

# Analýza a zpracování EEG signálu

Autor: Aleksandra Bajičová, bajicale@fel.cvut.cz

## Cíle laboratorní úlohy

1. Zaznamenat 4kanálový EEG signál
2. Na základě vybraných časově-frekvenčních analýz identifikovat změny v EEG
3. Vyhodnotit změny a konektivitu mezi oblastmi mozku
4. Ověřit možnosti klasifikace stavů pomocí jednoduchých algoritmů

## Naše data

V našem experimentu jsme změřili 5 EEG signálů Kamila. Dvakrát jsme změřili otevírání a zavírání očí po 10 sekundách, celý záznam trval 40 sekund. Dále jsme změřili mačkání pěsti opět po 10 sekundách, poté myšlení na mačkání pěsti po 10 sekundách a nakonec pohyb ušima. Záznam byl proveden pomocí 4 kanálů: 1 F3 – C3 (levá hemisféra – motorika), F4 – C4 (pravá hemisféra – motorika), C3 – Cz, C4 – Cz.

## Metodika

### A. Filtrace EEG signálu

- Načíst a zobrazit signál
- Provést filtraci:
  - 0.5–30 Hz (základní EEG pásmo)
  - 8–13 Hz (alfa pásmo)
- Zobrazit filtrované signály a obálku alfa v jednotlivých kanálech

### B. Segmentace signálu

- Identifikovat přechody mezi stavy
- Označit časové body segmentace

### C. Spektrální analýza

- Rozdělit signál na segmenty dle výsledků detekce
- Vykreslit spektrogramy pro všechny 4 kanály a různé stavy
- Vyčíst a zobrazit výkon v EEG pásmech:
  - delta (0.5–4 Hz)
  - theta (4–8 Hz)
  - alfa (8–13 Hz)

- beta (13–30 Hz)

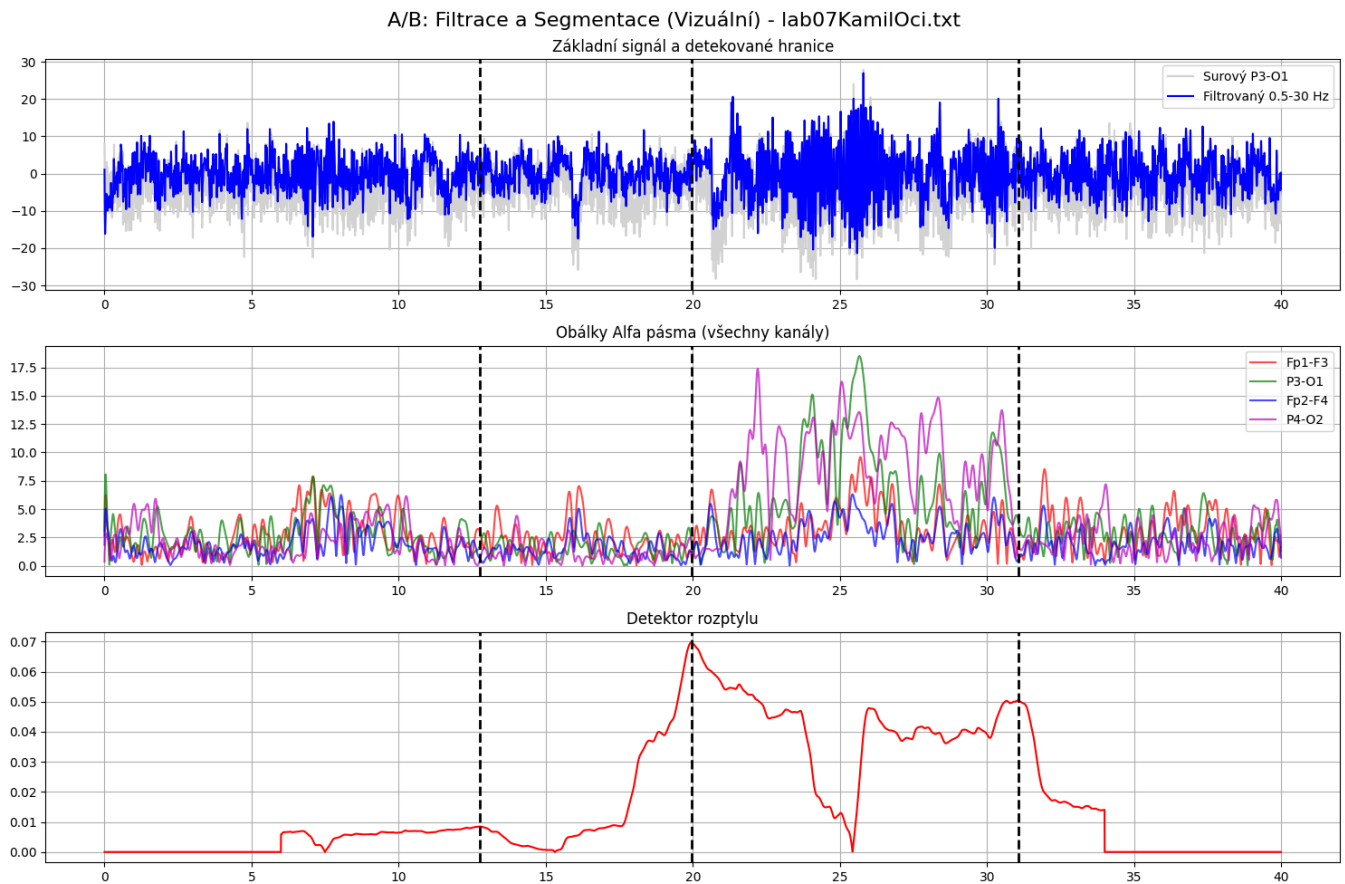
- Porovnat výkon jednotlivých pásů mezi stavy
- Porovnat spektra získaná různými metodami

## D. Konektivita (CSD)

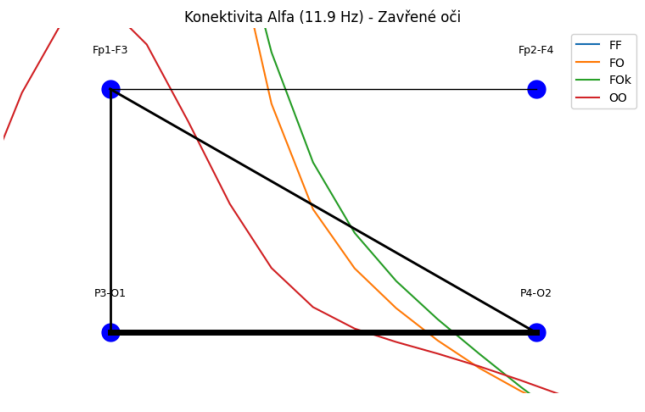
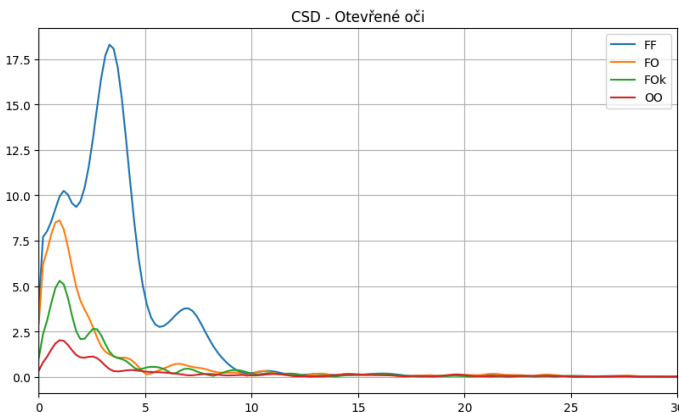
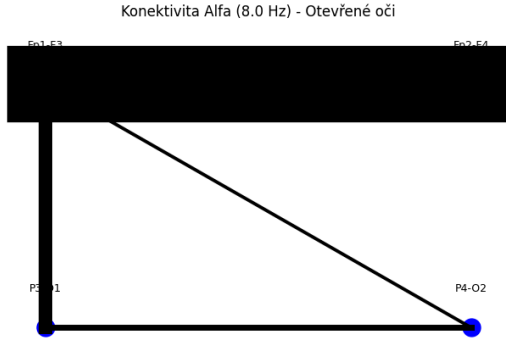
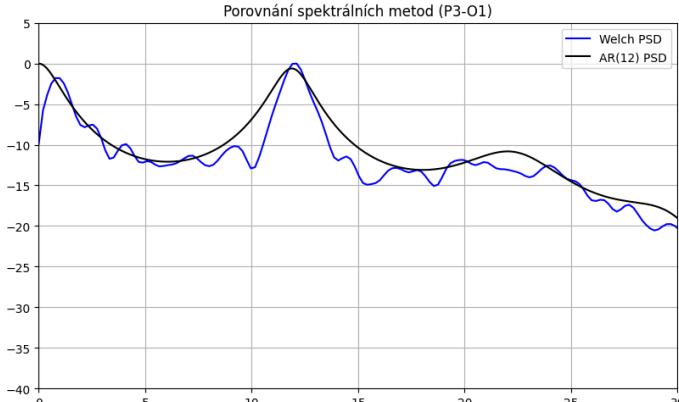
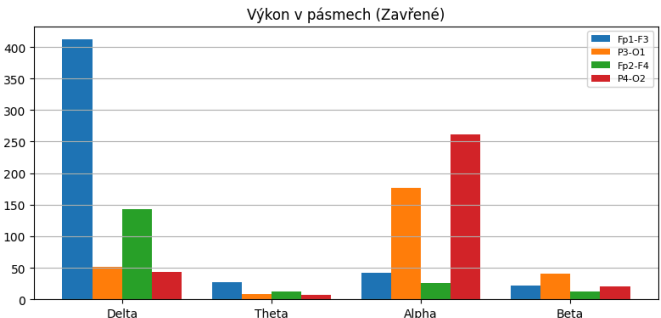
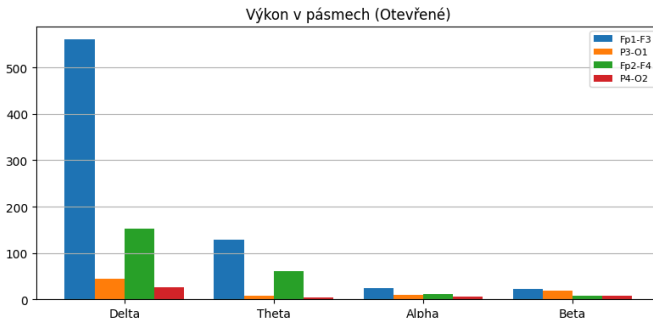
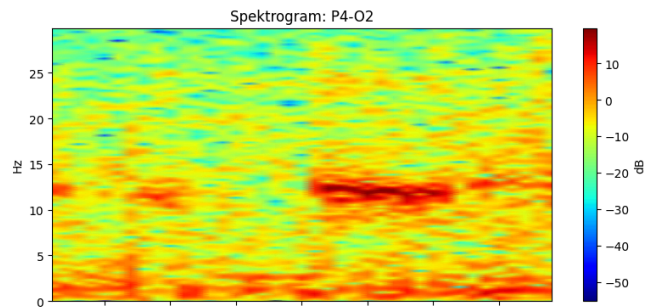
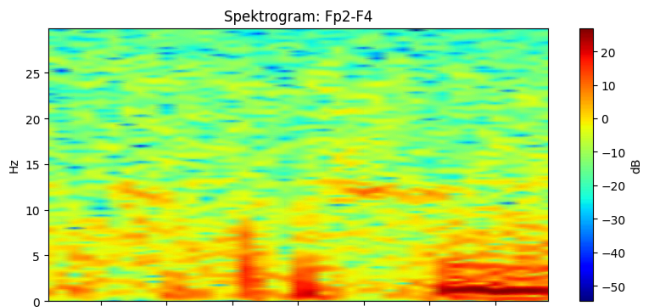
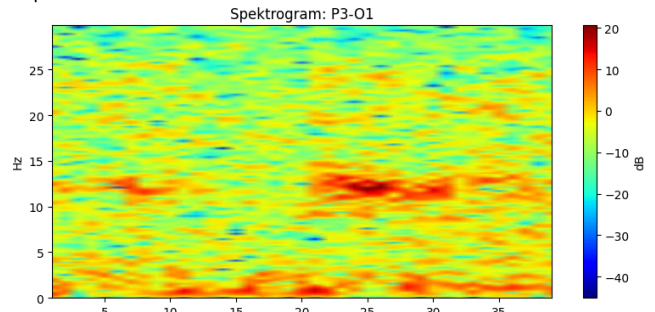
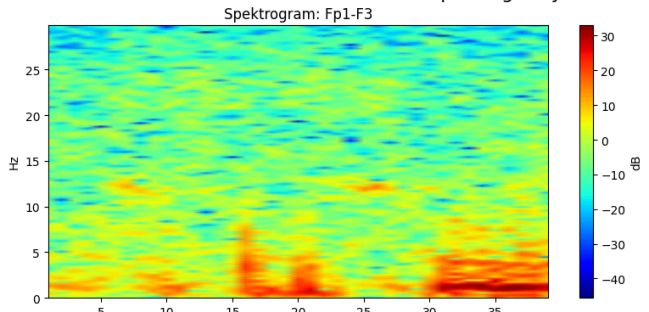
- Spočítat vzájemná spektra mezi oblastmi (např. FF, FO, OO nebo CC)
- Vykreslit CSD během různých stavů
- Sledovat změny konektivity mezi oblastmi
- Pomocí procentuálního zastoupení alfa výkonu vizualizovat změny pomocí grafu konektivity

