

UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN
INFORMACIJSKE TEHNOLOGIJE

ZAKLJUČNA NALOGA

**KLASIFIKACIJA PODATKOV
MOŽGANSKE AKTIVNOSTI**

TADEJ MAGAJNA

UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN
INFORMACIJSKE TEHNOLOGIJE

Zaključna naloga

Klasifikacija podatkov možganske aktivnosti

(Classification of Brain-Computer Interface Data)

Ime in priimek: Tadej Magajna

Študijski program: Računalništvo in informatika

Mentor: doc. dr. Jernej Vičič

Somentor: doc. dr. Branko Kavšek

Koper, avgust 2014

Ključna dokumentacijska informacija

Ime in PRIIMEK: Tadej MAGAJNA

Naslov zaključne naloge: Klasifikacija podatkov možganske aktivnosti

Kraj: Koper

Leto: 2014

Število listov: 49 število slik: 7 število tabel: 3

Število prilog: 3 število strani prilog: 4 število referenc: 27

Mentor: doc. dr. Jernej Vičič

Somentor: doc. dr. Branko Kavšek

Ključne besede: podatkovno rudarjenje, elektroencefalograf, elektroencefalografija, klasifikacija, vmesnik med računalnikom in možgani

Izvleček:

Zaključna projektna naloga obravnava tematiko podatkovnega rudarjenja podatkov možganske aktivnosti v povezavi z nizkocenovnimi vmesniki med računalnikom in možgani.

Podrobno in celovito je predstavljeno teoretično področje problema. Razložene so osnove elektroencefalografije v povezavi z delovanjem možganov. Našteti so frekvenčni pasovi in njihove pripadajoče lastnosti. S pomočjo matematične teorije je razložen proces spektralne dekompozicije s poudarkom na Fourierjevi transformaciji. Opisani so primeri aplikacije podatkovnega rudarjenja možganske aktivnosti na številna področja. Omenjene so pomembnejše raziskave in izpostavljene so dobre prakse. Našteti so glavni klasifikatorji, ki prinašajo najboljše rezultate.

S pomočjo po meri izdelane programske opreme je nazadnje v okviru zaključne naloge izvedena raziskava, ki vključuje 50 subjektov. Namen je izgradnja modela, ki uspešno klasificira dve mentalni stanji, ki jih dosežemo z opazovanjem dveh različnih slik. Modeli so zgrajeni na podlagi uspešne prakse, ki prinaša zanesljive rezultate v že obstoječih raziskavah. Modeli so zgrajeni na dva načina. Pri prvem so modeli zgrajeni za vsakega posameznika posebej, pri drugem pa samo enkrat, z uporabo združenih meritev vseh subjektov. Izkaže se, da je splošni model, ki je grajen na

osnovi meritev vseh sodelujočih v raziskavi, neuspešen. Posledično spoznamo, da je izgradnja tovrstnega splošnega modela, ki velja za večje število posameznikov, kompleksen problem. Vseeno pa modeli posameznih podatkovnih zbirk prinašajo dovolj dobre rezultate, da lahko trdimo, da je z uporabo nizkocenovne EEG naprave možno uspešno klasificirati omenjena mentalna stanja.

Key words documentation

Name and SURNAME: Tadej MAGAJNA

Title of final project paper: Classification of Brain-Computer Interface Data

Place: Koper

Year: 2014

Number of pages: 49 Number of figures: 7 Number of tables: 3

Number of appendices: 3 Number of appendix pages: 4 Number of references: 27

Mentor: Assist. Prof. Jernej Vičič, PhD

Co-Mentor: Assist. Prof. Branko Kavšek, PhD

Keywords: EEG, data mining, classification, electroencephalography, electroencephalogram, Brain-Computer interface

Abstract:

This work discusses the topic of classification of the brain activity data in relation to the use of low cost brain-computer interface devices.

Theoretical background is firstly thoroughly discussed. The fundamentals of electroencephalography are explained in relation to how the brain works. Basic frequency bands are studied and their properties are pointed out. Moreover, with the help of mathematical theory we study the process of spectral decomposition with emphasis on Fourier transform. Basic examples of the numerous fields of application are explained. The main studies are pointed out and best practice is noted along with the best resulting classifiers.

Using the software explicitly made for this case, we do a study involving 50 subjects. The purpose of this study is to build a model that is able to successfully distinguish between two mental states which we achieve by looking at two different photos. Models are based on best practice that brings reliable results in already existent studies. On one hand, the models are built individually for each separate person, while on the other hand, we also build a single model using all the data from the full set of subjects. As it turns out, the full, general model is unsuccessful. Knowing this, we can assume that building such a general model and applying it to

numerous individuals is a complex problem. Nevertheless, the models built explicitly for each individual turn out to perform well enough for us to be able to say that it's possible to classify those states using low cost electroencephalography devices.

Zahvala

Zahvaljujem se mentorju doc. dr. Jerneju Vičiču in somentorju doc. dr. Branku Kavšku za strokovno pomoč pri izdelavi zaključne naloge.

Prav tako se zahvaljujem študentom UP FAMNIT in Fakulteti za socialno delo Univerze v Ljubljani za sodelovanje pri raziskavi. Posebna zahvala gre Lei Rojec za pomoč pri zbiranju kandidatov za raziskavo.

Kazalo vsebine

1	Uvod	1
2	Pregled področja	2
2.1	Elektroencefalograf in elektroencefalografija	2
2.1.1	Zgradba možganov	3
2.1.2	Postavitev elektrod	3
2.2	Frekvenčni pasovi možganske aktivnosti	4
2.3	Dekompozicija podatkov	6
2.3.1	Fourierjeva transformacija	7
2.4	Tehnike podatkovnega rudarjenja	8
2.4.1	Raziskovalna podatkovna analiza (Exploratory Data Analysis)	8
2.4.2	Opisno modeliranje (Descriptive Modeling)	8
2.4.3	Napovedno modeliranje (Predictive Modeling)	9
2.5	Ocenjevanje točnosti modelov podatkovnega rudarjenja	9
2.6	Proces in standardi podatkovnega rudarjenja	10
2.7	Področja in primeri uporabe EEG podatkovnega rudarjenja	12
2.7.1	Klasifikacija inicializacije gibov	12
2.7.2	Klasifikacija čustev	12
2.7.3	Klasifikacija mentalnih stanj	13
2.7.4	Klasifikacija vizualizacije premikanja in rotacije objektov	14
2.7.5	Povzetek rezultatov raziskav	14
3	Metodologija	16
3.1	Opis problema	16
3.1.1	Neurosky Mindwave	17
3.2	Programska oprema	19
3.2.1	Uporaba in grafični vmesnik	19
3.2.2	Povezovanje	20
3.2.3	Branje	20
3.2.4	Shranjevanje	21
3.2.5	Faze razvoja	21

3.3	Zbiranje podatkov	22
3.4	Izgradnja modela	23
3.4.1	Izgradnja posamezniku specifičnih modelov	23
3.4.2	Izgradnja splošnega modela	24
4	Rezultati	25
4.0.3	Rezultati posamezniku specifičnih modelov	25
4.0.4	Rezultati splošnega modela	26
5	Zaključek	28
6	Literatura	30

Kazalo tabel

1	Lastnosti frekvenčnih pasov možganske aktivnosti	6
2	Rezultati posamezniku specifičnih modelov	25
3	Rezultati splošnega modela	26

Kazalo slik

1	Mednarodni standard 10-20 za postavitev elektrod	4
2	Proces in rezultat Fourierjeve transformacije EEG podatkov	7
3	Cikel faz CRISP-DM standarda	11
4	Upravljanje invalidskega vozička z mislimi	13
5	Naprava Neurosky Mindwave	17
6	Grafični vmesnik programske opreme	20
7	Rezultati klasifikatorja bagging za posamezne podatkovne zbirke	26

Kazalo prilog

Zaslonski posnetki testnega programa	33
QR kodi za dostop do programske opreme in podatkovnih zbirk	34
Primer .arff datoteke iz zbirke podatkov	35

Seznam kratic

<i>KDD</i>	Knowledge Discovery in Databases
<i>EEG</i>	Elektroencefalograf
<i>CRISP – DM</i>	Cross Industry Standard Process for Data Mining
<i>DFT</i>	Discrete Fourier Transformation
<i>CTF</i>	Continious Fourier Transformation
<i>FFT</i>	Fast Fourier Transformation
<i>SVM</i>	Support Vector Machines
<i>KNN</i>	K Nearest Neighbor
<i>LDA</i>	Linear Discriminant Analysis
<i>BCI</i>	Brain Computer Interface
<i>USB</i>	Universal Serial Bus

1 Uvod

Elektroencefalografija je proces merjenja možganske električne aktivnosti, ki jo oddajajo živčne celice, imenovane nevroni. Čeprav je bil izumitelj omenjene tehnologije (Richard Caton) fizik, jo je prvi v praktične namene uporabil nevrolog Hans Berger. Ugotovil je, da lahko z analizo EEG podatkov izvemo določene informacije o subjektu. Elektroencefalograf (naprava za merjenje možganskih valov) je bila do sedaj večinoma uporabljana le na področju medicine, kjer je služila za diagnosticiranje epilepsije, določanje stopenj spanja in ugotavljanje motenj v delovanju možganov. Zmožnosti raziskav so bile omejene zaradi težke dostopnosti naprav in števila pacientov z motnjami možganov. Zaradi relativno visoke cene naprav je bilo izvajanje raziskav z elektroencefalografom dostopno le manjšini [22].

Prihod nove tehnologije, namenjene zgolj za raziskovalne namene, ki omogoča direktno povezavo z osebnimi računalniki, je prinesel nove možnosti na področju EEG raziskav. Nizka cena naprav (od 99\$ naprej) omogoča večjo dostopnost. Zmožnost brezžične komunikacije pa omogoča zajemanje podatkov med fizičnimi aktivnostmi, ki bi jih žična povezava ovirala [13].

Podatkovno rudarjenje je v knjigi J. Hana [11] definirano kot proces odkrivanja zanimivih vzorcev in znanja iz velikih količin podatkov. Na medicinskem področju se aktivno uporablja za določanje stopnje nagnjenosti k epileptičnim motnjam [20], v novejših raziskavah pa se uporablja za napovedovanje gibov in stanj zavesti.

Naloga se osredotoča na predstavitev področja podatkovnega rudarjenja meritev, pridobljenih s pomočjo elektroencefalografa in izpostavi prednosti ter slabosti novih, komercialno dostopnih EEG naprav.

