



Elektroencefalogramų analizės metodų tyrimas

Ataskaita už trečiuosius doktorantūros metus (2015
spalis – 2019 spalį)

Doktorantas: Andrius Vytautas Misiukas
Misiūnas

El. paštas: andrius.misiukas@mif.vu.lt

Vadovas: prof. dr. Tadas Meškauskas

FORMALIŲ REIKALAVIMŲ VYKDYMAS

Egzaminų laikymas

- Daugiamačių duomenų vizualizavimo metodai. Komisija: prof. habil. dr. Gintautas Dzemyda, prof. dr. Olga Kurasova, dr. Viktor Medvedev. Įsakymas Nr. (1.3)580000-DV-31. Egzamino data: 2017.10.26. Įvertinimas: **9 balai (labai gerai)**.
- Visi likę egzaminai išlaikyti ankstesniais metais (**10, 10, 9 balai**).

Dalyvavimas konferencijose

- Suplanuota: Tarptautinės konferencijos, skirtos matematikos ir informatikos taikymams medicinoje arba signalų analizėje (2017-2019 metais).
 - Įgyvendinta: 9th International Conference on Numerical Methods and Applications (NM&A'18). 2018.08.20-2018.08.24, Borovets, Bulgarija.
- Suplanuota: Lietuvos Matematikų Draugijos konferencija (2017-2019 metais)
 - Įgyvendinta ankstesniais metais (dalyvauta LMD 56 ir 57 konferencijose)
- Suplanuota: Data Analysis Methods for Software Systems konferencija (2017-2019 metais)
 - Planuojama įgyvendinti šiais mokslo metais, užpildyta registracijos forma ir išsiųsta santrauka 2018.10.7

Publikacijos šiais mokslo metais

- Publikacija, suplanuota 2016/2017 mokslo metams:
 - Algorithm for automatic EEG classification according to the epilepsy type: benign focal childhood epilepsy and structural focal epilepsy. *Biomedical Signal Processing and Control, Elsevier IF=2.783.*
 - Gautas sprendimas apie publikacijos **priėmimą** 2018 spalio 6 dieną, suteiktas DOI: 10.1016/j.bspc.2018.10.006
- Plane nenurodyta publikacija:
 - Accuracy of different machine learning type methodologies for EEG classification by diagnosis. *9th International Conference Numerical Methods and Applications proceedings (ISI indeksuoti), Springer lecture notes in computer science.* Publikacija recenzuota, priimta, bet dar nesuteiktas DOI ir puslapių numeriai.
- Publikacija, suplanuota 2017/2018 mokslo metams:
 - Turime preliminarius mokslinius rezultatus, publikacija ruošiamą.

Ankstesnių metų publikacijos

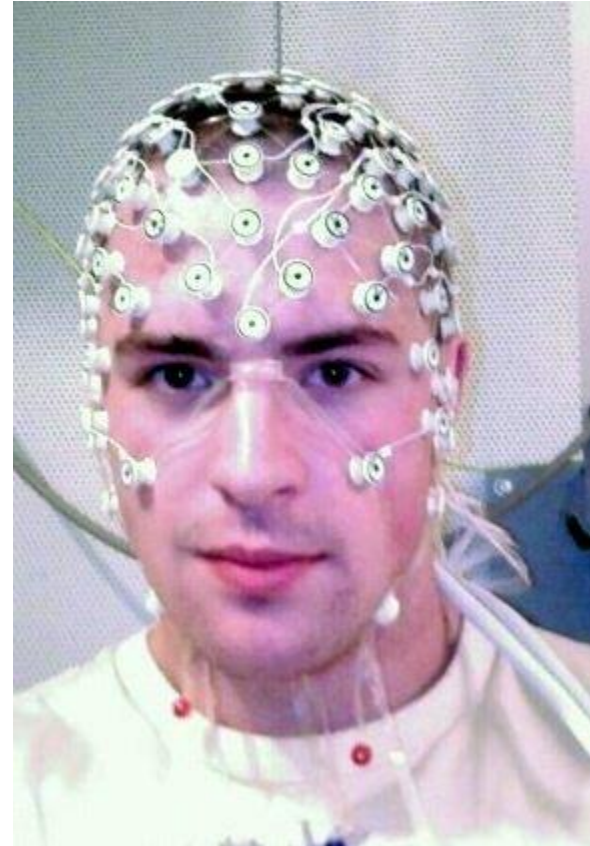
- A.V. Misiukas Misiūnas, T. Meškauskas, A. Juozapavičius
 - On the implementation and improvement of automatic EEG spike detection algorithm
 - Liet. matem. rink. Proc. LMS, Ser. A, 56, 2015, 60–65
- A.V. Misiukas Misiūnas, T. Meškauskas, R. Samaitienė
 - Derivative parameters of electroencephalograms and their measurement methods
 - Liet. matem. rink. Proc. LMS, Ser. A, 57, 2016, 47–52