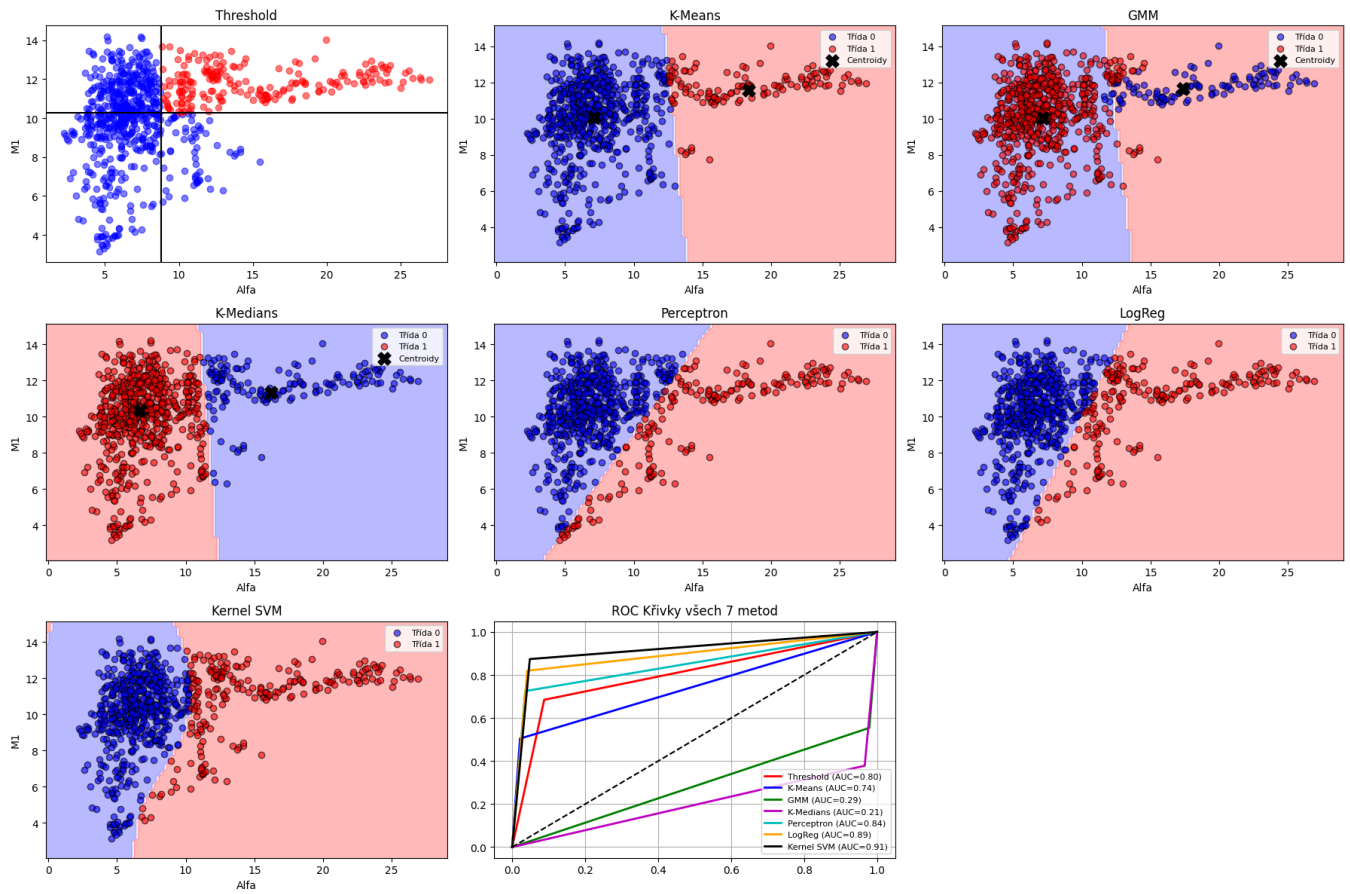
## E. Klasifikace stavů

- Vypočítat vhodné příznaky
- Provést klasifikaci pomocí alespoň tří jednoduchých metod
- Porovnat shody klasifikace mezi metodami a interpretovat výsledky



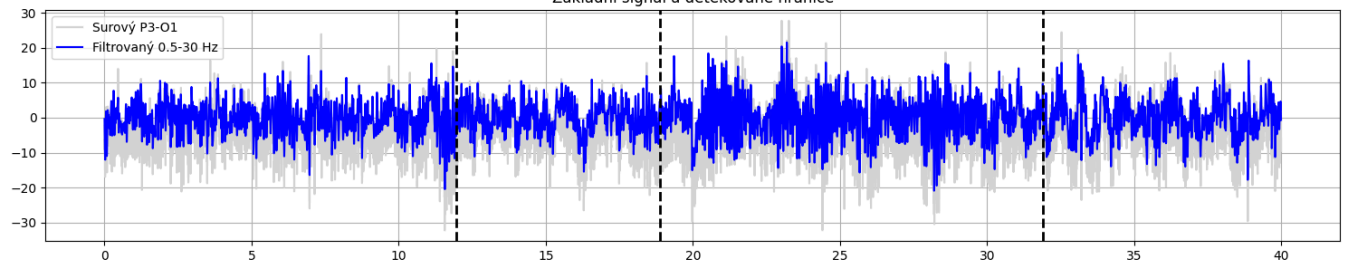
C: Spektrogramy a Porovnání Spekter - lab07KamilOci.txt



