



Univerzita Tomáše Bati ve Zlíně

Fakulta aplikované informatiky

Disertační práce

Rozpoznávání skrytých informací k určení intencionální lži pomocí elektroencefalografie

**Detecting Concealed Information to Identify Intentional
Deception Using Electroencephalography**

Autor: **Ing. Martina Žabčíková**

Studijní program: P3902 Inženýrská informatika

Studijní obor: 3902V023 Inženýrská informatika

Školitel: prof. Mgr. Roman Jašek, Ph.D., DBA

Oponenti: doc. RNDr. Martin Kotyrba, Ph.D.
prof. Ing. Zuzana Komínková Oplatková, Ph.D.
doc. MUDr. Michal Filip, Ph.D., MBA

Zlín, 2024

© Martina Žabčíková

Vydala **Univerzita Tomáše Bati ve Zlíně** v edici **Disertační práce**.
Publikace byla vydána v roce 2024.

Klíčová slova: elektroencefalografie, EEG, detektor lži, skryté informace, intencionální lež, ERP, P300, CIT, DWT, SVM, GA.

Key words: electroencephalography, EEG, lie detector, concealed information, intentional lie, ERP, P300, CIT, DWT, SVM, GA.

Práce je dostupná v Knihovně UTB ve Zlíně.

PODĚKOVÁNÍ

Poděkování patří mému školiteli prof. Mgr. Romanu Jaškovi, Ph.D., DBA, za jeho cenné rady, odborné vedení a konzultace, které mi poskytl během celého mého studia a při psaní této disertační práce.

Dále bych ráda poděkovala mé rodině, příteli a kolegům za jejich podporu a pomoc v průběhu celého mého studia.

PROHLÁŠENÍ

Prohlašuji, že jsem disertační práci na téma „Rozpoznávání skrytých informací k určení intencionální lži pomocí elektroencefalografie“ vypracovala samostatně pod vedením prof. Mgr. Romana Jaška, Ph.D., DBA, za použití literatury a zdrojů uvedených na konci disertační práce v seznamu použité literatury.

Prohlašuji, že při tvorbě této práce jsem použila nástroj generativního modelu AI ChatGPT 3.5 (<https://chat.openai.com/>) za účelem doplnění některých částí kódu v programu a ke kontrole gramatických chyb. Po použití tohoto nástroje jsem provedla kontrolu obsahu a přebírám za něj plnou zodpovědnost.

ABSTRAKT

Rozpoznávání skrytých informací za účelem odhalení intencionální lži představuje klíčovou výzvu v oblasti bezpečnosti. Současné detektory lži jsou nákladné, nepohodlné a náchylné k manipulaci. Detektory lži založené na elektroencefalografii (EEG) se staly populárnějšími oproti polygrafům, jelikož nabízejí vyšší odolnost vůči vědomému ovlivnění. Hlavním cílem této disertační práce je vyvinout systém pro rychlou a efektivní detekci intencionální lži pomocí analýzy EEG signálů získaných nízkonákladovým zařízením, s využitím pokročilých metod strojového učení a výběru vlastností. EEG signály byly snímány během vizuálního experimentu, který zkoumá neurofyzilogické změny při pokusu o lhaní prostřednictvím analýzy Event-Related Potential (ERP) komponenty P300. Následně byly analyzovány rozdíly v průběhu vln P300 účastníků v reakcích na vizuální podněty známých a neznámých tváří pomocí genetického algoritmu (GA) pro výběr vlastností a algoritmu Support Vector Machine (SVM) pro klasifikaci. Tento přístup dosáhl vysoké přesnosti klasifikace, což potvrzuje jeho účinnost a spolehlivost při rozpoznávání skrytých informací. Navržený systém je uživatelsky přívětivý a vhodný i pro uživatele bez hlubších odborných znalostí, což umožňuje snadnou implementaci s možností efektivního využití v akademických i praktických scénářích. Nízké náklady na zařízení a zjednodušení procesu detekce lži pomocí EEG naznačují potenciál pro širší využití, včetně forenzních a bezpečnostních oblastí.

ABSTRACT

Detecting concealed information to identify intentional deception represents a key challenge in security. Current lie detectors are expensive, uncomfortable, and susceptible to manipulation. Lie detectors based on electroencephalography (EEG) have become more popular than polygraphs, as they offer higher resistance to conscious influence. The primary aim of this dissertation is to develop a system for high-quality lie detection using the analysis of EEG signals obtained from a low-cost device, employing advanced machine learning methods and feature selection techniques. EEG signals were recorded during a visual experiment that examines neurophysiological changes during attempts to lie by analyzing Event-Related Potential (ERP) component P300. Differences in the P300 waveforms of participants' responses to visual stimuli of familiar and unfamiliar faces were then analyzed using a Genetic Algorithm (GA) for feature selection and a Support Vector Machine (SVM) for classification. This approach achieved high classification accuracy, confirming its effectiveness and reliability in detecting concealed information. The proposed system is user-friendly and suitable even for users without deep expertise, facilitating easy implementation with the possibility for effective use in both academic and practical scenarios. The low cost of the device and the simplification of the lie detection process using EEG indicate the potential for broader applications, including forensic and security fields.

OBSAH

ÚVOD.....	9
1. SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY.....	11
1.1 Úvod do problematiky	11
1.2 Počátky detekce lži pomocí EEG	11
1.3 Současný stav detekce lži pomocí EEG	12
1.3.1 Vizualní podněty.....	12
1.3.2 Techniky identifikace podvodu	12
1.3.3 Nejčastěji využívané pozice elektrod	13
1.3.4 Přehled současného stavu detekce lži pomocí EEG	13
1.4 Zhodnocení současného stavu řešené problematiky	16
2. CÍLE DISERTAČNÍ PRÁCE.....	17
3. TEORETICKÝ RÁMEC	18
3.1 Centrální nervový systém	18
3.2 Mozek a jeho funkce.....	18
3.2.1 Typy signálů.....	20
3.3 Elektroencefalografie.....	21
3.3.1 Event-Related Potential	23
3.3.2 P300 komponenta	23
3.3.3 Artefakty v EEG signálu.....	25
3.4 Detekce lži	25
3.4.1 Konvenční detektory lži.....	25
3.4.2 EEG detektory lži.....	26
3.5 Concealed Information Test	27
3.6 Zařízení pro snímání EEG signálů.....	29
3.6.1 Použité snímací zařízení	29
3.7 Softwarové vybavení pro zpracování dat EEG	31

3.7.1 EmotivPro.....	31
3.7.2 Matlab.....	31
3.7.3 EEGLAB	31
3.7.4 ERPLAB.....	31
4. METODIKA VÝZKUMU	32
5. EXPERIMENTÁLNÍ ČÁST.....	35
5.1 Získávání dat (Experiment).....	35
5.1.1 Příprava experimentu	36
5.1.2 Experimentální scénář.....	37
5.1.3 Snímání dat.....	39
5.2 Předzpracování dat.....	42
5.2.1 Vizuální kontrola dat.....	50
5.3 Extrakce vlastností	52
5.3.1 Statistické vlastnosti.....	54
5.3.2 Discrete Wavelet Transform	57
5.4 Výběr vlastností	60
5.4.1 Genetický algoritmus	62
5.5 Klasifikace.....	65
5.5.1 Porovnání metod pro klasifikaci EEG dat	67
5.5.2 Gaussian Support Vector Machine	71
5.5.3 Postup klasifikace dat a tvorba modelu	73
5.5.4 Vyhodnocení modelu	76
5.6 Srovnání se stávajícími studiemi	80
6. VÝVOJ APLIKACE PRO DETEKCI LŽÍ.....	83
6.1 Popis hlavních funkcí aplikace	87
6.1.1 Načtení dat.....	87
6.1.2 Analýza dat.....	88

6.1.3 Detekce lži.....	92
6.2 Distribuce aplikace pro detekci lži pomocí EEG	94
7. VÝSLEDKY A DISKUSE	95
8. VYHODNOCENÍ CÍLŮ PRÁCE.....	99
9. PŘÍNOS PRÁCE PRO VĚDU A PRAXI	100
ZÁVĚR.....	102
SEZNAM POUŽITÉ LITERATURY	104
SEZNAM OBRÁZKŮ	114
SEZNAM TABULEK	116
SEZNAM ZDROJOVÝCH KÓDŮ	117
SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK	118
SEZNAM PŘÍLOH	120
PŘÍLOHA P I: SCHVÁLENÍ EXPERIMENTU ETICKOU KOMISÍ UTB.....	121
PŘÍLOHA P II: INFORMOVANÝ SOUHLAS	122
PŘÍLOHA P III: ÚČEL A PRINCIP EXPERIMENTU – VINNÍ....	123
PŘÍLOHA P IV: ÚČEL A PRINCIP EXPERIMENTU – NEVINNÍ	124
PŘÍLOHA P V: UMÍSTĚNÍ KANÁLŮ	125
PŘÍLOHA P VI: SEZNAM EVENTŮ.....	126
PŘÍLOHA P VII: SEZNAM UDÁLOSTÍ	127
PŘÍLOHA P VIII: DOKUMENTACE S NÁVODEM PRO DETEKCI LŽI POMOCÍ EEG	128
PŘÍLOHA P IX: DOKUMENTY A APLIKACE PRO DETEKCI LŽI POMOCÍ EEG	131
PUBLIKAČNÍ AKTIVITY AUTORA	132
ODBORNÝ ŽIVOTOPIS AUTORA.....	135

ÚVOD

S rostoucí mírou kriminality a podvodů se identifikace lži stává klíčovou oblastí výzkumu. V současnosti se k detekci lži nejčastěji používá polygraf, avšak jeho přesnost a spolehlivost jsou omezené kvůli možnosti vědomé kontroly fyziologických reakcí subjektem. Tato práce se zaměřuje na rozpoznávání skrytých informací pro detekci intencionální lži pomocí neinvazivní metody elektroencefalografie (EEG). EEG umožňuje detekovat neurofyziologické signály v mozku, což je klíčové pro odhalení informací relevantních pro rozpoznání podvodů a tím překonává problémy spojené se současnými detektory lži.

V posledních letech je metoda EEG často využívána, zejména v lékařství, kde slouží ke zkoumání mozkových signálů a diagnóze různých poruch a nemocí. Nicméně EEG nachází uplatnění i v jiných oblastech, jako je komunikace, ovládání, zábava či detekce lži. Pro snímání elektrické aktivity mozku bylo využito nízkonákladové bezdrátové zařízení EEG umístěné na povrchu pokožky hlavy. Toto zařízení je levné, přenosné a jednoduché na použití, což je výhodou pro výzkum v různých prostředích.

Cílem této disertační práce je vývoj systému pro rychlou a efektivní detekci lži prostřednictvím analýzy behaviorálních a nervových korelátů klamání s využitím pokročilých metod strojového učení a nízkonákladového zařízení pro snímání EEG signálů. Výzkum se zaměřuje na rozpoznání lži pomocí Event-Related Potential (ERP) komponenty P300, která je aktivována při intencionální lži. Za tímto účelem byl navržen experiment pro identifikaci podvodů s využitím Concealed Information Test (CIT), během kterého byly prezentovány vizuální podněty s cílem vyvolat reakce ERP P300. V rámci zpracování dat EEG byly analyzovány funkční a strukturální změny v jednotlivých mozkových oblastech při reakci na vizuální stimuly známých a neznámých tváří. Výsledné reakce byly klasifikovány do dvou tříd, vinných a nevinných, pomocí různých algoritmů strojového učení s cílem dosáhnout co nejvyšší přesnosti a spolehlivosti při detekci lži.

Tato práce přináší nový přístup k detekci lži s využitím moderních metod a zařízení, které mohou přispět ke zlepšení bezpečnostních opatření a efektivnější detekci podvodů v různých oblastech, přičemž zdůrazňuje potenciál nízkonákladových technologií pro širší využití v praxi.

Struktura disertační práce je rozdělena do několika kapitol. Kapitola 1 uvádí současný stav řešené problematiky, po níž následují cíle disertační práce. V kapitole 3 je zpracován teoretický rámec, včetně základů o mozku, elektroencefalografii, komponentě ERP P300, detekci lži, protokolu CIT, zařízeních pro snímání EEG signálů a softwarovém vybavení pro zpracování dat

EEG. Další kapitola popisuje metodiku výzkumu včetně zvolených metod zpracování, na kterou navazuje kapitola 5 s experimentální částí. Celý proces detekce lži pomocí EEG je detailně popsán, od návrhu a provedení samotného experimentu, přes analýzu dat a klasifikaci, až po vyhodnocení modelu a srovnání se stávajícími studiemi. Výsledky z experimentální části slouží jako základ pro komplexní vývoj aplikace pro detekci lži, podrobně popsány v kapitole 6. Poslední čtyři kapitoly se věnují výsledkům, diskusi, vyhodnocení cílů práce, přínosu práce pro vědu a praxi a samotnému závěru disertační práce. V některých případech, kde neexistují vhodné české překlady metod či algoritmů, jsou použity všeobecně používané anglické termíny.

1. SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

V současnosti existuje mnoho výzkumných článků a vědeckých prací zabývajících se rozpoznáváním skrytých informací pro detekci lži pomocí elektroencefalografie (EEG), které jsou realizovány vědeckými skupinami po celém světě. Tato kapitola se věnuje důkladné rešerši odborných a vědeckých prací publikovaných v posledních letech, zaměřených na odhalování podvodů s využitím EEG signálů. Byly zde identifikovány použité nástroje, postupy a metody, a následně byly analyzovány pozadí jednotlivých studií.

1.1 Úvod do problematiky

V posledních letech vzrostl zájem vědecké komunity o metody rozpoznávání lži. Tradiční zařízení pro odhalení podvodů, jako je polygraf, měří reakce autonomního nervového systému. Polygraf však nedokáže přesně určit, zda subjekt lže, nebo je pouze ve stresu. Pro lepší rozpoznání lži se vědci zaměřili na zkoumání centrálního nervového systému pomocí různých metod, jako jsou EEG, funkční magnetická rezonance (fMRI) a funkční blízká infračervená spektroskopie (fNIRS) [1]. Nejčastěji využívanou metodou je EEG, jelikož vykazuje velmi slibné výsledky [2]. Z toho důvodu je hlavním zaměřením této práce využití EEG signálů pro detekci lži. Problematika rozpoznávání lži pomocí EEG se rychle rozvíjí, přičemž výzkumníci vyvíjejí různé metody pro vylepšení klasifikace a přesnější identifikaci lži. EEG signály mohou odhalit mnoho důležitých aspektů myšlení, což z nich činí účinný nástroj pro detekci podvodů. I když se tato myšlenka objevila již před několika lety, stále existuje mnoho možností pro vylepšení, jako jsou výkonnější klasifikační algoritmy, lepší dostupnost nebo nižší cena [2].

Pokroky v lékařských zobrazovacích technikách vedly k lepšímu porozumění mozkové činnosti, což umožnilo vytvářet různé aplikace. Podobně jako DNA či otisky prstů, které jsou úspěšné při identifikaci pachatele, může být zkoumání mozku další vhodnou možností. Nedávný výzkum ukázal, že elektrická aktivita mozku může být spolehlivým ukazatelem zpracování informací v mozku pro identifikaci pachatele trestného činu. Tato metoda by mohla být velmi užitečná a ušetřit mnoho času při výsleších svědků a podezřelých, a má velký potenciál v trestních vědách jako nový vyšetřovací nástroj pro propojení důkazů o zločinu s informacemi uloženými v mozku pachatele. EEG má velkou výhodu oproti klasickým vyšetřovacím metodám, protože ji lze použít u každého případu [3].

1.2 Počátky detekce lži pomocí EEG

Využití EEG signálů pro detekci podvodu se zkoumá již od konce 20. století, kdy se na tuto oblast zaměřil Farwell a jeho kolegové. Vyvinuli systém MERMER, někdy označovaný jako „Brain Fingerprinting“, který vyhodnocuje přítomnost či nepřítomnost informací v mozku a je vyvolán pouze podněty, které

jsou pro subjekt pozoruhodné či významné [3]. Vyvinuli také Guilty Knowledge Test (GKT) s použitím Event-Related Potential (ERP) komponenty P300 pomocí EEG. ERP a komponenta P300 jsou blíže specifikovány v podkapitolách 3.3.1 a 3.3.2. K analýze a porovnání reakcí mozku na různé typy podnětů použili statistickou techniku bootstrapping. Ve všech studiích pomocí techniky MERMER byla dosažena vysoká míra přesnosti. Tato technologie má však určitá omezení jako vysoká cena, zranitelnost vůči protiopatřením, limitovaná aplikovatelnost, lidský faktor, únik informací či existence dvou podezřelých. Práce Farwella a jeho kolegů dále odhalila, že neurokognitivní spojení může být aktivováno automaticky mimo kontrolu osoby, bez ohledu na to, zda člověk něco skrývá či nikoliv [3].

1.3 Současný stav detekce lži pomocí EEG

Současný výzkum detekce lži zahrnuje mnoho experimentů a studií zaměřených na zkoumání možností využití EEG k odhalování lži. Účelem experimentů bylo zjistit, s jakou mírou přesnosti je metoda užitečná při detekci lži. Výzkumníci v tomto oboru zkoumají různé aspekty, provádějí různé testy, používají různé experimentální metody a metody extrakce vlastností, a aplikují různé metody strojového učení pro binární klasifikaci EEG dat na vinné (vyšetřovaný subjekt lže) a nevinné (vyšetřovaný subjekt nelže). Jedním z hlavních směrů výzkumu je analýza ERP, zejména ERP P300, která se ukázala jako velmi účinná pro rozpoznání lži. Diverzita těchto přístupů nabízí širokou škálu technik a metod, které lze zkoumat a aplikovat za účelem dosažení co nejpřesnější identifikace podvodu [P.1].

1.3.1 Vizuální podněty

Největší část prací v oblasti detekce lži pomocí EEG se zaměřuje na zkoumání vizuálních podnětů, zejména rozpoznávání obličejů [4-15]. Vzhledem k tomu, že se jedná o nejrozšířenější přístup, byl rovněž aplikován i v této práci. Přehled současného stavu zahrnuje primárně studie zabývající se touto problematikou, aby bylo možné provést srovnání na základě shodných vstupních podmínek a parametrů. Existují však také studie, které se zabývají identifikací předmětů trestného činu [16-24], autobiografickými informacemi [25-28], rozhovory [29-32], rozpoznáváním jmen [2, 33-35] či audiovizuálními stimuly [36, 37]. Byly vytvořeny různé typy experimentů s různými scénáři falešných zločinů, které zahrnovaly obličej oběti, vražednou zbraň, jméno spolupachatele či ukradený předmět. Kontrolovalo se, zda se subjekt účastnil dané události, nebo zda poznává místo činu či daný předmět [3].

1.3.2 Techniky identifikace podvodu

Vědci v současnosti využívají různé techniky identifikace podvodu k rozlišení vinných od nevinných. V kontextu detekce lži pomocí EEG při zkoumání vizuálních podnětů se jedná o metody jako Concealed Information Test (CIT)

[4-8, 10, 12], GKT [11], Deceit Identification Test (DIT) [9] a Complex Trial Protocol (CTP) [14]. Nejužívanější metodou pro analýzu chování jedince při lhaní je CIT, založený na paradigmatu ERP P300, při němž se zkoumají reakce na jednotlivé stimuly. Přítomnost P300 naznačuje, že subjekt lže. Tuto metodu je obtížnější oklamat, ovládat nebo potlačit ve srovnání s tradičním polygrafem [P.1]. V souladu s tím byl v tomto výzkumu představen systém detekce lži využívající ERP P300 v rámci přizpůsobeného protokolu CIT. Podrobnější informace o CIT jsou uvedeny v podkapitole 3.5.

1.3.3 Nejčastěji využívané pozice elektrod

Složka P300 se nejčastěji měří v místě elektrod Pz, Fz a Cz, které jsou umístěny na střední linii lebky, viz Obr. 1. V předchozích studiích zaměřených na analýzu EEG signálů bylo zjištěno, že maximální amplituda této složky je v parietálním laloku (Pz), minimální ve frontálním laloku (Fz) a střední hodnoty nabývá v centrálním laloku (Cz). Někteří vědci se zaměřili především na zkoumání kanálu Pz v parietální oblasti, kde je amplituda ERP P300 nejvyšší [11, 17, 18]. Proto byla v této práci využita elektroda Pz, jelikož je nejvíce informativní pro proces zatajování informací [13, 38].