V drugem poglavju je podrobno opisano delovanje elektroencefalografa in pomen shranjenih podatkov. Navedeni so uveljavljeni standardi podatkovnega rudarjenja in razložene so posamezne faze. Dosedanje raziskave na področju EEG rudarjenja so primerjane med seboj in rezultati so uporabljeni kot izhodišče za novo raziskavo. Tretje poglavje opisuje potek opravljene raziskave, izvedene v okviru zaključne naloge z nizko-cenovno komercialno dostopno napravo Neurosky [18]. Cilj raziskave je zgraditi model, ki uspešno klasificira dve mentalni stanji, ki jih dosežemo z opazovanjem dveh različnih slik. V četrtem poglavju so opisani rezultati raziskave. Sledi jim zadnje poglavje kjer je povzeta celotna vsebina naloge skupaj z ugotovitvami.

2 Pregled področja

Vmesnik med računalnikom in možgani je naprava za zajemanje, procesiranje in tolmačenje podatkov možganske aktivnosti z namenom krmiljenja zunanje naprave. S pomočjo tehnik podatkovnega rudarjenja je uporabo vmesnikov možno aplicirati na številna področja, omenjena v sledečem poglavju.

Za popolno razumevanje problema je potrebno razložiti osnove elektroencefalografije, ki skupaj z detajlnim opisom tehnik podatkovnega rudarjenja predstavljajo temelje projektne naloge.

Sprva so v poglavju razložene osnove elektroencefalografije, ki so prisotne v večini raziskav na omenjenem področju. Razloženi so pomen in značilnosti posameznih frekvenčnih pasov. Opisane so osnove spektralne dekompozicije podatkov s poudarkom na Fourierjevi transformaciji. Sledi jim obravnava tehnik podatkovnega rudarjenja s poudarkom na klasifikaciji. Ta se nadaljuje z razlago procesov in uveljavljenih tehnologij. Nazadnje se obe veji združita. Omenjene so temeljne raziskave, v katerih se uporablja podatkovno rudarjenje možganske aktivnosti na različnih področjih. Navedeni so najuspešnejši algoritmi in predstavljene so najboljše prakse.

2.1 Elektroencefalograf in elektroencefalografija

Možganska aktivnost, na katero lahko vplivajo stanje zavesti, čustveno stanje, motorična ali miselna operacija, se odraža z aktivnostjo električnih signalov na zunanji strani lobanje. Naprava za zajemanje teh signalov se imenuje elektroencefalograf (EEG). Je ena izmed komponent vmesnika med računalnikom in možgani. Kljub temu, da obstajajo vmesniki med možgani in računalnikom, pri katerih je elektroencefalograf vgrajen v notranjost lobanje, bo poglavje obravnavalo le naprave, ki podatke zajemajo na zunanji strani lobanje. Sestavlja jo ena ali več elektrod, tipično nameščenih na površino lobanje, ki merijo jakost signala na določenem možganskem področju. Merska enota posamezne elektrode se meri v μV oziroma mV . Zaradi majhne napetosti signala se ta ojača s pomočjo ojačevalnikov, ki ga posredujejo v analogno digitalni pretvornik. Določene sodobne naprave, omenjene v članku [21], vsebujejo integrirane filtre za odstranjevanje šuma, saj lahko neprečiščen signal, pridobljen s pomočjo elektroencefalografa vsebuje šume zaradi gibanja mišic in zunanjih elektronskih naprav.

Elektroencefalografija se na medicinskem področju med drugim uporablja za: analizo pacientov v komi, iskanje lokacije možganskih poškodb, ugotavljanje epilepsije in lociranja vira za vzrok napada [22]. Na nemedicinskem področju pa EEG vmesniki med možgani in računalnikom uporabljajo znanja računalništva in informatike za namen: krmiljenja zunanjih naprav (invalidski vozički, proteze), klasifikacije mentalnih stanj, klasifikacije čustvenih stanj, analize odvisnikov od drog in alkohola, ugotavljanje stopnje koncentracije in meditacije [13] [17] [1].

2.1.1 Zgradba možganov

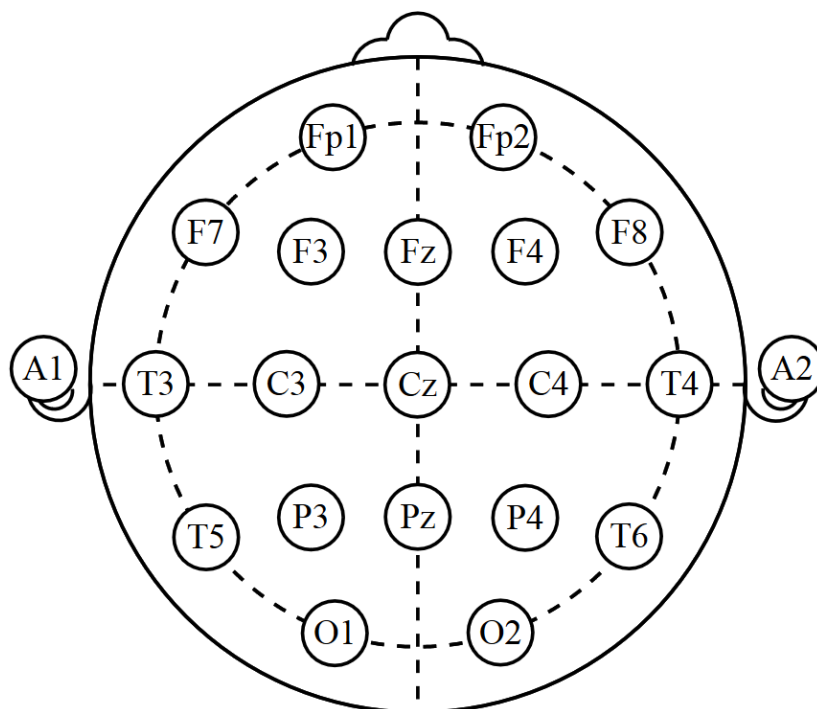
Z anatomskega vidika lahko možgane razdelimo na tri glavne dele: veliki možgani, mali možgani in možgansko deblo. Na področju klasifikacije podatkov možganske aktivnosti so najbolj relevantne meritve aktivnosti velikih možganov. Ti so sestavljeni iz leve in desne hemisfere. Zunanji sloj velikih možganov se imenuje možganska skorja in je glavni vir EEG signalov. Veliki možgani vsebujejo centre za [22]:

- iniciacijo gibov,
- čustva,
- stanje zavesti.

Zaradi porazdelitve omenjenih možganskih centrov distribucija možganske aktivnosti po površini lobanje ni enakomerna. Zato je za doseganje želenih rezultatov ključnega pomena ustrezna postavitvev elektrod [22].

2.1.2 Postavitvev elektrod

Pri vmesnikih med možgani in računalnikom se pogosto uporablja mednarodni sistem 10-20, ki ga sestavlja postavitvev 21 elektrod, ki zajemajo področje leve in desne hemisfere. Elektrode niso postavljene v bližino oči, da se izognemo nastanku motenj očesnih mišic. Kljub temu, da 10-20 sistem uporablja 21 elektrod, se pri vmesnikih med možgani in računalnikom pogosto uporablja le podatke nekaterih elektrod [22]. V zadnjih letih so na trgu prisotne naprave, ki jih sestavlja manjše število elektrod. S tem posledično izboljšajo tako udobje uporabe naprave, kot tudi zmanjšajo ceno izdelka. Primer bralnika, ki uporablja le eno elektrodo, je Neurosky Mindwave [21]. Namesti se jo na levo stran čela. Postavitvev je ekvivalentna elektrodi Fp1 pri postavitvi 10-20 (slika 1).



Slika 1: Mednarodni standard 10-20 za postavitev elektrod

2.2 Frekvenčni pasovi možganske aktivnosti

Signal posamezne elektrode elektroencefalografa je mešanica valovanj na več frekvenčnih pasovih. Ker je ta signal tipično zapleten in se z vizualno analizo podatkov ugotavljajo le določene karakteristike signala, se za globlje razumevanje signal tipično razbije na več frekvenčnih pasov [19]. Nemški znanstvenik Hans Berger je odkril korelacijo med amplitudo posameznih frekvenc in različnimi mentalnimi stanji [13]. Glede na frekvenčne pasove poznamo tipično pet različnih vrst signalov:

Delta valovi so valovi z najdaljšo valovno dolžino. To so signali možganske aktivnosti s frekvenco 0,5-4hz. Njihova prisotnost se pogosto opaža pri globokem spanju. Prav tako jih zaznamo v fazi prebujanja iz spanja. Če se delta valovi pojavljajo v fazi zbuđenosti, je to lahko znak fizičnih poškodb možganov [13]. Šum, ki ga proizvajajo vratne in čeljustne mišice, je najpogosteje prisoten ravno na frekvenčnem področju delta valov, zato je v fazi prečiščevanja podatkov največji poudarek na filtriranju tega pasu [19] [22].

Theta valovi se nahajo v območju od 4-7,5hz z amplitudo tipično več kot $20\mu V$. Povečana aktivnost theta valov je pogosto posledica čustvenega stresa, natančneje frustracije in razočaranja. Igrajo pomembno vlogo pri dojenčkih in otrocih. Pri odraslih so manj prisotni. Podatke theta valov se pogosto uporablja pri raziskavah, povezanih s čustvi [22] [13].

Alfa valovi ležijo med 8 in 13hz z amplitudo $30 - 40\mu V$. Najlažje jih zaznamo na

zadnjem delu možganov na področju elektrod O1 in O2 (slika 1). Je najbolj izrazita vrsta valov, ki v primerjavi z ostalimi tipi pasov poda največ informacij o subjektu. Alfa valovi indicirajo sproščeno stanje brez koncentracije in pozornosti. Alfa valove najlažje prikličemo tako, da zapremo oči in pomislimo na nekaj sproščujočega. Ko odpremo oči, zaslišimo neznan zvok ali pa smo zaskrbljeni, se aktivnost alfa valov hitro porazgubi [13] [19]. Kljub temu, da signali možganske aktivnosti v splošnem uspešno nakazujejo stanje zavesti, se pri izjemnih posameznikih izkažejo za nezanesljive. Sodeč po medicinskih EEG posnetkih, naj bi bil Albert Einstein zmožen reševanja zahtevnih matematičnih problemov v alfa stanju, kljub temu, da so pri aktivnem razmišljanju bolj izraziti beta valovi [22].

Beta valovi so valovi možganske aktivnosti med 14 in 26Hz z amplitudo $5 - 30\mu V$. Pogosto se razdelijo na dva pasova z oznakama α_1 in α_2 . So tipi signalov, močno povezani z aktivnim razmišljanjem, pozornostjo, koncentracijo in osredotočenostjo na zunanji svet. Najlažje jih prepoznamo na frontalnem delu lobanje na področju elektrod F_p1 in F_p2 (slika 1) [22] [19].