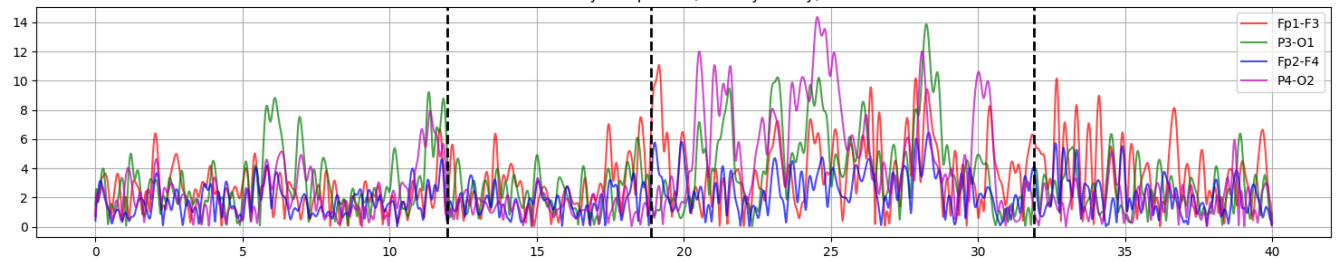


### A/B: Filtrace a Segmentace (Vizuální) - lab07KamilOci2.txt

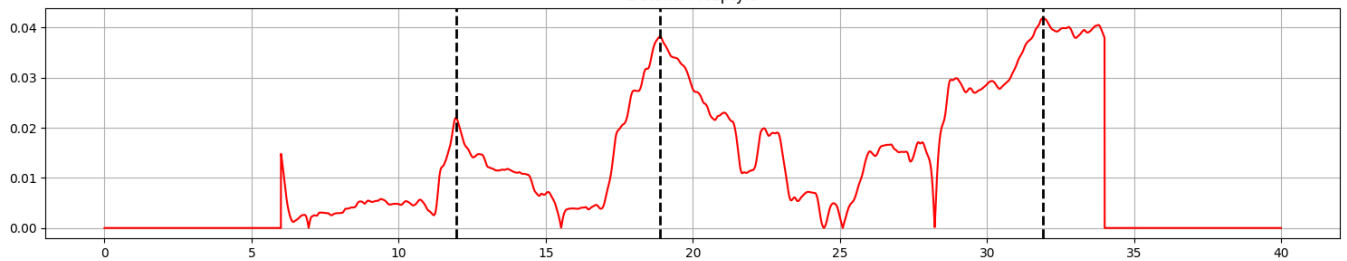
Základní signál a detekované hranice



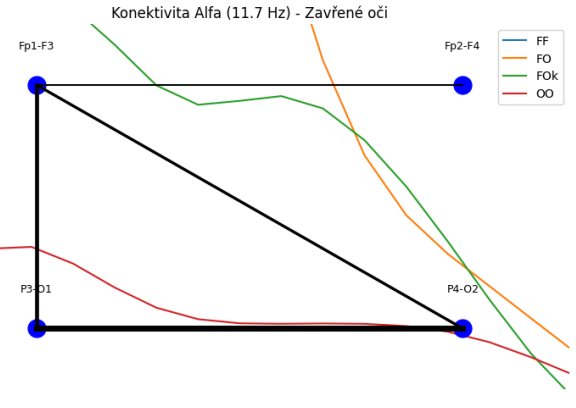
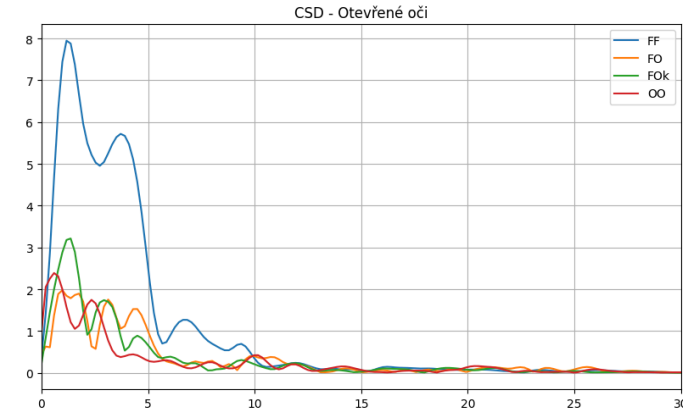
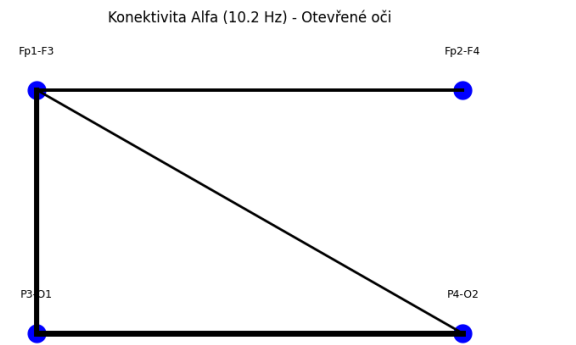
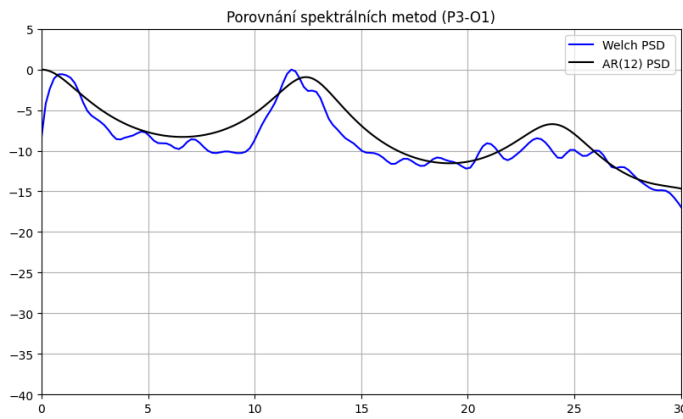
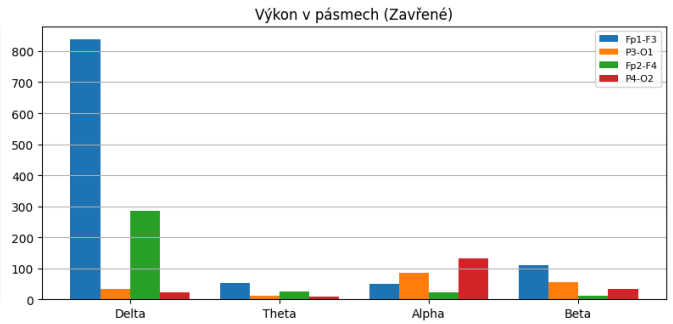
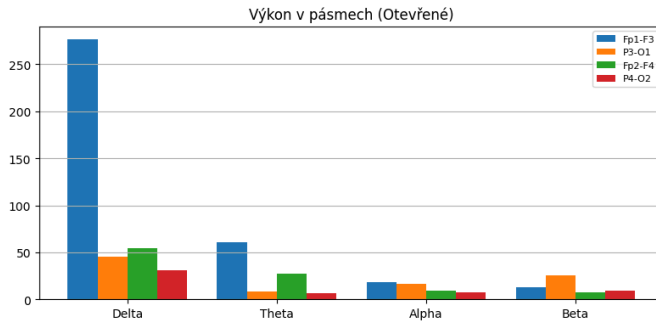
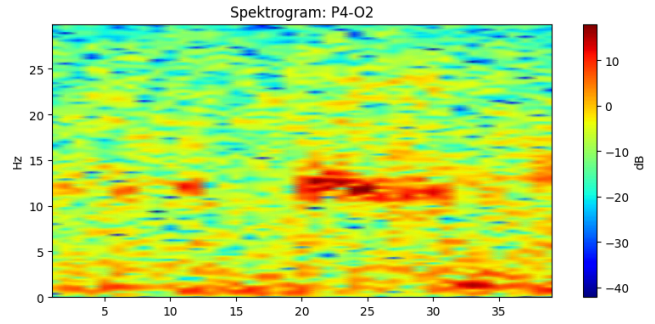
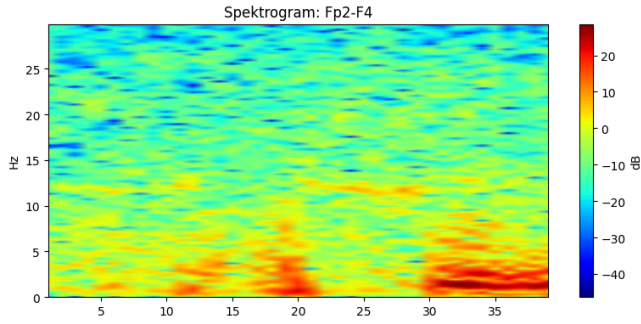
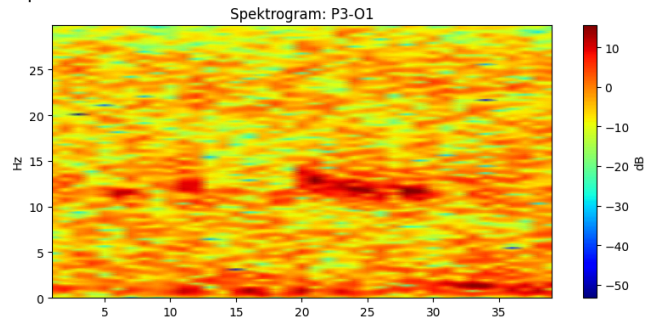
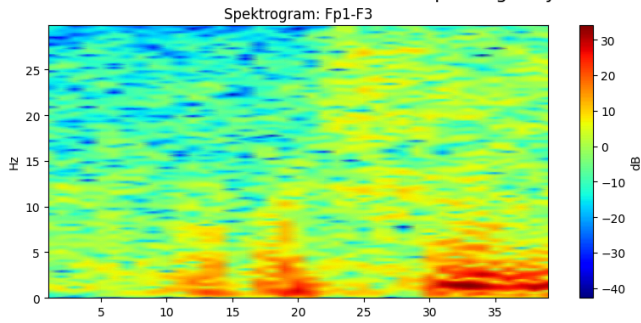
Obálky Alfa pásma (všechny kanály)

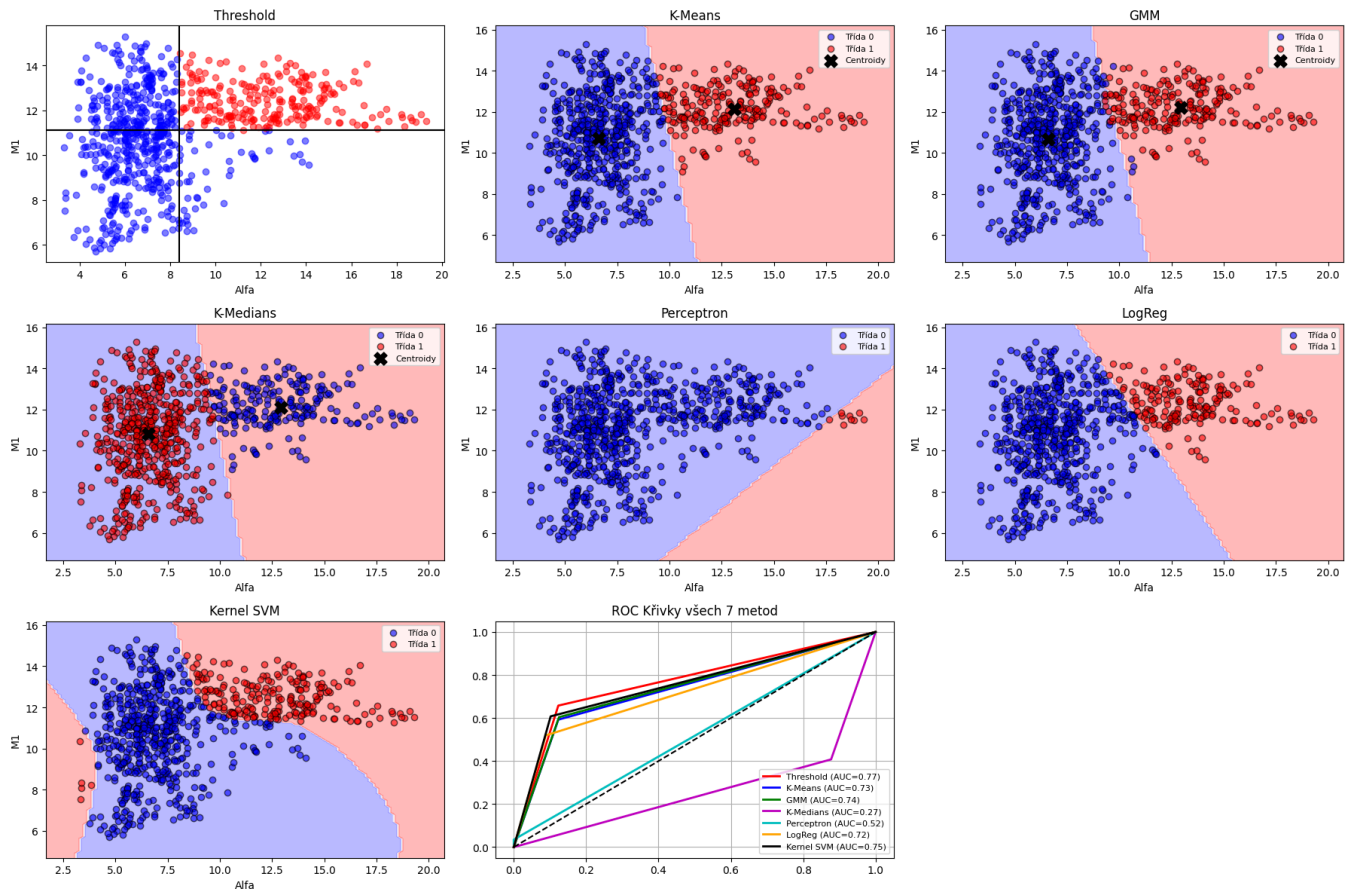


Detektor rozptylu



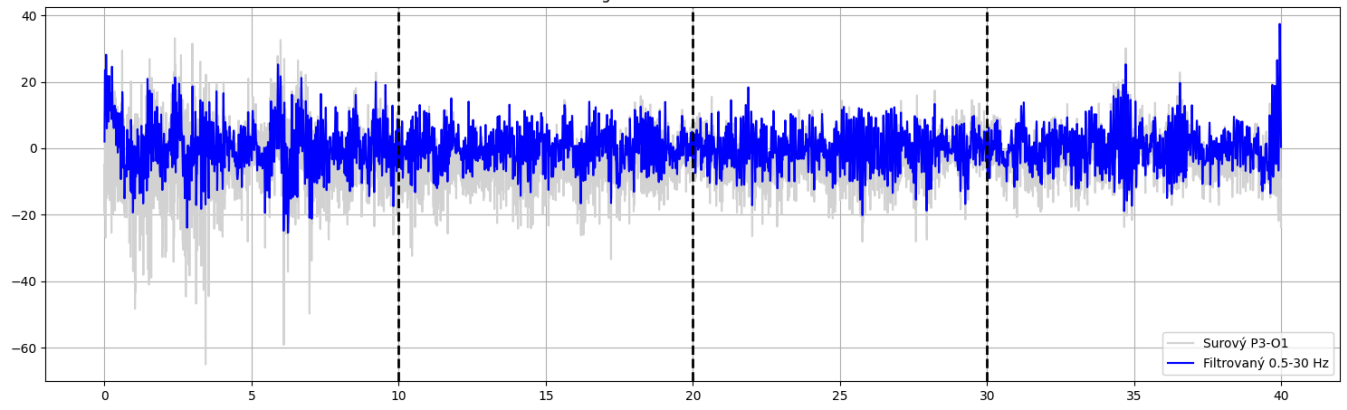
C: Spektrogramy a Porovnání Spekter - lab07KamilOci2.txt