1.3.4 Přehled současného stavu detekce lži pomocí EEG

Následující odstavce poskytují přehled nedávných studií zaměřených na odhalování skrytých informací pro detekci lži pomocí EEG v kontextu ERP P300 při reakci na známé a neznámé tváře [4-15]. Na tuto oblast se nejvíce zaměřili Bablani a kolegové, kteří vytvořili několik významných publikací.

Bablani a kolegové navrhli přístup k identifikaci lži pomocí CIT a EEG v kontextu ERP P300 při reakci na známé a neznámé tváře. Data byla získána od 10 subjektů pomocí 16kanálového zařízení EasyCap a předzpracována pomocí Band-Pass Filter (BPF). Parametry Hjorth (aktivita, mobilita a složitost) byly použity pro extrakci vlastností a K-Nearest Neighbors (KNN) jako klasifikátor. S využitím tohoto přístupu dosáhli průměrné přesnosti klasifikace dat na vinné a nevinné 81,9 % [8].

V další studii Bablani a kolegové analyzovali chování jedince při lhaní pomocí ERP P300 a vytvořili nový scénář pro CIT. Tato práce zahrnovala scénář simulovaného zločinu, využívající 16kanálové zařízení EasyCap pro získávání EEG dat od 10 subjektů rozpoznávajících známé a neznámé tváře. K odstranění šumu smíchaného se signálem použili BPF. Pro analýzu EEG dat využili různé techniky extrakce vlastností. Rámec byl vyvinut agregací výsledků tří nejlepších klasifikátorů, Support Vector Machine (SVM), Linear Discriminant Analysis (LDA) a Multi-Layer Feed Forward Neural Network (MLFFNN) pomocí přístupu Weighted Voting (WV). Použitím navrženého přístupu (3-WV) s využitím Discrete Wavelet Transform (DWT) pro extrakci bylo dosaženo průměrné přesnosti klasifikace dat na vinné a nevinné 84,7 % [10].

Následně Bablani a kolegové navrhli systém identifikace podvodu, kde použili EEG data od 10 subjektů při provádění CIT pro experimentální vyhodnocení ERP P300 u rozpoznávání známých a neznámých tváří. Pro analýzu využili 16 kanálů a signály extrahovali pomocí různých metod. Mezi různými přístupy k extrakci vlastností se nejlépe osvědčila Wavelet Transform (WT) kombinovaná s klasifikátorem SVM. Navrhli zde novou nákladovou funkci, kde se algoritmus BAT používal pro optimalizaci parametrů SVM ke zvýšení přesnosti klasifikace a Binary BAT (BBA) se používal pro výběr kanálů EEG. Po odstranění nefunkčních kanálů umístěných v okcipitálním laloku mozku se přesnost systému zvýšila na průměrných 96,8 %. Tento přístup poskytl nejlepší výsledky v binární klasifikaci tříd vinných a nevinných v kontextu ERP P300 pro CIT při reakci na známé a neznámé tváře [4].

Dále Bablani a kolegové navrhli hybridní třístupňový klasifikační přístup CIT, kombinující výhody WT, k-means a MLFFNN. Test byl vyvinut analýzou složky ERP P300 během falešného zločinu při rozpoznávání známých tváří. Pro předzpracování dat byl využit BPF. Data EEG od 10 účastníků byla zaznamenávána pomocí 16kanálového zařízení EasyCap pro implementaci navrhovaného rámce. Navržený přístup poskytoval přesnost 83,1 % [12].

V další práci vyvinuli Bablani a kolegové systém CIT využívající ERP P300, kde se subjektům během experimentu zobrazovaly snímky známých a neznámých tváří. Z dat 10 subjektů bylo 7 použito pro trénování a 3 pro testování. Pro předzpracování dat EEG získaných 16kanálovým zařízením byl využit BPF a pro extrakci vlastností byl použit Common Spatial Pattern (CSP). Systém fuzzy integrátoru byl vyvinut s použitím ukazatelů výkonnosti klasifikátorů jako předchůdců (LDA, MLFFNN, SVM, KNN, Naive Bayes (NB)). Experimentální výsledky demonstrovaly účinnost fuzzy systému založeného na CIT s průměrnou přesností klasifikace 86,7 % pro 3 subjekty pomocí přístupu WV [7].

Bablani a kolegové dále použili hluboké učení s omezeným Boltzmannovým strojem a WT k získání informací v časové a frekvenční oblasti signálů. Byl proveden experiment na datech EEG provedením CIT pomocí 16kanálového zařízení EasyCap pro zkoumání vlny ERP P300, kde byly subjektům prezentovány obrázky známých a neznámých tváří. Signály EEG byly předzpracovány pomocí BPF a analyzovány pomocí WT. Pro klasifikaci EEG dat na vinné a nevinné vyvinuli Deep Belief Network (DBN), kde bylo dosaženo průměrné přesnosti klasifikace 81,03 % u 10 subjektů [5].

Ve všech šesti studiích od Bablani a kolegů bylo dosaženo relativně vysoké přesnosti klasifikace 81,03 % až 96,8 %, přičemž publikace [4] dosáhla celkově nejvyšší přesnosti mezi všemi studiemi zabývajícími se touto problematikou.

Mezi další významné přispěvatele v této oblasti patří Dodia a kolegové, kteří navrhli přístup ke klasifikaci EEG signálů na vinné a nevinné provedením DIT založeném na ERP P300 při reakci na známé a neznámé tváře. Předzpracování bylo provedeno pomocí BPF. Signály EEG byly získány ze 16kanálového zařízení EasyCap u 20 subjektů. Vlastnosti byly extrahovány z podrobných koeficientů získaných z Wavelet Packet Transform (WPT) a poté zadány jako vstup do klasifikátoru LDA. Navržený přístup využívající WPT k extrakci vlastností a LDA jako klasifikátor dosáhl přesnosti klasifikace 91,67 % [9].

V další studii Dodia a kolegové navrhli CIT zkoumající ERP P300. Zde byly získané signály od 20 subjektů detekované 16kanálovým zařízením EasyCap předzpracovány pomocí BPF. Experiment zahrnoval reakce na snímky známých osobností a přátel. K extrakci vlastností ze signálů EEG využili metodu Short-Time Fourier Transform (STFT) a získaný soubor vlastností byl dán jako vstup do klasifikátoru Extreme Learning Machine (ELM) pro trénování vinných a nevinných. Pro výběr optimální podmnožiny vlastností byl použit BBA. Výsledná přesnost detekce lži dosáhla 88,3 % [6].

Mehrnam a jeho kolegové vyvinuli systém rozpoznávání vzorů v reakci na vlnu ERP P300, který klasifikuje vinné a nevinné subjekty za využití techniky GKT. Účelem bylo rozšířit sadu vlastností o nelineární prvky pro zlepšení klasifikace. Signály byly zaznamenány od 49 subjektů účastnících se testu rozpoznávání skrytého obličeje. Pro předzpracování použili BPF a pro extrakci vlastností z ERP extrahovali několik morfologických charakteristik, frekvenčních pásem a vlnkových koeficientů. K výběru nejlepší sady vlastností byl využit genetický algoritmus (GA). Výsledky ukazují, že metoda byla schopna správně zařadit 91,83 % subjektů díky kombinaci základních a nelineárních vlastností za použití klasifikátoru LDA a nového přístupu adaptivního prahu [11].

Chen a kolegové se zaměřili na návrh experimentu založeného na CTP pro rozpoznání známé a neznámé tváře. Data byla shromážděna od 147 subjektů. Navrhli zde novou metodu založenou na neuronové síti, EEG-based Face Recognition Model (EEG-FRM). Tato metoda kombinuje Multi-Scale Convolutional Neural Network (MSCNN) s mechanismem maximální pravděpodobnosti pro realizaci rozpoznávání jednotlivých obličejů. Předzpracování bylo provedeno pomocí BPF. MSCNN extrahovala časové informace a prostorové rysy z EEG dat. Channel and Time-Spatial Attention Module (CTSAM) a kontrolovaný kontrastivní výukový modul se poté používaly ke klasifikaci. Jejich model dosáhl přesnosti 85,64 % [14].

Výsledky naznačují, že tyto metody mohou dosáhnout vysoké přesnosti v identifikaci lži při rozpoznávání tváří, čímž představují významný pokrok ve využití neurovědeckých metod k odhalování skrytých informací a otevírají nové možnosti pro další výzkum v této oblasti.

1.4 Zhodnocení současného stavu řešené problematiky

Na základě provedené rešerše lze konstatovat, že nejčastěji využívanou metodou pro analýzu chování jedince při lhaní je CIT založený na ERP P300 při reakci na vizuální stimuly známých a neznámých tváří pomocí EEG [P.1, P.8]. Nejvyšší přesnosti klasifikace dat na vinné a nevinné dosáhli Bablani a kolegové, s hodnotou 96,8 % za využití WT a klasifikátoru SVM [4]. Důkazy naznačují, že nedávné pokroky v neurovědě umožňují detekovat informace uložené v mozku, což může vést k rychlejšímu, přesnějšímu a efektivnějšímu řešení případů, stejně jako k osvobození nevinných podezřelých [3].

Vysoká míra přesnosti některých studií naznačuje významný potenciál v oblasti detekce lži pomocí EEG zaměřené na analýzu ERP P300 při rozpoznávání tváří, avšak stále existuje prostor pro další zlepšení. Většina dosavadních studií využívá drahá a složitá EEG zařízení, což omezuje praktickou aplikaci těchto metod. Existuje potřeba vyvinout a ověřit metody, které jsou použitelné s nízkonákladovými a přenosnými zařízeními, což by umožnilo širší a dostupnější využití v praxi. Stávající metody často vyžadují komplexní a časově náročné procesy předzpracování a analýzy dat. Je nutné vyvinout systémy, které umožní rychlou a jednoduchou detekci lži, použitelnou i osobami bez odborných znalostí [39]. Klasifikační algoritmy dosahují různých úrovní přesnosti v závislosti na použitém datasetu a podmínkách. Je třeba vyvinout robustnější klasifikátory, které budou schopny udržet vysokou přesnost napříč různými experimenty a podmínkami a zároveň budou efektivní v reálném čase. Mnoho studií je prováděno na relativně malých vzorcích subjektů. Je nutné provést výzkum na větších a diverzifikovanějších populacích, aby se zajistila škálovatelnost a praktičnost navrhovaných metod v reálném světě [P.1].

Vzhledem k těmto výzvám byl výzkum zaměřen na vývoj systému pro detekci lži pomocí kombinace nízkonákladového zařízení Emotiv Insight s pokročilými metodami strojového učení a technikami výběru vlastností. Cílem je vyvinout systém pro rychlou, dostupnou, levnou a efektivní detekci intencionální lži pomocí EEG signálů, který bude přístupný i pro osoby bez odborných znalostí v oblasti EEG a neurovědy. Tento výzkum je zásadní pro další rozvoj technologií v oblasti detekce lži, protože jako první efektivně využívá nízkonákladová zařízení a pokročilé techniky strojového učení pro automatickou detekci lži prostřednictvím analýzy ERP P300 při rozpoznávání tváří. Navržený přístup přináší nové možnosti v dostupnosti a praktičnosti detekčních systémů, což může významně zlepšit detekci lži v reálných podmínkách a nabídnout praktická řešení pro aplikace v různých oblastech.

2. CÍLE DISERTAČNÍ PRÁCE

Hlavním cílem této disertační práce je vyvinout efektivní systém pro automatickou, rychlou a jednoduchou detekci intencionální lži. Při vývoji systému je kladen důraz na využití nízkonákladového zařízení pro měření EEG signálů v kombinaci s pokročilými metodami strojového učení a sofistikovanými technikami extrakce a výběru vlastností. Navržený přístup může přinést nové možnosti v oblasti informatiky a posílit efektivitu bezpečnostních opatření.

Pro dosažení hlavního cíle bylo nutno dosáhnout následujících dílčích cílů:

- **Vývoj a validace metodiky pro detekci lži pomocí EEG:** Vytvoření a ověření komplexní metodiky pro identifikaci intencionální lži pomocí analýzy EEG signálů.
- **Návrh a realizace vizuálního ERP experimentu:** Realizace experimentu, který měří ERP komponenty P300 získané z EEG signálů při vystavení účastníků vizuálním podnětům známých a neznámých tváří.
- **Analýza dat EEG a volba vhodných kritérií:** Identifikace a aplikace optimálních parametrů a kritérií pro zpracování a analýzu EEG dat zaměřených na detekci lži.
- **Klasifikace dat EEG pomocí strojového učení:** Využití pokročilých metod strojového učení pro dosažení vysoké přesnosti klasifikace EEG dat do dvou tříd (vinných a nevinných).
- **Vývoj uživatelsky přívětivého systému:** Vyvinutí snadno použitelného systému, který umožní rychlou a efektivní detekci lži pomocí nízkonákladového zařízení snímajícího EEG signál.
- **Ověření a validace navrženého systému:** Testování a ověření účinnosti a přesnosti navrženého systému.

Tyto cíle přispějí k rozvoji snadno dostupných a účinných nástrojů pro detekci lži, což má významné aplikace v oblastech jako forenzní věda a bezpečnost.

3. TEORETICKÝ RÁMEC

Tato práce se zabývá průzkumem v oblasti detekce skrytých informací měřených pomocí diagnostické metody EEG, kde se vyhodnocují míry přesnosti binární klasifikace dat na vinné a nevinné při reakci na vizuální podněty. Před zkoumáním skrytých informací je nejprve upřesněna terminologie centrálního nervového systému, zahrnující základní informace o mozku a různých typech signálů používaných v této oblasti. V další části je nastíněna metoda EEG, včetně jejích základních aspektů, způsobu snímání signálu, umístění elektrod a základních frekvencí. Dále je zde popsána podkapitola o ERP, včetně vlny P300 a artefaktů. Následně se diskutuje o různých metodách detekce lži a je představen princip protokolu CIT. Poslední teoretická část obsahuje přehled zařízení, programů, softwarových platforem a nástrojů využívaných pro práci s daty EEG.

3.1 Centrální nervový systém

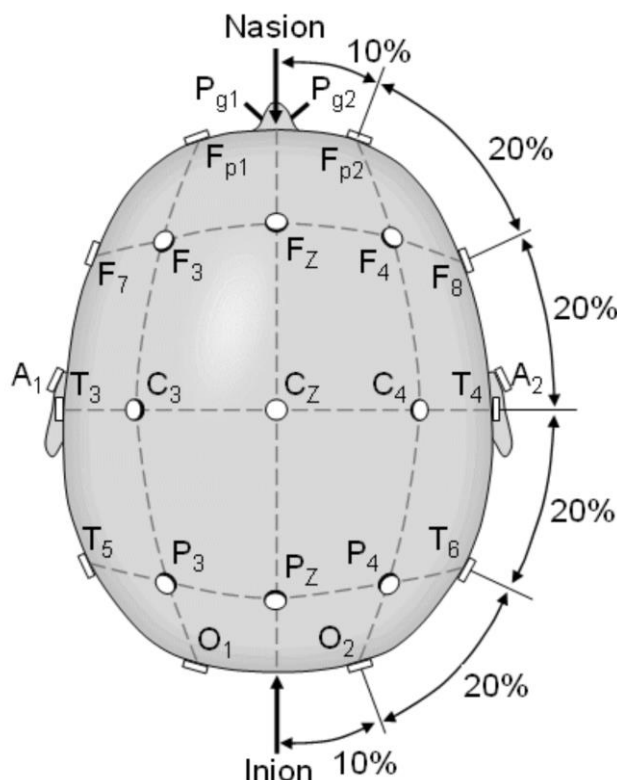
Během posledního desetiletí došlo v neurovědách ke značnému pokroku. Pomocí neurovědních technik vědci komunikují s nervovým systémem a snaží se porozumět mozkovým funkcím, a tím odhalit povahu lidského chování, včetně klamání [40]. Anatomické rozdělení nervového systému se dělí na centrální nervový systém (CNS), zahrnující mozek a míchu, a periferní nervový systém (PNS), zahrnující senzorycké neurony [41]. Součástí PNS je autonomní nervový systém (ANS). Pro správnou interpretaci frekvenčních a amplitudových změn v signálech EEG je rozpoznání dynamické činnosti CNS zásadní [42].

3.2 Mozek a jeho funkce

Mozek může být využit jako zdroj důkazů při vyšetřování trestných činů, jelikož je centrálně zapojen do každé lidské činnosti a zaznamenává vše, co člověk dělá. Mozkový signál jako první reaguje na jakékoli smyslové impulzy, což lze použít k identifikaci pravdy a lži. Mozkové vlny mají mnoho různých aspektů, které závisí na metodě pozorování a analýze dat. Vědci se domnívají, že by se mozek mohl stát ústředním bodem vyšetřování kriminality, když se propojí záznamy o akcích a myšlenkách uložených v mozku s důkazy o zločinu. Zatímco u některých zločinů mohou chybět fyzické důkazy, mozek lze použít u každého případu [3].

Při reakci mozku na událost či aktivitu vzniká rozdíl v potenciálu, který slouží k přenosu zprávy z jedné synapse do druhé prostřednictvím uvolňování chemické látky. Pohyb této chemické látky uvnitř mozkových buněk generuje mozkový signál, který je poté zaznamenáván pomocí speciálních zařízení, ve kterých jsou umístěny registrační elektrody. Existují dva typy elektrod, suché nebo s vodivým gelem, který snižuje přechodový odpor mezi lebkou a elektrodou. Každá elektroda dále obsahuje zesilovač nutný pro zachycení signálů, které jsou velmi nízké, jelikož musí procházet pokožkou hlavy, lebkou a dalšími vrstvami.

Potenciál se pohybuje v rozmezí od pěti do dvou set mikrovoltů (μV) [42]. Rozmístění elektrod se řídí pomocí antropometrického měření nazývaného mezinárodní systém 10-20, viz Obr. 1.

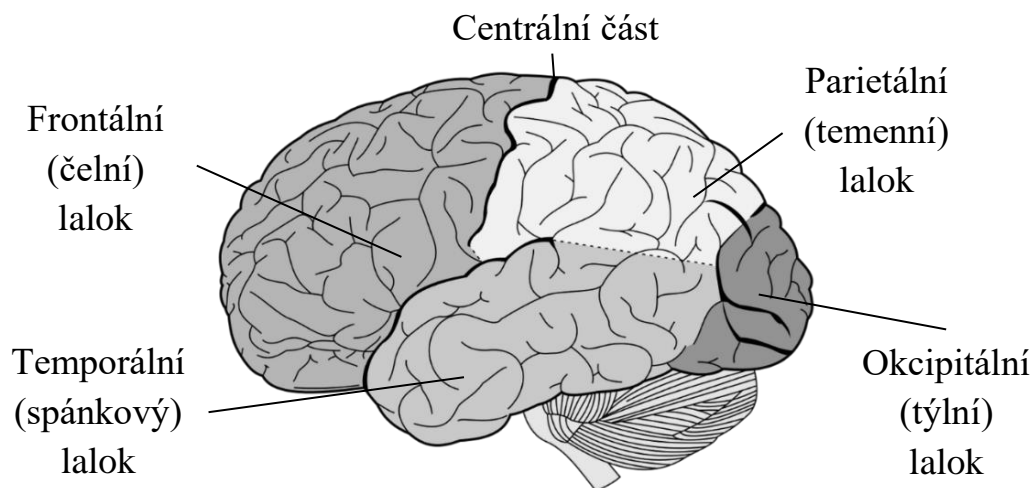


Obr. 1.: Mezinárodní systém 10-20 [39]

Jedná se o standardizovaný systém, který pro umístění jednotlivých elektrod využívá specifikované anatomické orientační body na lebce. Tento systém nese název dle procentuálních vzdáleností jednotlivých elektrod, a to 10 % či 20 % v rovině sagitální (nasion-inion) a frontální (mezi zvukovody). Tyto intervaly poté označují místo, kam bude elektroda umístěna [38]. Spojení všech elektrod je provedeno buď longitudinálně (podélně), nebo transverzálně (příčně). Následné spojení mezi dvěma elektrodami může být provedeno bipolárně (diferenčně), unipolárně (referenčně) či pseudounipolárně. V tomto výzkumu je spojení elektrod daného zařízení provedeno unipolárně a elektrody jsou zde propojeny s nulovou elektrodou, která spojuje všechny elektrody přes odpor k zemi. Signály EEG představují potenciální rozdíl mezi těmito dvěma typy elektrod [42]. Jednotlivé elektrody jsou označeny kombinací písmene a čísla, které určují pozici jejich umístění na lebce. Tato písmena označují lokalizaci jednotlivých elektrod. Elektrody jsou dále číslovány zleva doprava v jednotlivých rovinách, kde lichá čísla označují elektrody nad levou mozkovou hemisférou a sudá čísla nad pravou hemisférou. Označení „z“ odráží umístění středové čáry [38, 42].

