadanu vremensku seriju  $x(1), x(2), \dots, x(N)$  konstruira se  $k$  novih serija kao:

$$x_m^k = \left\{ x(m), x(m+k), x(m+2k), \dots, x\left(m + \left\lfloor \frac{N-m}{k} \right\rfloor k\right) \right\}$$

za  $m = 1, 2, \dots, k$  gdje  $m$  označava početni trenutak u vremenu, a  $k$  označava diskretni vremenski interval između točaka.

Za svaku od vremenskih serija ili krivih  $x_m^k$  dužina  $L_m(k)$  se računa kao:

$$L_m(k) = \frac{\sum_{i=1}^{\left\lfloor \frac{N-m}{k} \right\rfloor} |x(m+ik) - x(m+(i-1)k)| (n-1)}{\left\lfloor \frac{N-m}{k} \right\rfloor k} \quad (7.)$$

gdje su:

- $N$  – ukupna dužina vremenske serije  $x$
- $\frac{n-1}{\left\lfloor \frac{N-m}{k} \right\rfloor k}$  - faktor normalizacije

Prosječna dužina  $L(k)$  se računa kao srednja vrijednost  $k$  dužina  $L_m(k)$  za  $m = 1, \dots, k$ . Ova procedura se ponavlja za svako  $k$  iz opsega od 1 do  $k_{max}$  čime se dobiva prosječna dužina za svako  $k$ . Konačno, kad se  $\ln(L(k))$  nacrti na grafu nasuprot  $\ln\left(\frac{1}{k}\right)$ , gradijent prave koja najbolje odgovara paru  $\{\ln(L(k)), \ln\left(\frac{1}{k}\right)\}$  dobiven metodom najmanjih kvadrata je estimacija fraktalne dimenzije.

Algoritam računanja HFD-a [1] :

*HFD(s(j): vremenska serija duljine N,  $k_{min}, k_{max} \in \mathbb{N}$ : doseg duljina intervala)*

*for  $k = k_{min}, \dots, k_{max}$*

*definiraj  $k$  novih vremenskih serija,  $m = 1, 2, \dots, k$ :*

$$x_m^k = \left\{ x(m), x(m+k), x(m+2k), \dots, x\left(m + \left\lfloor \frac{N-m}{k} \right\rfloor k\right) \right\}$$

*izračunaj duljinu ovih krivulja:*

$$L_m(k) = \frac{\sum_{i=1}^{\left\lfloor \frac{N-m}{k} \right\rfloor} |x(m+ik) - x(m+(i-1)k)| (n-1)}{\left\lfloor \frac{N-m}{k} \right\rfloor k}$$

*definiraj  $L(k) = \text{mean}\{L_m(k) | m = 1, 2, \dots, k\}$*

*end*

*procjeni fraktalnu dimenziju kao gradijent pravca linearne regresije točaka*

*$\ln(L(k))$  i  $\ln\left(\frac{1}{k}\right)$  za  $k = k_{min}, \dots, k_{max}$ .*

Provođenje prethodnog algoritma u Matlabu je ostvareno funkcijom *hfd()* kojoj se kao ulazni parametar zadaje samo signal za koji se želi izračunati fraktalna dimenzija. Funkcija *hfd()* se nalazi u priloženoj dokumentaciji.

### ***ANOVA i odabir značajki***

Rezultati ANOVA-e su prikazani u tablici 6.

Tablica 6.: Rezultati ANOVA-e

| Kanal | Vrijednost p |
|-------|--------------|
| 1     | 0,21527719   |
| 2     | 0,001377525  |
| 3     | 5,52102E-05  |
| 4     | 0,039575699  |
| 5     | 0,849052474  |
| 6     | 0,034870594  |
| 7     | 0,001614782  |
| 8     | 0,173471614  |
| 9     | 0,546383876  |
| 10    | 0,457137428  |
| 11    | 0,001734491  |
| 12    | 0,329467224  |
| 13    | 0,068700608  |
| 14    | 0,41762136   |
| 15    | 0,385748637  |
| 16    | 0,033036155  |
| 17    | 0,00520182   |
| 18    | 0,004218291  |
| 19    | 0,079728652  |
| 20    | 0,006299697  |

Kao značajka će se koristiti fraktalna dimenzija EEG signala izračunatog u trećem kanalu.

#### 4.2.4. Skewness

*Skewness* je mjera asimetričnosti podataka oko njegove srednje vrijednosti. Ako je *skewness* negativna to znači da su uzorci signala više raspoređeni u lijevo u odnosu na srednju vrijednost, a ako je pozitivna onda su uzorci više raspoređeni u desnu stranu. *Skewness* za normalnu razdiobu (ili bilo koju drugo savršeno simetričnu razdiobu) je nula.

*Skewness* razdiobe se definira kao:

$$s = \frac{E(x - \mu)^3}{\sigma^3} \quad (8.)$$

gdje su:

- $\mu$  - srednja vrijednost signala  $x$
- $\sigma$  - standardna devijacija od  $x$
- $E(t)$  - očekivana vrijednost od  $t$

Konkretno, *skewness* se računa pomoću jednadžbe:

$$s = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left( \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \right)^3} \quad (9.)$$

Za računanje u Matlabu koristi se implementirana funkcija *skewness()* koja ima jedan ulazni parametar, a to je signal za koji se *skewness* računa.

#### 4.2.5. Kurtosis

*Kurtosis* je mjera koja pokazuje da li je signal izdužen (šiljast) ili pljosnat. *Kurtosis* za normalnu razdiobu iznosi tri. One razdiobe koje su izduženije imaju *kurtosis* veći od tri, dok one pljosnatije imaju *kurtosis* manji od tri.

*Kurtosis* se definira kao:

$$k = \frac{E(x - \mu)^4}{\sigma^4} \quad (10.)$$

gdje su:

- $\mu$  - srednja vrijednost signala  $x$
- $\sigma$  - standardna devijacija od  $x$
- $E(t)$  - očekivana vrijednost od  $t$

Konkretno, *kurtosis* se računa pomoću jednadžbe:

$$k = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2)^2} \quad (11.)$$

Za računanje u Matlabu koristi se implementirana funkcija *kutosis()*.

### ***ANOVA i odabir značajki***

Rezultati ANOVA-e se nalaze u tablici 7.

Tablica 7.: Rezultati ANOVA-e

|       | <i>Kurtosis</i> | <i>Skewness</i> |
|-------|-----------------|-----------------|
| Kanal | Vrijednost p    | Vrijednost p    |
| 1     | 0,654181778     | 0,813888371     |
| 2     | 0,111816183     | 0,137505904     |
| 3     | 0,639769077     | 0,006860485     |
| 4     | 0,743566215     | 0,029755834     |
| 5     | 0,01678458      | 0,133216754     |
| 6     | 6,42289E-05     | 0,000120065     |
| 7     | 0,011212661     | 3,17063E-05     |
| 8     | 0,409697533     | 0,803843141     |
| 9     | 0,076738834     | 0,531868696     |
| 10    | 0,021837959     | 0,033390313     |
| 11    | 0,651083887     | 0,071965329     |
| 12    | 0,526571214     | 0,664508522     |
| 13    | 0,006076878     | 0,776735187     |
| 14    | 0,088232726     | 0,031026853     |
| 15    | 0,895455062     | 0,013418599     |
| 16    | 0,590402126     | 0,382442296     |
| 17    | 0,169234633     | 0,127497375     |
| 18    | 0,211777985     | 0,622000217     |
| 19    | 0,831903994     | 0,924185395     |
| 20    | 0,351068854     | 0,384087235     |

Iz priloženog se vidi da će se koristiti šesti kanal za računanje *kurtosisa*, a sedmi kanal za računanje *skewnessa*.

#### 4.2.6. Hjorthovi parametri

Tri su *Hjortova* parametra koji opisuju signal  $x$  duljine  $N$  i to:

##### 1. Aktivnost

Aktivnost je jednaka varijanci signala  $x$ :

$$ACTIVITY(x) = Var(x) = \sigma_x^2 = \frac{\sum_{i=1}^N (x(n) - \bar{x})^2}{N} \quad (12.)$$

gdje su:

- $\bar{x}$  - srednja vrijednost signala  $x$
- $\sigma_x$  - standardna devijacija signala  $x$

##### 2. Mobilnost

Mobilnost je mjera srednje frekvencije signala i definira se kao drugi korijen omjera aktivnosti prve derivacije signala i aktivnosti signala.

$$MOBILITY(x) = \sqrt{\frac{ACTIVITY(x')}{ACTIVITY(x)}} = \sqrt{\frac{Var(x')}{Var(x)}} = \frac{\sigma_{x'}}{\sigma_x} \quad (13.)$$

gdje je  $x'$  prva derivacija signala  $x$ .

##### 3. Kompleksnost

Kompleksnost se definira kao omjer mobilnosti prve derivacije signala i mobilnost signala.

$$COMPLEXITY(x) = \frac{MOBILITY(x')}{MOBILITY(x)} = \frac{\frac{\sigma_{x'}}{\sigma_x}}{\frac{\sigma_x}{\sigma_x}} \quad (14.)$$

gdje je  $\sigma_x''$  druga derivacija signala  $x$ .

Kompleksnost mjeri odstupanje signala od sinusnog oblika.

Provođenje prethodnog algoritma u Matlabu je ostvareno funkcijom *hjorth()* koja se koristi za računanje stacionarnih *Hjortovih* parametara i koja se nalazi u priloženoj dokumentaciji.

## ANOVA i odabir značajki

Rezultati ANOVA-a se nalaze u tablici 8.