Na istem frekvenčnem področju kot alfa valove zaznamo tudi MU valove. Aktivni so na motoričnem (gibalnem) predelu možganske skorje. So indikacija inicializacije gibov. To pomeni, da jih zaznamo tako ob dejanskem gibanju, kot tudi ko pomislimo na izvajanje gibov. Zaznava MU valov v povezavi z gibanjem je tako mogoča tudi za gibalno ovirane [13] [19] [21].

Najvišji frekvenčni pas zavzemajo **gamma** valovi in sicer frekvence nad 30Hz. Amplituda teh je šibka glede na ostale frekvenčne pasove, zato z omenjenega področja dobimo manj informacij kot pri ostalih [22]. Kljub temu se prisotnost gamma valov, tako kot beta valov, povezuje z močno koncentracijo in pozornostjo [13].

Tabela 1: Lastnosti frekvenčnih pasov možganske aktivnosti

Oznaka	Frekvenčni pas	Amplituda	Značilnosti
Delta	0,5-4Hz	100 – 200 μV	Aktivni pri globokem spanju in prebujanju. Prisotnost v budni fazi je lahko znak fizičnih poškodb možganov. Šum, ki ga povzročajo vratne mišice, pogosto zamenjamo za delta valove.
Theta	4-7,5Hz	20 μV	Pokazatelj čustev, čustvenega stresa, frustracije in razočaranja.
Alpha	8-13Hz	30 – 40 μV	Najbolj aktivni na zadnjem delu možganov. So pokazatelj sproščenosti brez koncentracije in pozornosti. Prikličemo jih, ko zapremo oči.
Beta	14-26Hz	5 – 30 μV	Povezani z aktivnim razmišljanjem, pozornostjo, koncentracijo, osredotočenostjo na zunanji svet. Najbolj aktivni na frontalnem delu lobanje. Na istem frekvenčnem področju delujejo MU valovi, ki so pokazatelj inicializacije gibanja.
Gamma	nad 30Hz	+ + 3 – 5 μV	Pokazatelj močne koncentracije.

Ker je signal posamezne elektrode elektroencefalografa mešanica omenjenih valov na večih frekvenčnih pasovih, je za ugotavljanje navedenih lastnosti potrebna priprava podatkov s pomočjo spektralne dekompozicije.

2.3 Dekompozicija podatkov

Surovi, neprocesirani podatki, pridobljeni iz posamezne elektrode elektroencefalografa, so niz časovno razporejenih vrednosti napetosti signala (v voltih), ki pa direktno ne podajajo informacije o frekvenčnih spektrih. V tem poglavju bodo opisane teoretične osnove spektralne dekompozicije, s pomočjo katere lahko izluščimo podatke o frekveni signala.

2.3.1 Fourierjeva transformacija

Uporaba Fourierjeve analize je eden najbolj uporabljanih procesov spektralne dekompozicije podatkov, s pomočjo katere iz EEG meritev v časovni domeni dobimo vrednosti moči v frekvenčni domeni [21].

Fourierjeva vrsta sprejme signal in ga dekompresira v seštevke sinusov in kosinusov z različnimi frekvencami. Če imamo signal, ki traja 1 sekundo, $0 < T < 1$, lahko s fourierjevo vrsto ta signal zapišemo kot neskončno vsoto [25]:

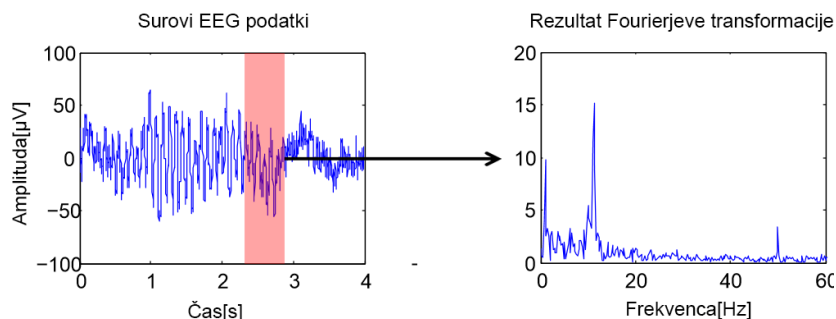
$$f(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(2\pi nt) + b_n \sin(2\pi nt)), \quad (2.1)$$

kjer je $f(t)$ časovno urejena funkcija, za katero je signal integrabilen na intervalu $[0, 1]$. Spremenljivka t predstavlja časovno spremenljivko, n frekvenco v hercih (Hz) in a_0, a_n in b_n Fourierjeve koeficiente, ki jih pridobimo z naslednjima formulama [25] [21]:

$$a_n = \frac{2}{T} \int_0^T x(t) \cos(\pi nt) dt, \quad n = 0, 1, 2, 3, \dots, N, \quad (2.2)$$

$$b_n = \frac{2}{T} \int_0^T x(t) \sin(\pi nt) dt, \quad n = 0, 1, 2, 3, \dots, N, \quad (2.3)$$

kjer T predstavlja mejo časovnega odseka, v našem primeru vrednost 1.



Slika 2: Proces in rezultat Fourierjeve transformacije EEG podatkov

S pomočjo Eulerjeve formule lahko vsoto $\cos(t) + j \sin(t)$ zapišemo kot $e^{jt} = \cos(t) + j \sin(t)$. Tako lahko s pomočjo zvezne Fourierjeve transformacije (CFT) prevedemo podatke iz časovne domene v frekvenčno.

Vrednosti EEG meritev pa tipično niso podane kot funkcija, temveč kot časovno urejen niz meritev, ki le aproksimira graf funkcije, ki bi jo lahko neposredno zapisali s Fourierjevo vrsto. Zato se za dekompozicijo EEG podatkov uporablja diskretna Fourierjeva transformacija (DFT) [21] [25]:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j2k\pi \frac{n}{N}}$$

Časovna zahtevnost izvajanja diskretne Fourierjeve transformacije je $O(n^2)$. Zaradi hitrosti izvajanja se pa v praksi uporablja hitro Fourierjevo transformacijo (FFT), ki je variacija DFT-ja. Ta ima časovno zahtevnost $O(n \log(n))$ [21] [26].

Kot alternativa Fourierjevi analizi se za ekstrakcijo detajlov in približkov signala, pri katerih FFT ne zadostuje, uporablja valčna transformacija [17].

2.4 Tehnike podatkovnega rudarjenja

Naslednja poglavja razlagajo teoretično podlago podatkovnega rudarjenja, ki jo zajema področje EEG raziskav. Vključujejo opis osnovnih tehnik podatkovnega rudarjenja. Sledi jim razlaga in opis korakov z vidika pridobivanja znanja iz podatkov in pregled uveljavljene metodologije CRISP-DM.

Podatkovno rudarjenje je interdisciplinarna smer, ki zajema znanje s področja računalništva in informatike, matematike in statistike [11]. Čeprav v citiranih knjigah ni enolične definicije, je za splošno razumevanje področja smiselno opisati osnovne tehnike podatkovnega rudarjenja.

2.4.1 Raziskovalna podatkovna analiza (Exploratory Data Analysis)

Namen raziskovalne podatkovne analize je raziskovanje brez točno določenega cilja in ugotavljanje, kaj lahko razberemo iz podatkov brez formalnega modeliranja in testnih metod. Tehnike so tipično interaktivne in/ali vizualne. Pri manjšem številu spremenljivk se pogosto uporabljajo grafične metode kot na primer tortni diagram [12].

Primer raziskovalne podatkovne analize na področju EEG rudarjenja bi bila grafična predstavitev razmerij med pasovi EEG signala.

2.4.2 Opisno modeliranje (Descriptive Modeling)

Cilj opisnega modeliranja je splošna predstavitev glavnih karakteristik podatkov. V knjigi [12] je opisana kot povzetek podatkov, ki nam omogoča študijo najpomembnejših aspektov podatkov. Poznamo dva glavna tipa opisnega modeliranja:

- segmentacija,
- razvrščanje v skupine.

Namen prve je razdeliti zapise v homogene skupine, tako da se podobni zapisi nahajajo v isti skupini. Število skupin pri segmentaciji je vnaprej določeno s strani raziskovalca. Pri razvrščanju v skupine pa skušamo najti naravno število skupin in

jih primerno uvrstiti [12]. Razvrščanje v skupine je pogost tip modeliranja pri EEG podatkovnem rudarjenju na področju določanja stopenj spanja.

2.4.3 Napovedno modeliranje (Predictive Modeling)

Cilj napovednega modeliranja je izgradnja modela za ugotovitev vrednosti ene spremenljivke, ki ji pravimo razred. Model je zasnovan na podlagi učne baze podatkov, kjer je spremenljivka razred že znana [11]. Dva osnovna tipa napovednega modeliranja sta:

- klasifikacija,
- regresija.

Cilj klasifikacije je preslikava vektorja meritev X v spremenljivko Y , ki je kategoričnega tipa. Iskani spremenljivki Y se tipično pravi razred. Regresija je sorodna vrsta napovednega modeliranja, kjer je razred kvantitativnega tipa. Funkcija, ki slika vektor X v razred, se imenuje model podatkovnega rudarjenja.

Večina raziskav, omenjenih v zaključni nalogi, rešuje problem napovednega modeliranja na podatkih EEG meritev.

2.5 Ocenjevanje točnosti modelov podatkovnega rudarjenja

Pri izbiri optimalnega modela podatkovnega rudarjenja je eden glavnih faktorjev ocena stopnje točnosti. Na podlagi te informacije lahko sklepamo, kakšno uspešnost bo imel model, ko se bo uporabljal v praksi. Zaradi tipično omejene množice podatkov se ocenjevanje točnosti izkaže za problematično fazo. Če za testiranje uspešnosti modela vzamemo isto množico podatkov kot za učenje, dobimo preveliko stopnjo natančnosti. Temu pojavu pravimo prekomerno prileganje. Zato moramo učno množico tipično ločiti od testne množice podatkov, tako da je testna množica dovolj velika za natančno oceno, hkrati pa učna množica ni okrnjena.

Razdeljevanje na testno in učno množico je najpreprostejši način testiranja uspešnosti klasifikacije. Deluje tako, da dodelimo določen delež podatkov učni množici, ki se uporablja za izgradnjo modela, neodvisen preostanek podatkov pa dodelimo testni množici. Uspešnost modela na testni množici nam poda informacijo o točnosti [11]. Slabost omenjenega pristopa je, da lahko zaradi omejene količine podatkov z razdeljevanjem učno množico okrnimo in posledično zmanjšamo uspešnost modela.