Mokslinės veiklos ataskaita

ELEKTROENCEFALOGRAMŲ ANALIZĖS METODŲ TYRIMAS

Tyrimų objektas

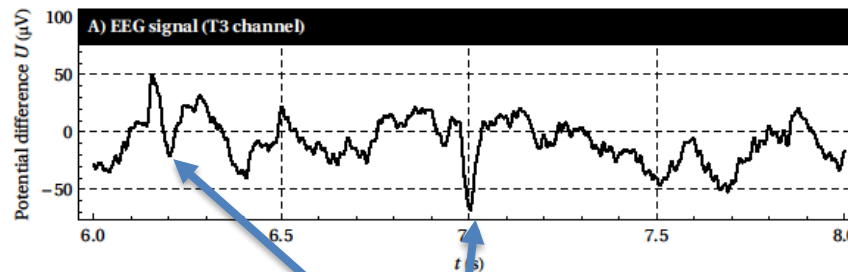
- Vaikų (3-17 metų), diagnozuotų su dviemis skirtingomis epilepsijos formomis elektroencefalogramos (EEG)
 - EEG – smegenų aktyvumo įrašas iš 21 elektrodo, uždėto ant paciento galvos odos



https://upload.wikimedia.org/wikipedia/commons/b/bf/EEG_cap.jpg

Tyrimų objektas

- EEG pikai visada rodo nenormalų smegenų aktyvumą
 - Daug centrinės nervų sistemos ligų pasižymi jais
 - Kai kurie lengvai randami ir atpažįstami, kiti – ne



EEG pikų pavyzdžiai

Pacientų grupės

RE Grupė

- Pacientai, kuriems diagnozuota Rolando epilepsija
 - Aiškūs EEG pikai
 - Pikių forma daug nesikeičia EEG eigoje

CP Grupė

- Struktūrinė židininė epilepsija, cerebrinis paralyžius ir kt.
 - Dažniausiai pikai aiškiai skiriasi savo forma nuo RE grupės
 - Kai kuriais atvejais pikai labai panašūs
 - Mes tiriamo tik tuos pikus, kurie yra labai panašūs (arba neatskiriami) neurologams

Pacientų grupės

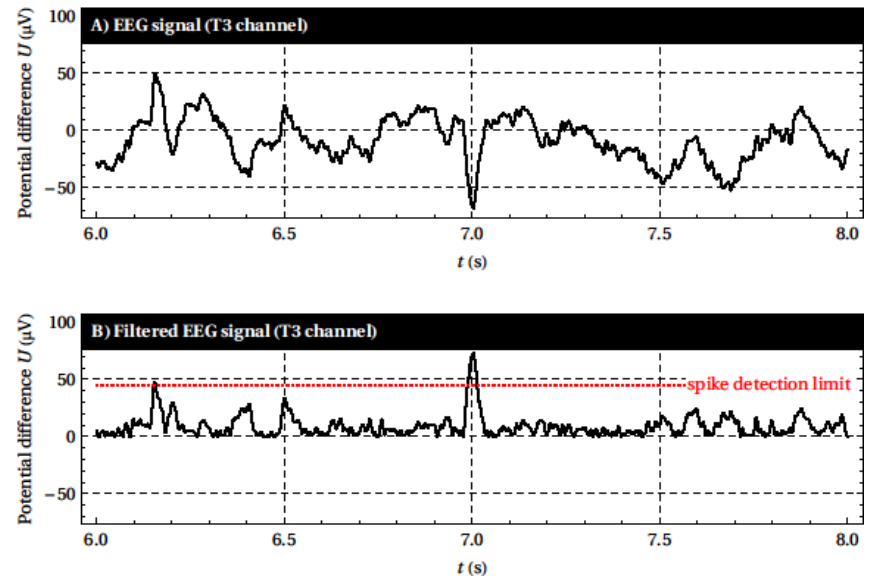
- 94 EEG iš 86 skirtingų pacientų
 - 56% berniukai, 44% mergaitės
 - 3-17 metų amžiaus
 - 62 EEGs iš RE Grupės, 32 EEG iš CP grupės
 - Apie 33% duomenų naudojami klasifikatoriaus apmokymui
 - Šį skaičių nulemia neurologų išvalytų duomenų kiekis
 - Tiksliai diagnozė žinoma iš ligos istorijos