Tablica 8.: Rezultati ANOVA-e

| Značajka | ACTIVITY     | MOBILITY     | COMPLEXITY   |
|----------|--------------|--------------|--------------|
| Kanal    | Vrijednost p | Vrijednost p | Vrijednost p |
| 1        | 0,526819527  | 0,974552393  | 0,00266755   |
| 2        | 0,005273681  | 0,004842305  | 0,152901128  |
| 3        | 0,00011993   | 0,052310336  | 0,587140858  |
| 4        | 0,008133822  | 0,134710312  | 0,763923943  |
| 5        | 0,068212591  | 0,997038722  | 0,004931429  |
| 6        | 4,64211E-06  | 0,0009355    | 2,59359E-07  |
| 7        | 2,55238E-08  | 0,003956884  | 0,002344088  |
| 8        | 1,32188E-05  | 0,625789821  | 0,001479408  |
| 9        | 0,008575589  | 0,646526873  | 0,002878421  |
| 10       | 0,003985845  | 0,311459392  | 0,005412599  |
| 11       | 1,90541E-07  | 0,831614316  | 0,881557286  |
| 12       | 0,00251456   | 0,019028783  | 0,576127648  |
| 13       | 0,00136047   | 0,747753739  | 4,68237E-05  |
| 14       | 0,005275746  | 0,077953376  | 0,042555347  |
| 15       | 1,49183E-05  | 0,354168117  | 0,623353481  |
| 16       | 2,47629E-06  | 0,769910574  | 0,061901115  |
| 17       | 0,043942448  | 0,630465746  | 0,843685567  |
| 18       | 0,006527631  | 0,789263964  | 0,655574441  |
| 19       | 0,000268067  | 0,185018629  | 0,691141307  |
| 20       | 4,61008E-05  | 0,971083879  | 0,866952717  |

Koristiti će se slijedeće značajke:

1. Parametar ACTIVITY izračunat za 7. kanal
2. Parametar MOBILITY izračunat za 6. kanal
3. Parametar COMPLEXITY izračunat za 6. kanal

Konačno, odabrano je 14 značajki koje će činiti vektor značajki pomoću kojeg će se trenirati klasifikator. Tih 14 značajki je:

1. PSD snaga  $\alpha$ -pojasa u 11. kanalu
2. PSD snaga  $\beta$ -pojasa u 15. kanalu
3. PSD snaga  $\delta$ -pojasa u 19. kanalu
4. PSD snaga  $\theta$ -pojasa u 11. kanalu
5. DWT snaga  $\alpha$ -pojasa u 11. kanalu
6. DWT snaga  $\beta$ -pojasa u 8. kanalu
7. DWT snaga  $\delta$ -pojasa u 6. kanalu
8. DWT snaga  $\theta$ -pojasa u 11. kanalu
9. fraktalna dimenzija 3. kanalu
10. *skewness* u 7. kanalu
11. *kurtosis* u 6. kanalu
12. Parametar ACTIVITY u 7. kanalu
13. Parametar MOBILITY u 6. kanalu
14. Parametar COMPLEXITY u 6. kanalu

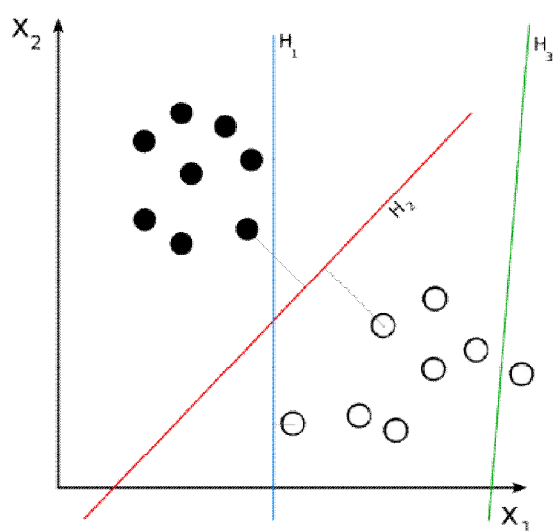
Navedene značajke daju najbolje rezultate kad se koriste zajedno. Gledajući ih pojedinačno ne uviđa se razlika između bolesnih i normalnih EEG-ova, ali zajedno SVM klasifikator uspjeva da nađe ravninu razdvajanja. Proučavanjem svake od 14 navedenih značajki pojedinačno primjetilo sa par zanimljivih stvari. Snage izračunate DWT-om u većini slučajeva, za sva četiri pojasa, poprimaju veće vrijednosti kad se računaju za normalne EEG-ove. Snage PTSP signala su niže. Isto to vrijedi i za PSD snage, a razlika je najizraženija za  $\alpha$ -pojas (značajka broj 1.). Također je primjećeno da značajke pod rednim brojevima 11., 12. i 14. poprimaju veće vrijednosti za normalne EEG-ove (ne u 100% slučajeva, ali u većini slučajeva), dok fraktalna dimenzija poprima malo veće vrijednosti za PTSP signale nego za normalne. Promatrajući značajku *skewness* bolesni EEG-ovi češće poprimaju negativne vrijednosti nego normalni.

### 4.3 SVM

Klasifikacija podataka je čest zadatak strojnog učenja. Recimo da postoje dvije klase podataka i cilj je donijeti odluku kojoj klasi pripada novi podatak koji se dovede na ulaz. Pomoću SVM-a klasifikacija se vrši tako da se kreira  $N$ -dimenzionalna ravnina (eng. *hyperplane*) koja optimalno razdvaja podatke u dvije klase. SVM modeli su slični neuronskim mrežama koji po svojoj prirodi također spadaju u algoritme strojnog učenja. SVM model se najprije treba trenirati sa skupom podataka koji se naziva skup za treniranje. Zajedno sa skupom za treniranje se mora imati određeno *a priori* znanje kojoj klasi pripada svaki od uzoraka iz skupa za treniranje na osnovu čega se dobije trenirani model koji bi trebao novi uzorak (koji nije iz skupa za treniranje) svrstati u odgovarajuću klasu.

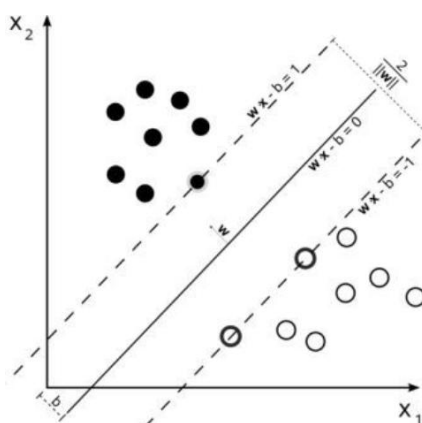
U slučaju SVM-a podatak se predstavlja i promatra kao  $p$ -dimenzionalan vektor (točka u prostoru) i želi se istrenirati model koji će dvije klase  $p$ -dimenzionalnih vektora moći razdvojiti  $(p-1)$ -dimenzionalnom ravninom. To se naziva linearnim klasifikatorom. Postoji nekoliko mogućih ravnina razdvajanja koje bi mogle klasificirati podatke, a jedna varijanta kojoj SVM teži je ona pomoću koje se postiže maksimalno razdvajanje odnosno maksimalna margina između klasa.

Dakle, SVM vrši klasifikaciju tako da konstruira  $N$ -dimenzionalnu ravninu koja optimalno razdvaja podatke u dvije kategorije (slika 17.).



Slika 17.: Razdvajanje podataka u dvije klase

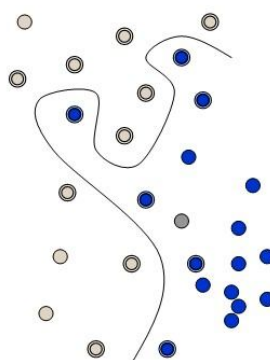
Postavlja se pitanje koja je razlika između tri pravca koji razdvajaju klase crnih i bijelih točaka. Vidljivo je da pravac H3 uopće ne razdvaja klase, tako da nam on nije interesantan, ali koja je razlika između pravaca H1 i H2 budući da oba odvajaju crne od bijelih točaka. Razlika je u tome da razdvajanjem klasa pravcem H2 se postižu maksimalne margine, odnosno udaljenosti od najbližeg člana pojedine klase do pravca H2. Stvar će možda biti jasnija uvođenjem još jedne slike (Slika 18.)



Slika 18.: Margine

Na ovoj slici pomoćni pravci (vektori potpore – eng. *support vectors*) su nacrtani iscrtkanom linijom i oni prolaze kroz onog člana klase koji je najbliži središnjem pravcu (punom linijom) koji je zapravo klasifikator klase.

Ovaj primjer je najjednostavniji koji se može pojaviti i najčešće se ne pojavljuje u stvarnim situacijama. Klasifikator na prethodne dvije slike je linearan, dok se mogu pojaviti i neki koji nisu linearni, npr. Slika 19.



Slika 19.: Nelinearan klasifikator

Najjednostavniji je linearni SVM pa ćemo njega detaljno opisati u nastavku da se uvede u svijet matematike koja se odigrava pri SVM učenju. Svi složeniji SVM-ovi od linearnog se zasnivaju na istom principu kao što će biti objašnjeno poslije.

Uzmimo skup za treniranje  $D$  s  $n$  točaka oblika

$$D = \{(x_i, y_i) | x_i \in \mathbb{R}^p, y_i \in \{-1, 1\}\}_{i=1}^n$$

gdje

- $y_i$  poprima vrijednost -1 ili 1 pokazujući time klasu kojoj pripada točka  $x_i$ , a
- svaki  $x_i$  je  $p$ -dimenzionalan realan vektor.