Prečno preverjanje delno rešuje zgoraj omenjeni problem. Sprva množico podatkov razdelimo na n disjunktih, enako velikih podmnožic. Nato n -krat ponovimo korak, kjer vsako podmnožico označimo za testno množico, preostanek podatkov pa za učno. Na učni množici zgradimo model in uspešnost preverimo na testni množici. Po n -kratni ponovitvi imamo izmerjeno točnost za vsako od n razdelitev. Povprečje vseh dobljenih vrednosti nam predstavlja oceno točnosti modela. V splošnem se uporablja razdelitev na 10 podmnožic ($n = 10$) [11].

2.6 Proces in standardi podatkovnega rudarjenja

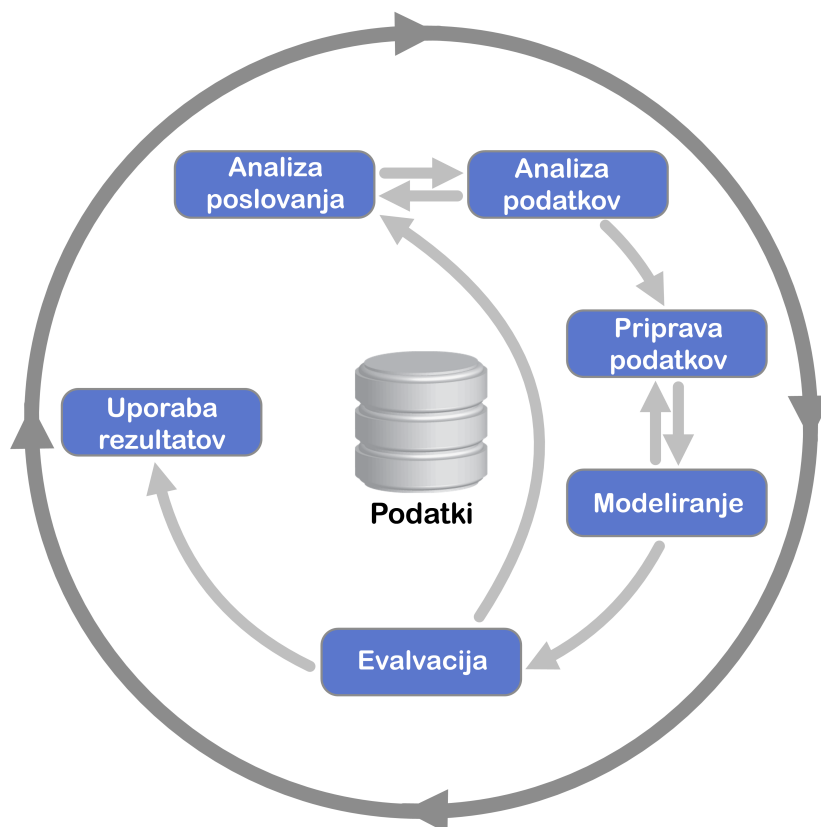
Čeprav za podatkovno rudarjenje ne obstajajo formalna pravila, ki bi določala potek izvajanja projekta, se za to področje vse bolj pogosto uporabljajo formalni standardi. Ti omogočajo boljše razumevanje, poenoteno izvajanje in določajo pravila izvajalcem novih projektov [5].

Pridobivanje znanja iz podatkov ali KDD (Knowledge Discovery in Databases) je proces, ki s pomočjo podatkovnega rudarjenja izvleče in predstavi znanje iz podatkov, shranjenih v podatkovnih bazah (Fayyad, 1996). Poznamo pet osnovnih faz pridobivanja znanja iz podatkov [11]:

1. *Čiščenje podatkov* - Izločevanje šuma in nekonsistentnih podatkov. Podatki EEG meritev vsebujejo veliko šuma, zato je čiščenje ključnega pomena za doseganje zanesljivih rezultatov [2].
2. *(Integracija podatkov)* - Združevanje večih virov podatkov v homogeno celoto. EEG meritve večih subjektov se pogosto združujejo v enotne nabore podatkov.
3. *Selekcija podatkov* - Selekcija relevantnih podatkov iz podatkovne baze
4. *Transformacija podatkov* - Obdelava z izvajanjem agregacije in povzemanja podatkov.
5. *Modeliranje* - Ključna faza, ki uporablja modele podatkovnega rudarjenja.
6. *Evalvacija modelov* - Identifikacija vzorcev, ki predstavljajo znanje.
7. *Predstavitev znanja* - Znanje, pridobljeno s pomočjo podatkovnega rudarjenja, je predstavljeno uporabnikom z vizualizacijo in predstavitvijo znanja.

Izraz pridobivanje znanja iz podatkov se pogosto uporablja kot sinonim za podatkovno rudarjenje [11].

CRISP-DM (Cross Industry Standard Process for Data Mining) je standard, ki opisuje pogosto uporabljene procese, ki jih uporabljajo izvajalci podatkovnega rudarjenja. Zasnovala so ga podjetja Daimler Chrysler, NCR in SPSS. Del procesa je namenjen razumevanju problema s poslovnega vidika.



Slika 3: Cikel faz CRISP-DM standarda

Poznamo šest osnovnih faz izvajanja projekta, ki sestavljajo cikel. Prva je analiza poslovanja, ki določa poslovne cilje in zahteve. Sledi ji analiza podatkov, ki se začne z zajemanjem in nadaljuje s procesom raziskovanja, opisovanja in razumevanja podatkov. Zaključni se z oceno kakovosti in stopnje šuma. Naslednja faza je priprava podatkov, ki zajema proces izdelave končnega nabora podatkov. Ta vključuje čiščenje, integracijo in formatiranje. V fazi modeliranja se za namen izgradnje modela izvedejo tehnike podatkovnega rudarjenja. Sledi ji faza evalvacije natančnosti modela in opravljene faze procesa, ki rezultate primerja z želenimi poslovnimi cilji. Nazadnje sledi faza uporabe podatkov. Tu se opravi pregled celotnega projekta, ustvari se poročilo in uporabniku se prikaže in predstavi rezultate tako, da jih bo lahko uporabljal [5] [23].

Na sliki 3 je prikazano zaporedje faz CRISP-DM standarda. Puščice znotraj kroga predstavljajo razpored faz. Zunanja puščica pa predstavlja iterativno naravo izvajanja projektov na področju podatkovnega rudarjenja, saj se v praksi izvajanje posameznih faz pogosto večkrat ponovi [23].

Sklicujoč se na dosedanje raziskave s področja EEG podatkovnega rudarjenja, omejnene v razdelku 2.7, lahko sklepamo, da se tovrstne raziskave striktno ne držijo standardov, kot je CRISP-DM.

2.7 Področja in primeri uporabe EEG podatkovnega rudarjenja

Čeprav lahko klasifikacijo EEG podatkov s pomočjo učne baze apliciramo na poljubna področja, se izkaže, da natančnost in uspešnost modelov opazno variira. Poznamo področja klasifikacije, ki se uporabljajo v praksi in prinašajo zanesljive rezultate [2]. Obenem se pa v zadnjih letih pogosto izvajajo raziskave na novih področjih, pri katerih se rezultati glede na dane pogoje ne izkažejo za uspešne.

V tem poglavju si bomo ogledali osnovna področja klasifikacije EEG podatkov, ki so se izkazala za uspešna.

2.7.1 Klasifikacija inicializacije gibov

Modeli klasifikacije inicializacije gibov podajo informacijo, ali je subjekt skušal izvesti določen telesni gib. Prednost področja je, da za uspešno klasifikacijo subjektu ni potrebno dejansko izvajati giba, temveč si mora izvajanje giba le zamisliti. To gibalno oviranim osebam omogoča možnost izboljšanja mobilnosti z upravljanjem invalidskih vozičkov ali drugih gibalnih naprav z mislimi [9].

V članku [9] je omenjena raziskava, v kateri so sodelovale tri univerze (medicinska, računalniška in elektrotehniška). Opisan je potek raziskave, v kateri subjekti z mislimi upravljajo invalidski voziček s pomočjo klasifikacije EEG podatkov. Izvedena je na tridesetih gibalno oviranih osebah. Uporabljenih je več klasifikacijskih algoritmov podatkovnega rudarjenja kot na primer Naivni Bayes, in SVM (glej: [27]). Rezultati so se izkazali za uspešne z natančnostjo 57% pri uporabi komercialnega vmesnika med možgani in računalnikom. V članku K. Liao [14], je raziskano področje klasifikacije gibanja posameznih prstov na roki. Za najbolj uspešen algoritem se je izkazal SVM, ki je upošteval spremembe na frekvenčnih spektrih. Povprečna natančnost modela je 77,11%.

2.7.2 Klasifikacija čustev

Podatkovno rudarjenje EEG podatkov s pomočjo učne baze omogoča klasifikacijo vnaprej določenih čustvenih stanj.



Slika 4: Upravljanje invalidskega vozička z mislimi

Primer uspešne raziskave na omenjenem področju je omenjen v članku [17]. V začetni fazi nosilci raziskave definirajo pet možnih čustev (gnus, sreča, presenečenje, strah in nevtralno), ki predstavljajo razrede učne baze. Nato z uporabo mednarodnega 10-20 sistema 1 zajamejo EEG podatke dvajsetim subjektom v omenjenih stanjih. S pomočjo valjčne transformacije se podatki razbijejo na tri frekvenčne pasove, ki nosijo največ informacij: alfa, beta in gama. Uporabljeni sta dve preprosti klasifikacijski metodi, KNN in LDA. Obe z nad 75% natančnostjo, pri čemer KNN dosega 83,26% povprečno uspešnost. Najlažje razberljivo čustvo je gnus z 91,67% povprečno natančnostjo klasifikacije.

2.7.3 Klasifikacija mentalnih stanj

Klasifikacija mentalnih stanj je področje, sorodno raziskavi, omenjeni v prejšnjem razdelku, ki pa ne klasificira čustvenih stanj, temveč mentalne naloge, ki jih opravlja subjekt. Za doseganje natančnosti klasifikacije je potrebna izbira mentalnih stanj, ki vzbudijo čim večjo asimetrijo aktivnosti možganskih hemisfer.

V članku [3] je razložen potek raziskave, v kateri se klasificira med petimi mentalnimi nalogami:

- osnovna naloga, kjer se subjekti skušajo sprostiti,
- naloga s pismom, kjer so morali subjekti v mislih sestaviti pismo prijatelju,
- matematična naloga, kjer so morali subjekti reševati zahtevne probleme množenja števil brez pisanja govora ali pomoči s fizičnim gibanjem,

- naloga z vizualnim štetjem, kjer so si subjekti morali vizualno predstavljati številke, ki jih nekdo piše na tablo,
- naloga z rotacijo geometrijske figure, kjer so si morali subjekti vizualno predstavljati rotacijo geometrijskega objekta.