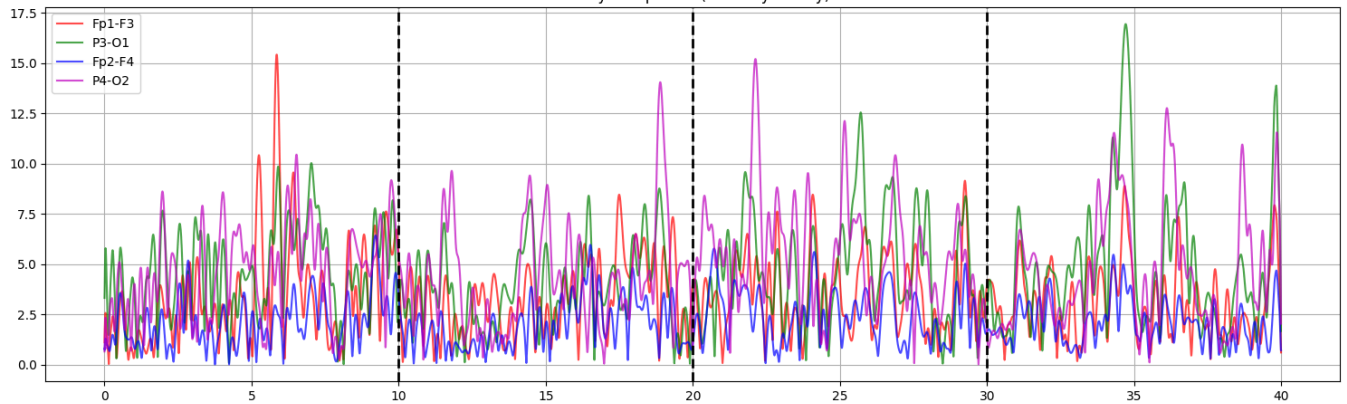


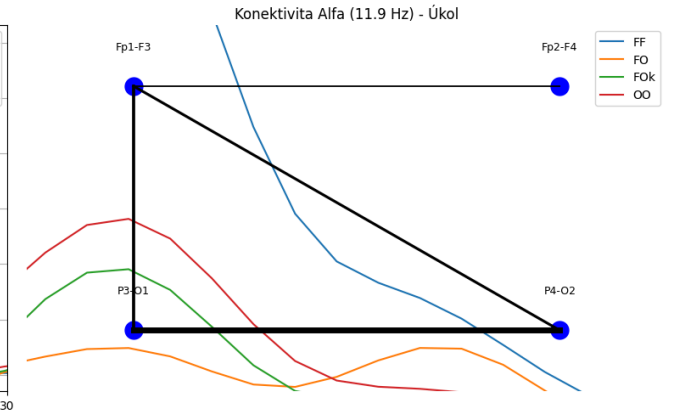
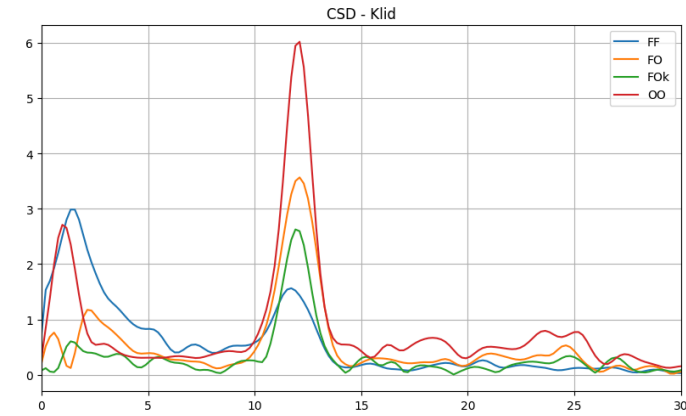
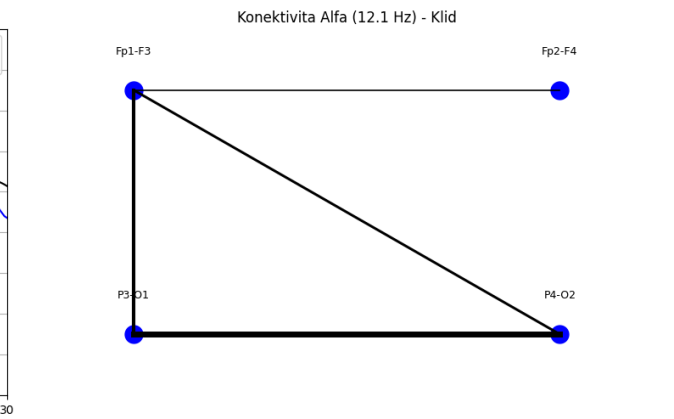
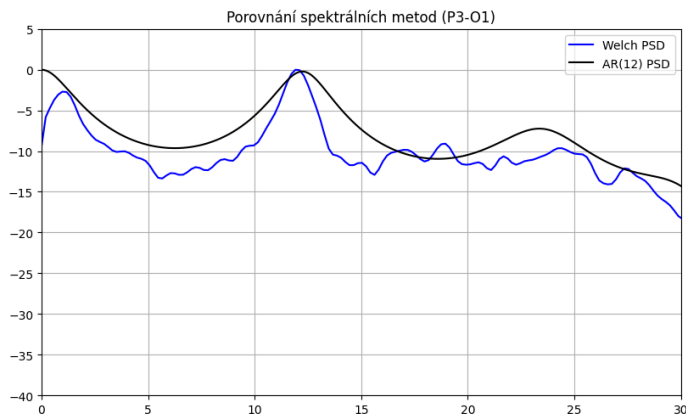
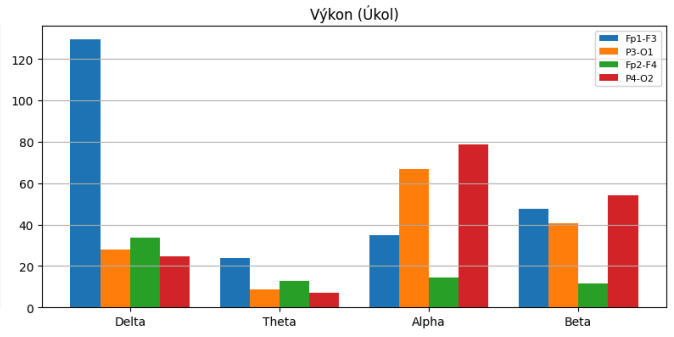
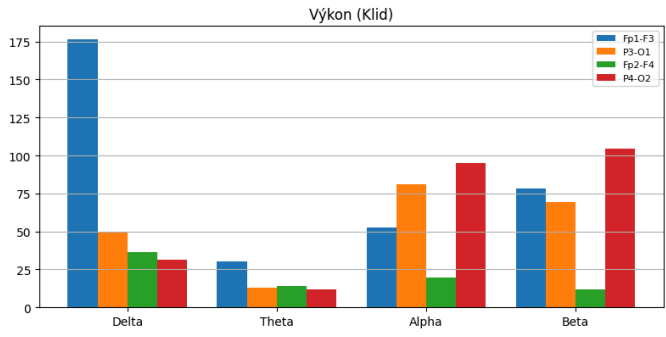
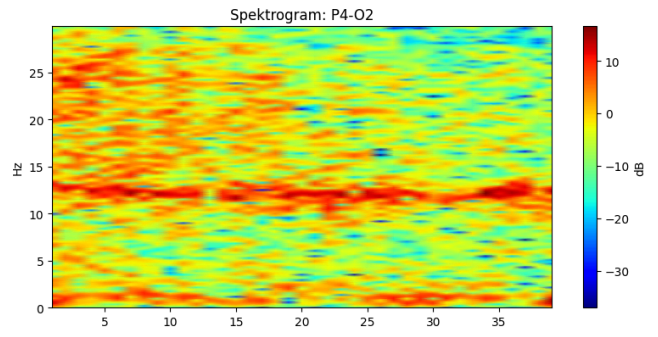
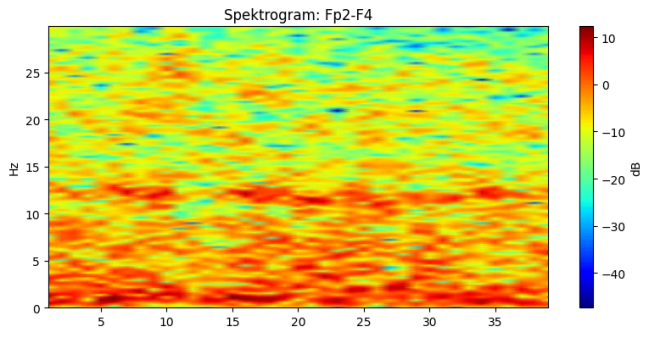
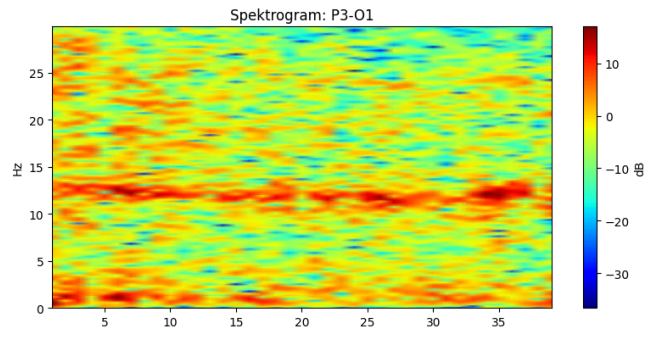
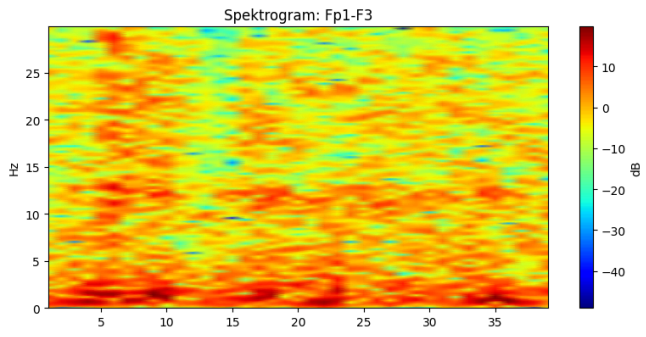
A/B: Filtrace a 10s BCI Protokol (Motorika) - lab07KamilMysleni.txt

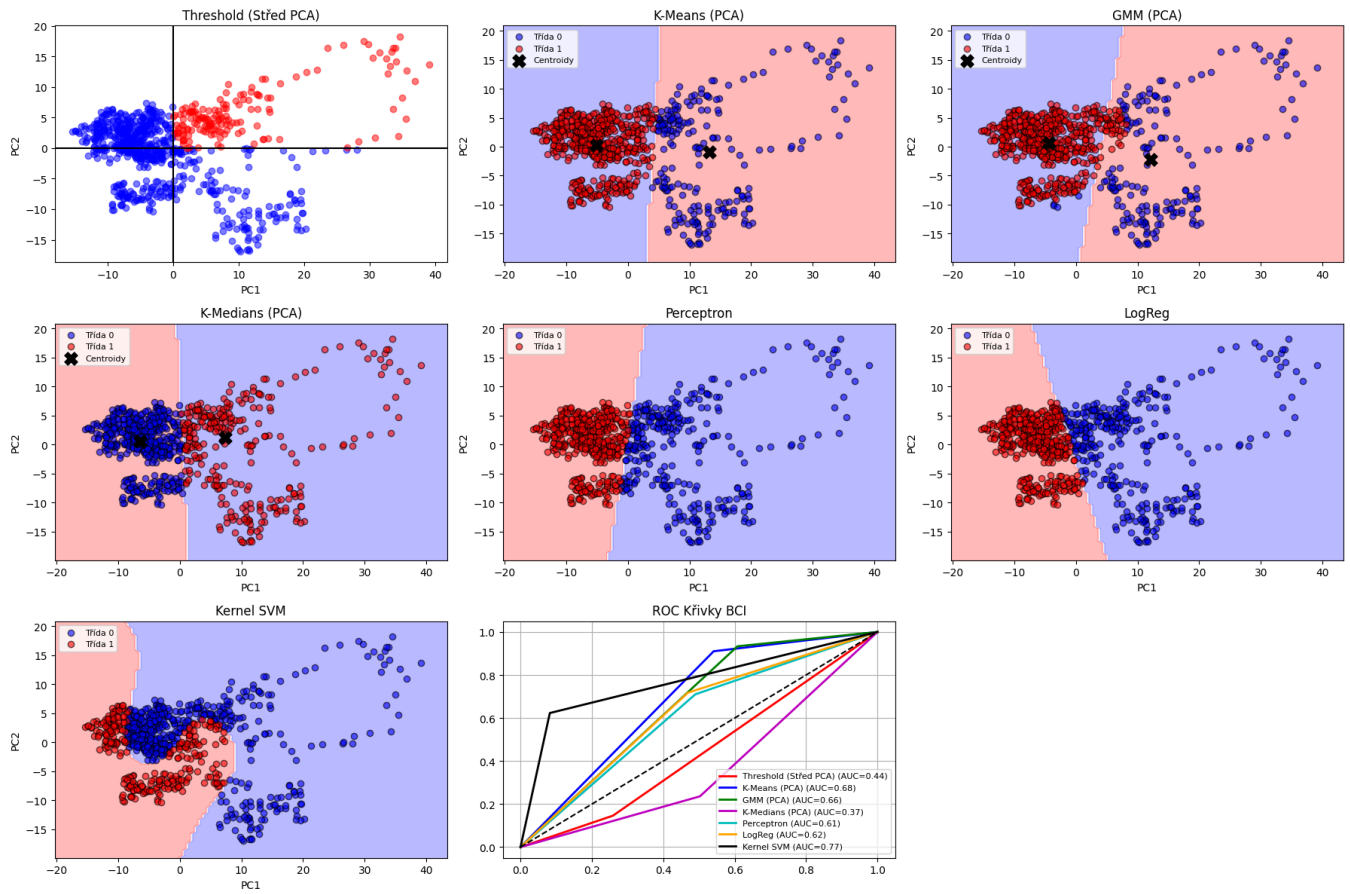
Základní signál a teoretické 10s hranice úkolu