Cilj je pronaći maksimalnu marginu koja odvaja točke za koje vrijedi  $y_i = 1$  ili  $y_i = -1$ . Bilo koja ravnina razdvajanja se može zapisati kao skup točaka  $x$  koje zadovoljavaju jednadžbu

$$w \cdot x - b = 0 \quad (15.)$$

gdje  $\cdot$  označava skalarni produkt vektora.

Parametar  $\frac{b}{\|w\|}$  predstavlja udaljenost ravnine razdvajanja od ishodišta.

Žele se odabrati parametri  $w$  i  $b$  takvi da se maksimizira margina, odnosno, udaljenost između paralelnih potpornih vektora, a da se pri tom odvajaju podaci.

Potporni vektori se mogu opisati jednadžbama:

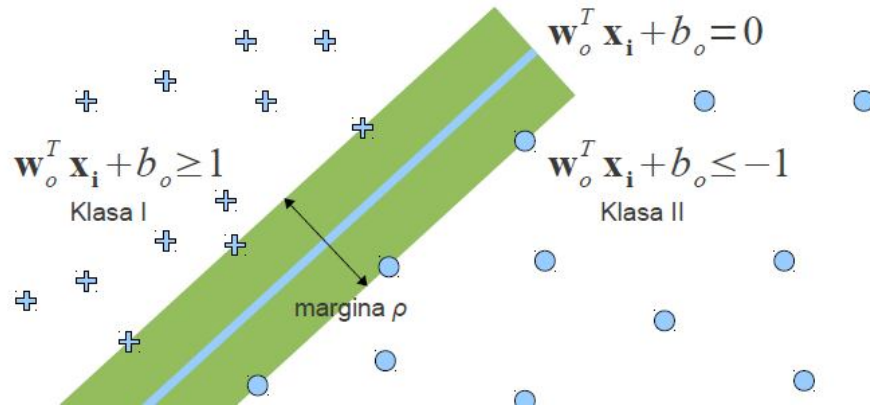
$$w \cdot x - b = 1 \quad (16.)$$

$$w \cdot x - b = -1 \quad (17.)$$

Ako su klase linearno separabilne onda se mogu odabrati dva potporna vektora tako da nema točaka između njih i onda maksimizirati udaljenost između njih. Ako se koriste jednadžbe 16. i 17. za opis vektora potpore dobije se da je udaljenost između njih, odnosno, da je širina margine  $\frac{2}{\|w\|}$ . Da bi maksimizirali širinu margine potrebno je minimizirati  $\|w\|$ . Također, da bi spriječili pojavu podataka (točaka) unutar margine dodaje se novi uvjet za svaki  $x_i$  (slika 20.):

$$w \cdot x_i - b \geq 1 \text{ za } x_i \text{ iz prve klase} \quad (18.)$$

$$w \cdot x_i - b \leq -1 \text{ za } x_i \text{ iz druge klase} \quad (19.)$$



Slika 20.: Klase podataka

Jednadžbe 18. i 19. se mogu zapisati kao:

$$y_i(w \cdot x_i - b) \geq 1 \quad (20.)$$

Kao što je spomenuto, maksimizacija margine ovisi o minimiziranju  $\|w\|$ . To stvara problem jer je to operacija modula koja u sebi sadrži korjenovanje koje je računski zahtjevno. Zbog toga se primjenjuju razni postupci optimizacije.

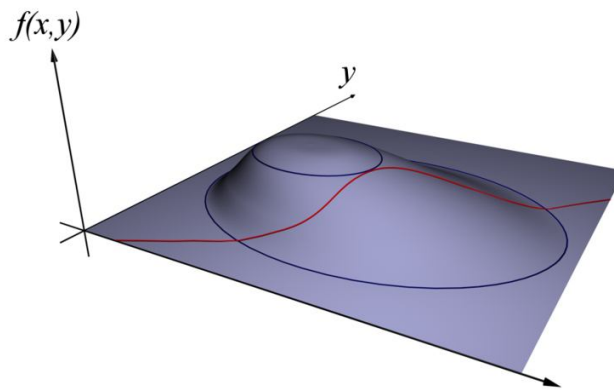
Problem korjenovanja se rješava substitucijom:

$$\|w\| \rightarrow \frac{1}{2} \|w\|^2$$

pri čemu se konačni rezultat neće promijeniti.

Da bi našli minimum od  $\frac{1}{2} \|w\|^2$  uvode se *Lagrangeovi* multiplikatori  $\alpha$ .

U području matematičke optimizacije, metoda *Lagrangeovih* multiplikatora pruža strategiju za pronalazak lokalnih maksimuma i minimuma neke funkcije. Tako npr. ako se traži maksimum funkcije  $f(x, y)$  koja ima neko ograničenje (uvjet) koji može biti algebarska jednadžba kao npr.  $g(x, y) = c$  kao na slici 21.



Slika 21.: Funkcija  $f(x,y)$  i uvjet  $g(x,y)=c$

Budući da je gradijent funkcije okomit na konture te funkcije može se reći da su gradijenti funkcija  $f$  i  $g$  paralelni. Stoga, traže se točke  $(x, y)$  gdje je  $g(x, y) = c$  i

$$\nabla_{x,y} f = -\alpha \nabla_{x,y} g$$

gdje su:

- $\nabla_{x,y} f = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$  - gradijent funkcije  $f(x, y)$
- $\nabla_{x,y} g = \left( \frac{\partial g}{\partial x}, \frac{\partial g}{\partial y} \right)$  - gradijent funkcije  $g(x, y)$

Konstanta  $\alpha$  je potrebna jer iako su dva gradijenta paralelna njihove magnitude generalno nisu jednake.

Kad se svi ovi uvjeti ukomponiraju u jednu jednadžbu dobije se tzv. pomoćna funkcija:

$$\Lambda(x, y, \alpha) = f(x, y) + \alpha(g(x, y) - c) \quad (21.)$$

Analogno prethodnom izvodu želi se dobiti pomoćna funkcija za konkretnu primjenu rješavanja optimizacijskog problema SVM-a. Uzimajući u obzir slijedeće:

- funkcija koja se želi minimizirati je  $\frac{1}{2} \|w\|^2$
- uvjet na tu funkciju je  $y_i(w \cdot x_i - b) \geq 1$

slijedi:

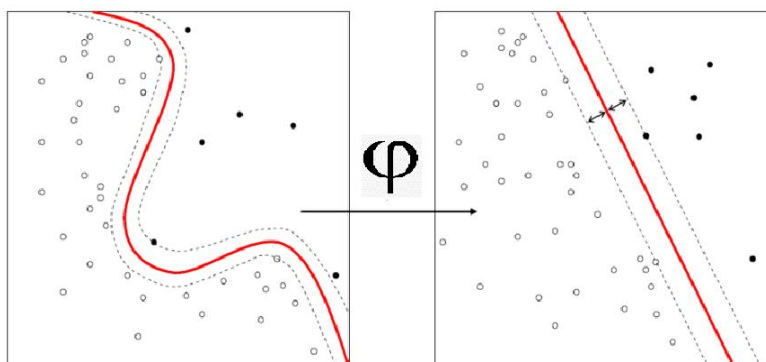
$$\Lambda(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i [y_i(w \cdot x_i - b) - 1] \quad (22.)$$

Metoda *Lagrangeovih* multiplikatora se onda svodi na rješavanje parcijalnih derivacija izjednačenih s nulom:

|                                           |                                           |                                                |
|-------------------------------------------|-------------------------------------------|------------------------------------------------|
| $\frac{\partial \Lambda}{\partial w} = 0$ | $\frac{\partial \Lambda}{\partial b} = 0$ | $\frac{\partial \Lambda}{\partial \alpha}$     |
| $w - \sum_{i=1}^n \alpha_i y_i x_i = 0$   | $\sum_{i=1}^n \alpha_i y_i = 0$           | $\sum_{i=1}^n [y_i (w \cdot x_i - b) - 1] = 0$ |
| $w = \sum_{i=1}^n \alpha_i y_i x_i$       |                                           | $\sum_{i=1}^n [y_i (w \cdot x_i - b)] = 1$     |

Samo nekoliko  $\alpha_i$  će biti različiti od nule i upravo ti *Lagrangeovi* multiplikatori koji su različiti od nule automatski odabiru potpore vektore. Odgovarajući  $x_i$  su vektori potpore koji leže na margini i zadovoljavaju jednadžbu  $[y_i (w \cdot x_i - b)] = 1$ .

Izvorni problem klasifikacije se može definirati u nekom konačnom dimenzionalnom prostoru, ali često se događa da klase koje se žele klasificirati nisu linearno separabilne u tom prostoru. Upravo zbog toga je predloženo da se originalni ulazni prostor mapira u prostor više dimenzionalnosti. Ta ideja potiče od *Coverovog* teorema koji kaže da prelaskom u višedimenzionalni prostor, raste vjerojatnost linearne separabilnosti. Osnovna ideja se sastoji iz dva koraka. Najprije se vrši nelinearno mapiranje ulaznog prostora u novi prostor značajki više dimenzionalnosti i zatim se konstruira optimalna ravnina razdvajanja u tom novom (višedimenzionalnom) prostoru značajki (slika 22.).



Slika 22.: Konstrukcija optimalne ravnine razdvajanja u novom prostoru značajki

Nelinearni klasifikatori se javljaju onda kad klase nisu linearno separabilne i nije moguće između njih povući linearni klasifikator. Kreiranje nelinearnih klasifikatora se vrši korištenjem *kernel* funkcija u svrhu maksimiziranja margine razdvajanja. Algoritam pronalaska klasifikatora je isti kao i kod linearnih osim što se u ovom slučaju svaki skalarni produkt zamjeni s nelinearnom *kernel* funkcijom.