Model je bil zgrajen s pomočjo nevronske mreže [15]. Kljub temu, da so rezultati močno variirali glede na posameznika, je bila na podlagi prečnega preverjanja v povprečju ocenjena 71% natančnost.

2.7.4 Klasifikacija vizualizacije premikanja in rotacije objektov

Raziskave so pokazale, da je možno klasificiranje premikanja in rotacije objektov s pomočjo vizualiziranja približevanja, oddaljevanja in rotacije geometrijskega objekta. Področje odpira nove možnosti predvsem z apliciranjem področja na igranje računalniških igravic, kjer bi igralec svoj lik lahko upravljal z mentalno rotacijo. Kljub temu, da na trgu že obstajajo komercialne igre, ki jih je možno upravljati s pomočjo EEG naprav [13], te upravljajo igro na podlagi mentalnih stanj in ne na podlagi vizualizacije premikanja objekta.

V članku C. W. Andersona in Z. Sijerčića [3] se del raziskave posveti klasifikaciji tridimenzionalnega objekta okoli osi.

2.7.5 Povzetek rezultatov raziskav

Glede na rezultate raziskav se izkaže, da je za zagotavljanje uspešnosti raziskave ključnega pomena ustrezna izbira modela. Iz validacijskih rezultatov lahko razberemo, da so se za uspešne izkazali naslednji algoritmi [2] [3] [17] [14] [9]:

- SVM,
- KNN,
- bagging,
- naive Bayes.

Razen za algoritem SVM, ki je bil optimalen pri dveh omenjenih raziskavah, se izkaže, da je pri vsaki raziskavi izbira najboljšega algoritma drugačna. Iz rezultatov torej lahko sklepamo, da je izbira najustrežnejšega modela odvisna od posameznega primera raziskave in da optimalen algoritem, ki bi deloval za vse EEG klasifikacijske probleme, ne obstaja.

Hkrati je v raziskavah opazna velika razlika uspešnosti modelov pri različnih posameznikih. Iz tega lahko sklepamo, da je pri izgradnji čim boljše učne baze ključnega pomena ustrezno upoštevanje navodil za subjekte, ki sodelujejo v raziskavi.

3 Metodologija

Prihod komercialnih vmesnikov med računalnikom in možgani je odprl nove možnosti raziskav in posledično povečal in pospešil stopnjo razvoja na omenjenem področju. Cenovno dostopne naprave, ki kot kompromis pri ceni zmanjšajo število elektrod, so sedaj na voljo že za 99\$ [21].

V okviru zaključne naloge se je v okviru zaključne naloge izvedla raziskava klasifikacije možganske aktivnosti. Opisana je v naslednjih poglavjih.

3.1 Opis problema

Iz prej navedenih raziskav je razvidno, da se različna mentalna stanja subjekta lahko odražajo z različnimi možganskimi aktivnostmi, ki jih je z določeno natančnostjo možno uspešno klasificirati [3].

Tako je namen raziskave zgraditi model, ki bo uspešno klasificiral tip slike, ki jo opazuje subjekt. Torej bo znal klasificirati med prvim stanjem, ko subjekt opazuje prvo sliko, in drugim stanjem, ko subjekt opazuje drugo sliko, z uporabo nizkocenovnega vmesnika med računalnikom in možgani. Idealno naj bi vsaka slika različne subjekte spravila v enako oziroma kar se da podobno stanje, saj bi tako model lahko posplošili na populacijo in ne samo posameznika.

Izbira slik je bila v fazi načrtovanja raziskave ključnega pomena, saj je bilo potrebno izbrati dve sliki, ki bosta ob opazovanju idealno reflektirali dve kar se da različni mentalni stanji. Tako sta bili izbrani dve sliki, kjer prva vsebuje veliko število podrobnosti in motečih nepravilnosti, druga pa vsebuje malo podrobnosti in deluje sproščujoče in umirjeno (priloga 1). Izbrani sta bili kot rezultat začetnih meritev, kjer se je pokazala možnost uspešne klasifikacije. Meritvam torej pripadata dva razreda: stanje1 za prvo sliko in stanje2 za drugo sliko. Poleg razredov, ki pripadata slikam, je v zbirki podatkov prisoten tudi evidenčni razred, poimenovan stanje0, ki pa pripada meritvam, ko uporabnik še ne opazuje slik in se ne zaveda, da je sneman. Ta služi kot evidenca za nadzorovanje kakovosti signala.

Sklicujoč se na delo [19], je bil postopek raziskave, zasnovan po principu tipičnega procesa raziskav z vmesniki med računalnikom in možgani:

- Zbiranje meritev,

- Procesiranje podatkov,
- Klasifikacija,
- Interakcija z zunanjo napravo.

Z izjemo, da faza interakcije z napravo ni bila izvedena. Nadomestilo jo je testiranje s prečnim preverjanjem, ki je opisano v razdelku 2.5. Prav tako je bila pred izvajanjem raziskave razvita programska oprema za zajemanje podatkov, shranjevanje podatkov in povezavo z osebnim računalnikom, ki je kompatibilna z napravo Neurosky Mindwave [18].

3.1.1 Neurosky Mindwave

Mindwave podjetja Neurosky je nizkocenovni vmesnik med možgani in računalnikom, ki vsebuje preprost elektroencefalograf z eno samo elektrodo. Namestimo jo na področje čela blizu elektrode Fp1 pri sistemu 10-20 1.



Slika 5: Naprava Neurosky Mindwave

Uporablja se v povezavi z osebnim računalnikom, ki mu v realnem času pošilja pet vrst podatkov: surove EEG meritve, nivoje frekvenčnih spektrov, stopnjo koncentracije, stopnjo meditacije in jakost mežikanja. Hkrati naprava pošilja tudi ostale podatke, ki niso vezani na elektroencefalograf: stanje baterije, kakovost signala in stanje povezave.

Glede na vrsto podatkov je v osnovi izdelana za dva različna namena.

Lahko se uporablja kot zabavna naprava za igranje preprostih računalniških igrice, ki jih je moč upravljati s koncentriranjem, sproščanjem in mežikanjem. Namen teh je spoznavanje uporabnikov (predvsem otrok) z delovanjem elektroencefalografov in vmesnikov med računalnikom in možgani [19].

Hkrati pa je naprava namenjena za opravljanje raziskav (primer: [21]). V tej situaciji se tipično zajema ali surove podatke (s 512 meritvami na sekundo) ali pa meritve frekvenčnih pasov, omenjenih v razdelku 2.2, ki jih že mikrokontroler, integriran v napravo, razbije na frekvenčne spektre. V tem primeru meritve vsebujejo numerične vrednosti meritev osmih frekvenčnih pasov [16]:

- Delta - $1 - 3Hz$,
- Theta - $4 - 7Hz$,
- Alpha1 - $8 - 9Hz$,
- Alpha2 - $10 - 12Hz$,
- Beta1 - $13 - 17Hz$,
- Beta2 - $18 - 30Hz$,
- Gamma1 - $31 - 40Hz$,
- Gamma2 - $41 - 50Hz$.

Za razumevanje problema in uspešen razvoj so na uradni spletni strani podjetja NeuroSky javno dostopna programska orodja, sestavljena iz primerov programov in dokumentov, ki služijo kot pomoč razvijalcem pri razvijanju aplikacij z napravo MindWave.

Podatki se iz naprave pošiljajo brezžično preko Wifi omrežja do USB modema, za katerega je potrebno predhodno namestiti ustrezne (priložene) gonilnike. USB modem v sistemu simulira napravo, ki se povezuje preko serijskih vrat. Naprava se lahko poveže s hitrostjo 1200, 2400, 4800, 9600, 57600, 115200 bitov na sekundo. Zaradi brezžične povezave ima naprava samostojno napajanje v obliki AAA baterij, ki zadostujejo za osem ur snemanja. [21].

Zaradi srčnega utripa generira čelna elektroda veliko šuma, ki bi lahko zmanjšal kakovost meritev. Zato vsebuje naprava senzor utripa v obliki ščipalke, ki si jo namestimo na uhelj. S podatki o srčnem utripu naprava izničuje posledice šuma, povzročene zaradi srčnega utripa na čelu.

MindWave sestavlja samo ena elektroda, ki zajema podatke s frontalnega dela možganov, kjer se motorične aktivnosti ne odražajo. Iz tega lahko sklepamo, da naprava ni primerna za raziskave, ki vključujejo analizo inicializacije gibov, saj se motorični del možganov nahaja na področju okoli temena (glej: [22]).

3.2 Programska oprema

V okviru raziskave je bila pred začetkom razvita programska oprema za zajemanje podatkov, shranjevanje podatkov in povezovanje z napravo MindWave. Zasnovana je na način, ki omogoča masovno snemanje podatkov brez potrebe po ponovnem zagonu v primeru prazne baterije ali izgube povezave. Meritve vsebujejo informacijo o razredih, ki jih je možno predhodno nastaviti in poimenovati. Podatki so shranjeni v formatu, ki je kompatibilen s programsko opremo za podatkovno rudarjenje.

V začetni fazi je bilo zaželeno aplikacijo razviti v okolju, s pomočjo katerega bi se lahko program prevedel tako za operacijski sistem Windows kot tudi Linux. A gonilniki MindWave naprave operacijskega sistema Linux ne podpirajo. Tako je programska oprema razvita v okolju Visual Studio in sprogramirana v programskem jeziku C++. Kompatibilna je le z operacijskimi sistemi Microsoft Windows XP, Microsoft Windows 7, Microsoft Windows Vista in Microsoft Windows 8 [18].

3.2.1 Uporaba in grafični vmesnik

Uporabnik s pomočjo grafičnega vmesnika upravlja z aplikacijo, s čimer dostopa do osnovnih funkcionalnosti programske opreme. Poleg interakcije s programom se mu med snemanjem prikazuje tudi grafični prikaz jakosti frekvenčnih spektrov.

Postopek uporabe programske opreme je sestavljen iz večih korakov. V začetni fazi je potrebna namestitev prižgane naprave na glavo, pri čemer moramo biti pozorni, da se senzor trdno dotika čela ter da je ušesna ščipalka pritrjena na uhelj. Nato zaženemo program.

S klikom na gumb *Connect* inicializiramo postopek povezovanja z napravo, ki je opisan v razdelku 3.2.2. V primeru uspešne povezave se indikator povezave, ki se nahaja v kotu zgoraj levo na sliki 6, obarva zeleno. V tej fazi program že prejema podatke iz naprave in preverja njihovo kakovost. V primeru, da je naprava uspešno nameščena in da je kakovost signala zadovoljiva, se obarva še drugi indikator.