Obálky Alfa pásma (Všechny kanály)

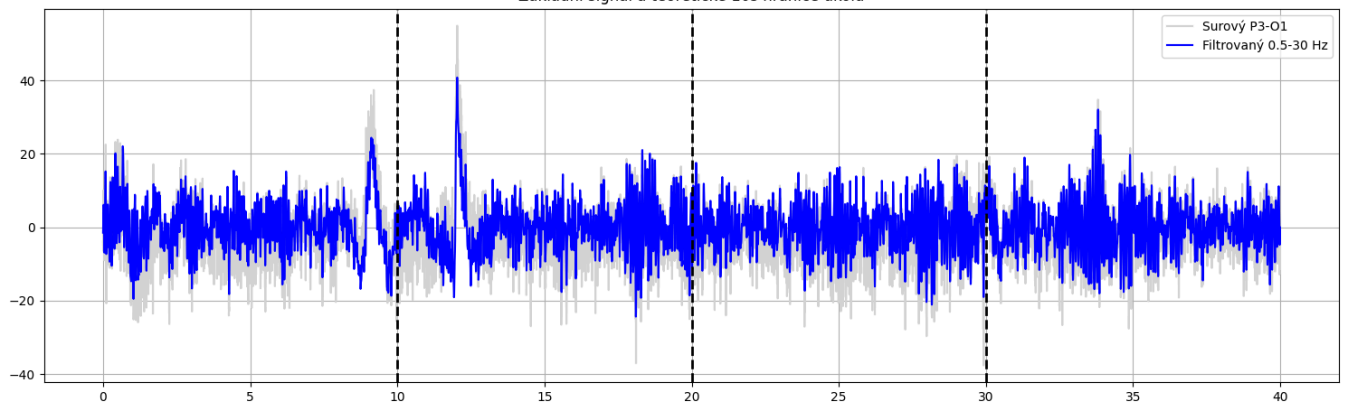




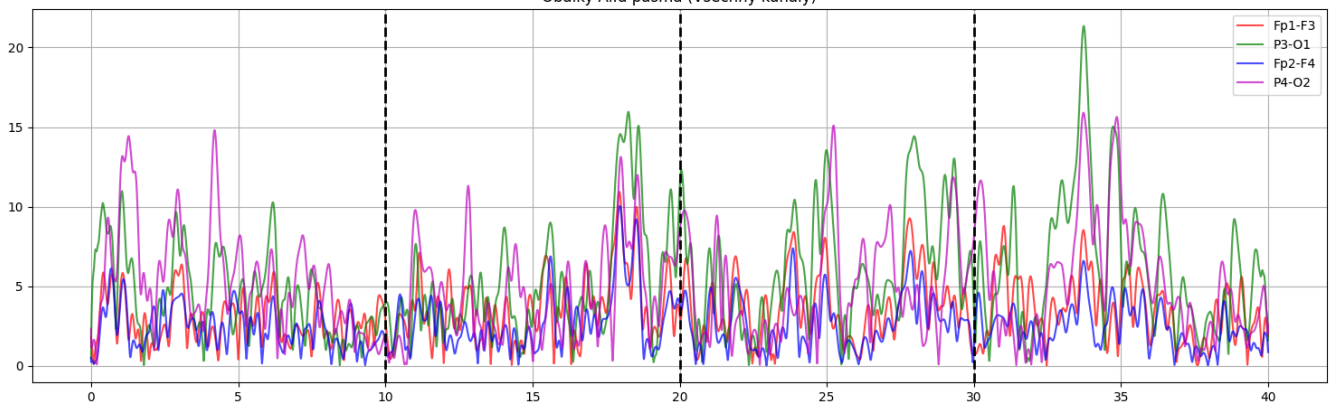


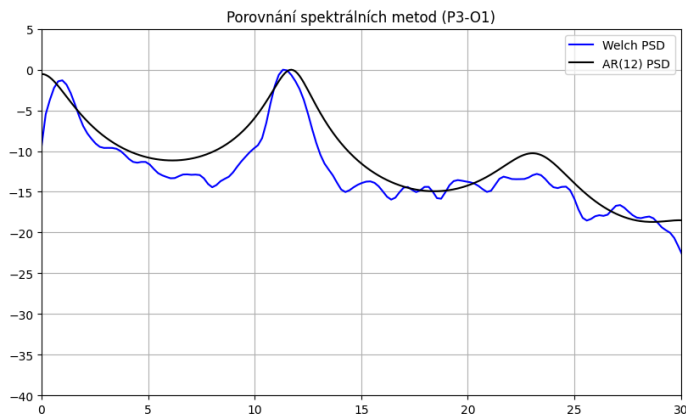
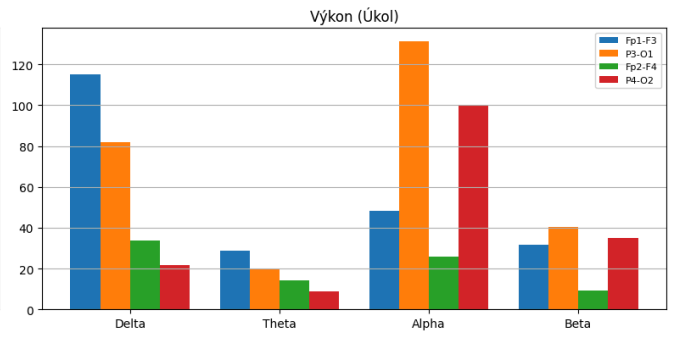
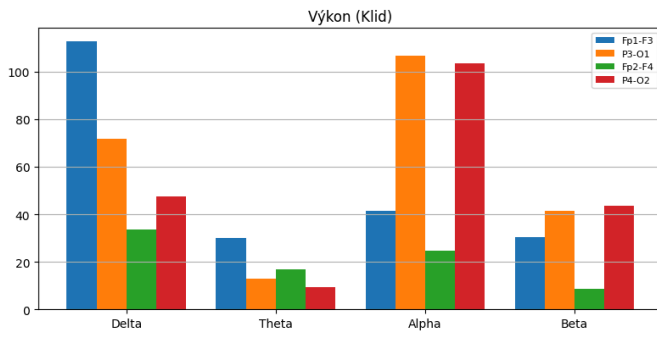
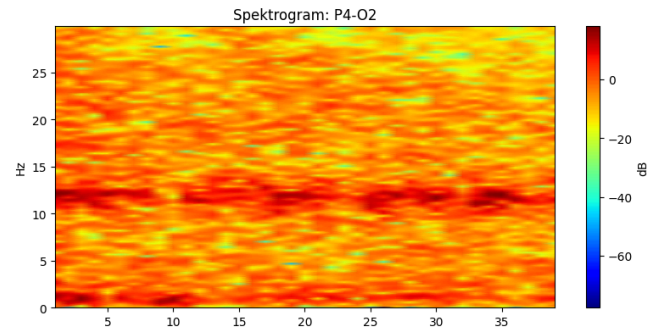
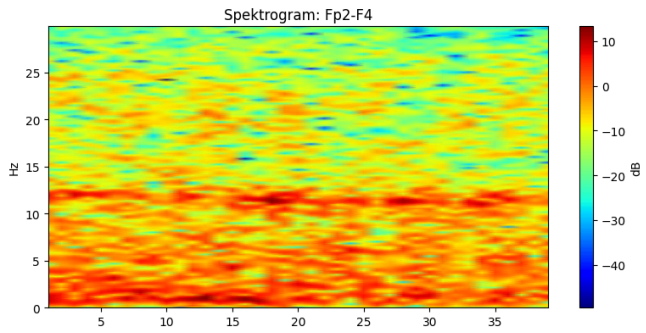
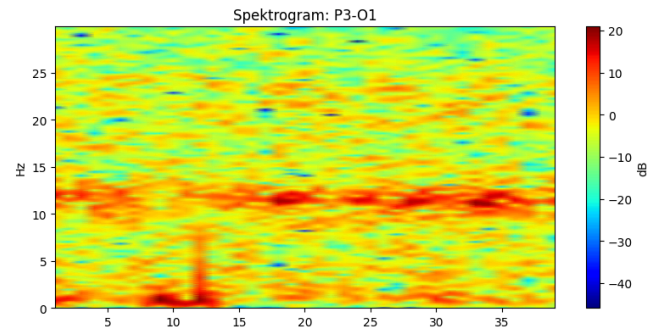
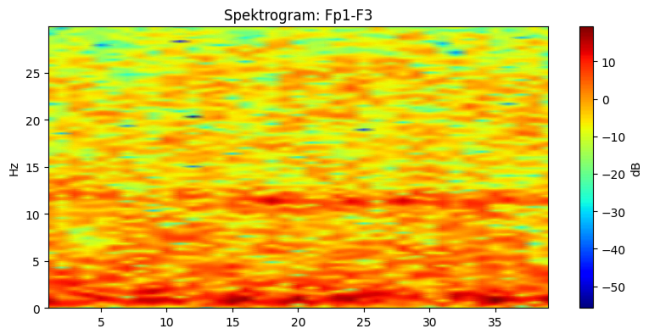
A/B: Filtrace a 10s BCI Protokol (Motorika) - lab07KamilRuka.txt

Základní signál a teoretické 10s hranice úkolu

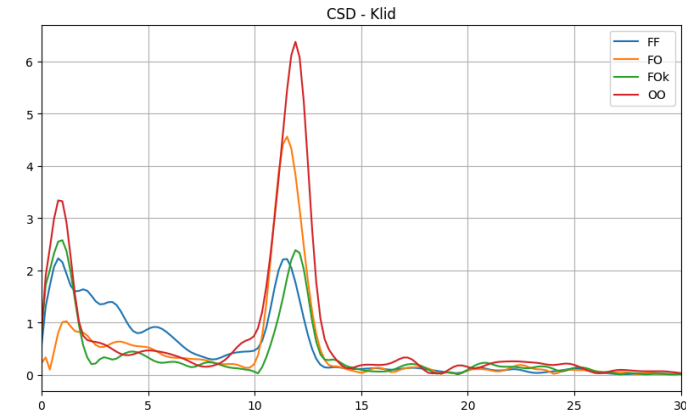
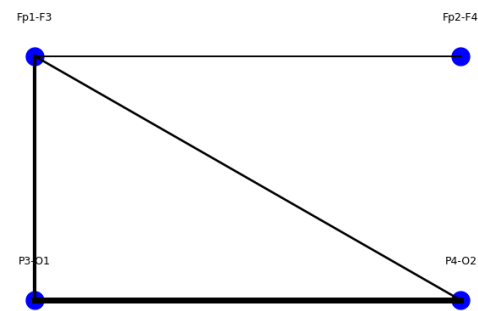


Obálky Alfa pásma (Všechny kanály)