Dakle, jednačba kojom se opisuje ravnina razdvajanja poprima novi izgled i to:

$$\sum_{j=1}^m w_j \varphi_j(x) + b = 0 \quad (23.)$$

Iz ove jednačbe se vidi da je jedina razlika u usporedbi sa prethodnom jednačbom koja je opisivala linearni klasifikator ta da se umjesto  $x$  koristi  $\varphi(x)$  što zapravo predstavlja vektor  $x$  u novom prostoru značajki.

Cijeli postupak određivanja klasifikatora i *Legendreovih* multiplikatora je identičan onom za linearni klasifikator uz substituciju  $x \rightarrow \varphi(x)$ .

Budući da se *kernel* funkcije koriste onda kad klasifikator nije linearan i prelazi se u novi prostor značajki onda je *kernel* funkcija direktno povezana sa transformacijom  $\varphi$  pri čemu je

$$K(x_i, x_j) = \varphi(x_i) \varphi(x_j) \quad (24.)$$

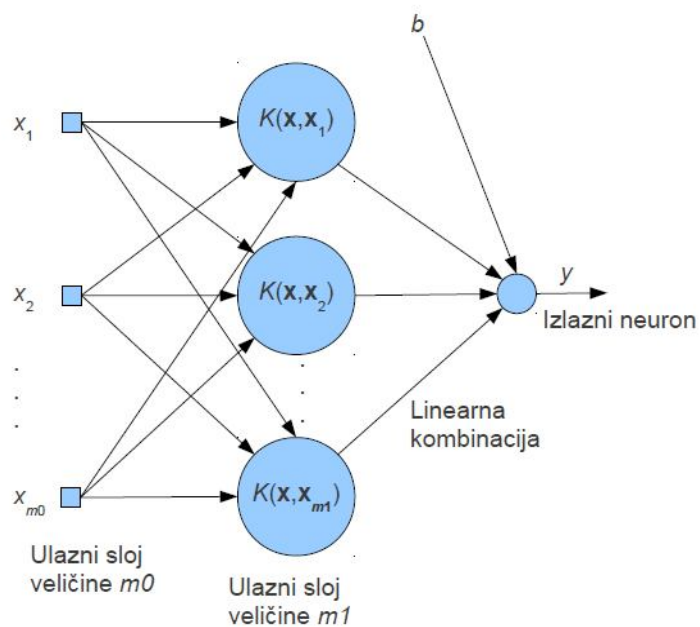
i tako dobivena funkcija  $K$  se naziva jezgrom unutarnjeg produkta. Postoji određena sloboda u izboru funkcije jezgre, ali je potrebno da zadovoljava *Mercerov* teorem koji kaže da je  $K(x_i, x_j)$  simetrična funkcija jezgre definirana na zatvorenim intervalima od  $x_i$  i  $x_j$  ako se može rastaviti na slijedeći niz:

$$K(x_i, x_j) = \sum_{i=1}^{\infty} \alpha_i \varphi_i(x_i) \varphi_i(x_j) \quad (25.)$$

Korištenjem *kernelovih* funkcija se omogućava da algoritam pronađe ravninu razdvajanja sa maksimalnom marginom u transformiranom prostoru značajki. Neki od najkorištenijih *kernela* su:

- Polinomni homogeni *kernel* –  $K(x_i, x_j) = (x_i \cdot x_j)^d$
- Polinomni nehomogeni *kernel* –  $K(x_i, x_j) = (x_i \cdot x_j + 1)^d$
- Gaussov radijalni *kernel* -  $K(x_i, x_j) = e^{\left(-\frac{1}{2\sigma^2}\|x_i - x_j\|^2\right)}$
- Hiperbolični *kernel* –  $K(x_i, x_j) = \tanh(\kappa x_i \cdot x_j + c)$

Konačno, arhitektura SVM-a se može prikazati dijagramom na slici 23.



Slika 23.: Arhitektura SVM-a

## Implementacija

U Matlabu postoje implementirane funkcije za treniranje i testiranje SVM-a kao dio *Bioinformatics Toolbox*-a. Za treniranje se koristi funkcija *svmtrain()* kojoj se kao parametri zadaju signali za treniranje. Pored signala za treniranje, ukoliko se ne žele koristiti defaultne vrijednosti, moguće je zadati i dodatne parametre. Dodatni parametar koji se koristio u ovom radu je '*kernel\_function*' pomoću kojeg se mijenja *kernel* funkcija koja se koristi u treniranju klasifikatora. Za treniranje klasifikatora najboljom se pokazala *kernel* funkcija *quadratic*<sup>4</sup> koja se opisuje izrazom:

$$K(x_i, x_j) = (1 + x_i \cdot x_j)^2 \quad (26.)$$

Za testiranje dobivenog SVM klasifikatora koristi se funkcija *svmcclassify()* kojoj se zadaju dva parametra od kojih je jedan trenirani klasifikator dobiven funkcijom *svmtrain()* i signal koji se želi klasificirati.

U ovom radu za treniranje SVM-a korišteno je ukupno 37 EEG signala od toga 22 PTSP i 16 normalnih. Klase su označene sa vrijednostima „0“ što znači bolestan i „1“ što znači zdrav (moguće je i označavanje klasa sa „+1“ i „-1“).

---

<sup>4</sup> Pored *quadratic* testirani su još linearni, polinomni i Gausov radijalni *kernel*.

## 5. Rezultati i implementacija

Tip I pogreška (pogreška prve vrste) i pogreška tip II (pogreška druge vrste) su precizni tehnički pojmovi koji se koriste u statistici za opisivanje određenih nedostataka u procesu testiranja gdje je točna nulta hipoteza pogrešno odbijena (tip I) ili kad je pogrešna nulta hipoteza prihvaćena (tip II). Ovi testovi se primjenjuju u testiranju klasifikatora koji se realizirao i koristio u procesu klasifikacije EEG signala.

U teoriji statističkog testiranja najprije je potrebno postaviti nultu hipotezu koja obično odgovara nekom uobičajenom stanju u prirodi kao npr. „osoba je zdrava“, „proizvod je čitav“, „optuženik je nevin“. Alternativna hipoteza je negacija nulte hipoteze kao npr. „osoba je bolesna“, „proizvod je slomljen“, „optuženik je kriv“. Ovdje se testiraju rezultati klasifikacije osoba oboljelih od PTSP-a na osnovu EEG signala tako da će nulta hipoteza za ovaj slučaj biti upravo  $H_0 = \text{„pacijent je zdrav“}$ . Rezultat testa može biti negativan u odnosu na nultu hipotezu (bolestan, slomljen, kriv) ili pozitivan (zdrav, nije slomljen, nevin). Ako rezultat testa odgovara stvarnosti onda je donesena ispravna odluka, a ako ne odgovara stvarnosti onda se desila pogreška. Kao što je već rečeno, razlikuju se dva tipa pogrešaka i to tip I i tip II.

Pogreška prve vrste se javlja kad je nulta hipoteza istinita, ali je odbijena, tj. kad dovedeni uzorak za testiranje pripada zdravoj osobi, a klasifikator kao rezultat vrati da je pacijent bolestan. Pogreška prvog tipa se još naziva lažno pozitivna (eng. FP – *False Positive*) i odlika dobrog klasifikatora bi bila da je FP što manji, tj. da što manje zdravih osoba klasificira kao bolesne.

Pogreška druge vrste se javlja kad je nulta hipoteza pogrešna, ali se prihvaća kao istinita. Konkretno, dovedeni uzorak za testiranje pripada bolesnoj osobi (PTSP), a klasifikator vraća vrijednost „zdrav“. Pogreška drugog tipa se još naziva lažno negativna (eng. FN – *False Negative*) i za dobar klasifikator će vrijediti da je FN, ujedno kao i FP, što niži, tj. da što manje bolesnih osoba klasificira kao zdrave. Naravno, idealno bi bilo da su FN i FP jednaki nuli, tj. da nema pogrešnih klasifikacija, ali to se rijetko dešava u praksi.

Pored FN i FP još se klasificiraju dvije vrste pogrešaka koje zapravo označavaju točan rezultat (tablica 9.) i to:

- TN (eng. *True Negative*) – kad se zdrava osoba klasificira kao zdrava
- TP (eng. *True Positive*) – kad se bolesna osoba klasificira kao bolesna

Dobar klasifikator će imati visoke vrijednosti za TN i TP, tj. veliki postotak točno klasificiranih uzoraka.

Tablica 9.: Pogreške

|                           | Nulta hipoteza ( $H_0$ ) je istinita | Nulta hipoteza ( $H_0$ ) nije istinita |
|---------------------------|--------------------------------------|----------------------------------------|
| Odbijena nulta hipoteza   | Tip I<br>False Positive – FP         | Točan rezultat<br>True Positive - TP   |
| Prihvaćena nulta hipoteza | Točan rezultat<br>True Negative - TN | Tip II<br>False Negative – FN          |

Prilikom istraživanja i odabiranja značajki koje bi se koristile za klasifikaciju testiralo se nekoliko vrste i kombinacija značajki. Najprije su se koristile snage dobivene PSD-om i DWT-om posebno za svaki pojas, pa kombinacije pojaseva. Najprije su se koristile značajke izračunate za sve kanale jer se htjelo utvrditi da li se dobiju bolji rezultati korištenjem svih kanala ili korištenjem samo onih kanala dobivenih ANOVA-om.