Mozek je rozdělen na dvě hlavní hemisféry, z nichž každá obsahuje jednotlivé laloky. Rozdělení jednotlivých částí mozku v transverzální rovině je na frontální (F, čelní), centrální (C, střední), parietální (P, temenní), okcipitální (O, týlní)

a temporální (T, spánkový), viz Obr. 2. Frontální lalok souvisí s myšlením, plánováním, emocemi, řešením problémů a kontrolou chování. Centrální lalok obstarává pohyb a smysly. Parietální lalok souvisí s myšlením, pamětí a vnímáním. Okcipitální lalok zpracovává převážně zrakové podněty a temporální lalok sluchové podněty, paměť a porozumění [38].



Obr. 2.: Anatomické rozčlenění jednotlivých částí mozkové kůry [43]

Z mozkových signálů lze pochopit a získat mnoho užitečných informací, jelikož umožňují nahlédnout do různých procesů, které jsou základem lidského chování a jednotlivých reakcí. Pokud subjekt ukrývá nějaké informace, lze je detekovat pomocí mozkových signálů. Elektrická aktivita mozku představuje složitý biosignál, ve kterém jsou uloženy informace o subjektu. K extrakci ukrytých informací od subjektů ze složitých biosignálů musí být dodržena určitá metodická pravidla pro získání smysluplných a náležitých informací [42]. Z tohoto důvodu je v dalších částech práce popsán přesný metodický postup zpracování signálů EEG.

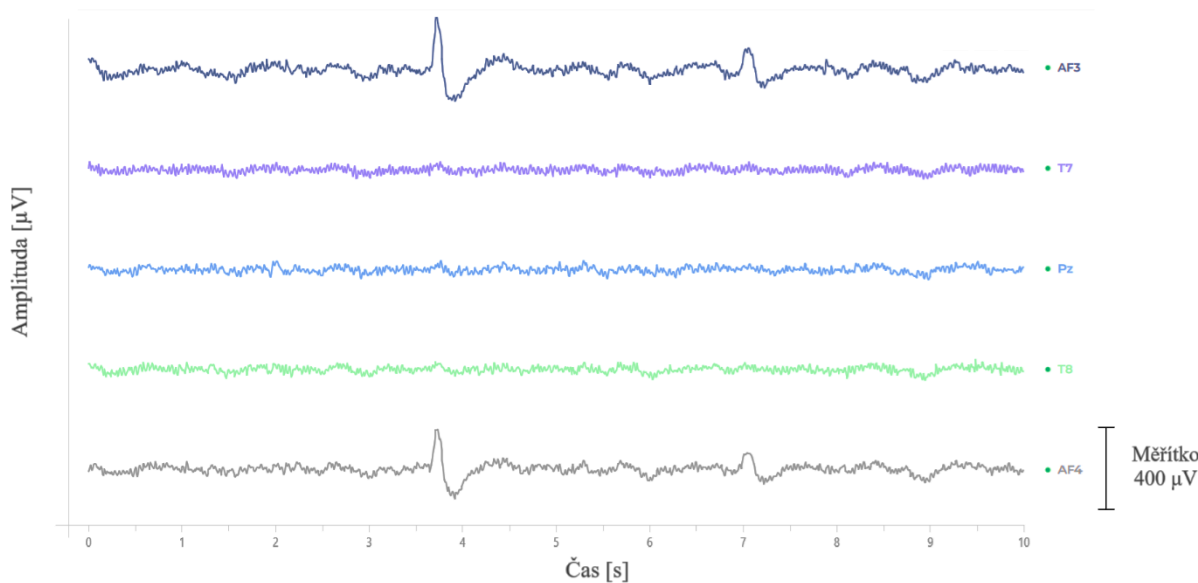
3.2.1 Typy signálů

Existují různé invazivní a neinvazivní metody měření a zaznamenávání mozkových aktivit. Neinvazivní metody poskytují široký přehled elektrické aktivity v obou mozkových hemisférách s elektrodami umístěnými na povrchu pokožky hlavy. Mezi běžně používané neinvazivní techniky pro snímání mozkových signálů patří EEG, fMRI a fNIRS. Invazivní techniky, jako je elektrokortikografie (ECoG), poskytují podrobnější záznam elektrické aktivity přímo z povrchu mozkové kůry prostřednictvím chirurgicky implantovaných elektrod zaměřených na specifické oblasti mozku [38]. Neinvazivní metody sice poskytují signály nižší kvality než invazivní, ale nevyžadují chirurgický zákrok. Z tohoto důvodu se ve výzkumu častěji používají neinvazivní metody pro zkoumání elektrické aktivity mozku, jelikož jsou bezpečné, snadno použitelné a mají relativně nízké náklady [10]. Tato práce se zaměřuje na rozpoznávání skrytých informací pro detekci lži pomocí neinvazivního nízkonákladového zařízení EEG.

3.3 Elektroencefalografie

Elektroencefalografie je funkční neurovizuální metoda, která zaznamenává elektrickou aktivitu mozku neinvazivně pomocí elektrod umístěných na pokožce hlavy. EEG signál lze definovat jako součet postsynaptických potenciálů korových neuronů. Termín elektroencefalogram označuje záznam naměřených signálů a elektroencefalograf je přístroj pro měření těchto signálů [42]. Za zakladatele elektroencefalografie je považován německý neuropsychiatr Hans Berger, který provedl první záznam lidského mozku v roce 1924 [44]. Během následujícího století byla tato metoda studována a vylepšována a dnes je široce využívána v lékařství pro monitorování a diagnostiku epilepsie, mrtvice, záchvatů, poruch spánku a dalších mozkových abnormalit [3]. EEG má však širší oblast využití, včetně komunikace, ovládání, zábavy, rozpoznávání emocí, bezpečnosti a detekce lži [P.9]. V posledních desetiletích se EEG rovněž stalo nástrojem kognitivní neurovědy ke zkoumání mozkových funkcí, poskytující neinvazivní měření s milisekundovým časovým rozlišením.

EEG zaznamenává neuronální aktivitu ve formě signálů, které se liší v závislosti na činnostech a stavech jednotlivce. Během každodenních aktivit reagují různé části mozku na specifické stimuly či aktivity. EEG se zaměřuje na mapování těchto mozkových vln a jejich vztahu k různým podnětům a činnostem. Tyto vzory jsou následně rozpoznávány a analyzovány [11]. Pro zaznamenávání specifických aktivit jsou elektrody umístěny na konkrétní místa na hlavě, což generuje EEG vlny s odlišnými charakteristikami [4], viz Obr. 3.



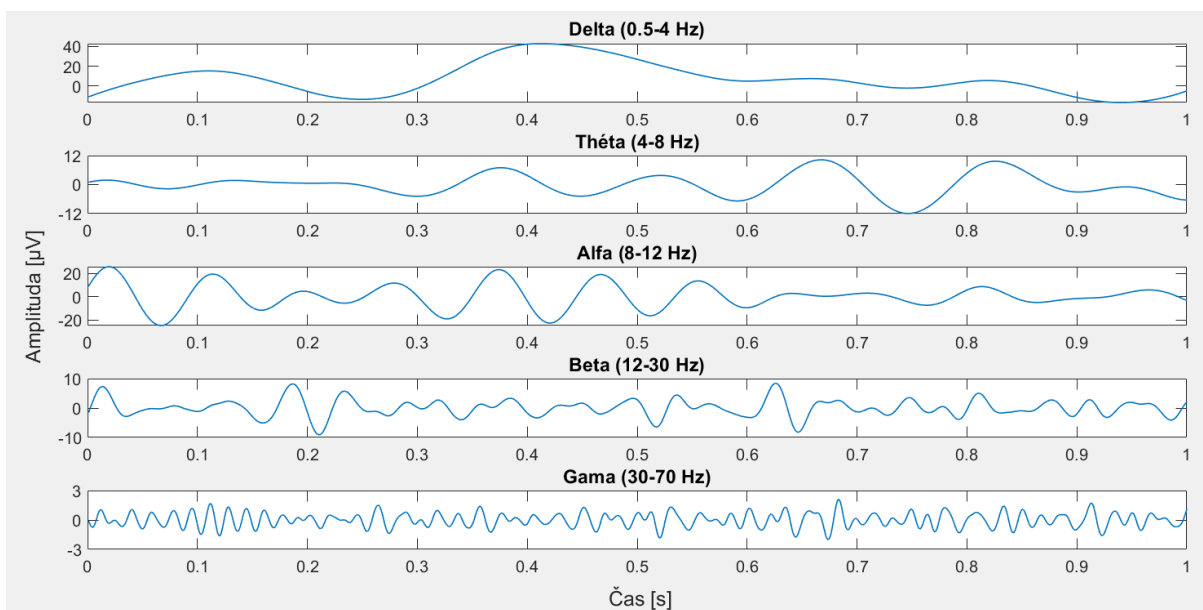
Obr. 3.: EEG signály

EEG signály lze kategorizovat podle frekvenčních pásem, jak ukazuje Tabulka 3.1 [9, 42]. EEG aktivita je definována jako součet frekvencí alfa, beta, gama, delta a théta, které se mění v závislosti na vědomí a stavu jedince. Frekvence rytmů je vyjádřena v cyklech za sekundu (Hz) a potenciál v mikrovoltech (μV).

Tabulka 3.1 Rozdělení frekvenčních pásem EEG [42]

Pásmo	Frekvence [Hz]	Amplituda [μV]	Stav vědomí	Význam v detekci lži
Alfa	8-12	20-80	Uvolněný bdělý stav, zavřené oči, relaxace	Snížené alfa vlny mohou indikovat zvýšenou kognitivní aktivitu a stres spojený s klamáním.
Beta	12-30	10-20	Aktivní bdělý stav, zvýšená pozornost, mentální námaha	Zvýšená beta aktivita je často spojována s mentálním stresem a zvýšenou kognitivní námahou, což může být indikativní při detekci lži.
Gama	30-70	<10	Vysoká soustředěnost, komplexní zpracování informací, koncentrace	Gama vlny se mohou objevit při intenzivním soustředění a mohou indikovat zvýšenou mentální aktivitu spojenou s pokusy o podvod.
Delta	0,5-4	až 200	Hluboký spánek, regenerace	Zvýšená delta aktivita může být spojena se stresem a kognitivními konflikty, což může být relevantní při detekci lži.
Théta	4-8	30	Lehký spánek, snění, zvýšené emoční stavy, pozornost, paměť	Zvýšená théta aktivita může indikovat emocionální stres a mentální úsilí, což může být relevantní při pokusech o lhaní.

Studie ukázaly, že různé typy mozkových vln odrážejí různé fyziologické stavy. Snížené alfa vlny a zvýšené beta vlny významně souvisí s pracovní zátěží, přemýšlením a pozorností, přičemž nadměrná beta aktivita naznačuje, že subjekt lže [45, 46]. Některé studie také prokázaly pokles aktivity alfa vln a zvýšenou aktivitu delta, théta a beta vln při působení stresových podnětů [45].



Obr. 4.: Frekvenční pásma EEG

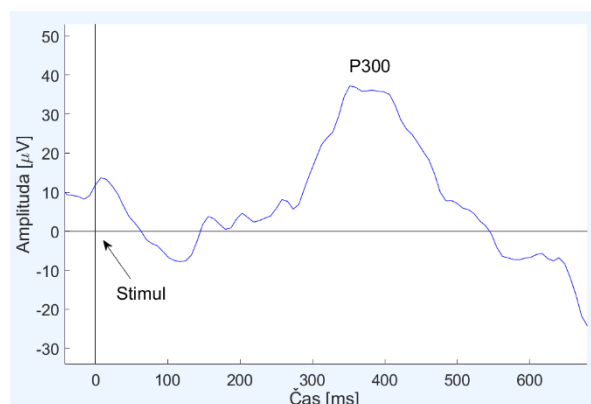
3.3.1 Event-Related Potential

Různé typy mozkových potenciálů jsou generovány v závislosti na typu podnětu. Jedním z nich je ERP, který představuje podvědomou psychologickou reakci vznikající v důsledku reflexu generovaného v lidském mozku při rozhodování, vykonávání mentální činnosti, pozorování něčeho známého nebo v reakci na vnější události [10, 17]. ERP analyzuje specifické krátkodobé segmenty dat mozkových vln vyvolaných mozkovou aktivitou zpracovávající informace a zkoumá různé aspekty těchto dat [3]. Při použití ERP byla identifikována aktivace mozku spojená s informacemi o podvodu, což činí ERP vhodným pro detekci skrytých informací. ERP je základní a nejpoužívanější metodou pro zkoumání reakce mozkové aktivity měřené jako výsledek motorické, kognitivní či sensorické události při zpracování informací z dat EEG [47].

3.3.2 P300 komponenta

Vlna P300 je pozitivní komponentou ERP, kterou lze identifikovat jako pozitivní výchylku v EEG signálu s typickou latencí přibližně 300-1000 ms po prezentaci stimulu (viz Obr. 5). Latence těchto výchylek se liší podle rychlosti zpracování informací v mozku jednotlivých subjektů. Tato odezva je vyvolána v mozku pouze v reakci na vzácné a smysluplné podněty v řadě irelevantních podnětů, které generují odlišnou reakci v mozku subjektu [4, 11]. Vzácná událost se poté objeví jako ERP s pozitivním vrcholem v záznamech EEG, běžně známým jako P300. Pravděpodobnost výskytu stimulu silně ovlivňuje velikost P300. Méně časté a výrazné stimuly v pravidelné sekvenci předvídatelných stimulů vyvolávají větší složky P300, což odráží zapojení kognitivních procesů. Tato složka je spojena s mnoha procesy, jako jsou pozornost, rozpoznávání a pracovní paměť [2]. Odezva P300 se objevuje s různou amplitudou na různých místech hlavy [16]. Zkoumáním amplitudy vlny P300 se zjišťuje, zda jednotlivci skrývá nějaké

informace [10, 17, 47]. Komponenta P300 představuje aktualizaci reprezentace stimulu v pracovní paměti. Tento proces je vysoce závislý na kontextu a je ovlivněn jak bezprostřední historií stimulů, tak požadavky na úkoly, jako jsou předchozí znalosti, očekávání a selektivní pozornost [25].



Obr. 5.: Odezva P300

Studium ERP bývá většinou zaměřeno na odezvu P300, která je maximální v parietálním laloku na středové elektrodě Pz. Výsledky výzkumů ukázaly, že podnět sondy (fotografie známé tváře) vyvolal signifikantně vyšší amplitudu P300 v parietálním místě Pz ve srovnání s irelevantními podněty u účastníků vinné skupiny, což naznačuje, že subjekt spáchal trestný čin [2]. To však není případ nevinné skupiny, ve které ERP složka P300 nevykazuje žádný významný rozdíl mezi sondou a irelevantními podněty, což dokazuje, že je subjekt nevinný. Toto klíčové pozorování lze využít k identifikaci lži u jednotlivých jedinců [9]. Detekce skrytých informací založená na analýze komponenty P300 je jednou z neúspěšnějších metod využívajících zpracování nervové aktivity zaznamenané EEG signály [25]. ERP jsou uznávány jako velmi výkonné nástroje pro detekci lži, přičemž amplituda P300 prokázala svůj významný vztah k rozpoznávání skrytých informací spojených s vnímáním obličeje [2, 13].

Průměrování ERP se v nedávných studiích osvědčilo jako efektivní metoda pro redukci šumu a zajištění konzistentních výsledků při analýze P300, zůstává tedy standardním přístupem ve výzkumu EEG pro studium mozkových a kognitivních procesů [3, 13, 22, 37, 48, 49]. Tento přístup výrazně zlepšuje kvalitu EEG signálů tím, že snižuje úroveň šumu a zvyšuje detekovatelnost specifických komponent ERP, jako je P300. I když může ztrácet některé variace signálu mezi jednotlivými pokusy, významně zvyšuje robustnost a spolehlivost výsledků, což je klíčové pro přesnou detekci mozkových reakcí a kognitivních procesů, jako je rozpoznávání známých tváří. Hlavní výhodou je izolace jednoduchých vzorců mozkové aktivity spojených se zpracováním konkrétního stimulu, které by nebyly rozpoznatelné v surových EEG datech. Stimul je prezentován opakovaně a reakce mozkových vln jsou následně zprůměrovány, což umožňuje sledovat změny v mozkových vlnách probíhajících během velmi krátkých časových úseků, které jsou nezbytné pro zkoumání rychle probíhajících mozkových procesů [3].

3.3.3 Artefakty v EEG signálu

Artefakt je nežádoucí jev, který má fyziologický nebo technický původ mimo samotná data EEG. Signály EEG jsou biologická data, a proto obsahují mnoho artefaktů a jsou ovlivněny značným šumem. Artefakty dělíme do dvou základních skupin, na technické a biologické. Mezi technické artefakty patří elektrostatické potenciály, impulsní rušení, šum elektronických obvodů, síťové napětí a nedostatečně navlhčené, zoxidované či poškozené elektrody. Biologické artefakty zahrnují pohyb, mrkání, pocení, pohyb očí, svalovou a srdeční činnost. Mnoho autorů se zabývalo odstraněním artefaktů pomocí různých metod, mezi které patří například průměrování signálu, adaptivní filtry a algoritmy časově-frekvenční analýzy, které zajišťují přesné odstranění artefaktů z EEG signálů. Metoda průměrování signálu redukuje náhodný šum a artefakty z dat získaných z opakovaných pokusů, čímž poskytuje čistší a lépe interpretovatelný signál pro následnou analýzu. Adaptivní filtry umožňují dynamicky oddělit artefakty od skutečných mozkových signálů na základě charakteristik šumu. Časově-frekvenční analýza pomáhá identifikovat a odstranit artefakty na základě jejich specifických frekvenčních vlastností [3, 8, 11]. Artefakty lze detekovat také pomocí dalších měření, jako jsou elektrookulografie (EOG) [5, 8, 11, 13, 15], elektromyografie (EMG) či elektrokardiografie (EKG). EOG slouží k měření pohybu očí, EMG k detekci svalových artefaktů a EKG k hodnocení srdečních aktivit. Artefakty je nutné odstranit, jelikož mohou způsobit zkreslení signálu či jiné neadekvátní informace [39, 50].

3.4 Detekce lži

Zkoumání současných technologických schopností detekovat lež je v této době velmi důležité kvůli zvýšeným bezpečnostním rizikům a situaci v oblasti zákona a pořádku v mnoha zemích [47]. Schopnost spolehlivě a přesně analyzovat skryté chování pro účely detekce podvodu je nezbytná a je používána v široké škále oblastí, jako jsou psychologie, soudní vědy, neurovědy, bezpečnost, hodnocení důvěryhodnosti, tajné služby, obrana, trestní řízení, boj proti terorismu a odhalování špionáže [2]. Cílem kriminalistické vědy je nejen správná identifikace pachatele, ale také zbavení nevinné osoby od podezření. Z hlediska lidských práv je minimalizace času a traumatu z vyšetřovacích postupů zásadní [3]. Výzkum skrytých informací se zaměřuje na detekci lži z biosignálů interpretací inherentních reakcí na jednotlivé stimuly. Identifikace podvodu je velmi náročný úkol a výsledky získané touto metodou musí být přesné, protože žádný nevinný by neměl být odsouzen za trestný čin, který nespáchal [8].

3.4.1 Konvenční detektory lži

Vyšetřovací agentury využívají mnoho detekčních technik pro podporu soudního systému odhalováním znalostí o vině v trestních případech [47]. Konvenční detektory lži založené na polygrafu jsou široce používanou technikou pro hodnocení pravosti skrytých informací, které analyzují lidské chování při

ústním výslechu pomocí sledování fyziologických ukazatelů jako je zvýšená tepová a dechová frekvence, vodivost kůže či krevní tlak [45]. Pozorování těchto nedobrovolných změn při odpovědi subjektu na soubor otázek je poté předmětem analýzy a subjektivní interpretace vyšetřovatelů, kteří určují pravdivost odpovědí, což může vést k falešně pozitivním či falešně negativním výsledkům [4, 8]. Současné přístupy k detekci lži jsou založeny na skutečnosti, že klamání může stimulovat různé fyziologické reakce zaznamenávané pomocí specializovaného vybavení, které může být časově a finančně náročné nebo nepohodlné [45, 47].