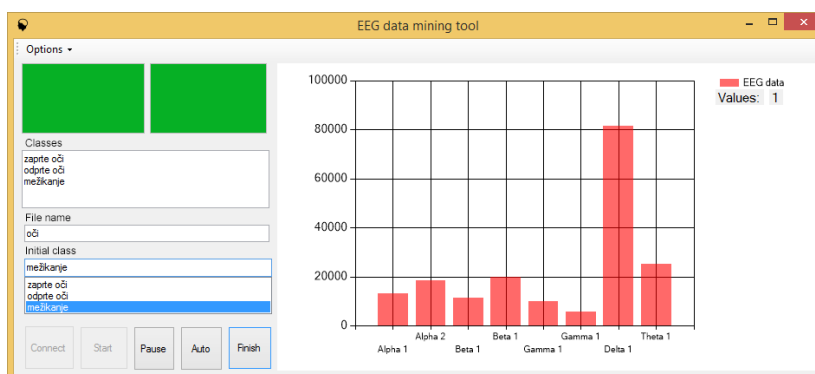
V polje *Classes* vnesemo imena vseh možnih razredov. V polje *File name* vnesemo brez končnice ime datoteke, ki bo vsebovala meritve. Nazadnje v padajočem meniju *Initial class* izberemo razred, ki bo pripadal meritvam ob začetku snemanja.

S pritiskom na gumb *Start* pričnemo zajemanje podatkov. V desnem predelu okna

se nam pričnejo vizualizirati jakosti frekvenčnih pasov meritev. Ob napisu *Values* se nam izpisuje število meritev. S pritiskom na gumb *Pause* začasno zaustavimo snemanje, tako ob potrebi spremenimo razred snemanja, ki ga nadaljujemo s ponovnim pritiskom na gumb *Start*. V primeru nekakovostnih meritev programska oprema avtomatsko ignorira meritve in čaka na meritve zadovoljive kakovosti.

Konec snemanja sprožimo s stiskom na gumb *Finish*, ki inicializira proces shranjevanja.

Celoten postopek lahko ponovimo s ponovnim pritiskom na gumb *Connect*.



Slika 6: Grafični vmesnik programske opreme

3.2.2 Povezovanje

V procesu povezovanja program rezervira dostop do serijskih vrat in skuša brati podatke. V primeru, da USB modem ni priključen, da je naprava izkljupnjena ali da je naprava že rezervirana za drug proces, program sporoči napako.

Ko lahko uspešno beremo podatke s serijskih vrat, se postopek povezovanja zaključi in indikator uspešne povezave se obarva zeleno. Del paketa podatkov vsebuje tudi informacijo o kakovosti meritev. V primeru zadovoljive kakovosti se obarva še drugi indikator in omogoči se možnost pričetka snemanja.

3.2.3 Branje

Ob inicializaciji procesa branja program v pomnilniku uporabniškega prostora rezervira prostor za shranjevanje meritev. Sestavlja ga matrika velikosti 28800x9 tipa Unsigned Int, kjer prvo število predstavlja maksimalno kapaciteto meritev. Program se stolmači za 32 bitno arhitekturo. Tako tip integer v programu zavzema 4 bajte oziroma 32 bitov. Posledično se za meritve rezervira 28800*9*32 bajtov prostora oziroma približno 1MB. To zadostuje za osem ur meritev.

Vsako sekundo branja se v polje zapiše devet integer vrednosti vseh frekvenčnih pasov in zaporedna številka pripadajočega razreda.

3.2.4 Shranjevanje

S pritiskom na gumb *Finish* se inicializira proces shranjevanja. Namen tega je zapisati datoteko tipa *.arff* [4] na datotečni sistem, ki bo vsebovala podatke, shranjene v pomnilniku.

V prvi fazi program preveri, ali bo lahko uspešno pisal v datoteko z navedenim imenom. V primeru, da v datoteko ni mogoče pisati, program javi napako. Nato se v datoteko zapiše glava *.arff* datoteke, ki vključuje opise atributov vključno z imeni vseh razredov. Prvih osem atributov predstavlja meritve osmih frekvenčnih pasov (glej: 3.1.1), zadnji atribut pa razred.

Nazadnje se iz spomina vse meritve preslikajo v *.arff* datoteko (glej prilogo 3). Kazalec na datoteko se zapre.

Omogoči se pritisk gumba *Connect*, s katerim lahko pričnemo snemanje nove zbirke podatkov.

3.2.5 Faze razvoja

Razvoj programske opreme je potekal v skladu s standardi programskega inženirstva. Po razporeditvi I. Sommervilla poznamo pet faz razvoja [24]: analiza zahtev, načrtovanje, implementacija, integracija in testiranje, obratovanje in vzdrževanje.

V fazi analize zahtev so bile navedene vse funkcijske zahteve interakcije uporabnika s sistemom. Navedene so bile tudi nefunkcijske zahteve s poudarkom na stabilnosti sistema v primeru izgube signala, težave z baterijo ali izvlečenega modema.

V fazi načrtovanja je bila definirana struktura programa, programskih komponent in relacij med njimi. Posebej je izstopal poudarek na večopravnosti programa, saj v fazi zajemanja podatkov program obenem bere podatke ter jih izrisuje na zaslon. Kljub temu, da bi bila izvedljiva tudi zaporedna rešitev, se je zaradi preprečevanja izgube podatkov izvedla vzporedna rešitev z uporabo niti.

Implementacija je potekala v okolju Visual Studio v programskem jeziku C++.

Testiranje je bilo izvedeno s pomočjo testnih primerov, kjer se je med snemanjem namerno ugasnila naprava, izklopil USB modem ali program večkrat vzporedno zagnal. Preverjeno je bilo tudi delovanje programa z ustvarjanjem umetnega šuma s premikanjem elektrode, ki bi ga moral program zaznati in posledično spregledati šumne meritve. Rezultati testiranja so bili zabeleženi in v primeru najdenih napak so bili izvedeni ustrezni popravki.

V zadnji fazi je bila aplikacija uporabljena v sklopu raziskave za zaključno nalogo. V namen širše uporabe in vzdrževanja je bil celoten projekt vključno z izvorno kodo, prevedenim programom in pripadajočimi gonilniki naložen na spletno shrambo GitHub kjer je prosto dostopen na naslovu <https://github.com/thelole/mindRecorder.git> kot alternativa plačljivim programom, dostopnim v spletni trgovini podjetja NeuroSky [8].

Skozi celoten proces razvoja se je uporabljal linearni model s povratki, saj se je bilo po testiranju večkrat potrebno vrniti nazaj v fazo načrtovanja.

3.3 Zbiranje podatkov

Za nemoten potek raziskave je bilo potrebno izvesti spremembe na programski opremi, specifične za problem zajemanja možganske aktivnosti ob gledanju dveh slik.

Oprema je bila za namen raziskave dodatno prilagojena, tako da lahko avtomatsko vodi subjekt skozi proces prikazovanja slik in sočasno upravlja potek branja podatkov z naprave.

Zajemanje podatkov poteka na sledeč način. Po uspešni namestitvi naprave in uspešni povezavi se ob pritisku na gumb *Auto*, ki je prikazan na sliki 6, prične proces zajemanja podatkov.

Prične se prva faza snemanja, kjer se v celozaslonskem načinu izrišejo navodila za potek raziskave. Subjektu je naročeno, naj pozorno opazuje fotografije, ki se bodo prikazovale na zasonu. Faza traja 10 sekund. Kljub temu, da se subjekt tega ne zaveda, se po prvih petih sekundah že prične snemanje meritev, ki pripadajo razredu stanje0, omenjenem v razdelku 3.1.

Nato se prične druga faza, ki traja 25 sekund. Na zaslonu se v celozaslonskem načinu prikazuje prva slika. Zaradi nenadne spremembe slike ob prehodu med fazami prvih pet sekund druge faze še ne beremo z naprave. Po petih sekundah, ko naj bi se misli že adaptirale na prikaz slike, se prične branje meritev, katerim pripada razred stanje1.

Sledi tretja faza, ki prav tako kot druga faza traja 25 sekund. Na zaslonu se v celozaslonskem načinu prikazuje druga slika. Snemanje poteka s podobnimi omejitvami, kot smo jih navedli zgoraj. Meritve pripadajo razredu stanje2.

V zadnji fazi se subjektu sporoči, da se je proces zajemanja podatkov zaključil. Izvede se proces shranjevanja, omenjen v razdelku 3.2.4, ki ustvari *.arff* datoteko.

Tako je zajemanje podatkov avtomatiziran proces, ki traja 60 sekund. Rezultat tega je datoteka s 45 meritvami, kjer 5 meritev pripada razredu stanje0, 20 razredu stanje1 in ostalih 20 razredu stanje2.

Raziskava je bila izvedena na petdesetih osebah, ki jih sestavlja 24 moških in 26

žensk. Potekala je v mirnem okolju brez večjih motenj. Pri vsaki meritvi je bil zabeležen tudi spol.

3.4 Izgradnja modela

Model za ugotavljanje stanj opazovanja slik je bil zgrajen na podlagi meritev, pridobljenih iz faze zbiranja podatkov.

V prvi fazi se je izvedlo prečiščevanje, kjer so se odstranili podatki, ki so pripadali razredu stanje0, saj bi prisotnost tretjega razreda zmanjševala kakovost in uspešnost modela.

Z uporabo večih klasifikatorjev so bili zgrajeni različni modeli, ki so bili nazadnje preverjeni s pomočjo prečnega preverjanja, omenjenega v razdelku 2.5.

Ker se je izkazalo, da je uspešnost klasifikacije meritev posameznega subjekta v povprečju veliko večja od klasifikacije zbirke, ki vsebuje meritve vseh sodelujočih v raziskavi, je model zgrajen na dva načina.

Prvi je zgrajen za vsakega posameznika posebej. Za testno in učno množico uporablja le meritve posameznika neodvisno od meritev ostalih sodelujočih v raziskavi. Proces izgradnje in validacije se ponovi petdesetkrat in izračuna se povprečje uspešnosti. Drugi je zgrajen na osnovi posamezne podatkovne zbirke, ki vsebuje meritve vseh petdesetih subjektov. Za vsak klasifikator se izgradnja in testiranje izvede le enkrat. Rezultati se lahko posplošijo in niso specifični le za en subjekt, kot pri prvem načinu izgradnje modela.