Budući da rezultati variraju mijenjanjem skupa uzoraka za treniranje, testiranje se provodilo na osam različitih varijanti matrice A koja se koristi za treniranje SVM-a. Dakle, korišteno je osam različitih skupova za treniranje pri tom pazeći da se uzorci za testiranje uvijek razlikuju od onih za treniranje. U svih osam varijacija veličina skupa za treniranje i testiranje je jednaka samo su se signali umutar skupina mijenjali. Mijenjanjem konfiguracija pronašle su se one skupine uzoraka za treniranje za koje je klasifikacija bila 100% točna, ali upravo zbog ovisnosti točnosti klasifikacije o skupu za treniranje na ovaj način se dobila određena srednja vrijednost točnosti. U tablici 10. su prikazani dobiveni rezultati testiranja različitih značajki gdje se neka, nazovimo ju, ukupna točnost označavala sa  $T$  što predstavlja točnost klasifikacije izražena u postocima izračunate pomoću izraza:

$$T = \frac{\text{broj\_pogrešno\_klasificiranih\_signala}}{\text{ukupan\_broj\_signala\_za\_testiranje}} \times 100\% \quad (27.)$$

Tablica 10.: Rezultati testiranja

| Vektor značajki                                                              | FP     | FN     | TP     | TN     | T      |
|------------------------------------------------------------------------------|--------|--------|--------|--------|--------|
| PSD snaga $\alpha$ -pojasa                                                   | 48,21% | 7,29%  | 92,71% | 51,79% | 77,63% |
| PSD snaga $\beta$ -pojasa                                                    | 39,29% | 17,71% | 82,29% | 60,71% | 74,34% |
| PSD snaga $\delta$ -pojasa                                                   | 60,71% | 20,83% | 79,17% | 39,29% | 64,47% |
| PSD snaga $\theta$ -pojasa                                                   | 64,29% | 47,92% | 52,08% | 35,71% | 46,1%  |
| PSD snaga svih pojaseva                                                      | 51,79% | 4,17%  | 95,83% | 48,21% | 78,29% |
| PSD snaga $\alpha, \beta$ i $\delta$ -pojasa                                 | 41,07% | 2,08%  | 97,92% | 58,93% | 83,55% |
| DWT snaga $\alpha$ -pojasa                                                   | 42,86% | 2,08%  | 97,92% | 57,14% | 82,89% |
| DWT snaga $\beta$ -pojasa                                                    | 42,86% | 11,46% | 88,54% | 57,14% | 76,97% |
| DWT snaga $\delta$ -pojasa                                                   | 17,86% | 3,13%  | 96,87% | 82,14% | 91,45% |
| DWT snaga $\theta$ -pojasa                                                   | 46,43% | 0      | 100%   | 53,57% | 82,89% |
| DWT snaga svih pojaseva                                                      | 58,93% | 2,08%  | 97,92% | 41,07% | 76,97% |
| DWT snaga $\alpha$ -pojasa i<br>PSD snaga $\alpha, \beta$ i $\delta$ -pojasa | 37,5%  | 1,04%  | 98,96% | 62,5%  | 85,53% |

Iz gornje tablice se vidi da se klasifikator najprije testirao koristeći značajku snage  $\alpha$ -pojasa dobivena PSD-om koja se pokazala dobrom za klasifikaciju PTSP signala što se zaključuje na osnovu niske vrijednosti FN i/ili visoke vrijednosti TP za tu značajku, ali vrijednost FP je relativno visoka što nam govori da klasifikacija nije jednako uspješna i za zdrave signale što nam sveukupnu klasifikaciju čini neučinkovitom. Dalje je ispitivanje provedeno koristeći samo snagu  $\beta$ -pojasa, pa zatim samo snagu  $\delta$ -pojasa. Te značajke nisu zadovoljile ni jedan kriterij, tj. klasifikacija je loša u svakom pogledu. Najgorom značajkom se pokazala snaga  $\theta$ -pojasa koja je ubjedljivo dala najgoru klasifikaciju od svih ostalih. Zatim se došlo na ideju da se kao značajka uzme kombinacija pojaseva što se pokazalo kao dobrom idejom jer koristeći snage svih pojaseva se dobio bolji rezultat nego kad su se koristili posebno. Zatim se ista stvar ponovila za snage DWT pojaseva i kao što se iz tablice vidi snaga  $\delta$ -pojasa izračunata iz DWT koeficijenata se pokazala najboljim klasifikatorom do sad. Kasnije se čak probala i kombinacija snaga dobivenih PSD-om i DWT-om što se isto pokazalo dobrom idejom. Općenito, sve snage kao značajke su

dale puno bolju klasifikaciju PTSP signala (označeno zelenom bojom u tablici) nego normalnih.

Iz ovih testiranja se naučilo da se kombiniranjem različitih značajki mogu dobiti bolji rezultati klasifikacije nego koristeći jednu značajku ili značajke koje su po svojoj prirodi slične. Isto tako se naučilo da preveliko miješanje značajki ne mora nužno dati dobre rezultati i opet se dolazi do zaključka da je odabir značajki najbitniji i ujedno najteži dio u klasifikaciji signala.

Koristeći značajke izračunate za cijeli EEG signal (sve kanale) se možda moglo utvrditi u kojem od karakterističnih valova se dešava najveća promjena na osnovu koje bi se mogla izvršiti klasifikacija, ali htjelo se pokazati još detaljnije, u kojim kanalima bi se te promjene trebale gledati. Tako se došlo do predprocesiranja značajki pomoću ANOVA-e i pronalaska 14 konačnih značajki (vidi poglavlje „Značajke“).

Za odabrani skup značajki, testiranjem klasifikatora dobili su se slijedeći rezultati (tablica 11.):

Tablica 11.: Rezultati klasifikacije za ANOVA značajke

| Vektor značajki            | FP    | FN    | TP     | TN     | T     |
|----------------------------|-------|-------|--------|--------|-------|
| Značajke dobivene ANOVA-om | 7,14% | 2,08% | 97,92% | 92,86% | 96,1% |

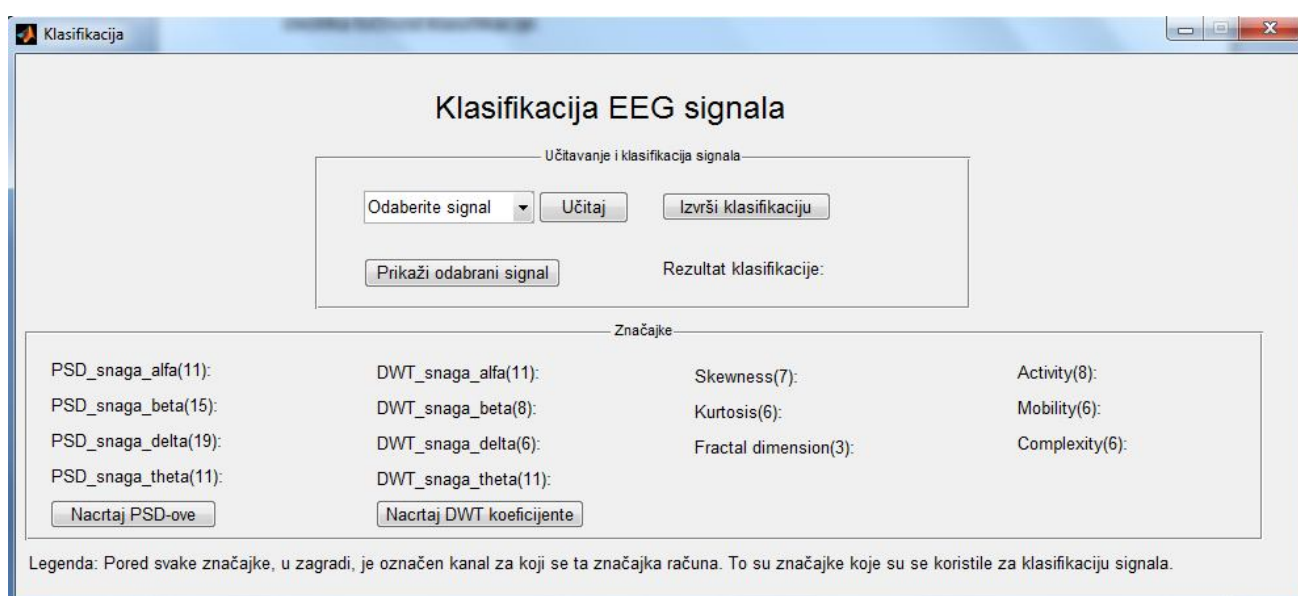
Iz priloženog se vidi da su dobivene sve željene karakteristike klasifikatora, niske vrijednosti za FP i FN, a visoke vrijednosti za TP i TN, te je čak i sveukupna točnost klasifikacije jednaka 96,1% što je vrlo zadovoljavajući rezultat. Zaključuje se da je predprocesiranje ANOVA-om itekako urodilo plodom i bez toga bi se teško dobila ovolika točnost klasifikacije.

## 5.1. Implementacija

Da bi rješenje ovog rada bilo pristupačnije korisniku i intuitivnije za korištenje implementiran je GUI koji je prikazan na slici 24. U nastavku će se, korak po korak, prikazati dobiveni rezultati i sva svojstva GUI-a te njihova implementacija.