Spolehlivost a platnost detekce založené na polygrafu je diskutabilní, protože fyziologické reakce mohou být ovlivněny z různých důvodů a mohou být kontrolovány vědomě [47]. Existuje mnoho způsobů, jak detektor lži oklamat, například vyvoláním stresu po celou dobu výslechu či ovládním fyziologických funkcí usměrněním dechu a tepu. Někteří jedinci se také dokáží dostat do takového duševního stavu, že věří své vlastní lži. Na druhou stranu, nevinný člověk může být shledán vinným ze strachu či nervozity ze samotného výslechu, měření hladiny stresu tedy není vhodnou měřicí technikou [8]. Výsledky této techniky nejsou považovány za dostatečně spolehlivé důkazy v soudním systému kvůli své subjektivitě a nízké přesnosti. Toto vyšetřování je v České republice pouze dobrovolné, takže pokud podezřelá osoba odmítne, nemá to na ni žádný negativní vliv. K překonání problémů s polygrafickým testem bylo v posledních letech provedeno mnoho výzkumů využívajících neurofyziologické signály pro detekci podvodu [9]. Signály EEG pomáhají tento problém vyřešit, jelikož jsou nedobrovolné a subjekt je nemůže ovládat [8].

3.4.2 EEG detektory lži

Nejmodernější metody behaviorálních studií založené na neurovědách značně inspirovaly vědce a výzkumníky v oblasti detekce lží [47]. Hledání metod pro rozpoznání lži je neustálým cílem a předmětem intenzivního výzkumu. Mnoho studií zohledňuje různé přístupy k detekci skrytých informací, které jsou generovány jak ANS, tak CNS [11]. K dnešnímu dni bylo vynaloženo značné úsilí na odhalování podvodů prostřednictvím měření neurofyziologických signálů CNS. Mezi techniky, které prokazují své výhody při identifikaci lži, patří EEG, fMRI a fNIRS, přičemž EEG je nejčastěji využívanou metodou [2].

Lež je definována jako komplexní kognitivní psychologický proces, který vyžaduje pozornost a zahrnuje činnosti v různých oblastech mozku zaměřené na utajení podvodů a zločinů [11, 40]. Rozsáhlé důkazy prokázaly velký potenciál EEG jako nástroje pro měření elektrických aktivit mozku v různých oblastech prostřednictvím analýzy ERP P300, kde se tato složka výrazně projevuje během klamání [2, 9, 49]. Pravdivost odpovědí subjektu na otázky lze posoudit pozorováním změn mozkových vln, kde pravda a lež vytvářejí významné variace mezi různými frekvenčními pásmy [45, 46]. EEG je přesnější než jiné metody,

protože neměří fyziologické indexy, ale přímo mozkovou aktivitu [8]. Pro odhalení skrytých informací jsou odezvy P300 hodnoceny v různých frekvenčních pásmech. Neurofyziologické odezvy jsou poté zpracovávány pomocí různých klasifikačních algoritmů nebo metod statistické analýzy k posouzení pravdivosti výpovědi subjektu [40, 45].

Výsledky ukazují, že signály EEG odrážejí kognitivní a pozornostní mechanismy mozku, které jsou při pokusu o klamání spojeny se zvýšenými hodnotami ERP. Tento fakt umožňuje vývoj rychlejšího a objektivnějšího nástroje pro detekci lži ve srovnání s klasickými detektory, které závisí na subjektivní interpretaci [10, 45]. Dále bylo prokázáno, že kognitivní aktivity provinilého jedince zahrnují pozornost, zpracování klamné paměti, vybavování podvodu, vědomí krádeže a vyjádření kognitivní kontroly nad klamnými reakcemi na podněty [15, 49]. ERP P300 jsou zvláště významné díky svému výraznému vrcholu při vzácných událostech, což nabízí možnost spolehlivého testu CIT odolného vůči protiopatřením [13, 47]. Využitím výhod ERP byly vyvinuty detektory lži, které analýzou variací mozkových signálů identifikují vinu subjektu [10]. Analýza komponenty ERP P300 prokázala, že pachatel je pozorný a vědomý si své viny prostřednictvím delta, alfa a beta vln z parietálního laloku, což potvrzuje účinnost neurálních korelátů viníků během klamání [49].

3.5 Concealed Information Test

Detekce skrytých informací je široce využívána v oblastech, kde je klíčové zjistit, zda subjekt mluví pravdu, nebo se snaží utajit určité informace [16]. Při rozpoznávání skrytých informací pomocí EEG pro detekci lži se využívají různé techniky identifikace podvodu, jako je GKT, CIT, DIT a CTP, které jsou teoreticky spolehlivé a eticky přijatelné [11]. Umožňují detekci psychofyziologických aktivit, přičemž detaily související s trestným činem jsou známy pouze pachateli [9]. Různé studie implementovaly tyto techniky prostřednictvím vytvoření scénářů falešného zločinu k identifikaci změn mozkového potenciálu v kognitivních složkách EEG [10].

Nejčastěji používanou technikou je CIT, která je jednou z klíčových metod v kognitivní psychologii a odhalování podvodů. CIT se provádí za účelem analýzy lidského chování při lhaní, aby se zjistilo, zda je daný subjekt vinný či nevinný [2, 8, 10]. Pro tento typ výzkumu je CIT účinnou technikou, která studuje aktivitu CNS, ANS nebo obojí [16]. Metoda CIT je založena na rozpoznání specifických podnětů, jako je například vražedná zbraň či fotografie oběti [25]. Klasické paradigma pro CIT obsahuje tři kategorie podnětů prezentovaných účastníkům:

- **Sondy** (P-probes): První kategorií jsou zřídka se vyskytující, vzácné a smysluplné podněty související s trestným činem, známé pouze vinnými účastníky. Nevinný subjekt si tyto podněty neuvědomuje. Obrázky sond

fungují jako stimul pro subjekt, generující vlnu P300, a zobrazují známou tvář z falešného zločinu, což vede k výrazným paměťovým stopám [2, 25].

- **Cíle (T-targets):** Cílové podněty se používají k zajištění pozornosti a kontrole, zda subjekt spolupracuje. Jde o položky nesouvisející s trestným činem, které jsou známy všem zúčastněným subjektům, vinným i nevinným, a generují odpověď P300 [11].
- **Irelevantní (I-irrelevants):** Třetí kategorií jsou irrelevantní podněty, které nesouvisejí s vyšetřovaným trestným činem a negenerují žádnou odpověď P300. Jedná se o sadu neznámých obrázků anonymních tváří. Počet irrelevantních položek je větší, než počet sond a cílů. Vnímaný rozdíl podnětů sond od irrelevantních je způsoben jejich smysluplností pro subjekt a větší vnímaný rozdíl vede k větším rozdílům v P300 [25].

Analýza komponenty ERP P300 byla úspěšně přijata jako neurální indikátor pro CIT a používá se pro detekci lži. Princip CIT založený na ERP P300 spočívá v tom, že rozpoznání vzácného a smysluplného podnětu sondy, který je relevantní pouze pro provinilé účastníky, vygeneruje vyšší amplitudu P300 než irrelevantní podněty [2, 10]. Předpokládá se, že podezřelý rozpozná podnět relevantní pro trestný čin pouze v případě, že se jej účastnil. Osoba tedy může být odsouzena za trestný čin, pokud má specifické informace o zobrazeném podnětu, které záměrně popírá, přičemž je vyvolána odpověď P300 dokazující její vinu. Poskytnutým důkazem je, že pokud vinná osoba vidí známý podnět sondy, generuje se P300, i když viník tvrdí opak, zatímco u nevinného subjektu se P300 nevygeneruje. Na druhé straně nevýrazné podněty zůstávají nehlášené, takže je pro účastníka obtížné použít strategii protiopatření, aby metodu oklamal [2, 7]. Testy založené na ERP využívají amplitudu P300, která se opírá o hypotézu vyvolání odlišných odpovědí při předložení známých podnětů v sérii velkého počtu homogenních neznámých podnětů [7]. Z tohoto důvodu byl využit protokol CIT, kde je měřena aktivita CNS v reakci na ERP, přičemž P300 je dominantním potenciálem.

Předchozí studie dále prokázaly, že obličeje mohou být efektivně použity jako podněty v CIT založeném na ERP k implementaci účinného systému pro odhalování lži, jelikož je složka P300 citlivá na skryté rozpoznání obličejů. Výzkumy ukázaly, že výrazné podněty známých obličejů pronikají do povědomí a generují výrazné elektrické odpovědi v mozku účastníků, zatímco nové podněty nejsou dostatečně vnímány, aby byly zakódovány do paměti [13]. Vizualní stimuly známých a neznámých tváří v CIT založeném na P300 vyvolávají odlišné mozkové reakce a pomáhají tak v identifikaci vinné osoby, například zda podezřelý zná obličej konkrétní osoby (oběť, spolupachatel, člen teroristické skupiny) [11, 13, P.1]. Pro tento typ experimentu byl využit koncept zahrnující reakce na vizualní podněty pomocí analýzy vlny P300, která je odolná vůči protiopatřením. Tato metoda byla demonstrována na třídě podnětů, u kterých se dá očekávat, že poskytnou výrazné známé reakce, konkrétně obrázky známých tváří. Jednotlivé podněty jsou rozděleny do tří kategorií: obrázky sond, které znají

pouze účastníci zločinu, cílové obrázky pro kontrolu spolupráce subjektu a irelevantní obrázky, které nejsou podstatné ani pro viníky, ani pro nevinné.

3.6 Zařízení pro snímání EEG signálů

Díky vývoji nositelných zařízení obsahujících EEG senzory je tato technologie dostupnější a uživatelsky přívětivější [8]. EEG signály mapují mozkovou aktivitu jednotlivce prostřednictvím různých zařízení. Klasické dlouhodobé EEG vyžadovalo videozáznam, velký počet kabelů, záznamovou jednotku a odborníka pro nastavení elektrod. Zmenšením objemných jednotek EEG na malé náhlavní soupravy se snadnou manipulací se práce s touto technologií značně zjednodušila. V současné době existuje několik zařízení s různým počtem kanálů pro získávání signálů EEG [47]. Mezi tato zařízení patří například náhlavní soupravy od společnosti Emotiv, jako jsou Emotiv EPOC X, Emotiv Insight a Emotiv EPOC Flex [51]. Emotiv Insight, cenově nejdostupnější z těchto variant, obsahuje pět senzorů včetně nejvíce využívané elektrody Pz. Cena tohoto zařízení je přibližně 10 000 Kč. Emotiv EPOC X a Emotiv EPOC Flex nabízejí větší počet kanálů a širší funkčnost, avšak za vyšší cenu, která se pohybuje od cca 20 000 do 50 000 Kč [52]. Dalšími typy zařízení, která vědci v posledních letech používali v oblasti detekce lži, jsou EasyCap, Biosemi a Neuroscan. EasyCap, využívaný v několika studiích [5, 6, 8-10, 12], nabízí flexibilitu při umístění elektrod a jeho cena se pohybuje kolem 20 000 až 70 000 Kč v závislosti na konfiguraci [53]. Biosemi [13], další populární volba v EEG výzkumu, nabízí velmi přesné a stabilní měření, avšak jeho cena je výrazně vyšší, začínající na 350 000 Kč [54]. Neuroscan [15], známý pro svou vysokou přesnost a široké využití v neurovědách, je finančně nejnáročnější, s cenami od 450 000 Kč [55]. Tradiční polygrafy, které se běžně používají v praxi pro detekci lži, mají ceny v rozmezí od 70 000 do 230 000 Kč [56]. Byly také provedeny různé studie v této oblasti využívající levné náhlavní soupravy [21, 40, 45, 47, 49], nicméně žádná z nich se nezaměřila přímo na rozpoznávání obličejů pomocí ERP P300.

Na základě tohoto srovnání je zřejmé, že Emotiv Insight představuje cenově nejdostupnější variantu pro aplikace založené na EEG v oblasti detekce lži ve srovnání s jinými EEG zařízeními a tradičními polygrafy, což podporuje cíl této práce využít nízkonákladové zařízení pro detekci lži.

3.6.1 Použité snímací zařízení

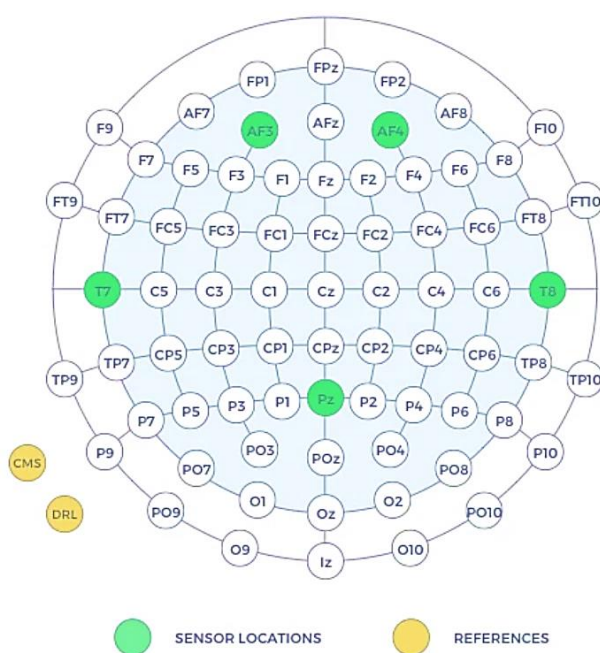
V tomto výzkumu bylo tedy ke snímání elektrické aktivity mozku využíváno komerční nízkonákladové zařízení Emotiv Insight (San Francisco, USA) [51] vyvinuté společností Emotiv Inc. (Obr. 6). Toto zařízení poskytuje pět EEG registračních elektrod, jmenovitě AF3, AF4, T7, T8 a Pz, snímajících elektrickou aktivitu, a dvě referenční elektrody, které slouží jako uzemňovací a jsou umístěny na levém mastoidu. Náhlavní souprava spolu s přidruženým softwarem umožňuje detekci a analýzu mozkových signálů. Zařízení disponuje senzory z hydrofilního polosuchého polymeru, které je nutné navlhčit fyziologickým roztokem pro

dosažení lepší kvality signálu. Podporovanými platformami pro zařízení jsou Windows, macOS, iOS a Android, což z něj dělá zařízení nezávislé na platformě. Data jsou přenášena bezdrátově pomocí protokolu Bluetooth, který slouží ke komunikaci s mobilními zařízeními, notebooky a stolními počítači [51]. Toto zařízení je neinvazivní, mobilní, bezdrátové, nízkonákladové, s nízkou časovou náročností, poskytující přenositelnost a snadné použití ve srovnání s klinickým nastavením EEG [P.5, P.10, P.12, 21, 29].



Obr. 6.: Zařízení Emotiv Insight [51]

Úloha lhaní vyvolává v některých specifických oblastech mozku uspořádanější mozkovou aktivitu než úloha říkat pravdu [48]. Umístění jednotlivých elektrod na pokožce hlavy podle mezinárodního systému 10-20 u zařízení Emotiv Insight je znázorněno na Obr. 7. Správný výběr kanálů s užitečnými informacemi poskytuje lepší výkon, šetří výpočetní čas, zvyšuje efektivitu získávání dat a může zvýšit přesnost klasifikace [4]. Jak již bylo zmíněno v předchozích částech, kanál Pz poskytuje cenné informace související s detekcí lži a nejlepší výkon v paradigmatu ERP P300 [17]. Z tohoto důvodu byla pro analytický postup v této studii použita právě elektroda Pz umístěná v parietální části pokožky hlavy.



Obr. 7.: Umístění elektrod u zařízení Emotiv Insight [51]

3.7 Softwarové vybavení pro zpracování dat EEG

Pro zpracování dat EEG bylo v této práci využito několik softwarových platforem a nástrojů, konkrétně EmotivPro, Matlab, EEGLAB a ERPLAB.

3.7.1 EmotivPro

EmotivPro je placený software sloužící k získávání dat ze zařízení Emotiv. Tento software poskytuje komplexní řešení společnosti Emotiv pro neurovědecký výzkum. EmotivPro umožňuje vytvářet a publikovat neurovědecké experimenty, stejně jako získávat a analyzovat data EEG v jednom integrovaném softwarovém prostředí. Naměřená data lze exportovat do různých formátů, jako je EDF, EDF+, BDF+ či CSV [57]. V této práci byla využita verze EmotivPro 4.2.0.541.

3.7.2 Matlab

MATLAB (MATrix LABoratory) je platforma pro programování a numerické výpočty. Využívá se k analýze dat, vývoji algoritmů a vytváření modelů. Implementace navrženého přístupu byla provedena v programovém prostředí Matlab verze R2020b (MathWorks Inc., Natick, USA).

3.7.3 EEGLAB

EEGLAB je bezplatná interaktivní sada nástrojů v Matlabu pro zpracování EEG dat. Zahrnuje mimo jiné časovou a frekvenční analýzu, odstranění artefaktů, statistiky související s událostmi a další funkce. EEGLAB běží na systémech Linux, Unix, Windows a macOS. Tento nástroj poskytuje interaktivní grafické uživatelské rozhraní (GUI), které uživatelům umožňuje flexibilně a interaktivně zpracovávat EEG data pomocí různých metod. EEGLAB obsahuje rozsáhlé výukové programy, okno nápovědy a funkci historie příkazů, které usnadňují přechod z průzkumu dat založeného na GUI k vytváření a spouštění vlastních skriptů pro analýzu dat. Poskytuje strukturované programovací prostředí pro ukládání, přístup, měření a vizualizaci EEG dat souvisejících s událostmi [58]. V této práci byla použita verze EEGLAB 2023.1. Dokumentace pro EEGLAB je dostupná online [59].

3.7.4 ERPLAB

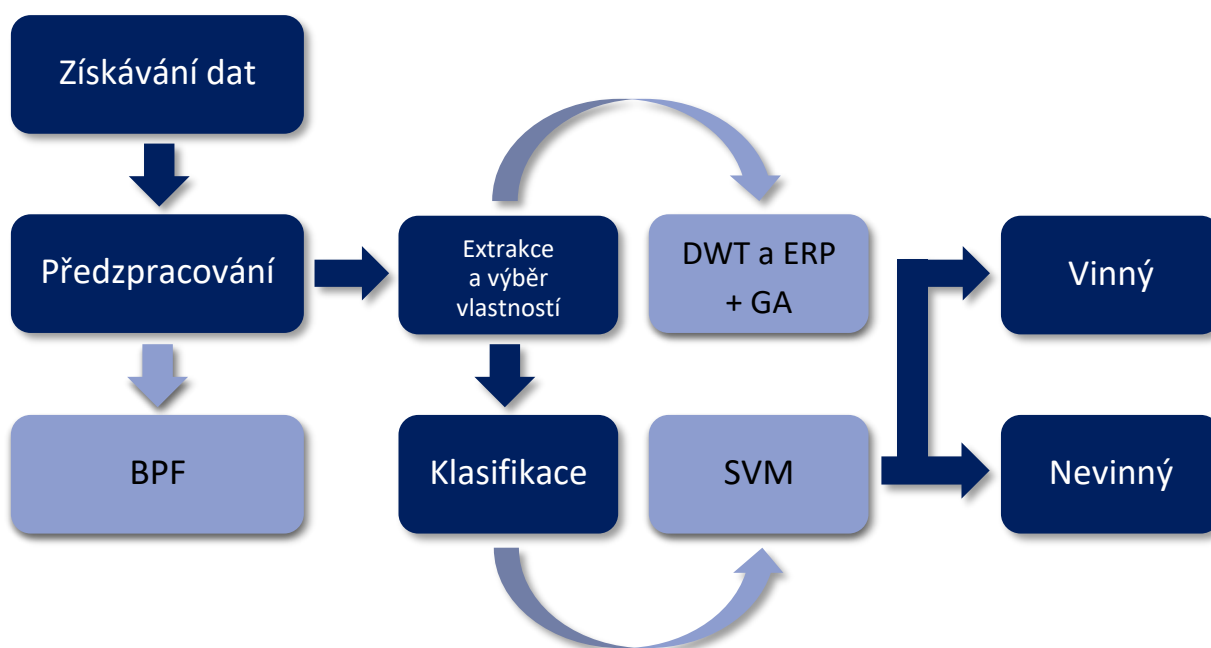
ERPLAB je bezplatná open-source sada nástrojů pro vytváření, zpracování, prohlížení a měření ERP z EEG dat. Využívá programovací prostředí Matlab a některé funkce z EEGLAB. Je úzce integrován s nástrojem EEGLAB a rozšiřuje jeho možnosti poskytnutím robustních a silných nástrojů pro zpracování, vizualizaci a analýzu ERP. GUI usnadňuje začátečnickům učení a skriptování v Matlabu a poskytuje vysoký výkon pro středně pokročilé a pokročilé uživatele. Pro zpracování a analýzu složky ERP zde byla použita verze ERPLAB 10.04 [60]. Více informací o tomto nástroji je dostupných v online dokumentaci [61].