EEG klasifikavimo algoritmo esminiai žingsniai

1. EEG pikų aptikimas
2. EEG pikų parametrų nustatymas
3. EEG klasifikavimas pasinaudojant mašinių mokymosi metodais

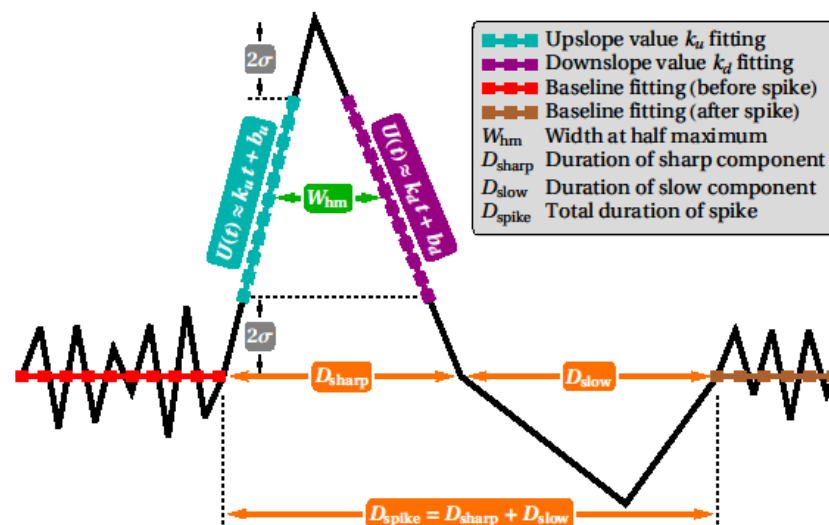
1 žingsnis. Pikų aptikimo algoritmas

- Paremtas matematinė morfologija
- Nufiltruoja žinomą normalų smegenų aktyvumą palikdamas nenormalų
- Algoritmas sugalvotas kitų autorių, patobulintas ir įgyvendintas doktoranto



2 žingsnis. EEG pikų parametrai

- Įmanoma nustatyti tam tikrus pikų parametrus
 - Galbūt tikslesnis ir objektyvesnis būdas, negu vizuali analizė
- Analizė rodo, kad reikšmingiausi diagnozei parametrai yra piko pakilimo ir nusileidimo greitis



Hipotezė

- Neurologai pastebėjo, kad RE grupės pacientų EEG pikai kinta daugiau, negu CP grupės pacientų EEG tyrimo eigoje
- Tai reiškia, kad šie pasikeitimai galbūt galėtų būti aptikti mašinių mokymusi pagrįstų klasifikavimo metodų

Duomenų strategija

- Tos pačios mokymosi ir testavimo duomenų imtys buvo naudojamos visiems tirtiems klasifikavimo algoritmams
 - Išvalyti duomenys – mokymuisi, visi kiti – testavimui
 - 34% mokymosi imtis, 66% - testavimo
 - Esant kelioms to paties paciento EEG, jos priskiriamos vienai (mokymosi arba testavimo) aibei

Duomenų strategija

- Visos EEG buvo sukarpytos į serijas, turinčias po 100 pikų
 - Daugelis mašinių mokymosi metodų reikalauja fiksuoto įvesties duomenų kiekio
 - Didelis serijų skaičius padeda susidoroti su problema, kad turima palyginti mažai EEG
 - Visų testuotų klasifikatorių veikimo patikimumas įsisotina pasiekus 100 pikų EEG atkarpos ilgį arba greičiau