Zbirke podatkov so prosto dostopne na spletni shrambi podatkov GitHub naslovu <https://github.com/thelole/datasets.git>

3.4.1 Izgradnja posamezniku specifičnih modelov

Ker se je izkazalo, da so rezultati prečnega testiranja v povprečju boljši, če izgradimo model za vsakega posameznika posebej, se je najprej izvedla faza izgradnje posameznih modelov. V prvi fazi se je izvedel postopek, ki nam pove, kateri atribut nosi največ informacij in z njim najlažje rešujemo problem. To dosežemo z uporabo OneR algoritma [7], ki za izgradnjo modela uporablja le en atribut. Ta za vsako vrednost meritev prešteje, kolikokrat se določen razred ponovi. Ko najde najbolj pogost razred, ustvari pravilo, ki vrednosti pripiše najpogostejši razred [27]. Algoritem je bil izveden, za vsako zbirko podatkov posebej. Rezultati pokažejo, da je bil največkrat izbran atribut Gamma2. Izbran je bil za deset podatkovnih zbirk. Razporeditev po spolih je enakomerna, zato pri izbiri OneR atributa ni opazne razlike med spoloma.

Nato je bil za vsakega posameznika zgrajen model z uporabo večih klasifikatorjev.

Glede na algoritme, omenjene v razdelku 2.7, je bilo izbranih šest klasifikatorjev, ki so se izkazali za uspešne v primerjavi z ostalimi:

- OneR,
- SVM,
- PART,
- NaiveBayes,
- DecisionStump,
- Bagging (PART).

Model je bil zgrajen z uporabo programske opreme WEKA [10] s pomočjo zaganjanja ukazov z ukazne vrstice. Zaradi velikega števila podatkovnih zbirk je bila razvita programska oprema, ki je avtomatizirala proces izgradnje modela za vseh petdeset zbirk. Izgradnja SVM modela je trajala eno minuto. Vsi ostali modeli so bili zgrajeni v manj kot eni sekundi.

3.4.2 Izgradnja splošnega modela

Podatki vseh podatkovnih zbirk so bili združeni v eno *.arff* datoteko za namen izgradnje splošnega modela.

Uporabljeni so bili klasifikatorji, navedeni v razdelku 3.4.1. Izgradnja SVM modela je trajala 294 sekund. Vsi ostali modeli so bili zvedeni v manj kot eni sekundi.

4 Rezultati

Poglavje predstavlja pregled uspešnosti raziskave, vrednosti različnih pristopov ter njihove prednosti in slabosti.

Uspešnost modelov se je preverjala s pomočjo prečnega preverjanja. Uporabljala se je razdelitev na 10 podmnožic, z namenom, da je število podmnožic delitelj števila meritev, torej delitelj števila 40.

4.0.3 Rezultati posamezniku specifičnih modelov

Na podlagi rezultatov preverjanja smo za pravilnosti vseh podatkovnih zbirk izračunali povprečje. Ker pa je informacija, kako oddaljene so vrednosti uspešnosti od povprečja, pomembna za raziskavo, smo izračunali tudi standardno deviacijo. To je mera statistične razpršenosti spremenljivke [15], ki smo jo vključili v rezultate.

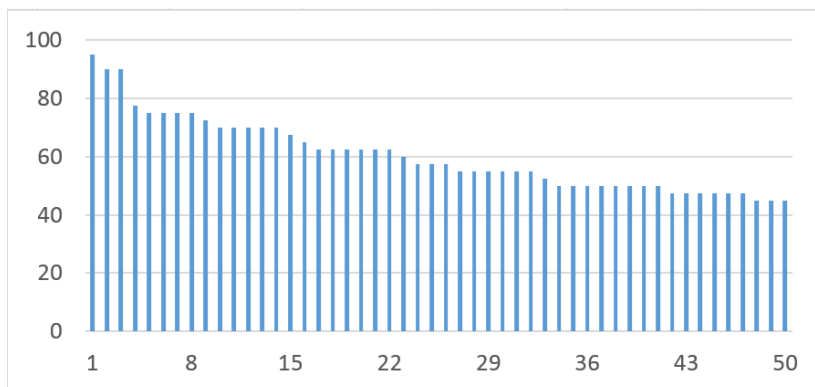
Tabela 2: Rezultati posamezniku specifičnih modelov

Klasifikator	Povprečna uspešnost	Standardna deviacija σ
OneR	57,35%	14,942
SVM	57,55%	13,039
PART	58,65%	12,516
NaiveBayes	58,05%	10,883
DecisionStump	59,05%	13,429
Bagging (PART)	60,35%	12,485

Za najuspešnejšega se je izkazal klasifikator bagging z uporabo PART algoritma. Povprečna uspešnost klasifikatorja s prečnim preverjanjem je 60,35%. Sklicujoč se na sliko 7, je 22 subjektov možno klasificirati z več kot 60% pravilnostjo.

Nazadnje smo preverili povprečno uspešnost najboljšega klasifikatorja glede na spol. Izkazalo se je, da pri uspešnosti klasifikacije ni opazne razlike med spoloma. Pri 24 podatkovnih zbirkah moških je uspešnost klasifikacije 61,05%, pri 26 podatkovnih zbirkah žensk pa 59,71%.

Slabost omenjenega pristopa je potreba po izgradnji učne baze za vsakega posameznika, ki uporablja sistem. Torej bi vsak uporabnik naprave, ki bi želel uporabljati



Slika 7: Rezultati klasifikatorja bagging za posamezne podatkovne zbirke

storitev klasifikacije možganske aktivnosti, moral predhodno izgraditi svojo učno bazo.

Kljub temu lahko na podlagi 61% natančnosti sklepamo, da model uspešno prepozna razliko med mentalnimi stanji. Tako lahko trdimo, da obstaja možnost, da bi lahko s pravo napravo in snemanjem v pravem okolju izgradili model, ki bi bil dovolj uspešen za praktično uporabo.

4.0.4 Rezultati splošnega modela

Model zbirke podatkov, ki vsebuje združene meritve vseh subjektov, je bil testiran s pomočjo prečnega preverjanja. Izkazalo se je, da je model neuspešen, saj uspešnost ni veliko višja od stopnje pravilnosti ugibanja med dvema razredoma, ki je 50% pri enakomerni porazdelitvi števila meritev.

Najbolje se je odrezal klasifikator SVM s 53,2% uspešnostjo. Ostali se gibajo med 50-53%.

Tabela 3: Rezultati splošnega modela

Klasifikator	Uspešnost
OneR	50,65%
SVM	53,20%
PART	53,00%
NaiveBayes	50,65%
DecisionStump	53,01%
Bagging (PART)	52,15%

Iz rezultatov lahko sklepamo, da je izgradnja modela, ki uspešno deluje na splošni populaciji, zahteven problem.

Obstaja tudi možnost razlage, da gre rezultat pripisati posamezniku unikatnemu načinu razmišljanja in doživljanja okolice, ki posledično oddaja možganske valove, ki so meritvam ostalih subjektov tuji. Pravilnost te teze bi lahko preverili v nadaljnjih raziskavah z ustrezno izbiro testne množice pri izgradnji splošnega modela. Ker je program sam razdeljeval testno množico na podmnožice, obstaja možnost, da so v učni množici prevladovali podatki enega ali le nekaj posameznik, kar posledično zmanjša uspešnost modela. Tako bi bilo v nadaljnjih raziskavah priporočeno uporabljati učno množico, ki bi vsebovala del meritev vsakega subjekta.

5 Zaključek

Podatkovno rudarjenje je interdisciplinarno področje, ki združuje znanje iz računalništva in informatike, medicine, statistike in matematike. Za celostno razumevanje problema je potrebno razumevanje osnov z vseh omenjenih znanosti.

V začetni fazi podatkovnega rudarjenja podatkov možganske aktivnosti je ključnega pomena znanje s področja medicine. Pomembno je razumevanje postavitve možganskih centrov, ki vpliva na lokacijo in število elektrod. Ta naj bi ustrezala zelenemu cilju raziskave. Elektrode se morajo nahajati na področjih, ki odražajo stanja klasifikacijskih razredov.

Pri pripravi meritev je poudarek na znanju s področja matematike. Za uspešno pripravo podatkov je potrebno znanje teorije spektralne dekompozicije.

Po zaključeni fazi priprave podatkov se izvede izgradnja modela in testiranje. S pomočjo znanja računalništva in informatike na področju podatkovnega rudarjenja se izvede izgradnja ustreznega modela z upoštevanjem števila atributov, tipov atributov, števila razredov in velikosti podatkovne zbirke.

Prednost klasifikacije podatkov možganske aktivnosti je zmožnost podajanja informacij z razmišljanjem brez fizične interakcije. Ta presega omejitve izražanja z uporabo telesa in kljub še zelo začetni, skoraj samo akademski fazi, sugerira številna področja aplikacije in odpira nove možnosti interakcije z napravami. Dokazano je, da se klasifikacija podatkov možganske aktivnosti lahko uspešno uporablja za krmiljenje invalidskih vozičkov, igranje računalniških igrice ali klasifikacijo osnovnih čustev.

V sklopu zaključne naloge se je izvedla raziskava klasifikacije dveh stanj, ki jih dosežemo z opazovanjem različnih fotografij z uporabo nizkocenovnega vmesnika med možgani in računalnikom. V raziskavi je sodelovalo 50 subjektov. Kljub majhni uspešnosti modela, ki v praksi ni zadovoljiva, smo ugotovili, da lahko s podatki zgradimo model, ki bo prepoznal razliko med mentalnimi stanji, v katerih so subjekti v času snemanja podatkov.

Modeli, ki so uporabljali združene meritve vseh petdesetih subjektov, so bili neuspešni. Iz rezultatov lahko sklepamo, da je izgradnja modela, ki bi uspešno deloval za splošno populacijo, zahteven problem. Ob pogojih, v katerih je bila izvedena raziskava, ni mogoče razviti splošnega modela klasifikacije stanj, ki bi uspešno ločeval med razredi.

Kljub temu se je pokazalo, da so v meritvah posameznikov podatki, iz katerih je s pomočjo podatkovnega rudarjenja možno izluščiti informacijo o mentalnem stanju subjekta, če izgradimo model za vsakega posameznika posebej. Slabost tega pristopa je potreba po predhodni izgradnji učne baze za vsakega posameznika. Z omenjenim pristopom se je izkazalo, da lahko s povprečno 61% natančnostjo klasificiramo meritve podatkovnih zbirk raziskave.