4. METODIKA VÝZKUMU

Primárním cílem disertační práce je vytvořit nízkonákladový systém pro rychlou a efektivní detekci intencionální lži pomocí EEG signálů s využitím pokročilých metod strojového učení. Tato část práce se zaměřuje na stanovení metodiky pro analýzu EEG dat, zahrnující získávání dat, předzpracování dat, extrakci a výběr vlastností a následnou klasifikaci. Analýza EEG dat je komplexní proces, kde každý krok je nezbytný pro úspěšné zpracování dat a musí být řešen konsektivně [P.1]. Proto byla pro tuto analýzu navržena přesná metodika.

Nejčastější a nejúspěšnější metodou pro odhalení skrytých informací je komponenta ERP P300, která je generována při výskytu smysluplného podnětu. Tento výzkum byl realizován pomocí měření neurofyziologických změn během experimentu, kde byla zkoumána vlna P300 v reakci na vizuální podněty známých a neznámých tváří. V navrhované metodice jsou pro analýzu uvažovány především odezvy mozku subjektů na sondy a irelevantní podněty v parietálním laloku v oblasti Pz, protože tato část mozku je během mentálních úkolů neaktivnější. Na základě interpretace výzkumných výsledků se poté hodnotí přesnost klasifikace dat do dvou tříd a zjišťuje se, zda rozpoznání vlny P300 při reakcích účastníků dokáže odhalit nepravdivou odpověď s vysokou mírou přesnosti. Rozhodování o vině či nevině bylo prováděno pomocí strojového učení, konkrétně SVM s využitím GA pro výběr vlastností z dat DWT a ERP.

V současné studii je experimentální postup a metodika zpracování dat navržena tak, aby poskytovala optimální výsledky z hlediska přesnosti s důrazem na efektivní a rychlý systém. Schématický přehled navrhované metodiky pro detekci lži pomocí EEG signálů je znázorněn v typickém blokovém schématu na Obr. 8.



Obr. 8.: Schéma navržené metodiky pro detekci lži pomocí EEG

Pro dosažení cílů disertační práce byly vybrány následující vědecké metody, které umožňují efektivní zpracování a analýzu EEG dat. Tyto metody byly pečlivě vybrány na základě jejich vhodnosti pro detekci lži pomocí EEG a schopnosti přispět k dosažení cílů této práce.

Získávání dat (Experiment)

Získání kvalitních EEG signálů je základem pro jakoukoli analýzu. V tomto výzkumu bylo použito nízkonákladové zařízení s malým počtem kanálů, vybavené elektrodou Pz, která je klíčová pro detekci lži pomocí ERP P300.

- **Metoda experimentu:** Testování a ověření formulovaných hypotéz za definovaných podmínek.
- **Experimentální design:** Navržení vlastního experimentu pro detekci lži pomocí EEG, včetně vytvoření specifických scénářů a stimulů pro vyvolání odpovídající mozkové reakce.
- **Získání dat EEG:** Data byla získána pomocí nízkonákladové náhlavní soupravy Emotiv Insight a softwaru EmotivPro.

Předzpracování dat

Předzpracování dat zajišťuje, že signály budou v dostatečně čisté formě pro následnou analýzu. Data byla předzpracována pomocí nástrojů EEGLAB a ERPLAB v Matlabu.

- **Tvorba epoch:** Data byla rozdělena do jednotlivých epoch pro sledování časové události.
- **Filtrace a odstranění artefaktů:** Pro odstranění šumu a nežádoucích složek signálu byly použity různé filtrační techniky, které zajistily čistotu dat.
- **Výběr elektrod:** Pro analýzu byla vybrána elektroda Pz, která je nejvhodnější pro tento typ výzkumu.
- **Průměrování Event-Related Potential (ERP):** Data ERP z jednotlivých pokusů pro každý podnět byla zprůměrována.

Extrakce a výběr vlastností

Byly testovány různé metody pro extrakci relevantních vlastností souvisejících s detekcí lži na základě jejich přínosu pro analýzu EEG dat.

- **Statistické vlastnosti:** Pro komplexní analýzu byly vybrány různé statistické vlastnosti aplikované na data z časové i časově-frekvenční domény.
- **Discrete Wavelet Transform (DWT):** DWT umožňuje dekompozici EEG signálu na různá frekvenční pásma, což usnadňuje identifikaci specifických frekvenčních komponent souvisejících s různými mentálními stavy.

- **Genetický algoritmus (GA):** GA je optimalizační technika inspirovaná evolučními procesy, která byla využita k výběru nejrelevantnějších vlastností pro klasifikaci.

Klasifikace

Pro klasifikaci EEG dat byly použity různé klasifikační algoritmy v nástroji Classification Learner v Matlabu, přičemž nejvyšší přesnosti dosahoval SVM.

- **Support Vector Machine (SVM):** Medium Gaussian SVM s jádrovou funkcí byl zvolen pro jeho schopnost efektivně zpracovávat data s vysokou dimenzionalitou.
- **Testování modelu:** Zde byly zkoumány závislosti parametrů EEG na úspěšnosti identifikace jednotlivých subjektů do dvou tříd, vinných a nevinných.
- **Metoda dedukce:** Byla využita pro ověření platnosti stanovených závěrů a správnosti navržené metodiky.

Vyhodnocení výsledků

Vyhodnocení přesnosti klasifikace zahrnovalo analýzu správnosti identifikace vinných a nevinných jedinců.

- **Metoda komparace:** Výsledky přesnosti klasifikace byly porovnány se současnými studii zabývajícími se touto problematikou.
- **Metoda indukce:** Byla použita pro vyvození závěrů na základě získaných výsledků a poznatků.
- **Metoda syntézy:** Byla aplikována k propojení poznatků získaných analýzou do komplexního modelu, který integruje různé aspekty EEG a poskytuje robustní nástroj pro detekci lži.
- **Tvorba aplikace:** Byla vyvinuta vlastní aplikace pro automatickou detekci lži pomocí EEG.
- **Archivace dat a dokumentace:** Data byla archivována ve formátu CSV pro další analýzu a byla vytvořena dokumentace s návodem k aplikaci a celému procesu detekce lži.

Vybrané metody předzpracování, extrakce a výběru vlastností a následné klasifikace dat jsou klíčové pro dosažení cílů této disertační práce. Každá metoda byla pečlivě zvolena na základě její schopnosti přispět k analýze EEG signálů pro detekci lži. Důkladné vyhodnocení výsledků a jejich porovnání se současnými studii poskytuje hluboký vhled do efektivity těchto metod.

5. EXPERIMENTÁLNÍ ČÁST

Tato část podrobně popisuje analýzu dat pro detekci lži pomocí EEG, rozdělenou do několika fází, konkrétně získávání dat, předzpracování dat, extrakce a výběr vlastností a klasifikace dat k odlišení vinných subjektů od nevinných. Byla zde zkoumána složka P300 v reakci na různé typy vizuálních podnětů. Výzkum se zaměřuje na detekci skrytých informací z dat získaných od účastníků pomocí komerčně dostupné náhlavní soupravy využívající signály EEG k vyhodnocení účinnosti a přesnosti rozpoznávání lži. Po snímání dat následuje předzpracování EEG dat, na něž jsou aplikovány různé techniky extrakce a výběru vlastností. Pro analýzu dat byly využity různé metody k dosažení co nejlepšího výsledku. Výsledky této práce byly porovnány se stávajícími studiemi založenými na ERP odpovědi s cílem zhodnotit přesnost, vhodnost a použitelnost navrženého systému využívajícího nízkonákladové zařízení EEG pro detekci lži.

5.1 Získávání dat (Experiment)

Jedním z cílů této práce je ověřit, zda skryté informace vyvolané vizuálními stimuly mohou odhalit lež s vysokou mírou přesnosti pomocí analýzy EEG signálů získaných nízkonákladovým zařízením. K zodpovězení této otázky byl proveden ERP experiment podle protokolu Concealed Information Test (CIT), kde byla sledována mozková aktivita v reakci na vizuální stimuly známých a neznámých tváří. Subjekty se snažily skrýt své vědomosti o účasti na dané události. Soubor dat se skládal z 50 jedinců (35 mužů, průměrný věk 30,7 let, a 15 žen, průměrný věk 34,2 let). Experimentální design zahrnoval studenty, učitele a pracovníky z Univerzity Tomáše Bati ve Zlíně. Cílem bylo zjistit, zda navržený postup dokáže detekovat skryté informace týkající se skutečných událostí uložených v mozku během experimentálních sezení.

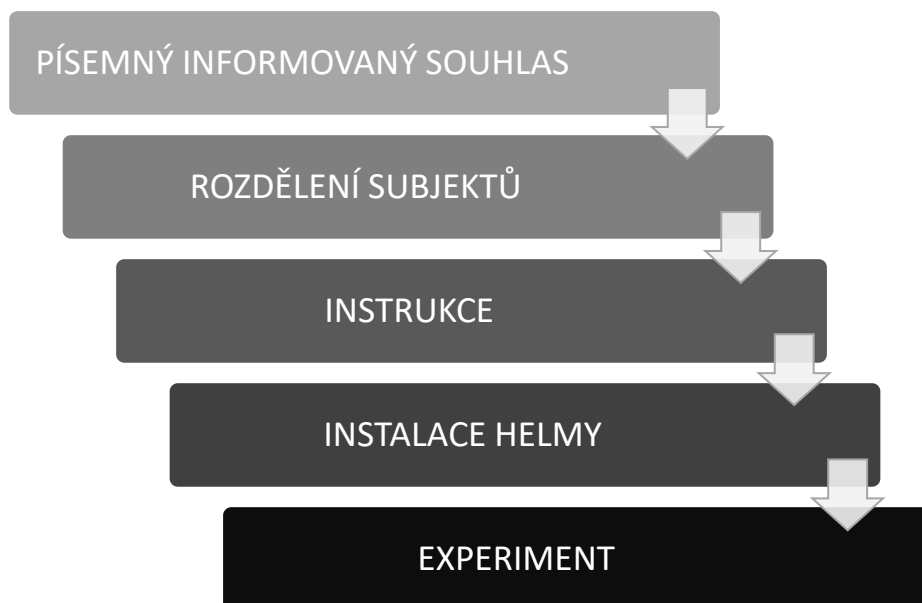
V rámci této disertační práce byl navržen nový přístup k detekci skrytých informací pomocí nízkonákladového zařízení EEG. Nově navržený experiment zahrnuje padesát jedinců pro prvotní ověření hypotézy, což je významně více než počet deseti subjektů běžně využívaných v současných studiích. Pro dosažení vyšší robustnosti výsledků bude však potřeba zahrnout větší a diverzifikovanější populaci, včetně jedinců s hlubšími znalostmi v příslušných oblastech.

Experimentální testování využívalo protokol založený na vyvolávání ERP s důrazem na složku P300. CIT je běžně využívanou technikou ke zkoumání rozdílů v mozkové aktivitě mezi vinnými a nevinnými subjekty. Namísto kladení otázek byly subjektům na obrazovce zobrazeny různé vizuální podněty, které fungovaly jako stimul a generovaly v mozku odlišné ERP reakce. Experiment zahrnoval fotografie osob souvisejících s kriminalitou, které se používají k vyvolání pravdivých a klamných odpovědí subjektů. V rámci experimentu byly prezentovány tři kategorie kritických podnětů, konkrétně sondy (známé tváře), cíle (známé tváře) a irelevantní (neznámé tváře). Podněty sond byly prezentovány

zřídka a týkaly se skrytých informací souvisejících s vyšetřovaným trestným činem, o kterých ví pouze vinný subjekt. Sonda je fotografie osoby, kterou vinný subjekt dobře zná. Irelevantní podněty nejsou rozpoznávány vinnými, ani nevinnými, protože nesouvisejí s vyšetřovaným trestným činem a negenerují odpověď P300. Irelevantní podněty tvořily fotografie získané z databáze neznámých neutrálních tváří a sloužily jako distraktory [62]. Cílové podněty byly známé všem zúčastněným subjektům a generovaly odpověď P300. Tyto podněty jsou používány k udržení koncentrace v experimentu a nesouvisejí s trestným činem. Pro položky cílů byly použity snímky známých osobností, politiků a vědců z celého světa, které byly shromážděny z různých internetových zdrojů.

5.1.1 Příprava experimentu

Před zahájením experimentu byl všem subjektům předložen písemný informovaný souhlas s účastí ve výzkumu. Účastníci byli následně náhodně rozděleni do dvou skupin, na vinné a nevinné. Každá relace zahrnovala specifickou prezentaci připravenou v PowerPointu. Všechny signály subjektů byly hodnoceny s jejich písemným souhlasem. Výzkum byl schválen Etickou komisí výzkumu Univerzity Tomáše Bati ve Zlíně. Anonymizace a důvěrnost dat byla zajištěna v souladu s ochranou osobních údajů. Následně byl účastníkům poskytnut stručný popis a princip experimentu. Účastníci byli instruováni, aby odložili mobilní telefony a minimalizovali pohyby těla, očí a mrkání během experimentu, aby se předešlo vzniku artefaktů. Byli také instruováni, aby odpovídali pravdivě či lhali, když čelí určitým podnětům, a aby měli oči fixované na střed obrazovky. EEG signály byly zaznamenávány pomocí náhlavní soupravy Emotiv Insight. Časové rozpětí experimentu bylo přibližně 15 minut, přičemž příprava trvala cca 10 minut a prezentace 5 minut. Navržený experimentální postup lze vidět na Obr. 9.



Obr. 9.: Experimentální postup

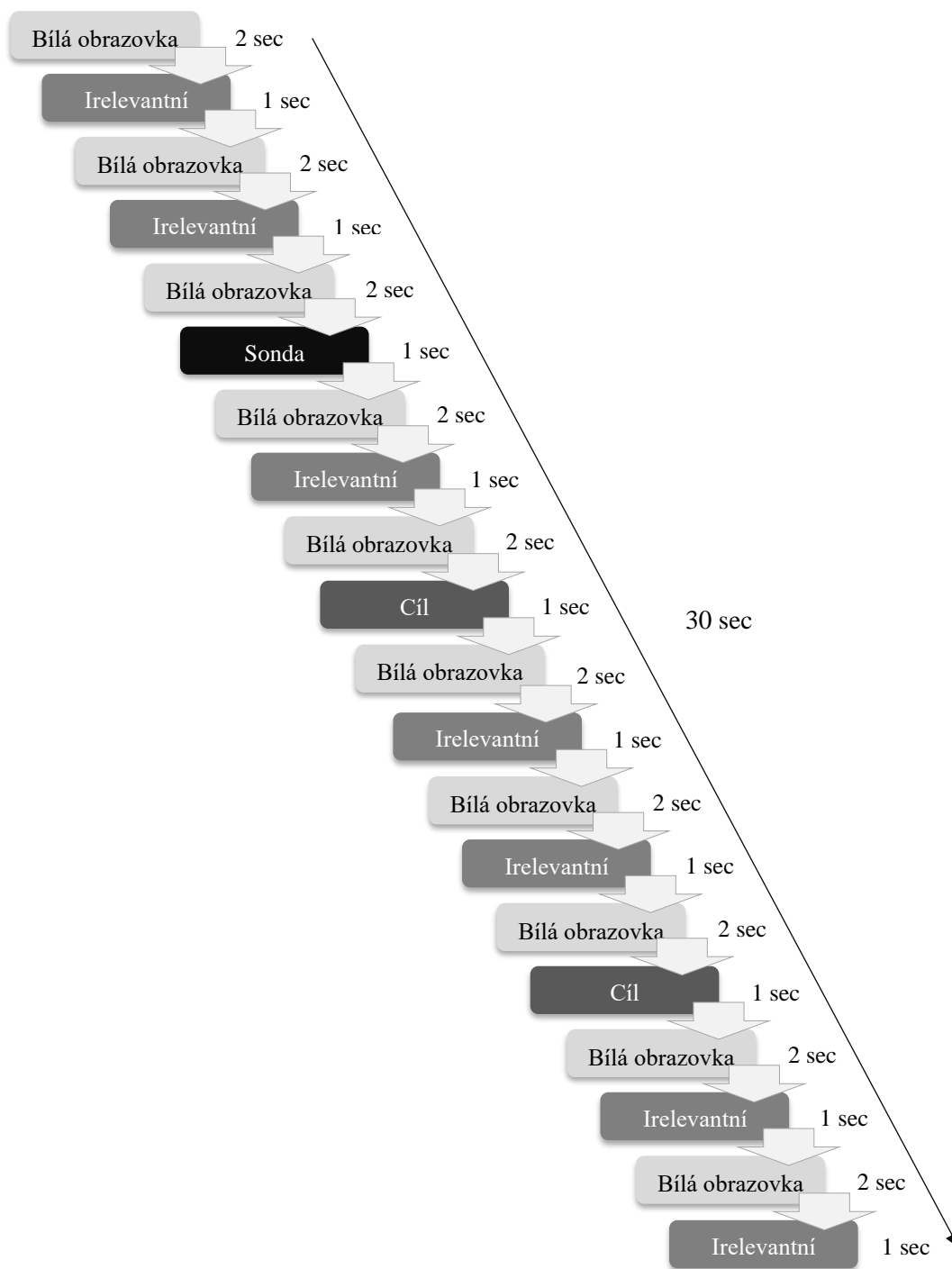
5.1.2 Experimentální scénář

Pro detekci lži byl vytvořen experimentální scénář simulující falešný zločin, konkrétně krádež na Fakultě aplikované informatiky na Univerzitě Tomáše Bati ve Zlíně. Tento scénář zahrnoval sérii vizuálních podnětů s cílem vyvolat vlnu P300. Účastníkům z vinné skupiny byla před experimentem předložena fotografie osoby, které měli údajně ukrást určitý předmět. Jejich úkolem bylo identifikovat tuto osobu jako neznámou a na ostatní podněty odpovídat pravdivě. Byli instruováni, aby se mentálně chovali jako nevinní, podobně jako při klasickém výslechu.

Odezva P300 generovaná v mozku viníků na sondu a cílový podnět je stejná, protože sondu znají, ale záměrně to popírají. Irelevantní obrázky neznámých osob vyvolávají jinou odpověď. Účastníci z nevinné skupiny odpovídali pravdivě na všechny podněty. Reakce generovaná v mozku nevinných subjektů je pro sondu a irelevantní podněty stejná, protože oba stimuly jsou pro ně neznámé. Pouze cílové stimuly vyvolají P300 v mozku nevinných subjektů. Všechny podněty vyžadovaly co nejrychlejší odpověď, která indikovala, že účastník viděl zobrazenou položku.

Během experimentu byly podněty prezentovány v 10 blocích. Každý blok obsahoval 10 vizuálních podnětů zahrnujících sondu, cíl a irelevantní podněty. Celkově se tedy paradigma skládalo ze 100 segmentů mozkových vln pro každý subjekt. Vzhledem k tomu, že vzorkovací frekvence EEG zařízení byla 128 Hz, bylo během každého experimentu získáno 12 800 datových bodů. Celkem bylo získáno 5000 reakcí na jednotlivé obrazové stimuly pro 50 subjektů, což představuje 640 000 datových bodů. Jednotlivé pokusy v prezentaci trvaly 30 sekund, s celkovou dobou trvání 300 sekund pro 10 bloků. Každé měření začínalo prezentací černého fixačního kříže na bílém pozadí pro záznam mozkové aktivity před stimulem, po kterém následoval snímek známé či neznámé tváře. Snímky byly prezentovány ve středu obrazovky po dobu 1000 ms, následované interstimulačním intervalem trvajícím 2000 ms, během něhož byl prezentován fixační kříž pro zaměření pozornosti zpět na střed obrazovky. Obrazovka byla umístěna ve vzdálenosti cca 1 m od účastníka.

Pořadí, ve kterém byly kritické stimuly prezentovány v rámci bloků, bylo rozloženo tak, aby mezi cílem/sondou a irelevantními podněty nebyly velké mezery. Sondy byly prezentovány omezený početkrát, aby se minimalizoval efekt naučeného chování. Naopak cílové a irelevantní stimuly byly prezentovány častěji, aby se dosáhlo dostatečné variability a aby účastníci nemohli spoléhat na naučené chování. Tento přístup zajišťuje validitu experimentu tím, že je detekce lži založena na skutečných mozkových reakcích, nikoliv na jednoduchém rozpoznávání stimulů. Časová osa experimentu pro jeden blok je znázorněna na Obr. 10.