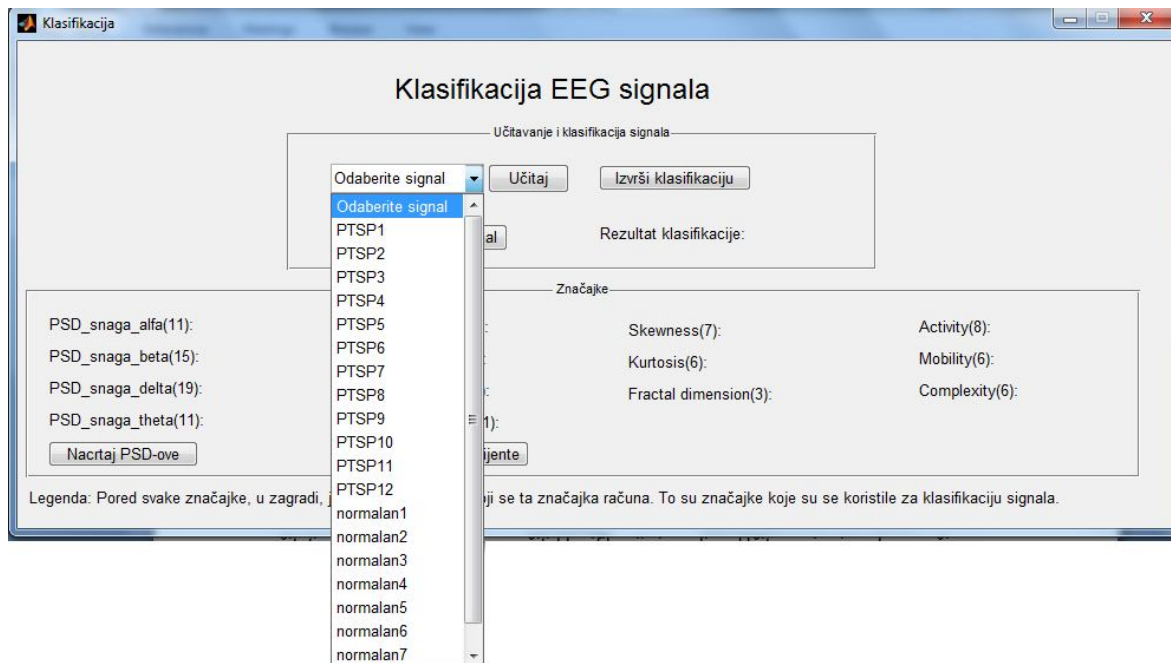
Za prikazivanje primjera u nastavku korišten je Matlab 7.11.0 (R2010b). Pokretanje GUI-a se izvodi utipkavanjem slijedeće naredbe u komandni prozor:

```
>> Klasifikacija
```



Slika 24.: GUI

Nakon ukucavanja navedene naredbe i pojavljivanja glavnog GUI prozora potrebno je učitati signal za koji će se vršiti klasifikacija i analiza. Učitavanje signala se vrši tako da se iz padajućeg menija u lijevom gornjem kutu (slika 25.) odabere signal koji se želi učitati.

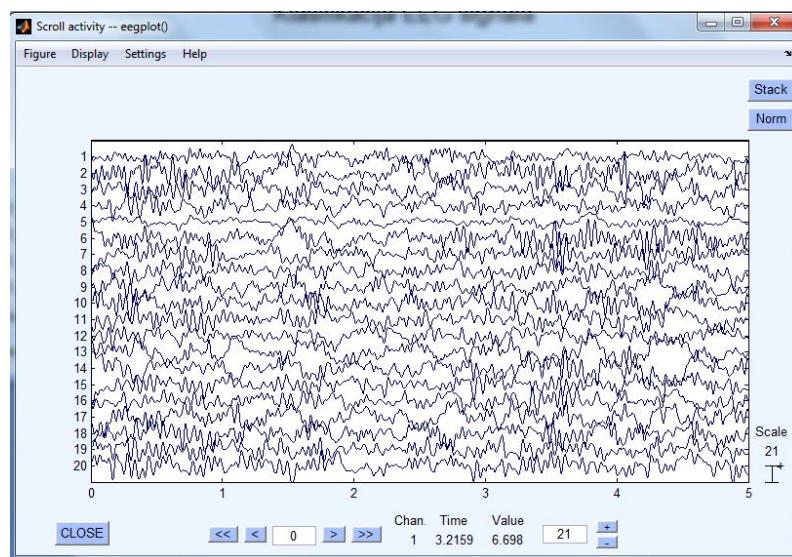


Slika 25.: Odabir signala u GUI-u

Pritiskom na gumb „*Učitaj*“ signal se učitava. Nakon što se odabere i učitava signal može se izvršiti jedna od dvije stvari:

1. prikazati odabrani signal po kanalima (svih 20)
2. pokrenuti klasifikaciju.

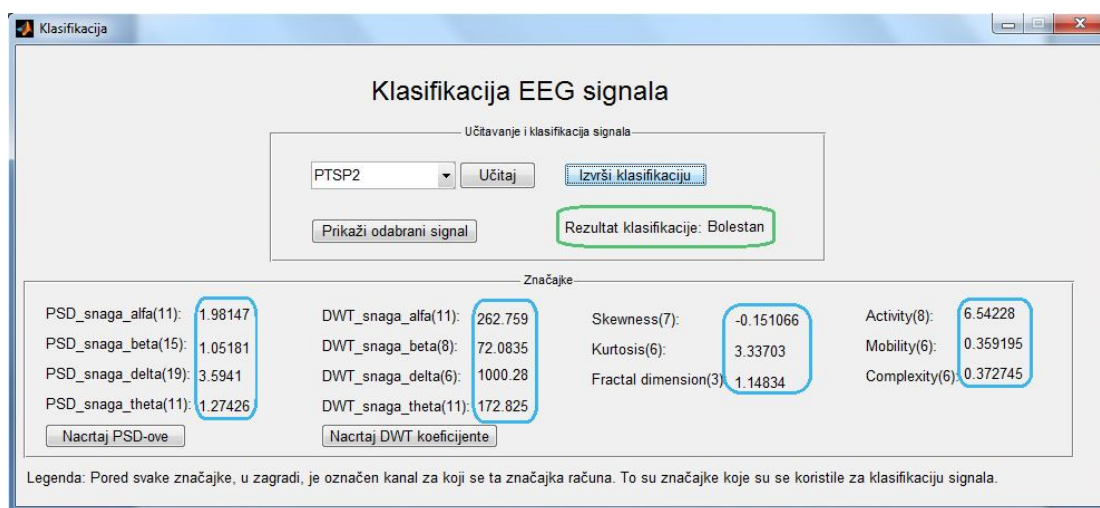
Ukoliko se odlučimo za prikaz signala to se postiže pritiskom na gumb „*Prikaži odabrani signal*“ nakon čega se dobije slijedeća slika (za npr. signal PTSP2):



Slika 26.: Prikazan signal PTSP2

Za prikaz signala na prethodni način koristi se funkcija *eegplot()* kojoj je jedini obavezni ulazni parametar EEG signal koji se želi prikazati. EEG signal može biti 2D matrica dimenzija *kanali x vremenske vrijednosti* ili 3D matrica dimenzija *kanali x vremenske vrijednosti x epohe*. Korišteni EEG signali su sačuvani kao 2D matrice gdje je EEG signal učitani u retke po kanalima, tj. koliko ima kanala toliko će matrica imati redaka, a stupaca ima onoliko koliko ima izmjerenih uzoraka.

Ukoliko se pak odlučimo za pokretanje klasifikacije to činimo pritiskom na gumb „Izvrši klasifikaciju“ nakon čega se dobije slijedeće (slika 27.):



Slika 27.: Izgled GUI-a nakon provedene klasifikacije

Na slici 27. je zelenim označen dobiveni rezultat klasifikacije koji, za odabrani signal PTSP2, glasi „*Bolestan*“ što znači da je klasifikacija točna. U slučaju da je signal klasificiran kao zdrav onda se na mjestu gdje na slici 27. piše „*Bolestan*“ ispiše „*Zdrav*“.

Pored rezultata klasifikacije ispisuju se sve izračunate značajke za odabrani signal (označeno plavim na slici 27.). Prikazane značajke su one koje se koriste u klasifikaciji signala. Pored naziva svake značajke u zagradi stoji broj kanala za koji je ta značajka izračunata (dobiveno ANOVA-om).

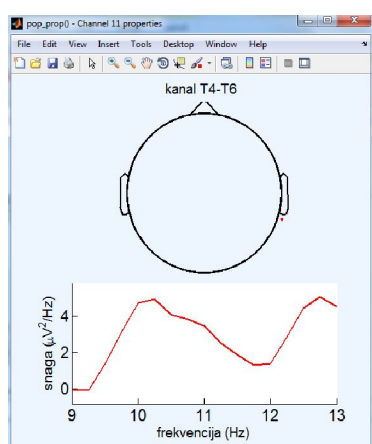
Klasifikacija se izvršava pozivanjem funkcije *klasificiraj()* koja se poziva na slijedeći način:

```
[rezultat Sample] = klasificiraj(signal)
```

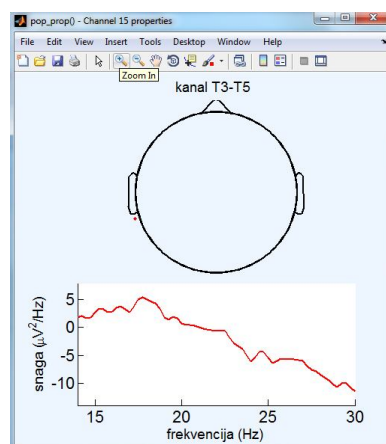
gdje su:

- rezultat – rezultat klasifikacije u obliku stringa „Bolestan“ ili „Zdav“
- Sample – izračunate značajke za zadani signal
- signal – EEG signal koji se želi klasificirati.