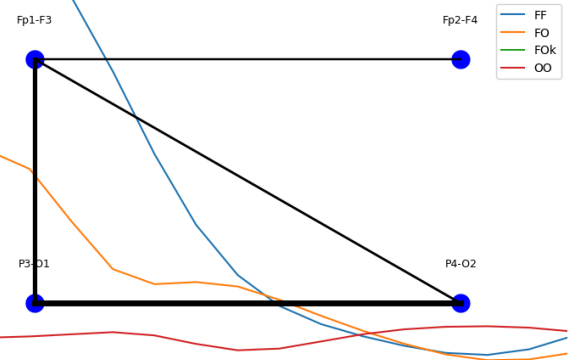


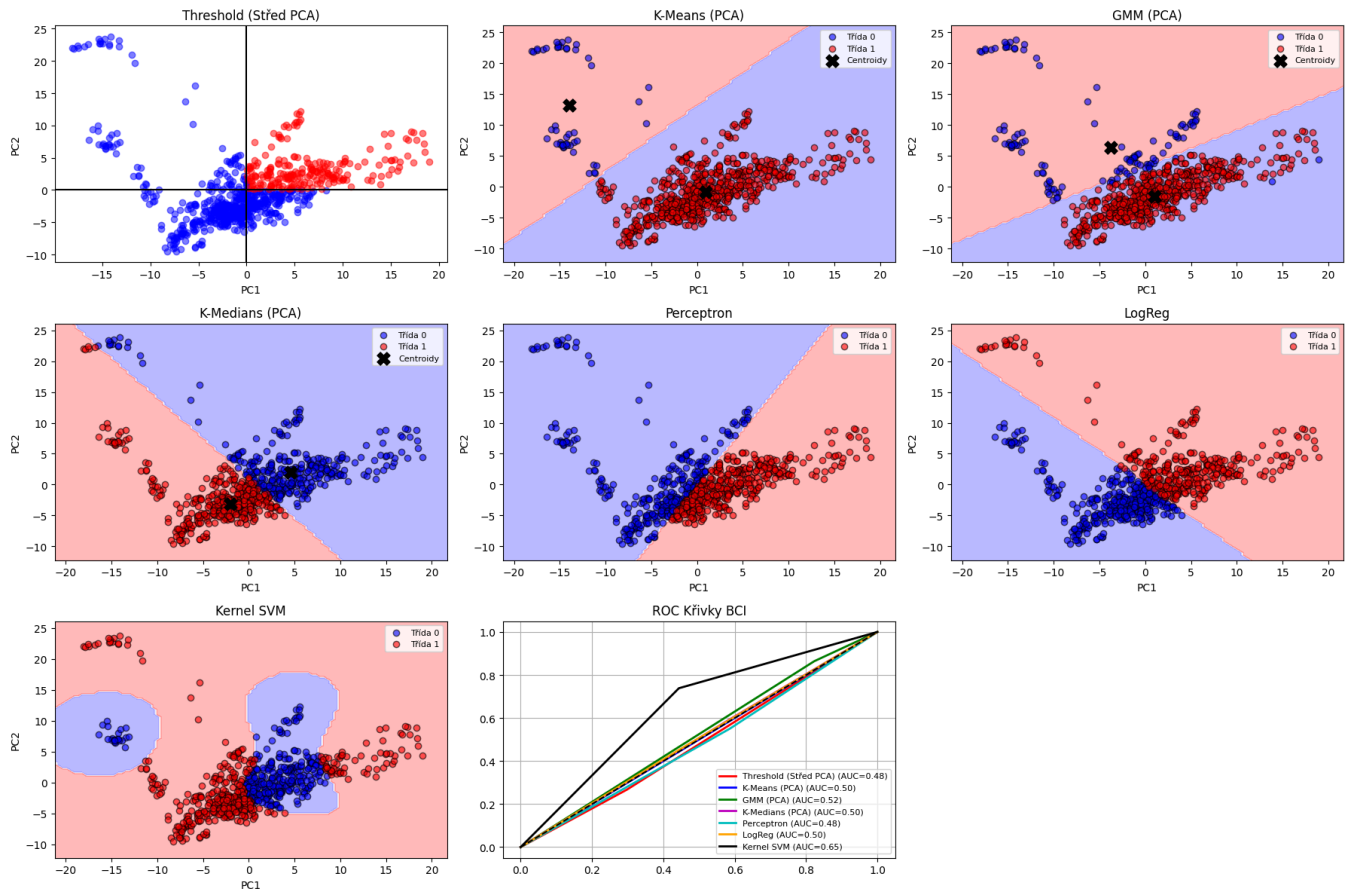


Konektivita Alfa (11.9 Hz) - Klid



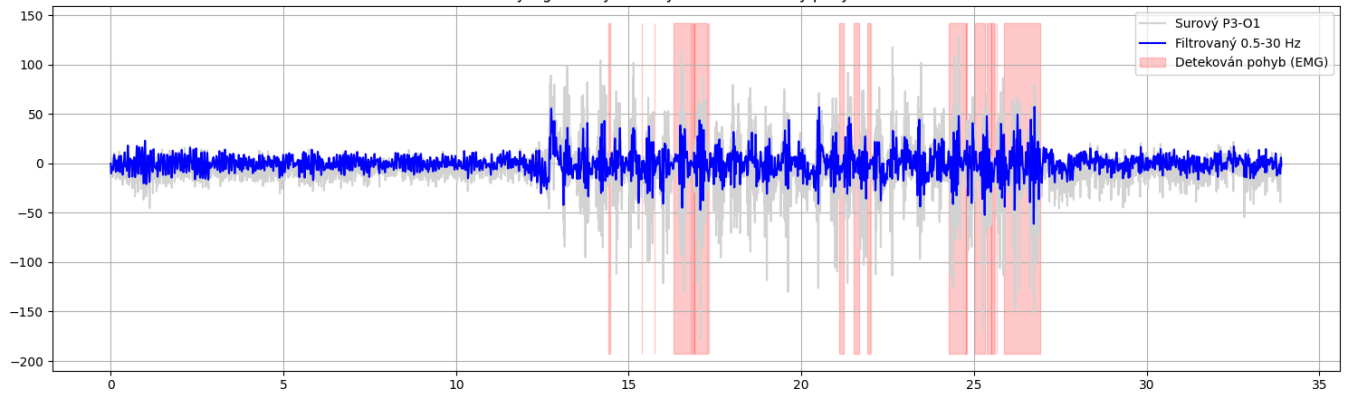
Konektivita Alfa (11.5 Hz) - Úkol



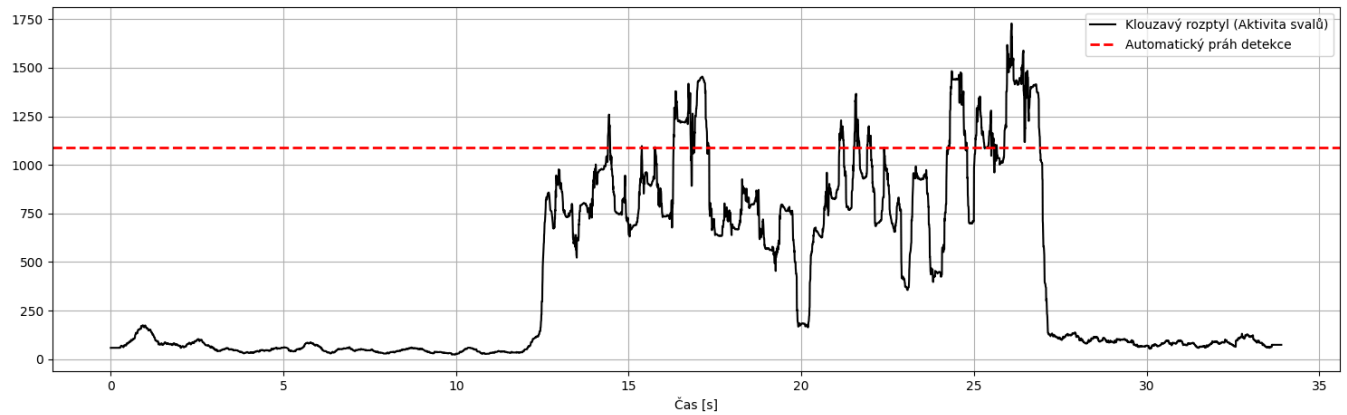


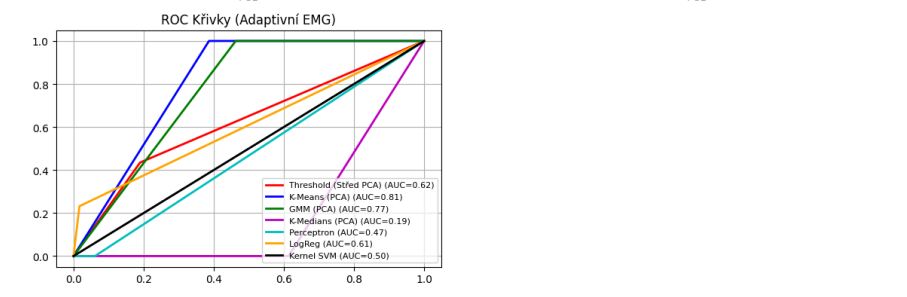
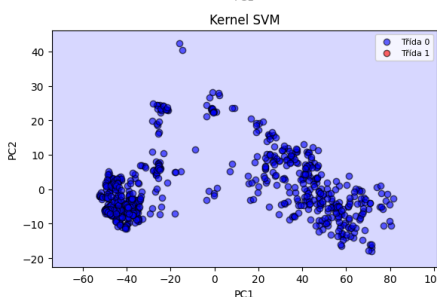
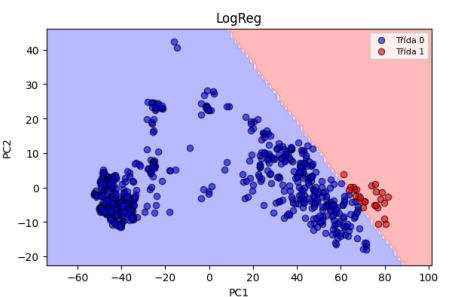
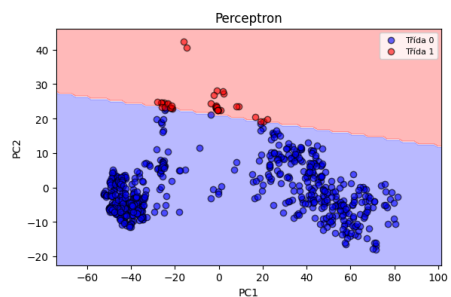
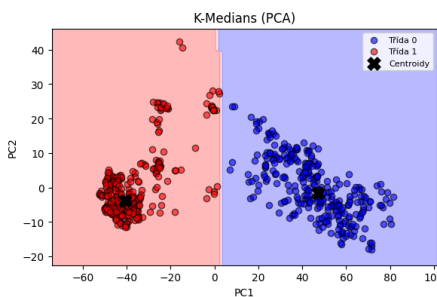
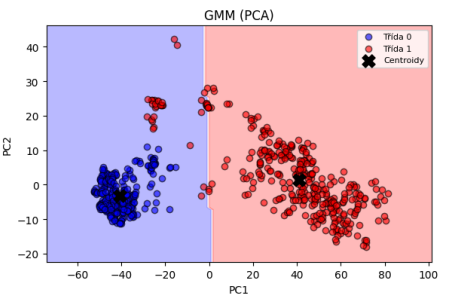
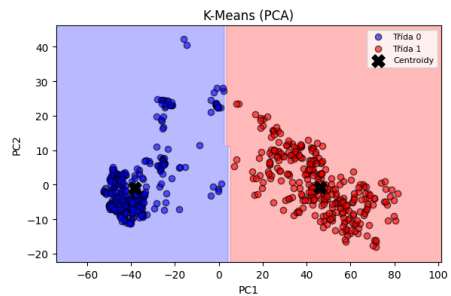
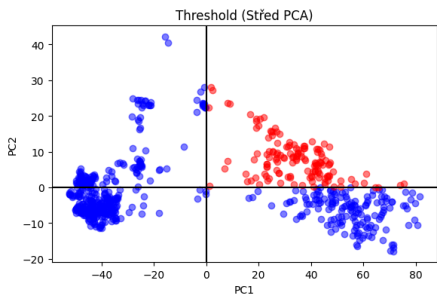
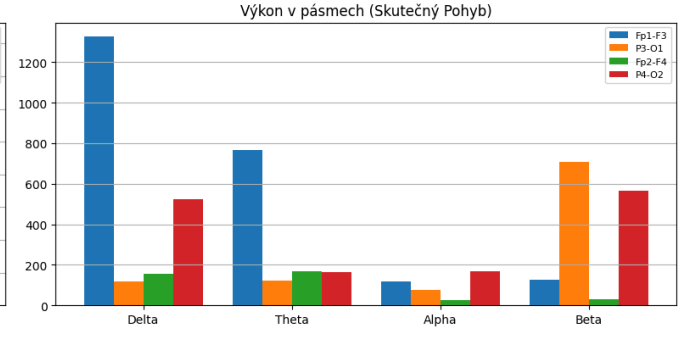
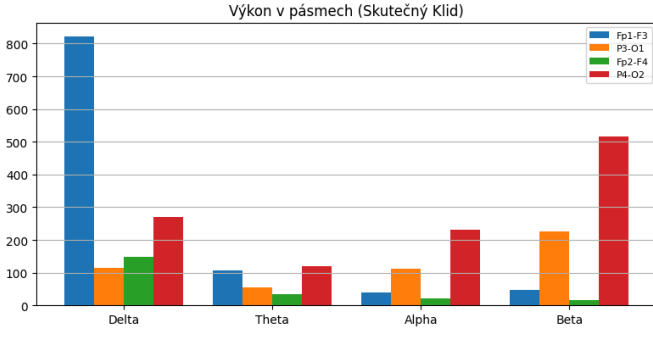
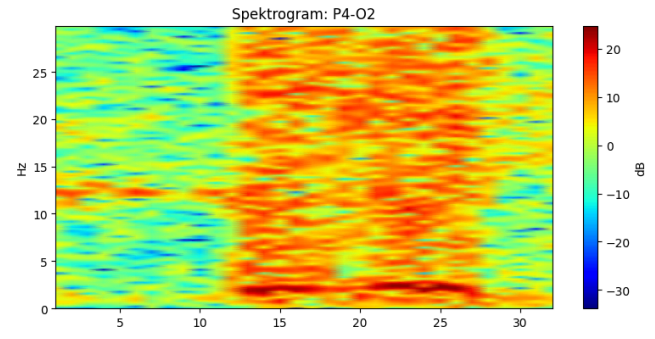
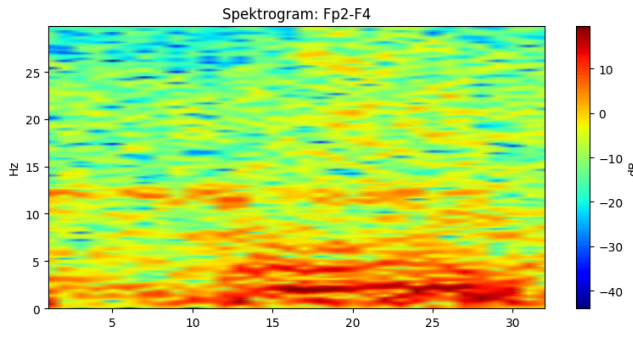
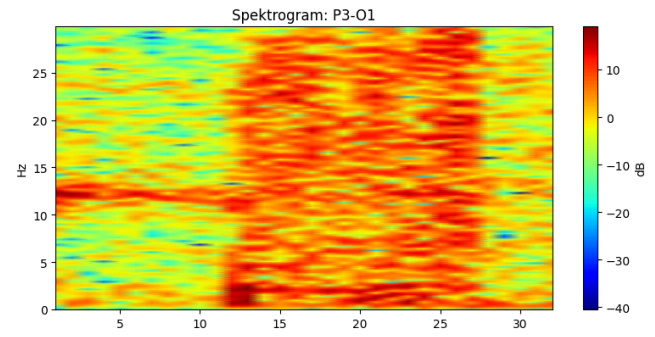
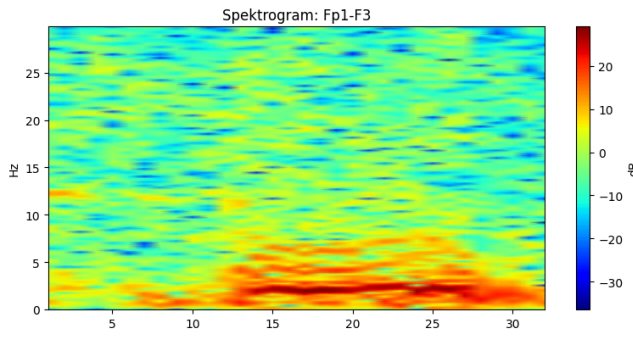
### A/B: Filtrace a Adaptivní detekce reálného pohybu (EMG) - lab07KamilUsi.txt