Obr. 10.: Časová osa experimentu pro jeden blok

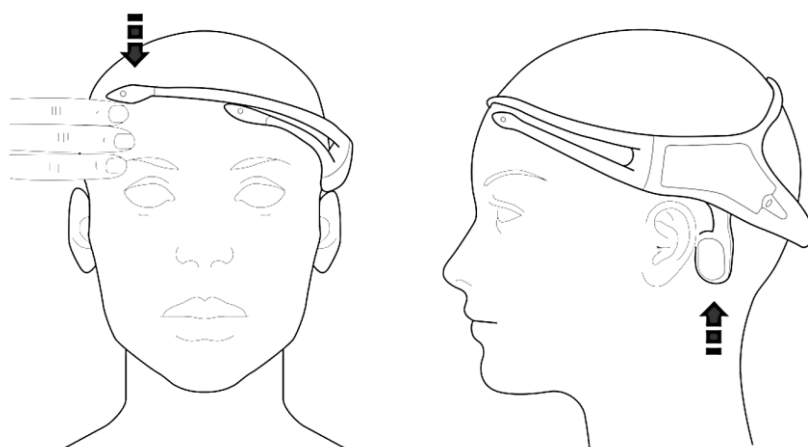
Experimentální scénář byl navržen pro sběr dat mozkových vln v klidném prostředí, což je zásadní pro tento typ experimentu. Účastníci byli instruováni, aby odpovídali „ANO“ či „NE“ po zobrazení každé fotografie, aby bylo možno porovnat změny mozkových vln při jednotlivých odpovědích. Data EEG byla shromážděna během prezentace snímků známých a neznámých tváří. Všechny fotografie neutrálních tváří s bílým pozadím byly upraveny tak, aby mezi nimi nebyly významné vizuální rozdíly. Fotografie byly vycentrovány zarovnáním oční linky do stejné horizontální polohy ukazující čelní pohledy na tváře, byly převedeny do odstínů šedi a byl upraven kontrast.

5.1.3 Snímání dat

EEG data byla získávána během experimentu za účelem detekce lži pomocí analýzy komponenty P300, která zahrnuje reakce na vizuální podněty v rámci přizpůsobeného protokolu falešného zločinu. Subjektům byly prezentovány obrázky promítané na obrazovce, přičemž bylo zjišťováno, zda jsou obeznámeni s tváří či nikoli. Fotografie známé osoby související s trestným činem fungovala jako sonda, což mělo vyvolat v mozku vinného subjektu reakci P300 s vysokou amplitudou. Tato složka P300 je vyvolána, když je jedinec konfrontován s podněty souvisejícími s jeho vzpomínkami, což je klíčové pro identifikaci vinných subjektů.

Umístění náhlavní soupravy

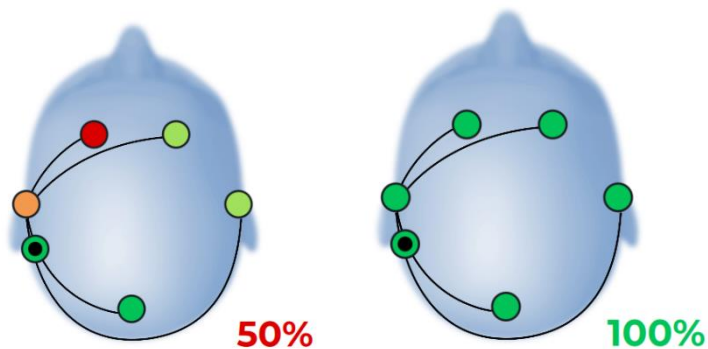
Nejprve byla nasazena náhlavní souprava s adekvátně navlhčenými elektrodami. Signály EEG byly zaznamenávány pomocí nízkonákladového neinvazivního 5kanálového zařízení Emotiv Insight během prezentace vizuálních podnětů. Toto zařízení obsahuje aktivní polosuché polymerové elektrody umístěné na povrchu pokožky hlavy podle mezinárodního systému 10-20 v oblastech AF3, AF4, T7, T8 a Pz. Signál byl zaznamenáván s frekvencí vzorkování 128 Hz. Správné nasazení zařízení je zásadní, jelikož každá elektroda měří specifickou část mozku, ze které se získávají jednotlivá data pro analýzu (viz Obr. 11).



Obr. 11.: Správné umístění zařízení Emotiv Insight [51]

Kontrola kvality signálu

Před zahájením záznamu bylo nutné dosáhnout vysoké kvality EEG signálu, což je klíčové pro následnou analýzu. Kvalita signálu byla monitorována a ověřována vizuálně pomocí softwaru EmotivPro, viz Obr. 12. Nejprve se několik sekund čekalo než se signál stabilizuje. Signály s nízkou kvalitou byly indikovány červenou barvou, zatímco optimální signály byly zobrazeny zeleně. Pokud senzory vykazovaly nízkou kvalitu signálu, byly dodatečně navlhčeny, aby se zajistila optimální vodivost.



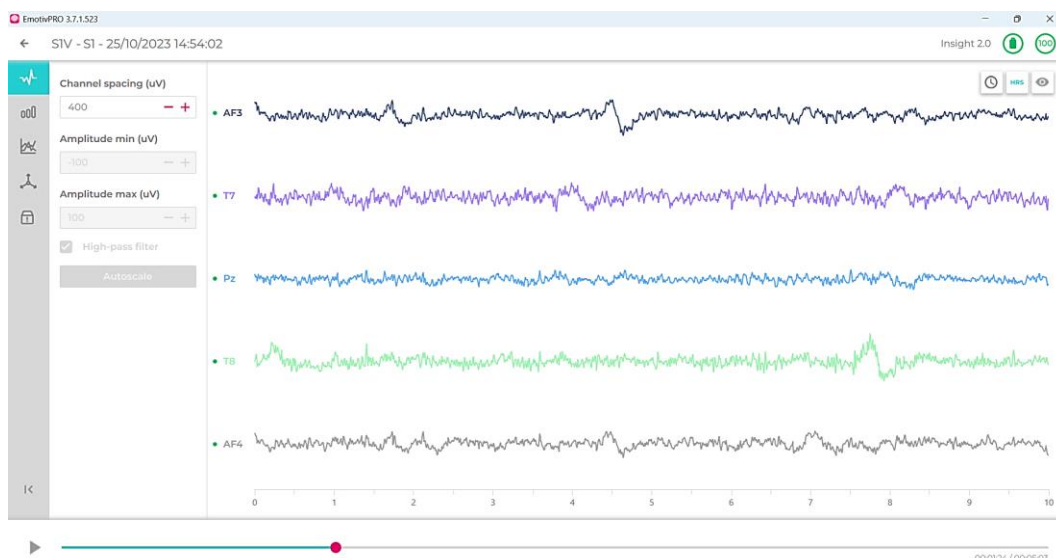
Obr. 12.: Kontrola kvality EEG signálů – nízká (vlevo) vs vysoká (vpravo) kvalita

Zahájení měření

Po dosažení vysoké kvality signálu, kdy všechny elektrody vykazovaly zelenou barvu indikující optimální kvalitu, byl zahájen proces záznamu signálu. Tato procedura zajišťuje, že pouze data s vysokou kvalitou jsou použita pro detailní analýzu a interpretaci, což minimalizuje riziko šumu a artefaktů v signálu a umožňuje přesnější detekci P300 v reakci na prezentované podněty. Následně byla spuštěna testovací prezentace, která nezahrnovala žádné podněty sond, pouze cílové a irelevantní podněty. Poté proběhl hlavní experiment.

Obr. 13.: Spuštění nahrávání záznamu v EmotivPro

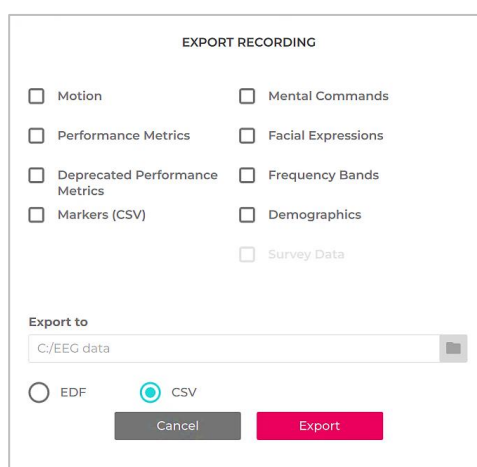
Experiment detekce lži začal prezentací snímků známých a neznámých tváří na obrazovce, během kterého byla zaznamenávána data EEG. Všechny položky byly prezentovány na stejném místě obrazovky. Subjekt musel tyto obrázky rozpoznat a po každém obrázku nahlas odpovědět „ANO“ nebo „NE“ co nejrychleji. Na začátku procesu sběru dat se účastníkovi zobrazila prázdná obrazovka po dobu dvou sekund, aby se získal základní klidový mozkový signál. Tím se zabránilo jakýmkoli mozkovým aktivitám souvisejícím s předchozími stimuly.



Obr. 14.: Průběh měření v EmotivPro

Export dat

Naměřené surové EEG signály byly získány pomocí softwaru EmotivPro, který je určen primárně pro práci s náhlavní soupravou od firmy Emotiv, a byly exportovány pro následnou analýzu. Jelikož je pro zpracování dat v nástroji EEGLAB potřeba upravit strukturu souboru, byl využit formát CSV, který tuto úpravu umožňuje.



Obr. 15.: Export dat z EmotivPro

Úprava dat

Před samotným předzpracováním proběhla úprava dat do specifického formátu, protože pro import dat do nástroje EEGLAB je potřeba mít správnou strukturu datového souboru. Byla odstraněna veškerá data, která by mohla zapříčinit chybu při nahrávání souborů, například nadpisy, hlavičky a další nepotřebná data.

Automatizace procesu

Úprava dat pro všechny subjekty proběhla pomocí kódu vytvořeného v Matlabu, který byl poté využit ve výsledném programu (viz podkapitola 6.1.1).

5.2 Předzpracování dat

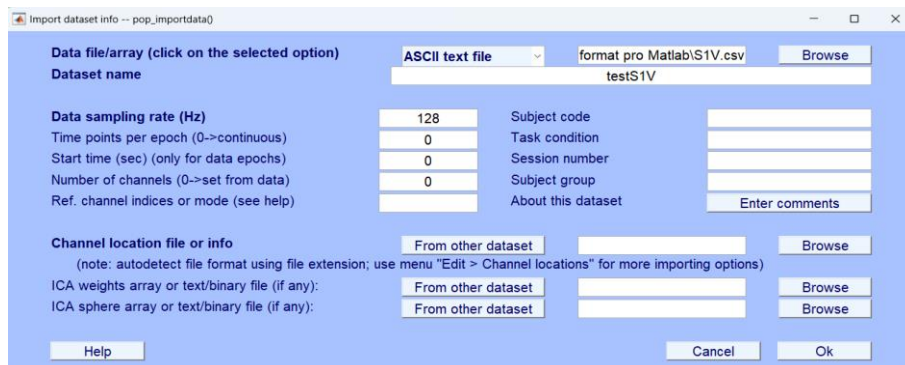
Před zahájením analýzy dat je nezbytné předzpracovat signály EEG za účelem odstranění artefaktů a šumu, které mohou ztížit detekci komponenty ERP P300 generované stimuly a snížit celkovou přesnost systému [10]. Mozkové vlny jsou slabé fyziologické signály a jsou náchylné k rušení, proto je odstranění nežádoucích signálů důležité pro následné zpracování a analýzu [8]. Artefakty mohou být odstraněny vizuálně nebo pomocí automatizovaných metod pro detekci nechtěných výkyvů. Vizuální kontrola a odstranění artefaktů je časově náročná, zvláště u velkého objemu dat, proto se využívají automatizované metody. Předzpracování EEG dat bylo provedeno pomocí nástrojů EEGLAB a ERPLAB v Matlabu.

Nastavení cesty pro EEGLAB

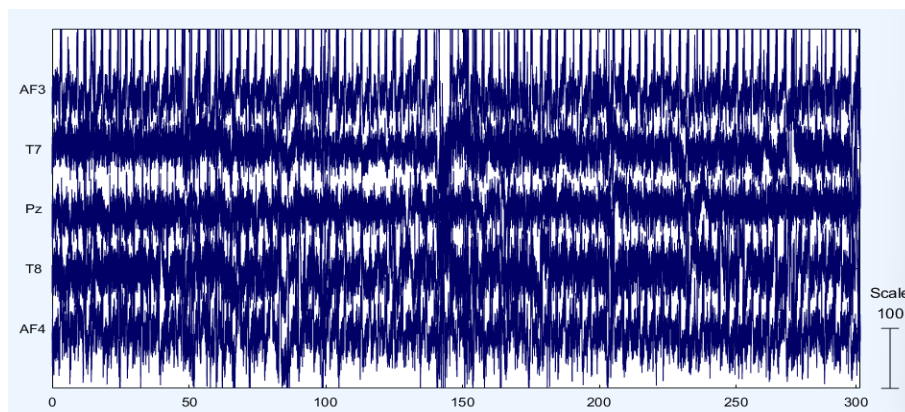
Před zahájením analýzy je nutné nastavit cestu ke složce EEGLAB. Postup v Matlabu: Home → Set Path → Add with Subfolders → „umístění složky EEGLAB“ → Save.

Import dat

Dalším krokem je import dat do EEGLAB. Signály byly vzorkovány při frekvenci 128 Hz. Postup v EEGLAB: File → Import data → Using EEGLAB functions and plugins → From ASCII/float file or MATLAB array.



Obr. 16.: Import dat v EEGLAB



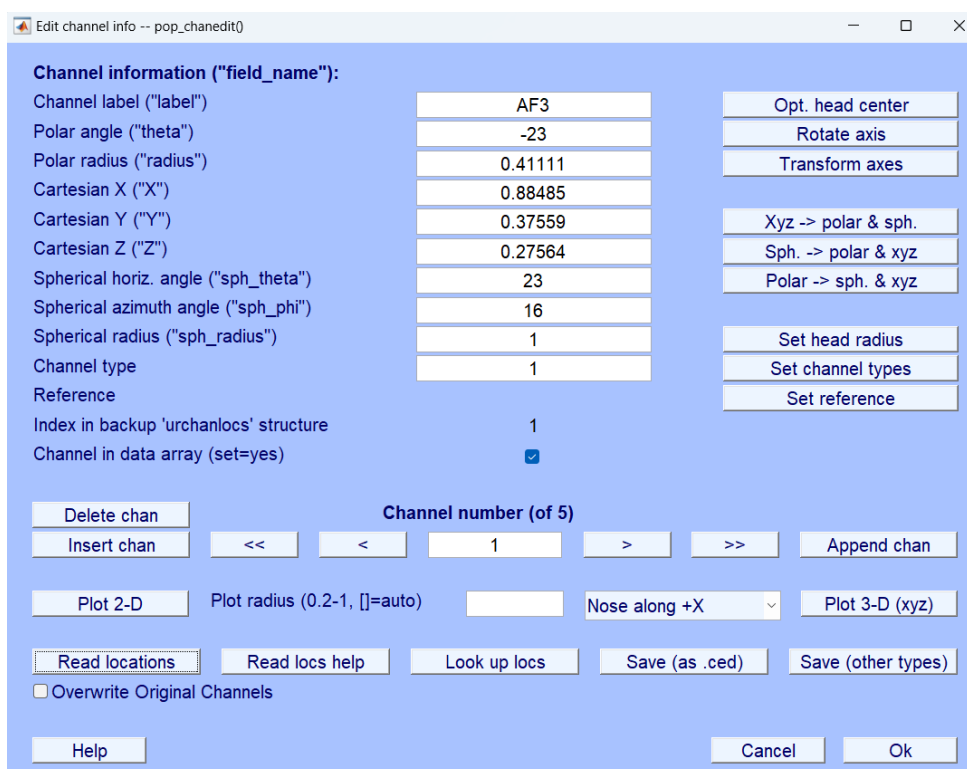
Obr. 17.: EEG signály

Výběr dat

Následně byl proveden výběr vhodných naměřených dat pro následnou analýzu. Čtyři subjekty byly odstraněny kvůli nadměrnému množství artefaktů nebo neplatnému měření, které byly způsobeny externím hlukem nebo systémovými chybami. Klidné prostředí je klíčové pro tento typ experimentu, protože zaznamenává i nejmenší výkyvy v mozkových signálech, což činí rušivé elementy velmi nežádoucími. Výběr dat byl založen na kontrole kvality měření signálu v EmotivPro, na vizuální kontrole ERP vln a na odstranění velkého množství artefaktů během předzpracování.

Umístění kanálů

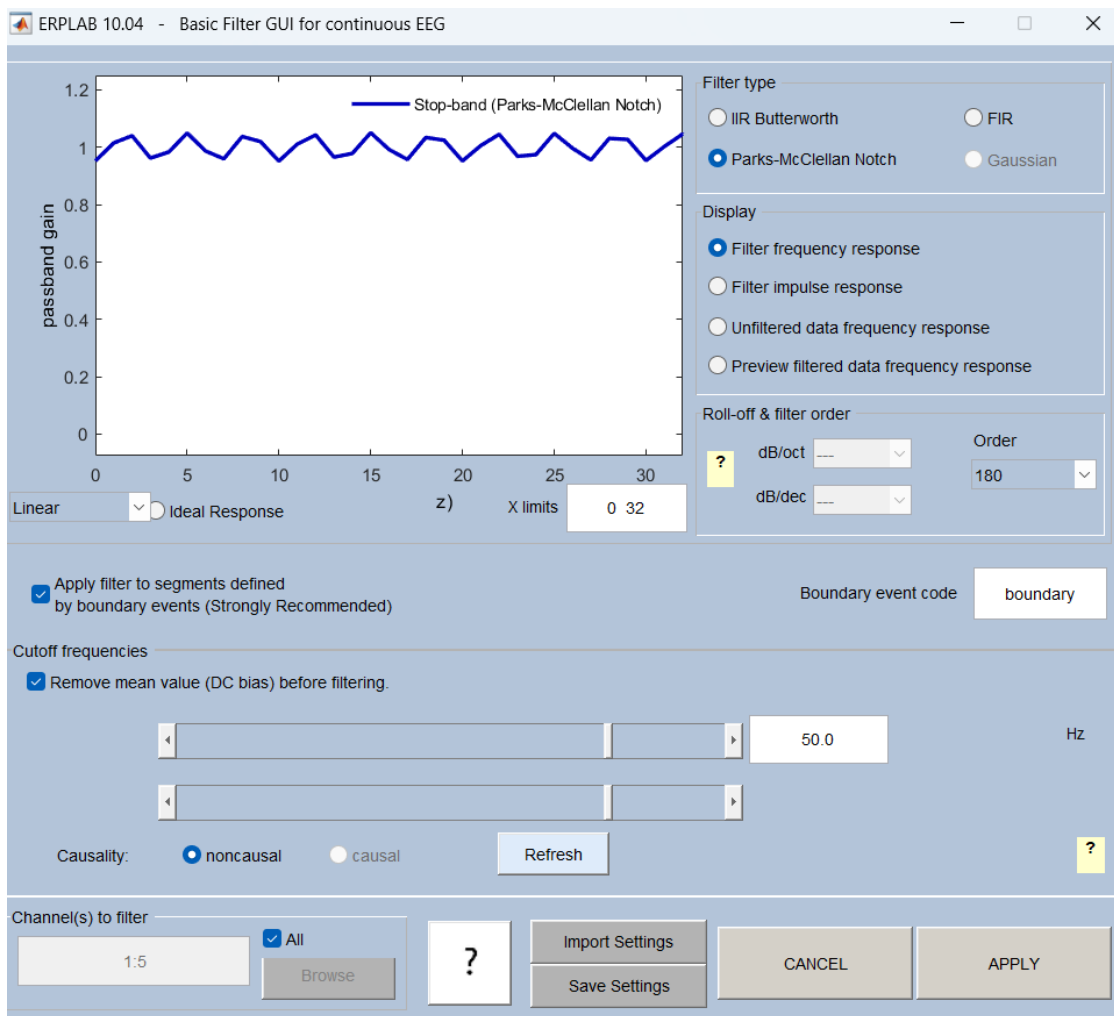
Po importu a kontrole dat se provádí nastavení umístění jednotlivých kanálů. Soubor pro umístění kanálů musí mít danou strukturu, viz PŘÍLOHA P V: UMÍSTĚNÍ KANÁLŮ. Postup v EEGLAB: Edit → Channel locations → Ok → Read locations → „výběr souboru Insight.ced“ → autodetect → Ok.