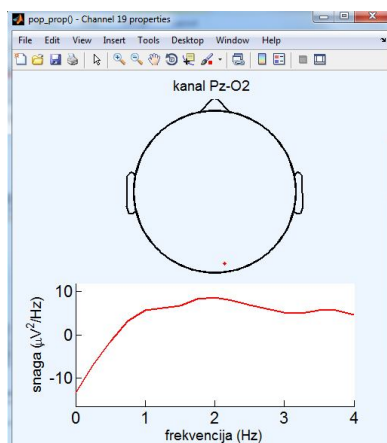
Nakon što se izvršila klasifikacija moguće je nacrtati PSD za sva četiri frekvencijska područja koja odgovaraju valovima  $\alpha$ ,  $\beta$ ,  $\delta$  i  $\theta$ . To se izvršava pritiskom na gumb „Nacrtaj PSD-ove“ nakon čega se kao rezultat dobiju slijedeće četiri slike:



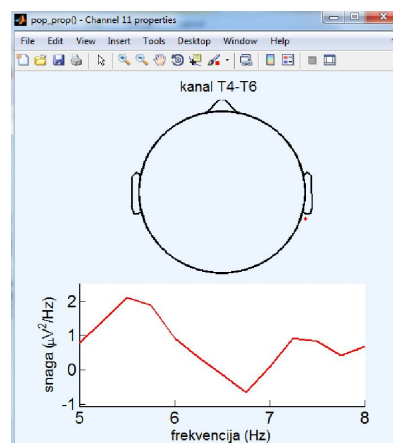
Slika 28.: PSD  $\alpha$ -pojasa za kanal T4-T6 (11. kanal)



Slika 29.: PSD  $\beta$ -pojasa za kanal T3-T5 (15. kanal)



Slika 30.: PSD  $\delta$ -pojasa za kanal Pz-O2 (19. kanal)



Slika 31.: PSD  $\theta$ -pojasa za kanal T4-T6 (11. kanal)

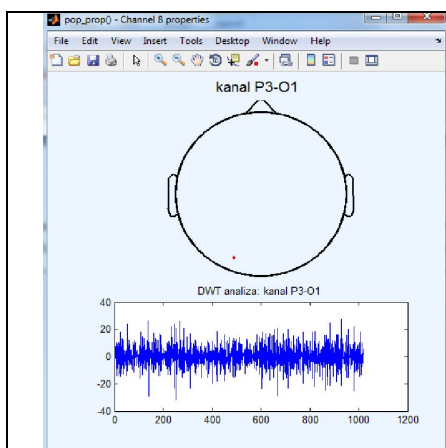
Na svakoj slici je u gornjem dijelu prikazan položaj elektrode koja mjeri odabrani kanal, a u donjem dijelu se nalazi PSD za frekvencijsko područje koje nas zanima.

Na slici 28. je prikazan PSD u frekvencijama od 9 Hz do 13 Hz što odgovara frekvencijama  $\alpha$ -valova. Značajka koja se koristila u klasifikaciji je PSD snaga  $\alpha$ -pojasa u 11. kanalu što je upravo i prikazano slikom 28. U gornjem dijelu slike piše kanal T4-T6 što je pravi naziv tog kanala.

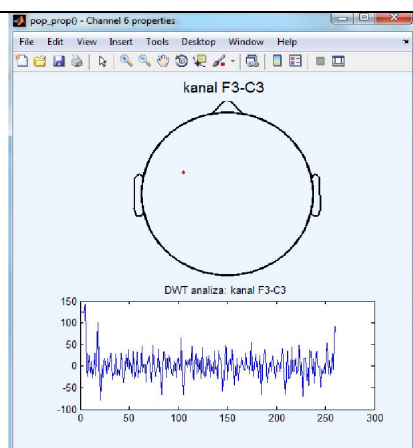
Na slici 29. je prikazan PSD od 14Hz do 30Hz za kanal T3-T5 što odgovara PSD-u  $\beta$ -valova u 15. kanalu. Analogno vrijedi i za ostale dvije značajke.

Crtanje ovih slika je izvršeno pomoću funkcije *pop\_prop()* koja je implementirana u *EEGlab toolboxu*. Funkciji *pop\_prop()* je, pored EEG signala, potrebno dati i frekvencijsko područje za koje se želi nacrtati odgovarajući PSD.

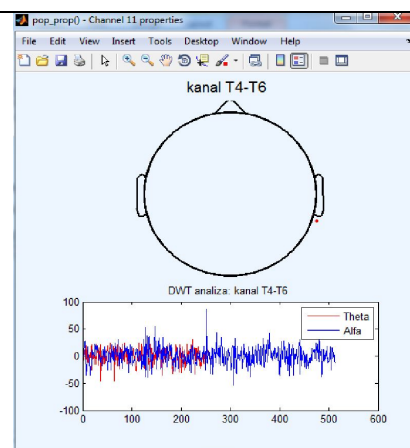
Pored crtanja PSD-a omogućeno je i crtanje DWT koeficijenata iz kojih se računala snaga. Pritiskom na gumb „Nacrtaj DWT koeficijente“ dobije se slijedeće:



Slika 32.: DWT koeficijenti  $\beta$ -pojasa



Slika 33.: DWT koeficijenti  $\delta$ -pojasa



Slika 34.: DWT koeficijenti  $\alpha$  i  $\theta$ -pojasa

Na slici 32. su prikazani DWT koeficijenti izračunati za osmi kanal (P3-O1) koji odgovaraju  $\beta$ -pojasu, a na slici 33. za šesti kanal (F3-C3) koji odgovaraju  $\delta$ -pojasu. DWT koeficijenti jedanaestog kanala za  $\alpha$  i  $\theta$ -pojas su prikazani na slici 34. Izračunata snaga svih pojedinih kanala se može vidjeti na slici 27. u plavim označenim kvadratima.

Za crtanje slika 32., 33. i 34. korištena je funkcija *pop\_prop2()* koja je u suštini ista kao i prethodno opisana funkcija *pop\_prop()* samo što je dio za crtanje PSD-a zamjenjen sa kodom za crtanje DWT koeficijenata.

## 6. Zaključak

U ovom radu je opisana razvijena metoda za klasifikaciju EEG signala. Cilj istraživanja i pronalaska odgovarajuće metode je bio pronalazak najboljeg skupa značajki i klasifikatora koji će uspješno razlikovati normalne od PTSP EEG-ova. Koristeći kombinaciju različitih značajki i SVM klasifikatora postigla se točnost klasifikacije od 96%. Iako se danas sve više razvijaju automatski sustavi za dijagnosticiranje raznih bolesti na osnovu biomedicinskih signala ovi sustavi ipak ne mogu raditi samostalno bez prisustva liječnika, a to vrijedi i za ovaj sustav. PTSP je takva bolest koja se u većini slučajeva utvrđuje psihoterapijom gdje je subjektivni doživljaj liječnika od velikog značaja u postavljanju dijagnoze. Pojavila se želja da se pomoću EEG signala proba ustanoviti neka razlika između PTSP i ne-PTSP EEG-ova što bi koristilo za sustave automatske detekcije, ali opet samostalan rad za takve sustave je nemoguć jer EEG varira od osobe do osobe i neke osobe mogu prirodno imati takav EEG koji bi neki klasifikator klasificirao kao bolestan, a zapravo nije nego ta osoba ima specifičan EEG signal.

Istraživanje provedeno u ovom radu dovodi do zaključka da se na osnovu značajki izvučenih iz EEG-a može uspostaviti razlika između PTSP i normalnih EEG-ova i pomoću statističkih alata kao što je ANOVA su se specificirali kanali u kojima su razlike najznačajnije.

Područje biomedicinskih signala i sustava je u sve većem zamahu i moguće je predvidjeti napredak u svim područjima istraživanja pa tako i u klasifikaciji PTSP-a.

## 7. Literatura

- [1] Paivinen, N., Lammi, S., Pitkanen, A., Nissinen, J., Penttonen M., Gronfors, T. "*Epileptic seizure detection: A nonlinear viewpoint*". Computer Methods and Programs in Biomedicine (2005.) 79, str. 151—159.
- [2] Mirowski, P., Madhavan, D., LeCun, Y., Kuzniecky, R. "*Classification of Patterns of EEG Synchronization for Seizure Prediction*". Courant Institute of Mathematical Sciences, New York University, 719 Broadway, New York, NY 10003, SAD. 2009.
- [3] Mohamed, N., Rubin, D. M., Marwala, T. "*Detection of Epileptiform Activity in Human EEG Signals Using Bayesian Neural Networks*". Neural Information Processing - Letters and Reviews Vol.10, No.1, siječanj 2006.
- [4] Štastny, J., Sovka, P., Stančák, A. "*EEG Signal Classification: Introduction to the Problem*". RADIOENGINEERING, VOL. 12, NO. 3, rujan 2003.
- [5] Štastny, J., Sovka, P. "*High-Resolution Movement EEG Classification*". Computational Intelligence and Neuroscience, vol. 2007, str. 12.
- [6] Štastny, J., Sovka, P., Stančák, A. "*EEG signal classification*". Biological Signal Lab., Faculty of Electrical Engineering, Češki tehnički univerzitet u Pragu, Republika Češka.
- [7] Doležal, J., Štastny, J., Sovka, P. "*Exploiting Temporal Context in High-Resolution Movement-Related EEG Classification*". RADIOENGINEERING, VOL. 18, NO. 1, travanj 2009.
- [8] Subasi, A., Ercelebi, E. "*Classification of EEG signals using neural network and logistic regression*". Computer Methods and Programs in Biomedicine (2005.) 78, str. 87—99.
- [9] Ansari-Asl, K., Chanel, G., Pun, T. "*A CHANNEL SELECTION METHOD FOR EEG CLASSIFICATION IN EMOTION: ASSESSMENT BASED ON SYNCHRONIZATION LIKELIHOOD*". 15th European Signal Processing Conference (EUSIPCO 2007.), Poznan, Poljska, rujan 3.-7., 2007.
- [10] Wu, L., Neskovic, P. "*Feature extraction for EEG classification: representing electrode outputs as a Markov stochastic process*". Department of Physics and Institute for Brain and Neural Systems Brown University, Providence, RI 02906.