Išbandyti mašinų mokymosi algoritmai

1. Sprendimų medis (angl. Decision tree)
2. Atsitiktinis miškas (angl. Random forest)
3. Labai atskiktinis medis (angl. Extremely randomized tree)
4. Adaptyvusis stiprinimas (angl. Adaptive boosting - AdaBoost)
5. Palaikomųjų vektorių mašina (angl. Supported vector machine - SVM)
6. Tiesinė diskriminantinė analizė (angl. Linear discriminant analysis - LDA)
7. Logistinė regresija (angl. Logistic regression)
8. Naivus Bayes klasifikatorius (angl. Naïve Bayes)
9. Dirbtiniai neuroniniai tinklai (angl. Artificial neural networks - ANN)

Dirbtinio neuroninio tinklo (ANN) savybės

- 100 įvesties neuronų
- 20 paslėptų neuronų
- 1 išvesties neuronas
 - Išvesties reikšmė suapvalinama iki artimiausio sveikojo skaičiaus
 - Jei išvestis yra 1 – virtuali EEG yra CP grupėje, kitu atveju – RE grupėje
- Softmax aktyvacijos funkcija
 - Toks pats veikimas kaip sigmoidinės, bet greitesnis mokymasis

Kai kurie algoritmai buvo atmesti dėl žemo patikimumo

- Prastu patikimumu pasižymėjo šie algoritmai
 - *LDA* (53%)
 - Logistinė regresija (59%)
 - *SVM* su tiesiniu ir kvadratinu branduoliais (56 ir 59% atitinkamai)

TNR ir TPR analizė

- Teisingas CP grupės detektavimas laikomas teisingu teigiamu atsakymu, RE- teisingu neigiamu
- Kai kurie algoritmai „sukčiauja“ pasinaudodami mūsų duomenų ypatybėmis
 - Naivus Bayes klasifikatorius, *SVM* su *RFB* ir Sigmoidiniu branduoliais visas įvestis klasifikuoja kaip RE Grupę
 - Kadangi 73% mūsų turimų pikų yra su RE diagnoze, todėl pasiekiamas 73% patikimumas
 - Šie algoritmai toliau nebuvo nagrinėjami

TNR and TPR analysis

Metrika / Algoritmas	Atsitiktinis miškas	Sprendimų medis	Labai atsitiktinis medis	AdaBoost	ANN	SVM N = 3
Tikslumas	0,78	0,76	0,80	0,81	0,75	0,69
TNR	0,79	0,76	0,83	0,90	0,79	0,79
TPR	0,74	0,77	0,71	0,52	0,74	0,48
F1 metrika	0,76	0,76	0,75	0,64	0,78	0,57
Cohen kapa	0,06	-0,01	0,12	0,38	0,28	0,26

Teisingas atsakymas gaunamas spėlioiant

Prastai aptinkamos CP EEG

F1 metrika atsižvelgia tiek į TPR tiek į TNR

Spėliojimas

- Kai kurie mašinų mokymosi algoritmai naudojami mūsų CP ir RE imčių skirtingu dydžiu
 - Greičiausiai šią informaciją naudoja spėliojant
 - Į problemą žiūri ne kaip į klasifikavimo, o kaip į nesubalansuotos monetos metimą
 - Įvesčių tvarka ar galimai kiti dalykai
 - Tie patys rezultatai gaunami iš *Cohen kappa*, Jaccard panašumo metrikos, *log loss* ir *zero one loss* metrikų

Išvados

- CP ir RE grupės EEG gali būti klasifikuojami mašinių mokymosi metodais
- Pasiektas 73% patikimumas
- Mūsų užduočiai šiuo metu tinkamiausias yra ANN klasifikatorius
 - Reikia tiek kuo artimesnio vienetai TPR tiek TNR

Tyrimų planas

- Tikimybės nustatymas, kad aptiktas EEG darinys yra pikas arba EEG artefaktas
- Morfologinio filtro parametrų parinkimo patobulinimas genetiniu algoritmu
 - Pagerėtų klasifikacijos tikslumas
 - Šie tyrimai jau vykdomi šiuo metu



Ačiū už dėmesį