Obr. 18.: Nastavení umístění kanálů v EEGLAB

Filtrace (Notch filtr)

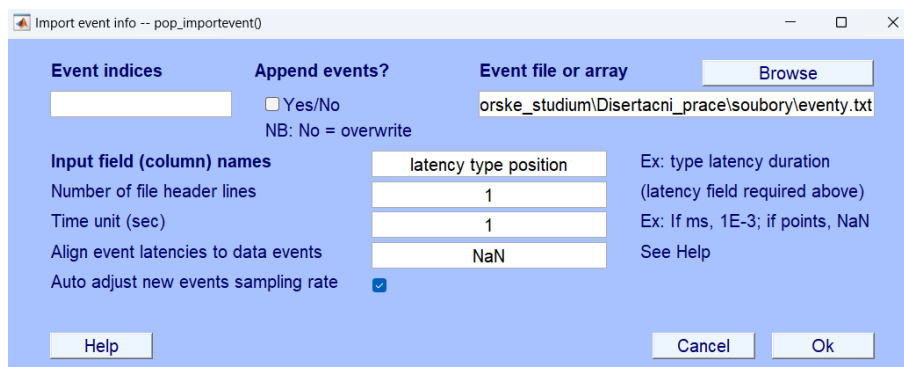
Artefakty a šum jsou vždy přítomny v záznamech EEG signálu, a proto je předzpracování signálu nutné pro zvýšení poměru signálu k šumu. Předběžná filtrace dat byla provedena pomocí Parks-McClellan Notch filtru v ERPLAB, který odstraní DC offset a horní hranici signálu nastavenou na 50 Hz. Postup v EEGLAB: ERPLAB → Filter & Frequency Tools → Filters for EEG data → Parks-McClellan Notch → Apply.



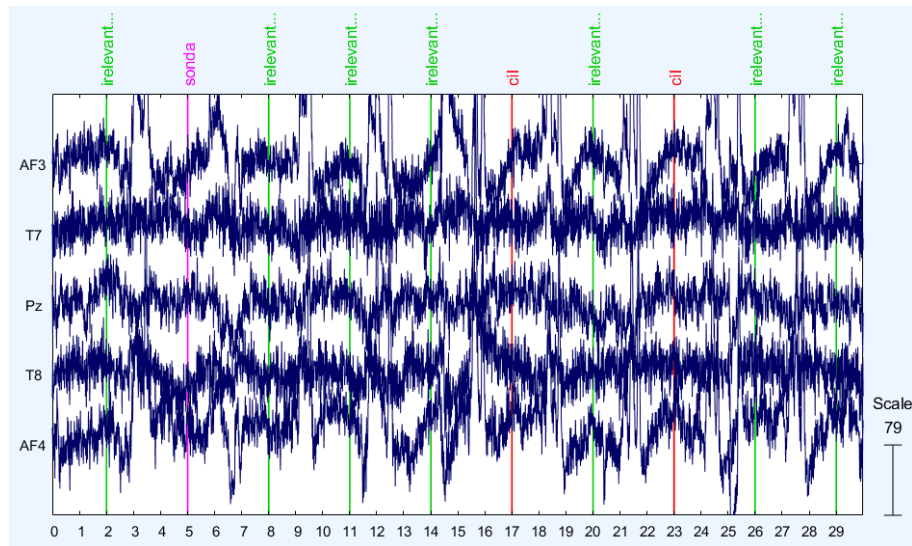
Obr. 19.: Nastavení filtrace pomocí Notch filtru v ERPLAB

Tvorba eventů

Dále byly importovány informace o jednotlivých eventech (sonda, cíl, irrelevantní), které se používají pro označení času nástupu jednotlivých stimulů a pomáhají při analýze EEG dat rozpoznat ERP komponenty očekávané jako reakce na konkrétní stimul. Soubor obsahující eventy musí mít konkrétní strukturu, viz PŘÍLOHA P VI: SEZNAM EVENTŮ. Postup v EEGLAB: File → Import event info → From MATLAB array or ASCII file → Ok.



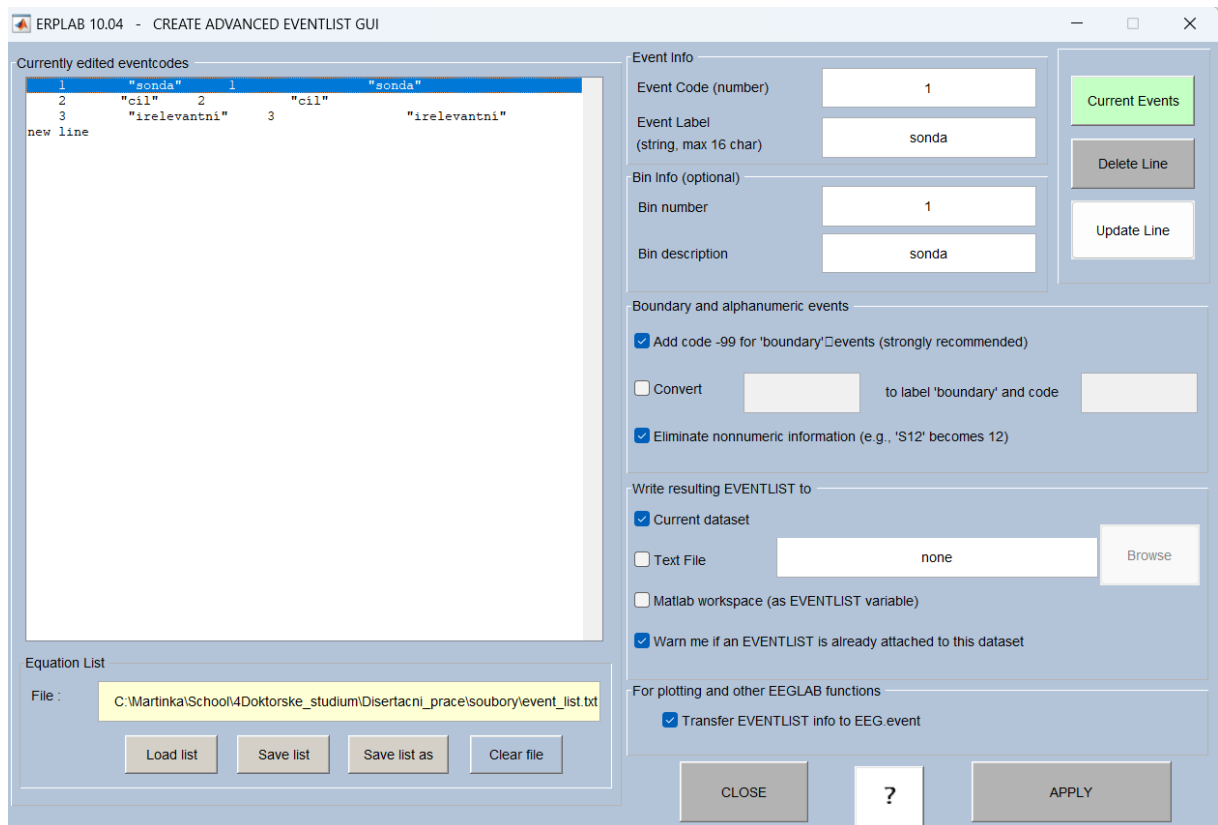
Obr. 20.: Import informací o eventech v EEGLAB



Obr. 21.: Rozdělení na jednotlivé eventy

Vytvoření seznamu událostí

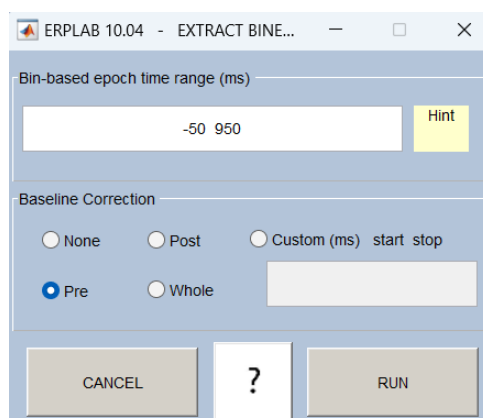
Následně byl vytvořen seznam událostí pro konkrétní eventy. Kontinuální data EEG byla rozdělena do tří kategorií dle sondy, cíle a irelevantních stimulů. Soubor musí obsahovat specifickou strukturu, viz PŘÍLOHA P VII: SEZNAM UDÁLOSTÍ. Postup v EEGLAB: ERPLAB → EventList → Create EEG EVENTLIST → Continue → Advanced → Load list → „výběr souboru event_list.txt“ → Apply → Ok.



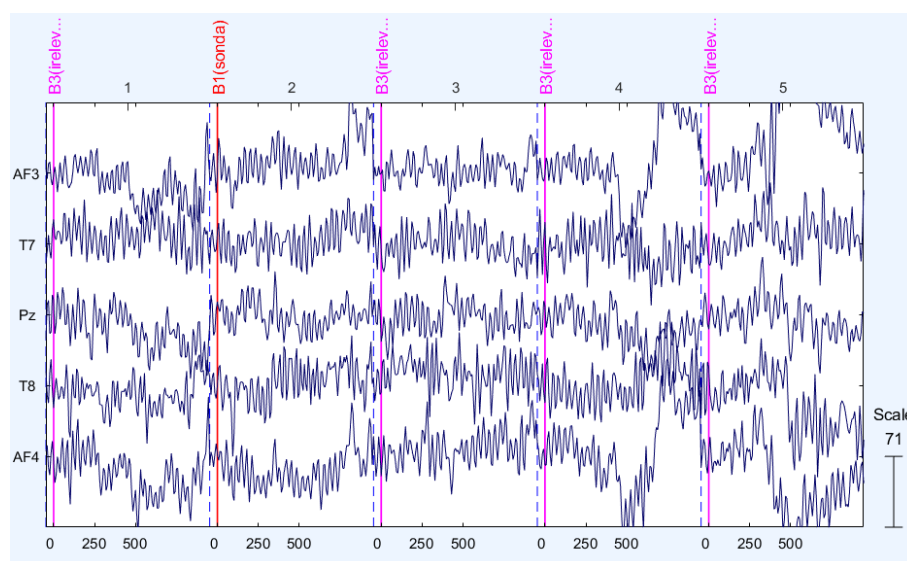
Obr. 22.: Vytvoření seznamu událostí v ERPLAB

Tvorba epoch

Dalším krokem bylo rozdělení dat do jednotlivých epoch, které umožňují sledování časové události od okamžiku před začátkem stimulu, během něj až po jeho ukončení. Každá epocha zahrnuje jeden segment obsahující konkrétní stimul. EEG data byla rozdělena do epoch trvajících 1000 ms, a to podle typu stimulu v segmentu. Epochy začínají 50 ms před kritickou položkou jako základní linie a končí 950 ms po výskytu stimulu. Bylo tedy vytvořeno časové okno s dobou trvání od -50 do 950 ms se vzorkovací frekvencí 128 Hz. Postup v EEGLAB: ERPLAB → Extract bin-based epochs → RUN.



Obr. 23.: Tvorba epoch v ERPLAB



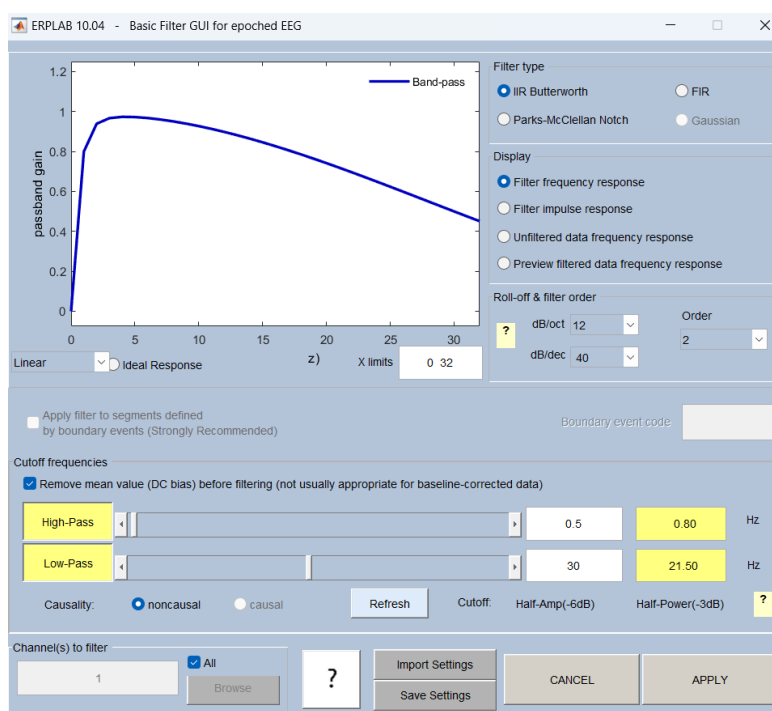
Obr. 24.: Rozdělení na jednotlivé epochy

Výběr kanálu

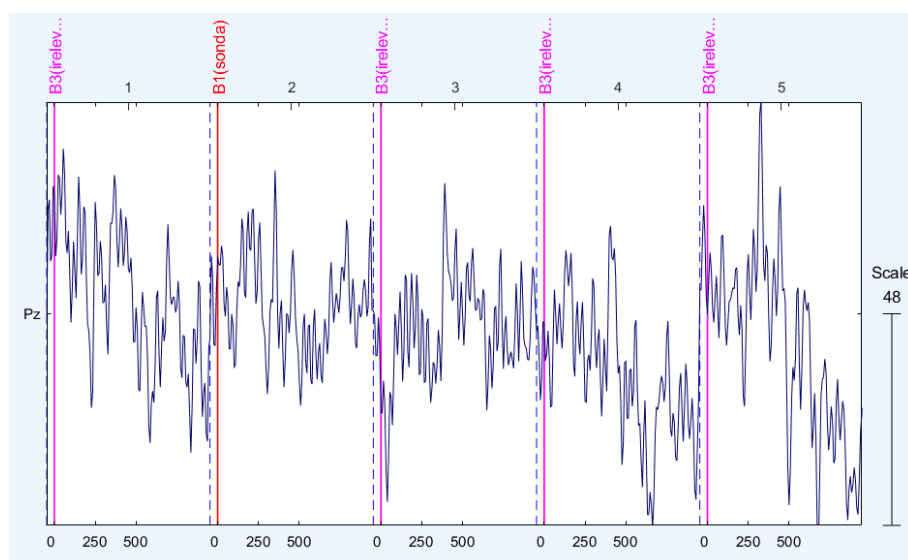
Výběr kanálů obsahujících relevantní informace může významně zvýšit přesnost klasifikace. Nejčastěji používané kanály jsou středové elektrody Fz, Pz a Cz, přičemž nejvyšší amplituda složky P300 byla zaznamenána na elektrodě Pz [11, 17, 18]. Kanál Pz byl vybrán pro analýzu kvůli své lokalizaci nad parietální částí mozku, která je spojena s kognitivními funkcemi a zpracováním informací. Postup v EEGLAB: Edit → Select data → Channel(s) → elektroda Pz → OK.

Filtrace (IIR Butterworth)

Pro odstranění šumu a artefaktů se nejčastěji využívá metoda Band-Pass Filter (BPF) [4-13, 15]. Užitečná složka EEG signálu je v rozmezí 0,5 Hz až 30 Hz [4, 7, 9-11], ostatní frekvence jsou považovány za šum a jsou eliminovány. Tento frekvenční rozsah je standardně využíván ve studiích zaměřených na P300 během mentálních úkolů [8, 10, 12]. BPF umožňuje propustit pouze určitý rozsah frekvencí bez snížení kvality signálu. Následná filtrace dat tedy byla provedena pomocí IIR Butterworth filtru v rozsahu 0,5-30 Hz, který odstraňuje neúčinná pásma a artefakty okolního prostředí. Postup v EEGLAB: ERPLAB → Filter & Frequency Tools → Filters for EEG data → IIR Butterworth → Apply.



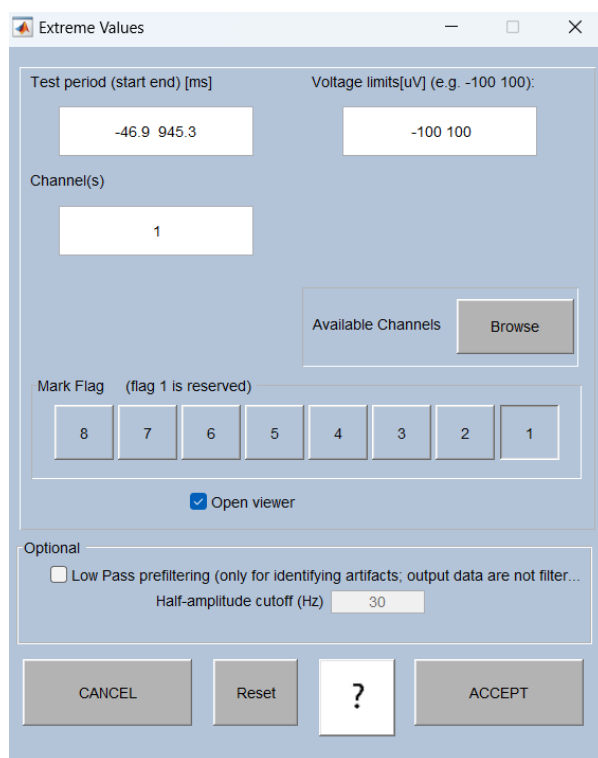
Obr. 25.: Nastavení filtrace pomocí IIR Butterworth filtru v ERPLAB



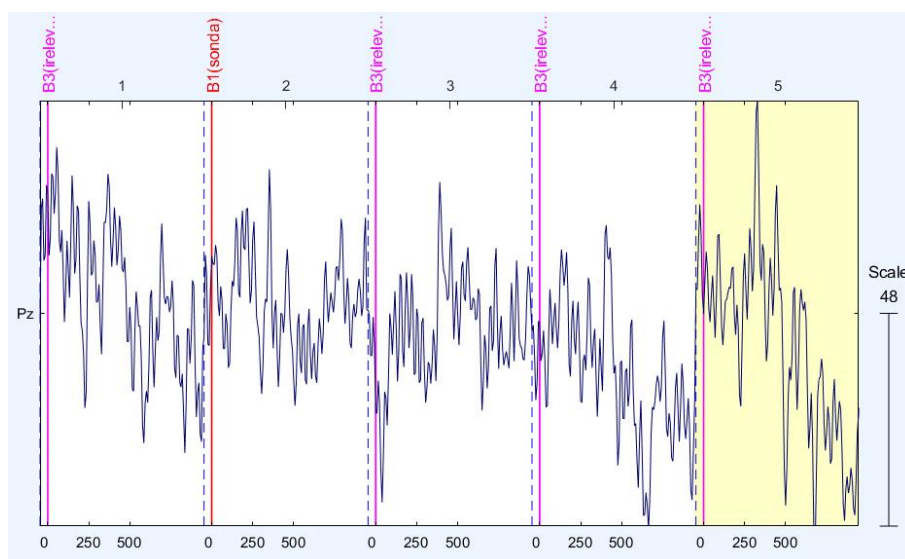
Obr. 26.: Filtrace pomocí IIR Butterworth filtru

Odstranění artefaktů

Po filtraci byla provedena detekce a odstranění artefaktů pomocí vestavěných funkcí v EEGLAB. Kritérium pro odstranění artefaktů bylo nastaveno na $\pm 100 \mu\text{V}$ pro jednotlivé epochy [39]. Pokusy přesahující tuto hodnotu byly definovány jako artefakty a automaticky zamítnuty. Následně byla provedena manuální kontrola, aby se ověřilo, že zamítnuté pokusy byly přesně detekovány. Postup v EEGLAB: ERPLAB → Artifact detection in epochs data → Simple voltage threshold → Accept → Update Marks → Ok → EEGLAB → Plot → Channel data (scroll) → Reject → Yes.