- [11] Latchoumane, C. V., Vialatte, F., Jeong, J., Cichocki, A. "*EEG CLASSIFICATION OF MILD AND SEVERE ALZHEIMER'S DISEASE USING PARALLEL FACTOR ANALYSIS METHOD: PARAFAC Decomposition of Spectral-Spatial Characteristics of EEG Time Series*"
- [12] Chiappa, S., Barber, D. "*Generative Independent Component Analysis for EEG Classification*". European Symposium on Artificial Neural Networks, Bruges, Belgija, (travanj 2005.).
- [13] Adlakha, A. "*Single trial EEG Classification*". Srpanj 2002.
- [14] Babiloni, F., Cincotti, F., Lazzarini, L., Millán, J., Mouriño, J., Varsta, M., Heikkonen, J., Bianchi, L., Marciani, M. G. "*Linear Classification of Low-Resolution EEG Patterns Produced by Imagined Hand Movements*". IEEE TRANSACTIONS ON REHABILITATION ENGINEERING, VOL. 8, NO. 2, Lipanj 2000.
- [15] Kerkeni, N., Alexandre, F., Bedoui, M. H., Bougrain, L., Dogui, M. "*Automatic classification of Sleep Stages on a EEG signal by Artificial Neural Networks*".
- [16] Guler, I., Ubeyli, E.D. "*Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients*". Journal of Neuroscience Methods. 2005.
- [17] Prochazka, A., Kukal, J. "*Wavelet Transform Use for Feature Extraction and EEG Signal Segments Classification*". Institute of Chemical Technology in Prague, Department of Computing and Control Engineering.
- [18] Anderson, C. W., Sijerčić, Z. "*Classification of EEG Signals from Four Subjects During Five Mental Tasks*". Department of Computer Science Colorado State University.
- [19] Yong, Y. P. A., Hurley, N. J., Silvestre, G. C. M. "*SINGLE-TRIAL EEG CLASSIFICATION FOR BRAIN-COMPUTER INTERFACE USING WAVELET DECOMPOSITION*". Department of Computer Science, University College Dublin Belfield, Irska.
- [20] Deng, S., Srinivasan, R., Lappas, T., D'Zmura, M. "*EEG classification of imagined syllable rhythm using Hilbert spectrum methods*". JOURNAL OF NEURAL ENGINEERING, 7 (2010.) 046006 (13pp).
- [21] Li, J., Zhang, L. "*Regularized tensor discriminant analysis for single trial EEG classification in BCI*". Pattern Recognition Letters (2009.).

- [22] Lee, F., Scherer, R., Leeb, R., Neuper, C., Bischof, H., Pfurtscheller, G. "A Comparative Analysis of Multi-Class EEG Classification for Brain Computer Interface". Institute for Computer Graphics and Vision Graz University of Technology.
- [23] Hill, N. J., Lal, T. N., Schroder, M., Hinterberger, T., Widman, G., Elger, C. E., Scholkopf, B., Birbaumer, N. "Classifying Event-Related Desynchronization in EEG, ECoG and MEG signals".
- [24] Obermaier, B., Guger, C., Neuper, C., Pfurtscheller, G. "Hidden Markov models for online classification of single trial EEG data". Pattern Recognition Letters 22 (2001.), 1299-1309.
- [25] Othmer, S., Othmer, S. F. "Post Traumatic Stress Disorder—The Neurofeedback Remedy". Biofeedback, vol. 37, issue 1., str. 24-31.
- [26] „Electroencephalography“. <http://en.wikipedia.org/wiki/Electroencephalography>. Svibanj 2012.
- [27] Šantić, A. "Biomedicinska elektronika". Školska knjiga; 1995. ISBN: 978-953-0-30664-6.
- [28] "PTSP (posttraumatski stresni poremećaj)", veljača 2009. <http://anksioznost123.blog.hr/2009/02/1625899331/ptsp-posttraumatski-stresni-poremecaj.html> . Svibanj 2012.
- [29] „Posttraumatski stresni poremećaj“, rujan 2011. [http://hr.wikipedia.org/wiki/Posttraumatski\\_stresni\\_poreme%C4%87aj](http://hr.wikipedia.org/wiki/Posttraumatski_stresni_poreme%C4%87aj) . Svibanj 2012.
- [30] „Neurofeedback“. <http://www.mensana.hr/default.asp?cms=GGGIII> . Svibanj 2012.
- [31] „ANOVA“. <http://www.mathworks.com/help/toolbox/stats/bgttcvf.html> . Svibanj 2012.
- [32] Kujundžić Tiljak, M., Ivanković, D. "Testiranje hipoteza". Zagreb. Svibanj 2012.
- [33] Bašić, B. D. „Analiza varijance: ANOVA, Analysis of variance, Analysis of means using variance“. Materijal za kolegij "Statističko učenje". Fakultet elektrotehnike i računarstva, Zagreb, 19.4.2005.
- [34] "Spectral density". [http://en.wikipedia.org/wiki/Power\\_spectral\\_density](http://en.wikipedia.org/wiki/Power_spectral_density) . Svibanj 2012.

- [35] <http://www.cygres.com/OcnPageE/Glosry/Spec.html> . Svibanj 2012.
- [36] "Welch's method", veljača 2012.  
[http://en.wikipedia.org/wiki/Welch%27s\\_method](http://en.wikipedia.org/wiki/Welch%27s_method) . Svibanj 2012.
- [37] Jančić, M. „Wavelet transformacija“.  
<http://www.vcl.fer.hr/dtv/jpeg/wave.htm> . Lipanj 2012.
- [38] „Discrete wavelet transform“, travanj 2012.  
[http://en.wikipedia.org/wiki/Discrete\\_wavelet\\_transform](http://en.wikipedia.org/wiki/Discrete_wavelet_transform) . Lipanj 2012.
- [39] Seršić, D. Materijal za kolegij "Napredne metode digitalne obrade signala". Fakultet elektrotehnike i računarstva.
- [40] Grbović, O. „Diplomski rad: Fraktalna dinamika hoda“. Elektrotehnički fakultet u Beogradu. 2007.
- [41] „Fraktali III dio - fraktalna geometrija i dimenzije“. 2008.  
<http://nauka.adsoglasi.com/fizikamatematika/fraktali3geometrija.php> . Lipanj 2012.
- [42] "Support vector machine".  
[http://en.wikipedia.org/wiki/Support\\_vector\\_machine](http://en.wikipedia.org/wiki/Support_vector_machine) . Lipanj 2012.
- [43] "Lagrange multiplier". [http://en.wikipedia.org/wiki/Lagrange\\_multiplier](http://en.wikipedia.org/wiki/Lagrange_multiplier) . Lipanj 2012.
- [44] Lončarić, S., Subašić, M. "Neuronske mreže: Stroj s potpornim vektorima (SVM)". Fakultet elektrotehnike i računarstva, Sveučilište u Zagrebu, materijali za kolegij "Neuronske mreže". Lipanj 2012.
- [45] "Type I and type II errors", svibanj 2012.  
[http://en.wikipedia.org/wiki/Type\\_I\\_and\\_type\\_II\\_errors](http://en.wikipedia.org/wiki/Type_I_and_type_II_errors) . Lipanj 2012.

## 8. Sažetak

### Metoda za automatsku klasifikaciju EEG signala za detekciju PTSP-a

Provedeno je istraživanje s ciljem pronalaska najbolje metode za klasifikaciju EEG signala. Signali koji su se klasificirali su pripadali dvjema skupinama i to skupina PTSP i skupina normalnih, ne-PTSP signala. Za odabir najboljih značajki se koristio statistički alat ANOVA pomoću kojeg su se identificirali konkretni kanali u kojima pojedine značajke pokazuju najveću razliku između dvije klase signala. Korištene značajke su snage izračunate pomoću PSD-a i DWT-a  $\alpha, \beta, \delta$  i  $\theta$ -pojava, fraktalna dimenzija izračunata Higuchijevim algoritmom, *skewness*, *kurtosis* i *Hjortovi* parametri (aktivnost, mobilnost i kompleksnost). SVM je korišten kao klasifikator i dobivena je točnost klasifikacije od 96%.

Ključne riječi: klasifikacija, EEG značajke, EEG, PTSP, SVM, ANOVA, PSD, DWT.

### The method for automatic classification of EEG signals to detect PTSD

Research was performed with the aim of finding the best methods for the classification of EEG signals. The signals that were classified belonged to two groups, a group of PTSD and a group of normal, non-PTSD signals. To select the best features a statistical tool called ANOVA was used to identify the specific channels in which the features had the largest difference between the two classes of signals. Features used in this paper were power calculated using PSD and DWT for all four EEG waves ( $\alpha, \beta, \delta$  and  $\theta$ ), Higuchi's fractal dimension, skewness, kurtosis and Hjorth parameters (activity, mobility and complexity). SVM was used as the classifier and a classification accuracy of 96% was achieved.

Keywords: classification, EEG features, EEG, PTSD, SVM, ANOVA, PSD, DWT.