Surový signál a dynamicky detekované zóny pohybu ušima



Detektor svalového artefaktu





# Analýza dat

Celou analýzu jsem rozdělila do tří odlišných větví.

Pro experimenty zaměřené na vizuální aktivitu, kde Kamil reálně otevíral a zavíral oči, jsem postavila algoritmus na klasickém Bergerově jevu = nástup alfa aktivity při zavření očí. Věděla jsem, že když zavře oči, globální alfa aktivita obrovsky vyskočí napříč celou mozkovou kůrou. Přestože jsme měli elektrody umístěné primárně nad motorickým a frontálním kortexem (C3, C4, F3, F4), šíření alfa vln bylo natolik silné, že se jasně projevilo i zde. Elektrody C byly totiž blízko k zomu, jak by byly umístěny elektrody O. K segmentaci jsem použila adaptivní detektor – algoritmus si sám "klouzal" po obálce signálu a skoky v alfa pásmu našel, aniž bych mu musela zadávat přesné časy. Zjistila jsem, že pro úspěšnou klasifikaci mi stačí sledovat jen jeden kanál a vzít z něj dva jednoduché příznaky: intenzitu alfa pásma a první spektrální moment. Klasifikátory pak měly docela jednoduchou práci, protože otevřené a zavřené oči tvořily dva oddělené shluky dat a šlo mezi nimi snadno nakreslit dělící hranici.

Úplně jiná situace ale nastala u motorických a kognitivních experimentů, tedy u myšlení na stisk ruky nebo u reálného stisku ruky. Tady měl Kamil oči celou dobu zavřené, takže původní vizuální algoritmus by selhal. Zde jsme se naopak plně opírali o naši topografii – elektrody C3 a C4 leží přesně nad motorickým kortexem. Hledala jsem takzvanou ERD (desynchronizaci rytmů) – mozek totiž při představě i provádění pohybu částečně utlumí své klidové frekvence. Jenže tento jev je mnohem slabší a jemnější než zavírání očí. Navíc se u reálného pohybu ruky a uší do signálu začaly míchat obrovské svalové artefakty z pohybu.

Musela jsem proto změnit přístup a přejít na multivariabilní metodu. Adaptivní detekci jsem u myšlení na pohyb nahradila striktní segmentací. Signál jsem natvrdo rozřezala na 10sekundové bloky přesně podle toho, jak se to při měření s Kamilem stopovalo. Aby z toho klasifikátory zachytily jemnou ERD, nemohla jsem se spolehnout jen na jeden kanál. Vzala jsem data ze všech čtyř dostupných elektrod a v každé z nich spočítala výkon pro všechna čtyři EEG pásma (Delta, Theta, Alfa, Beta). Tím mi pro každý časový úsek vzniklo hned 16 různých parametrů.

Protože se v 16D prostoru modely ztratí, promítla jsem data přes analýzu hlavních komponent (PCA). Ta mi z těch 16 čísel ukázala dvě nejdůležitější složky, ve kterých se skrývaly ty největší rozdíly mezi klidem a motorickým úkolem. Až na těchto PCA komponentách jsem učila klasifikátory. Díky tomuhle složitějšímu postupu dokázaly modely i v zašuměném signálu najít skryté vzorce spojené s motorickou aktivitou.

Třetí přístup byl uzpůsoben signálu u pohybu uší. Pohyb ušima znamená především fyzickou kontrakci svalů na hlavě. Svalová aktivita generuje elektromyografický šum, který má velkou amplitudu a klasické EEG vlny v podstatě "překřičí". Rozhodla jsem se tento nedostatek využít ve svůj prospěch a naprogramovala jsem adaptivní EMG BCI algoritmus s detektorem svalové aktivity. Algoritmus počítal klouzavý rozptyl napříč signálem s oknem o délce 0,5 sekundy. Protože EMG artefakt nese výkyvy amplitudy, klouzavý rozptyl v místech reálného pohybu hodně stoupl. Následně jsem algoritmus nechala vypočítat automatický práh. Vše nad tímto prahem algoritmus bezpečně

označil jako Skutečný pohyb, vše pod ním jako Skutečný klid. Teprve díky tomuto kroku jsem získala reálné hranice stavů (Ground Truth). Na takto očištěná data jsem nasadila multivariabilní extrakci příznaků a PCA. Klasifikátory díky tomu mohly reálně hledat rozdíl mezi svalovým artefaktem a klidovým EEG.

## Klasifikace

Rozhodla jsem se na vyextrahované příznaky nasadit celou řadu klasifikátorů. Nechtěla jsem se spoléhat jen na jeden model, protože biologická data mohou být pro různé modely různě složitá. Postupně jsem tak přešla od těch úplně nejjednodušších metod bez učitele (unsupervised) až po pokročilé modely s učitelem (supervised).

### 1. Prahování

Tohle byl můj absolutní základ, abych měla s čím srovnávat.

Model spočítá průměrnou hodnotu přes všechny body na ose X a ose Y. Vytvoří tak dva prahy. Pokud je nová hodnota větší než oba prahy, zařadí se do třídy 1, jinak do třídy 0.

Je to ideální přístup na triviální data, která jsou perfektně lineárně separabilní. U zrakových experimentů (kde Alfa vlny obrovsky vyskočí) to docela fungovalo, ale u složitějších kognitivních úkolů bylo potřeba využít jiné klasifikátory.

### 2. K-Means

Chtěla jsem zkusit klasické shlukování bez učitele, abych viděla, jestli se data přirozeně rozdělí do dvou skupin. Algoritmus hledá takové pozice dvou centroidů, aby minimalizoval součet druhých mocnin vzdáleností (rozptyl) mezi body a jejich nejbližším centroidem. Centroid se v každém kroku přepočítá jako obyčejný aritmetický průměr bodů, které mu náleží. K-Means je skvělý pro data, která tvoří kulaté, zhruba stejně velké shluky. U EEG má ale obrovský problém – stačí pár ustřelených artefaktů (třeba když Kamil silně mrkne a vytvoří obrovskou amplitudu) a protože se počítá průměr, tenhle extrém celý centroid odtáhne úplně špatným směrem.

### 3. K-Medians (s Weiszfeldovým algoritmem)

Právě kvůli artefaktům, které mi rozbíjely K-Means, jsem se rozhodla naprogramovat robustnější K-Medians. Místo druhých mocnin minimalizuje absolutní vzdálenosti (L1 normu). Hledá takzvaný geometrický medián. Protože to ale nejde spočítat jedním jednoduchým vzorcem jako průměr, musela jsem implementovat iterativní Weiszfeldův algoritmus. Ten v každém kroku přepočítává pozici centroidu tak, že bodům dává váhy nepřímo úměrné jejich vzdálenosti. Čím dál ten bod je, tím menší má váhu. Geometrický medián se nedá snadno strhnout jedním špatným měřením. K-Medians mi tak tvořil mnohem stabilnější rozhodovací hranice pro EEG data.

### 4. Gaussian Mixture Model (GMM)

Protože shluky v EEG datech po redukci přes PCA většinou nebyly hezky kulaté, ale spíš protažené elipsy, sáhla jsem po GMM. GMM předpokládá, že data jsou vygenerována směsí dvou Gaussových rozdělení. Každý shluk má svůj průměr  $\mu$ , ale navíc i kovarianční matici  $\Sigma$ , která určuje jeho tvar a natočení. K učení používá EM algoritmus, který iterativně odhaduje pravděpodobnost, s jakou bod patří do dané elipsy. GMM je perfektní pro data, která se překrývají a mají různý rozptyl (např. když je stav klidu stabilní a tvoří malý shluk, zatímco u myšlení data lítají všude možně a tvoří velkou protaženou elipsu). GMM si s těmito tvary poradí mnohem lépe než K-Means.

## 5. Perceptron

Tady jsem už přešla k učení s učitelem (supervised learning). Modelům jsem na první polovině dat řekla, co je co, a nechala je hledat hranice. Je to ta úplně nejjednodušší umělá neuronová síť. Snaží se najít nadrovinu, která data oddělí. Perceptron má jeden obrovský matematický limit, konverguje pouze tehdy, pokud jsou data dokonale lineárně separabilní. Jakmile se shluky trochu prolínají (což se u kognitivních EEG děje neustále), Perceptron se nikdy neustálí. Vzala jsem ho spíš jako ukázkou toho, proč potřebujeme složitější modely.

## 6. Logistická regrese

Chtěla jsem lineární model, který se ale na rozdíl od Perceptronu nezhroutí, když se data překrývají. Proloží data lineární funkcí, ale výsledek vloží do sigmoid funkce. Díky tomu nevrací jen tvrdé hodnoty 0 nebo 1, ale plynulou pravděpodobnost. Optimalizují pomocí gradientního sestupu, který minimalizuje cross-entropy. Pro lineární oddělení PCA komponent u jednodušších úkolů fungovala logistická regrese moc pěkně.