6 Literatura

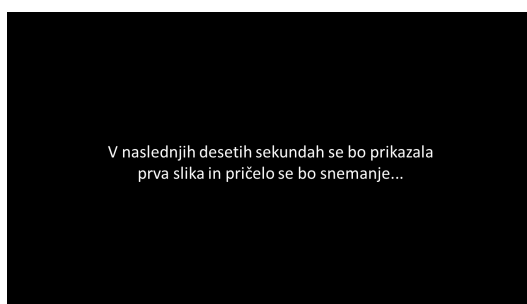
- [1] T. AL-ANI in D. TRAD, *Intelligent and Biosensors*, Signal Processing and Classification Approaches for Brain-computer Interface , UTIC-ESSTTS (2001), 28-29. (*Citirano na strani 3.*)
- [2] O. ALZOUBI in I. KOPRINSKA, Classification of Brain-Computer Interface Data, *Pre-processing and Feature Selection* AUSDM (2006), 123–132. (*Citirano na straneh 10, 12 in 14.*)
- [3] C.W. ANDERSON in Z. SIJERCIC, Proceedings of the Conference on Engineering Applications in Neural Networks, *Classification of EEG Signals from Four Subjects During Five Mental Tasks* EANN (1996), 407-414. (*Citirano na straneh 13, 14 in 16.*)
- [4] ATTRIBUTE-RELATION FILE FORMAT (ARFF), *Waikato Environment for Knowledge Analysis*, <http://www.cs.waikato.ac.nz/ml/weka/arff.html>. (*Citirano na strani 21.*)
- [5] A. AZEVEDO, *IADIS European Conf. Data Mining, KDD, SEMMA AND CRISP-DM: A Paralell Overview*, IADIS (2008), 182-185. (*Citirano na straneh 10 in 11.*)
- [6] P. BERKA, J. RAUCH in A.D. ZIGHED, *Data Mining and Medical Knowledge Management - Cases and Applications*, EEG Data Mining Using PCA, (2009), 161-180. (*Ni citirano.*)
- [7] S. CHAKRABARTI in E. COX, *Data Mining: Know It All*, Algorithms: The Basic Methods, Elsevier (2009), 131-188. (*Citirano na strani 23.*)
- [8] EEG RECORDER, *NeuroSky store*, <http://store.neurosky.com/products/eeg-recorder>. (*Citirano na strani 22.*)
- [9] B.M. FAIRA in L.P. REIS, N. LAU, Cerebral Palsy EEG signals Classification: Facial Expressions and Thoughts for Driving an Intelligent Wheelchair, *Brain Computer Interface* DETI/UA (2012), 2-5. (*Citirano na straneh 12 in 14.*)

- [10] M. HALL in E. FRANK, G. HOLMES, B. PFAHRINGER, P. REUTEMANN, I. WITTEN, *The WEKA Data Mining Software, An Update*, SIGKDD Explorations (2009), 8-35. (*Citirano na strani 24.*)
- [11] J. HAN, M. KAMBER in J. PEI, *Data Mining: Concepts and Techniques*, Data Mining Functionalities, Elsevier (2011), 21-27. (*Citirano na straneh 1, 8, 9 in 10.*)
- [12] D. HAND in H. MANNILA, *Principles of Data Mining, Models and Patterns*, MIT Press (2001), 165-183. (*Citirano na straneh 8 in 9.*)
- [13] E.A. LARSEN, *Classification of EEG Signals in a Brain-Computer Interface System*, Graphical sampling screen used to collect raw EEG samples, (2011), 28-36. (*Citirano na straneh 1, 3, 4, 5 in 14.*)
- [14] K. LIAO in R. XIAO, J. GONZALES, L. DING, Decoding individual finger movements from one hand using human EEG signals, *Data analysis PLoS One* (2014), 3-12. (*Citirano na straneh 12 in 14.*)
- [15] O. MAIMON in L. ROKACH, *Data Mining and Knowledge Discovery Handbook, Neural Networks For Data Mining*, Springer (2010), 419-445. (*Citirano na straneh 14 in 25.*)
- [16] MINDWAVE SCIFICATIONS, *Sensor Worksop*, <http://bit.ly/1jcJFEU>. (*Citirano na strani 18.*)
- [17] M. MURUGAPPAN, N RAMACHANDRAN in Y. SAZALI, Classification of human emotion from EEG using discrete wavelet transform, *Journal of Biomedical Science and Engineering* (2010), 390-396. (*Citirano na straneh 3, 8, 13 in 14.*)
- [18] NEUROSKY MINDWAVE SPECIFICATIONS, *NeuroSky MindWave store*, <http://store.neurosky.com/products/mindwave>. (*Citirano na straneh 1, 17 in 19.*)
- [19] J.B. OCHOA, *EEG Signal Classification for Brain Computer Interface Applications*, EEG Signal Classification, (2002), 36-49. (*Citirano na straneh 4, 5, 16 in 18.*)
- [20] P.M. PARDALOS, A. VAZACOPOULOS in V.L. BOGINSKI, *Data Mining in Biomedicine*, Data Mining in EEG : Application to Epileptic Brain Disorders, Springer (2007), 459-480. (*Citirano na strani 1.*)
- [21] W. SALABUN, *Processing and spectral analysis of the raw EEG signal from the MindWave*, , *Przegląd Elektrotechniczny* (2014), 169-173. (*Citirano na straneh 2, 3, 5, 7, 8, 16 in 18.*)

- [22] S. SANEI in J.A. CHAMBERS, *EEG signal processing*, EEG signal modeling, John Wiley Sons (2007), 940-965. (*Citirano na straneh 1, 3, 4, 5 in 19.*)
- [23] C. SHEARER, *The CRISP-DM Model: The New Blueprint for Data Mining*, Journal of Data Warehousing, (2000), 13-22. (*Citirano na strani 11.*)
- [24] I. SOMMERVILLE, *Software Engineering (9th Edition)*, Software Processes, Sommerville (2010), 27-46. (*Citirano na strani 21.*)
- [25] B.D. STOREY, *Computing Fourier Series and Power Spectrum with MATLAB*, (2006), 1-6. (*Citirano na strani 7.*)
- [26] A.B. SULENAM in T.A.F. FATEHI, Features Extraction Techniques of EEG Signal for BCI Applications, J. Biomedical Science and Engineering (2010), 1-4. (*Citirano na strani 8.*)
- [27] I.H. WITTEN in E. FRANK, *Data Mining: Practical Machine Learning Tools and Techniques*, Knowledge representation, Elsevier (2011), 61-84. (*Citirano na straneh 12 in 23.*)

Priloge

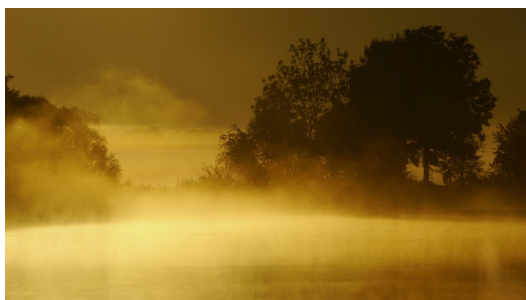
Priloga 1: Zaslonski posnetki izdelane programske opreme



Slika 8: Zaslonski posnetek prve faze snemanja



Slika 9: Zaslonski posnetek druge faze snemanja



Slika 10: Zaslonski posnetek tretje faze snemanja

Priloga 2: QR kodi za dostop do programske opreme in podatkov



Slika 11: QR koda za dostop do podatkovnih zbirk



Slika 12: QR koda za dostop do programske opreme

Priloga 3: Primer .arff datoteke iz zbirke podatkov

@RELATION EEG

@ATTRIBUTE Alpha_1 NUMERIC

@ATTRIBUTE Alpha_2 NUMERIC

@ATTRIBUTE Beta_1 NUMERIC

@ATTRIBUTE Beta_2 NUMERIC

@ATTRIBUTE Gamma_1 NUMERIC

@ATTRIBUTE Gamma_2 NUMERIC

@ATTRIBUTE Delta NUMERIC

@ATTRIBUTE Theta NUMERIC

@ATTRIBUTE class stanje0,stanje1,stanje2

@DATA

7895,9606,35537,17377,13139,8089,29218,28531,stanje0

8257,15862,18328,14208,6583,6017,51790,17066,stanje0

7680,4193,9606,11975,5252,3387,21023,12692,stanje0

6826,2474,6554,9151,5294,2098,69938,9411,stanje0

2127,12472,2496,17059,5562,2173,17232,20390,stanje0

2864,3034,1146,751,92,21,212108,8281,stanje1

2904,1380,1856,741,289,185,78654,5155,stanje1

2601,2534,2057,2145,1313,813,16245,4056,stanje1

30744,13394,10955,10556,11381,2841,46735,53709,stanje1

9259,20731,5309,13533,4901,2979,455713,18191,stanje1

18793,7675,3866,3626,8237,1126,29661,9105,stanje1

6085,24810,16704,7696,6991,3539,22363,12282,stanje1

31714,11053,13103,14075,6511,4071,17774,29669,stanje1

4849,20475,3745,9812,5951,2636,16120,27077,stanje1

8417,25102,17951,15166,5771,4574,24856,38464,stanje1

16418,8749,9334,6773,3761,1099,19049,5762,stanje1
18491,9755,2202,6766,3822,1614,12873,43756,stanje1
2243,5289,8693,5544,2990,2800,11665,8718,stanje1
17246,16741,8957,8035,5292,1743,21456,6469,stanje1
6551,7536,3422,8666,8427,2208,12476,53138,stanje1
6049,7363,4312,5278,4941,1258,8614,25263,stanje1
1301,10689,2731,8555,5256,5718,41480,19985,stanje1
10340,11615,2679,3985,5477,3908,7495,11092,stanje1
2579,3337,2865,4567,1562,2271,35753,6860,stanje1
34266,10532,5632,6110,3974,2457,6961,7832,stanje1
4880,10248,10251,14667,4549,1481,20180,19926,stanje2
4029,8527,8784,10541,3642,4624,36499,17274,stanje2
8595,6797,9707,9117,8114,2405,12785,30312,stanje2
7607,5339,13770,5739,4012,2271,23201,4641,stanje2
29185,507,5945,12435,5451,2281,6347,16176,stanje2
16547,7100,6361,3807,4740,1733,6876,19092,stanje2
11629,4383,4155,3636,2541,3013,4435,8697,stanje2
36975,10212,2625,5460,2800,2167,6419,4863,stanje2
9940,5117,2560,4234,6006,3616,1755,2431,stanje2
3669,6037,3068,3387,792,1578,6852,5969,stanje2
728,6082,4173,4996,3446,4154,16688,7440,stanje2
600,12224,6498,6644,2215,918,15160,3101,stanje2
1436,1257,1941,8224,7354,1597,10500,18616,stanje2
2335,4177,2840,5220,2848,4439,12590,11937,stanje2
2598,5396,1839,8762,2848,2578,9347,7555,stanje2
36651,9088,4442,12987,3976,1817,14358,8873,stanje2
24721,31552,15552,10799,6955,1595,127033,54799,stanje2
14249,15747,2897,7423,3391,1114,10457,13927,stanje2
22927,5111,10982,5803,5283,1481,13943,4219,stanje2
8349,6099,7876,4252,4136,1719,18433,16143,stanje2