## 7. Kernel SVM (Support Vector Machine s RBF jádrem)

Tohle byl můj nesložitější model na závěr pro případy (jako třeba ta motorická představivost), kdy lineární funkce k oddělení nestačila. Standardní SVM hledá nadrovinu, která maximalizuje margin mezi třídami. Já jsem použila kernel SVM s radiálním jádrem. RBF funkce promítne naše 2D data do nekonečně-dimenzionálního prostoru, tam je hladce oddělí rovnou plochou a když je promítne zpět do 2D, vytvoří nelineární hranice. SVM dosahuje největší přesnosti, ale často dochází k overfittingu, tady konkrétně i na našich datech.

# Jednotlivé signály

## Otevírání a zavírání očí (lab07KamilOci.txt)

V tomto experimentu jsme testovali reaktivitu mozkové kůry na vizuální podněty. Algoritmus, který jsme navrhli primárně pro extrakci alfa aktivity, na těchto datech fungoval a potvrdil klasické elektrofyziologické předpoklady.

Ve chvíli, kdy Kamil zavřel oči (kolem 20. sekundy), obálka alfa pásma ve svodech blíže k lobus occipitalis vystřelila nahoru. Náš detektor na bázi rozdílů rozptylů (červená křivka dole) na to okamžitě zareagoval nárůstem. Díky tomu algoritmus detekoval fyziologické hranice stavů (cca 13. s, 20. s a 31. s) bez nutnosti znát přesné stopování experimentu.

Když se podíváme na barploty výkonu v pásmech, ve stavu "Otevřené oči" je alfa aktivita v okcipitálních kanálech minimální (kolem 10–20  $\mu V^2$ ). Jakmile ale subjekt oči zavře, zrakový kortex ztratí vizuální fokus a začne generovat synchronní alfa vlny.

Ve spektrogramech se to projevilo jako jasně červený horizontální pruh v oblasti 10–12 Hz, který se objevil přesně mezi 20. a 31. sekundou.

Na frontálních elektrodách vidíme při otevřených očích obrovský výkon v pásmu Delta (přes 500  $\mu V^2$ ). Nejde o spánkovou aktivitu, ale o jasný artefakt z mrkání a pohybu očí, který při zavřených očích zmizí.

Obě spektrální metody identifikují dominantní frekvenci těsně pod 12 Hz - alfa aktivita.

Při otevřených očích dominovala spíše nízkofrekvenční konektivita ve frontální oblasti. Při zavřených očích se ale konektivita v alfa pásmu (11.9 Hz) dramaticky změnila a synchronizovala celou mozkovou kůru.

Na první pohled je vidět, že se data formují do dvou zřetelných shluků. Modré body (otevřené oči) mají nízkou alfu a vyšší spektrální moment. Červené body (zavřené oči) se táhnou doprava k vysokým hodnotám alfy.

Výsledky supervised modelů: Učení s učitelem zde fungovalo velmi dobře. Kernel SVM dosáhl přesnosti s AUC = 0.91 a Logistická regrese s AUC = 0.89.

Zajímavost u unsupervised modelů (GMM a K-Medians): V legendě ROC křivky si můžeme všimnout, že GMM a K-Medians mají AUC = 0.29 a 0.21. To neznamená, že selhaly. U shlukování bez učitele algoritmus nezná správné labely. Modely sice data perfektně rozdělily do dvou skupin, ale algoritmus pouze náhodně přiřadil shluku se zavřenými očima label 0 a otevřeným 1 (tedy naopak, než byla naše "Ground truth"). Kdybychom labely invertovali, jejich reálná přesnost (AUC) by byla  $1 - 0.21 = 0.79$ . To dokazuje, že i bez učitele dokáže algoritmus (zejména ten můj nový, robustní K-Medians) přirozeně a správně najít fyziologický rozdíl mezi otevřenými a zavřenými očima.

## Otevírání a zavírání očí podruhé (lab07KamilOci2.txt)

Zatímco první měření byla hezká ukázka reakce zrakového centra, tohle druhé měření se příliš nepovedlo. Je možné, že Kamil více hýbal očima, mluvil nebo se pohybovaly elektrody na jeho hlavě. Ukazuje se, že složité modely strojového učení (např. kernel SVM) nejsou všespásné a je potřeba mít kvalitní signál na vstupu. Alfa aktivita je však stále vidět.

## Myšlení na stisk ruky (lab07KamilMysleni.txt)

Do grafu jsem vynesla svislé černé čáry po 10 vteřinách. Ty reprezentují náš striktní protokol (0-10s klid, 10-20s úkol...). Kdybychom se spoléhali na automatickou detekci, algoritmus by se v tomhle šumu nejspíš ztratil.

Na spektrogramech je vidět nepřerušovaná horizontální linie kolem 12 Hz. To je důkaz toho, že Kamil měl oči celou dobu opravdu zavřené.

U frontálních elektrod vidíme obrovský výkon v pásmu Delta (přes  $175 \mu V^2$ ). To je čistý svalový/pohybový artefakt. Jakmile totiž Kamil začal usilovně myslet na mačkání pěsti, pravděpodobně u toho podvědomě napínal svaly v obličeji, mračil se nebo hýbal očima pod víčky, nebo se smál. To je u našich měření docela častý problém.

Očekávaný útlum alfy (ERD) se tu neprojevuje nijak drasticky. Motorická představivost se odehrává v centrální kůře, ale my měříme i blízko lobus occipitalis u zrakového centra.

Červená křivka (OO - Occipital-Occipital) absolutně dominuje se špičkou na 12 Hz. Zraková centra se synchronizují, když nemají stimul.

Protože nám zrakové elektrody moc nepomohly, vzala jsem všech 16 parametrů (4 kanály  $\times$  4 frekvenční pásma) a promítla je přes PCA do 2D prostoru.

Na scatter plotech vidíme, že data už netvoří dva oddělené oblaky, ale skládají se složitějších "ostrovů" a "prstenců".

Obyčejné prahování středem PCA tu absolutně selhalo s AUC pouhých 0.44.

Lineární modely jako Perceptron nebo Logistická regrese se snažily najít přímku, ale dosáhly jen průměrných AUC kolem 0.61–0.62.

Kernel SVM s RBF jádrem se tady dokázalo svými nelineárními hranicemi dosáhnout nejlepší přesnosti s AUC = 0.77. Dalším dobrým klasifikátorem pro toto měření byl můj implementovaný K-Medians s Weizfeldovým algoritmem.

## Reálný stisk ruky (lab07KamilRuka.txt)

Hned na prvním grafu surového a filtrovaného signálu je vidět zásadní rozdíl oproti všem předchozím experimentům. Přesně v místech našich 10sekundových hranic (např. kolem 10., 20. a 33. sekundy) se objevují ostré špičky. Tohle už nejsou jemné změny v mozkových vlnách, to jsou pohybové a svalové (EMG) artefakty. Když Kamil zatnul ruku, pravděpodobně u toho pohnul celým tělem nebo napnul svaly na krku. Tyto výkyvy se propisují i do obálky alfa pásma, takže se opět potvrzuje, že náš striktní 10s časový protokol byl pro segmentaci naprostou nutností.

Na spektrogramech je sice stále vidět slabší horizontální linka kolem 12 Hz (oči měl Kamil celou dobu zavřené), ale ve spektrogramu je vidět i šum i nižší frekvenci.

Sloupcové grafy ukazují velký výkon v pásmu Delta (vysoko nad  $100 \mu V^2$ ) napříč všemi kanály. Tento Delta výkon je podpisem fyzického pohybu a svalového napětí.

Nejvyšší špičky na frekvenci kolem 11.5–11.9 Hz (Alfa) jasně dominují na červené křivce (OO – Occipital-Occipital). Zraková centra si i během mačkání pěsti udržují svou klidovou synchronizaci, protože oči zůstaly zavřené.

Data z klidu (modrá) a z reálného pohybu (červená) se do sebe vpíjejí v jednom oblaku. Oproti experimentu s pouhým myšlením jsou tu data mnohem více rozlehlá, což je nejspíš způsobené

artefakty.

Všechny jednoduché a lineární modely tu neměly příliš dobrou performance. Prahování, Perceptron i Logistická regrese zůstávají na úrovni náhodného hádání s hodnotami AUC mezi 0.48 a 0.50.

Jediný model, který dokázal v tomto chaosu najít nějaký vzorec, byl opět Kernel SVM s RBF jádrem. Díky své schopnosti obalit data do nelineárních "bublin" dokázal oddělit specifické okraje shluků a dosáhl AUC = 0.65. Není to sice dokonalé, ale je to jediný statisticky významný výsledek. S velkou pravděpodobností se ale jedná o overfitting. Klasifikátor, který data dokázal smysluplněji rozdělit byl K-Medians.

Toto měření nepovažuji za příliš vypovídající kvůli artefaktům a neschopnosti klasifikátorů najít rozhodovací hranici.

## Pohyb uší (lab07KamilUsi.txt)

Hned první graf ukazuje, proč jsme museli přejít na adaptivní detektor svalového rozptylu. Kamil absolutně nedodržel 10sekundový protokol. S ušima začal hýbat až kolem 12,5 s a pokračoval v tom až do 27. sekundy. Nejdůležitější je ale samotná amplituda signálu – výkyvy dosahují extrémních hodnot od -150 do +150  $\mu\text{V}$ . Pravé mozkové (EEG) vlny jsou zde zcela přehlušeny svalovým (EMG) artefaktem. Není to chyba měření, je to dáno topografií: referenční elektrody pro EEG byly umístěny na mastoidy. Ve chvíli, kdy Kamil zapojil svaly pohybující ušima (musculi auriculares), vygeneroval elektrický potenciál přímo pod referenčními elektrodami.

Že neměříme mozek, ale svaly, dokazují i spektrální grafy. Typický EEG signál má energii koncentrovanou v určitých pásmech (např. Alfa). Tady ale ve fázi "Skutečný pohyb" vidíme na spektrogramech červené bloky napříč úplně všemi frekvencemi. Sloupcové grafy ukazují, že výkon vyletěl do absurdních hodnot – Delta přesahuje 1300  $\mu\text{V}^2$ , ale obrovsky narostla i Theta, Alfa a Beta. Svalový stah vygeneroval elektrický šum napříč celým frekvenčním spektrem.

Přestože (nebo spíše právě proto, že) byl EEG signál zničen svalovým artefaktem, modely strojového učení dokázaly stavy rozlišit.

Když se podíváme na rozložení dat v PCA prostoru, vidíme dva naprosto oddělené shluky. Modré shluky reprezentují klid (nízký rozptyl), zatímco červené shluky tvoří samostatný ostrov reprezentující svalový artefakt.

Klasifikátory bez učitele (unsupervised) si vedly nejlépe. K-Means dosáhl skvělého AUC = 0.81 a GMM AUC = 0.77. (K-Medians sice ukazuje 0.19, ale jak už víme, jde jen o invertované štítky, takže jeho reálná přesnost je  $1-0.19=0.81$ ). U tohoto měření bych označila K-Medians s Weizsfeldovým algoritmem za nejrobustnější variantu, jelikož EMG signál může vytvářet velké outliery.

## Závěr

Při zpracování EEG signálů neexistuje univerzální analytický postup. Výběr extrakce příznaků a segmentace musí striktně odpovídat fyziologii úkolu a poloze elektrod.

Vizuální experimenty: Silná změna alfa aktivity se při zavírání očí projevila natolik, že plně postačil 2D příznakový prostor.

Kognitivní a motorické úlohy: Přestože jsme měli elektrody ideálně umístěny (C3/C4 nad motorickým kortexem), změny (ERD) jsou velmi jemné. Bylo nutné využít striktní časový protokol, extrakci frekvenčních pásem napříč všemi kanály a PCA redukci, kde uspěly pouze pokročilé nelineární modely (Kernel SVM), ačkoli se mohlo jednat o overfitting.

Vliv svalových (EMG) artefaktů: U reálných pohybů končetin a uší se projevilo zarušení signálu svalovými potenciály. U pohybu uší jsme však ukázali, že pomocí adaptivní detekce rozptylu lze tento "šum" izolovat a využít jako vysoce přesný klasifikační znak, zejména při použití robustního K-Medians shlukování.

Nejlépe si moje klasifikátory vedly u měření otevírání a zavírání očí a u myšlení na stisk ruky. Nejlepších výsledků dosahoval můj robustní K-Medians a Kernel SVM s radial basis kernelem.