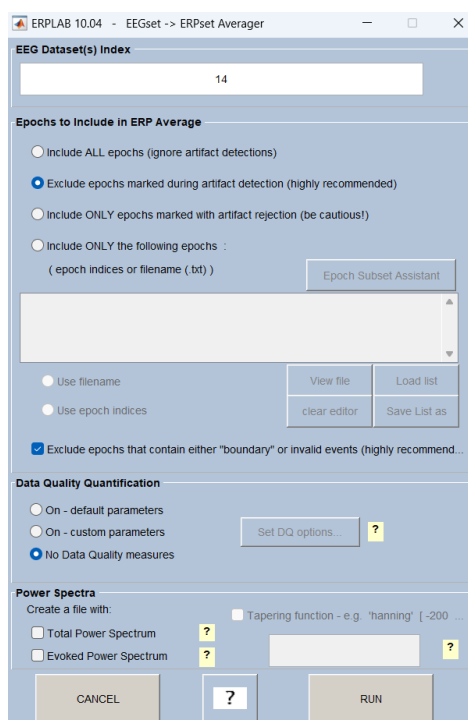
Obr. 27.: Nastavení odstranění artefaktů v ERPLAB



Obr. 28.: Odstranění artefaktů

Průměrování ERP

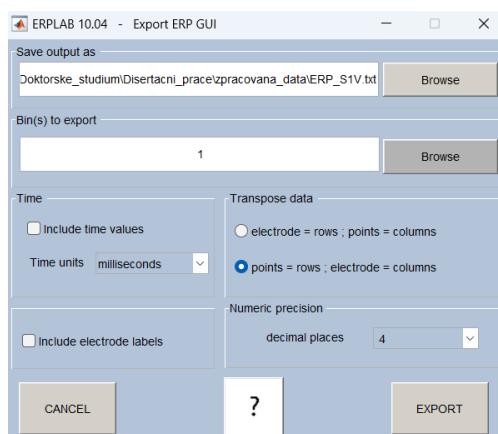
Závěrečným krokem bylo vytvoření zprůměrovaného ERP datasetu. EEG data obsahují mnoho šumu kvůli různým procesům probíhajícím v mozku, proto je vhodné provést více měření a data následně zprůměrovat [49]. Klíčové je srovnání mezi sondou a irelevantními ERP. Výsledkem předzpracování EEG signálů jsou průměrné hodnoty ERP pro sondy i irelevantní podněty na elektrodě Pz u každého účastníka, čímž se získává průběh vln P300 pro další analýzu. Postup v EEGLAB: ERPLAB → Compute averaged ERPs → RUN → Yes → Ok.



Obr. 29.: Průměrování signálu v ERPLAB

Export dat

Pro kontrolu výsledných dat byl proveden export dat do souboru CSV pro jednotlivé podněty. Postup v EEGLAB: ERPLAB → Export & Import ERP → Export ERP to text (universal) → Export.



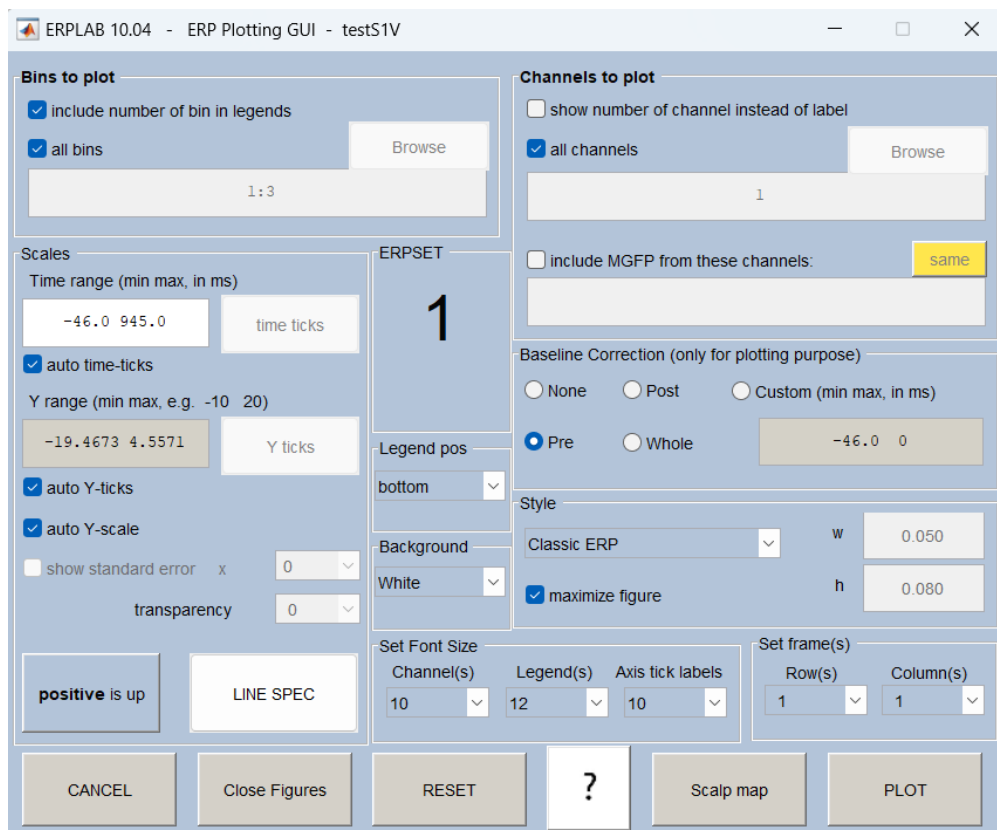
Obr. 30.: Export dat z ERPLAB

5.2.1 Vizuální kontrola dat

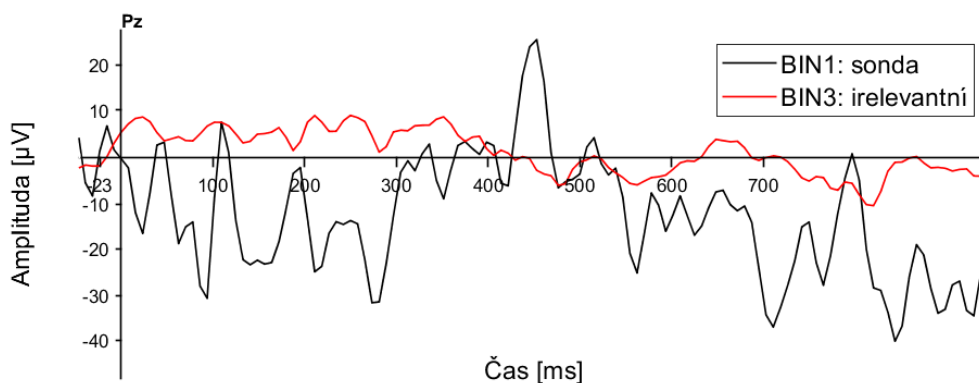
Vizuální kontrola předzpracovaných EEG dat je důležitým krokem pro ověření kvality signálů a správnosti jednotlivých kroků předzpracování. Umožňuje identifikovat a odstranit případné chyby nebo artefakty, které by mohly ovlivnit výsledky analýzy a pomáhá identifikovat rozdíly v EEG signálech mezi vinnými a nevinnými subjekty, což je klíčové pro přesnou klasifikaci.

Průběh ERP

Pro vizuální kontrolu předzpracovaných dat je možné využít ERPLAB k vykreslení jednotlivých grafů. Postup v EEGLAB: ERPLAB → Plot ERP → Plot ERP waveforms → Plot.



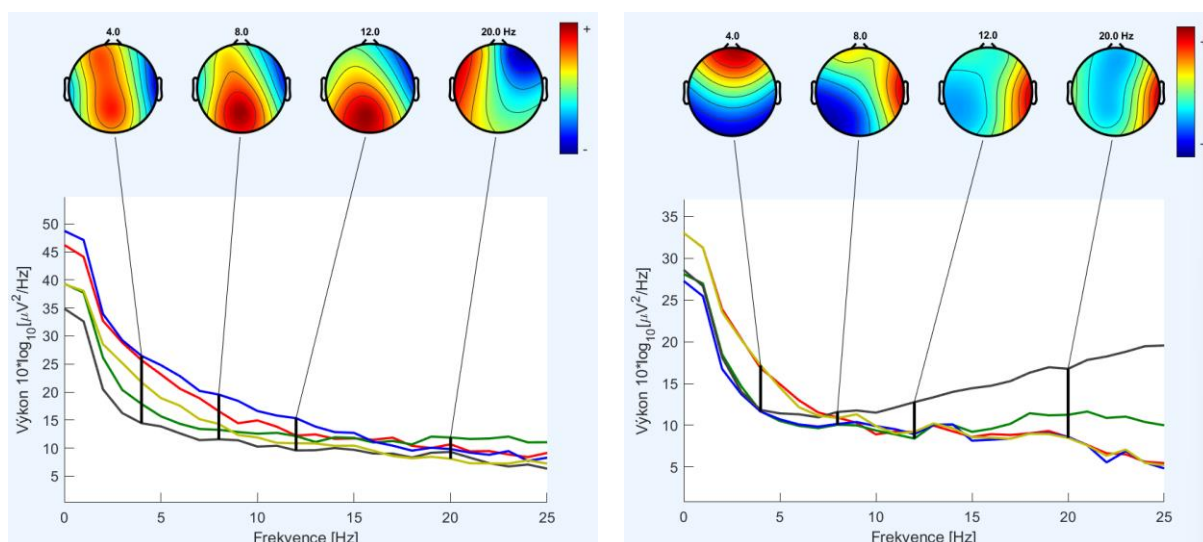
Obr. 31.: Nastavení parametrů pro vykreslení ERP dat v ERPLAB



Obr. 32.: Průběh vlny P300 na elektrodě Pz pro jednotlivé podněty

Výkonové spektrum

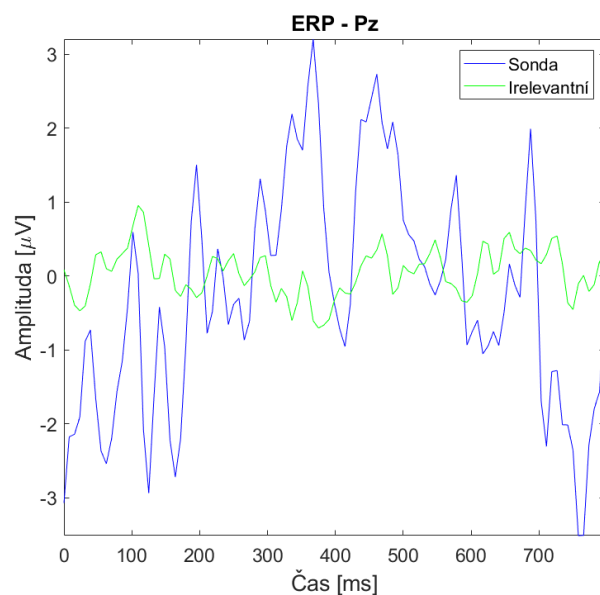
Power Spectral Density (PSD) je spolehlivou metodou pro identifikaci funkčnosti mozku a koherence mezi různými oblastmi mozku ze signálů EEG [4, 7, 9]. Pro analýzu rozdílů mezi signály generovanými vinnými a nevinnými subjekty byla vykreslena jejich výkonová spektra. Obr. 33 ukazuje PSD u vinného vs nevinného subjektu. Každý EEG kanál zobrazuje specifické hodnoty výkonu během provádění různých úkolů. Rozdíly mezi vinnými a nevinnými subjekty jsou patrné. Frekvence pro analýzu PSD zahrnují delta (4 Hz), théta (8 Hz), alfa (12 Hz) a beta (20 Hz). U vinného subjektu lze pozorovat rozložení výkonu převážně v parietálním laloku, což naznačuje vyšší mozkovou aktivitu. Toto zvýšení aktivity koreluje s vyšším výskytem P300 reakcí. U nevinných subjektů jsou naopak pozitivní hodnoty výkonu viditelné ve frontální oblasti mozku, často související s pohybem očí. Intenzita červené barvy u nevinných subjektů je menší než u vinných subjektů, což naznačuje nižší mozkovou aktivitu. Postup v EEGLAB: Plot → Channel spectra and maps → Ok.



Obr. 33.: PSD pro vinný (vlevo) vs nevinný (vpravo) subjekt

Průměr ERP všech subjektů

Byl vykreslen průměrný ERP pro dva typy stimulů na elektrodě Pz pro všechny subjekty (viz Obr. 34). Komponenta P300 představuje hlavní analytickou funkci, která přispívá k rozboru signálu a statistickým výpočtům, což umožňuje přesné určení přítomnosti či nepřítomnosti informací. V procesu předzpracování byly použity specifické kroky k získání průběhu zprůměrovaných vln ERP pro sondy a irelevantní podněty na elektrodě Pz. Tyto kroky zahrnovaly filtraci, tvorbu eventu a seznamu událostí, odstranění artefaktů, výběr časového okna a kanálu, a následné zprůměrování signálu.



Obr. 34.: Průměr ERP na elektrodě Pz pro všechny subjekty

Vizuální kontrola jednotlivých grafů potvrzuje rozdíly v reakci mozku na podněty sond a irelevantní podněty. Pro přesnou a spolehlivou detekci lži pomocí EEG je nezbytné provádět důkladné předzpracování dat. Tento proces zajišťuje, že výsledná data jsou čistá a kvalitní, což je zásadní pro následnou analýzu a interpretaci signálů EEG.

Automatizace procesu

Vzhledem k nutnosti předzpracovat velké množství dat byla využita automatizace pomocí naprogramovaných funkcí a skriptů vygenerovaných z historie EEGLAB. Tento kód byl následně použit ve výsledném programu pro jednotlivé části předzpracování, což výrazně urychlilo a zefektivnilo celý proces (viz podkapitola 6.1.2).

5.3 Extrakce vlastností

Tato část se zaměřuje na různé přístupy k extrakci vlastností sloužících k identifikaci složitých vzorců mozkových vln. Výběr užitečného signálu se provádí pomocí množiny parametrů, které se následně využívají ke klasifikaci. EEG signál je nestacionární a jeho frekvence se mění v závislosti na čase. Při detekci lži se frekvence mění při reakci subjektu na konkrétní podnět v určitém čase. Nezpracované EEG signály jsou zaznamenávány v časové oblasti, tedy jako funkce času. Pouze časová informace však nestačí k analýze nestacionárního signálu EEG, a proto se k získání užitečných informací používají různé techniky extrakce vlastností [2, 4, 10, 11, 17]. V předchozích studiích CIT založených na ERP P300 pro detekci lži byly pro analýzu EEG signálů použity různé metody extrakce vlastností založené na časové, frekvenční či časově-frekvenční doméně. Tyto metody a jejich kombinace se využívají ke zvýšení přesnosti a výkonu systému [P.1].

Časová doména

Mezi nejčastěji využívané vlastnosti v časové oblasti patří maximální a minimální amplituda [4, 6, 9-11, 13], statistické parametry Hjorth (aktivita, mobilita a složitost) [4, 6, 8-10], průměr [6, 9, 11], směrodatná odchylka [6, 9], výkon [4, 10], entropie [9], šikmost [6], špičatost [6] a energie [6]. Vlastnosti časové domény účinně zachycují variace ve tvaru a načasování křivek generovaných v mozku v důsledku vnějších událostí. Pomocí časových vlastností lze efektivně porozumět různým charakteristikám tvaru vlny komponenty P300 [4]. Špičkové hodnoty vln EEG byly v mnoha výzkumných pracích považovány za měřítko identifikace lži. Samotný vrchol signálu však nestačí k vysvětlení různých charakteristik ERP odpovědí generovaných při testování podvodu, a proto se využívají různé kombinace vlastností [10].

Frekvenční doména

Nejčastěji využívanými metodami ve frekvenční oblasti jsou Fast Fourier Transform (FFT) [4, 10] a PSD [9, 11]. FFT rozděluje spektra EEG do jednotlivých frekvenčních pásem transformací z časové do frekvenční domény, ze kterých se určuje celkové množství vln a nejvíce zastoupená frekvence v záznamu [17].

Časově-frekvenční doména

Mezi nejčastější metody v časově-frekvenční doméně patří Wavelet Transform (WT) [4, 5, 10-12, 15, 47], konkrétně Discrete Wavelet Transform (DWT). WT je široce využívána jako přístup k extrakci vlastností porovnávajících známé a neznámé tváře a mezi různými přístupy poskytuje nejlepší výkon [4, 10]. Extrahovat užitečné a relevantní informace je velmi náročný úkol [8]. Vzhledem k tomu, že se elektrody vzájemně překrývají a jsou těsně koncentrovány, je časově náročné analyzovat variace různých frekvencí. Proto se analýza signálů často provádí pomocí WT k separaci jednotlivých frekvenčních pásem [9]. Informace o nervových aktivitách lze odvodit pomocí analýzy v různých frekvenčních pásmech EEG na konkrétních místech pokožky hlavy. WT poskytuje informace o signálech překrývajících se v časové oblasti a současně pomáhá porozumět změnám charakteristik signálů s časem ve frekvenční oblasti [17]. Vzorce nervových složek, které se překrývají v čase, jsou poté přesně detekovány pomocí WT [9].

Vědci se v poslední době zaměřili také na kombinaci metod, přístupů a algoritmů pro dosažení vyšší přesnosti klasifikace a funkčnosti rozpoznávání skrytých informací [P.1]. Analýza dat zde byla provedena v časové oblasti na datech ERP a také v časově-frekvenční oblasti pomocí DWT. Podrobnější informace jsou uvedeny v podkapitole 5.5.1.

5.3.1 Statistické vlastnosti

Pro analýzu P300 v reakci na známé obličejce jsou vhodné různé charakteristiky signálu. Pro komplexnější prozkoumání byly vybrány různé statistické vlastnosti, které poskytují podrobný popis charakteristik a další pohled na distribuci a složitost signálu, konkrétně maximum, minimum, peak-to-peak, průměr, směrodatná odchylka, medián, šikmost, špičatost, entropie, parametry Hjorth (aktivita, mobilita, složitost), energie a výkon [6, 9, 10]. Tyto statistické vlastnosti byly vypočteny pro podněty sond a irelevantní podněty na datech ERP v časové oblasti a na koeficientech DWT v jednotlivých vlnových pásmech v časově-frekvenční oblasti pro všechny subjekty. Tyto vlastnosti slouží jako charakteristiky pro klasifikaci. Níže jsou uvedeny nejčastěji využívané statistické vlastnosti v této oblasti, které byly použity i v této práci:

Maximum

- Nejvyšší hodnota amplitudy signálu v průběhu časové osy.

$$\text{Maximum}(x) = \max(x_i), \quad (5.1)$$

kde x_i jsou jednotlivé hodnoty signálu.

Minimum

- Nejnižší hodnota amplitudy signálu v průběhu časové osy.

$$\text{Minimum}(x) = \min(x_i), \quad (5.2)$$

kde x_i jsou jednotlivé hodnoty signálu.

Peak-to-Peak (P2P)

- Rozdíl mezi maximální a minimální hodnotou signálu v průběhu časové osy, indikující celkovou amplitudu změny signálu.

$$P2P(x) = \max(x_i) - \min(x_i), \quad (5.3)$$

kde x_i jsou jednotlivé hodnoty signálu.

Průměr (Mean)

- Průměrná hodnota amplitudy signálu v daném časovém okamžiku přes všechna opakování události. Udává střední hodnotu amplitudy signálu a může odhalit celkové rozdíly v aktivitě mozku.

$$\text{Mean}(x) = \left(\frac{1}{N}\right) \sum_{i=1}^N x_i, \quad (5.4)$$

kde x_i jsou jednotlivé hodnoty signálu a N je počet vzorků.

Směrodatná odchylka (Std)

- Míra variability nebo rozptylu amplitudy signálu v daném časovém okamžiku. Ukazuje, jak se amplitudy liší od průměru, což může indikovat variabilitu signálu.

$$Std(x) = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - Mean(x))^2}, \quad (5.5)$$

kde x_i jsou jednotlivé hodnoty signálu, N je počet vzorků a $Mean(x)$ je průměrná hodnota signálu.

Medián (Median)

- Medián představuje střední hodnotu signálu v průběhu časové osy, která rozděluje setříděný signál na dvě stejně velké části. Je odolný vůči extrémním hodnotám, což může lépe reprezentovat centrální tendenci dat.

$$Median(x) = \begin{cases} x_{\left(\frac{N+1}{2}\right)}, & \text{pokud je } N \text{ liché} \\ \frac{x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}}{2}, & \text{pokud je } N \text{ sudé,} \end{cases} \quad (5.6)$$

kde x_i jsou jednotlivé hodnoty signálu a N je počet vzorků.

Šikmost (Skewness)

- Šikmost je míra asymetrie rozdělení signálu kolem jeho střední hodnoty.

$$Skewness(x) = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - Mean(x)}{Std(x)} \right)^3, \quad (5.7)$$

kde x_i jsou jednotlivé hodnoty signálu, N je počet vzorků, $Mean(x)$ je průměrná hodnota signálu a $Std(x)$ je směrodatná odchylka signálu.

Špičatost (Kurtosis)

- Špičatost je míra toho, jak moc jsou extrémní hodnoty signálu koncentrovány kolem jeho střední hodnoty.

$$Kurtosis(x) = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - Mean(x)}{Std(x)} \right)^4 - 3, \quad (5.8)$$

kde x_i jsou jednotlivé hodnoty signálu, N je počet vzorků, $Mean(x)$ je průměrná hodnota signálu a $Std(x)$ je směrodatná odchylka signálu.

Entropie (Entropy)

- Entropie měří náhodnost nebo složitost signálu. Normalizuje signál a je získána na základě pravděpodobnostního rozložení hodnot signálu.

$$Entropy(x) = - \sum p_i \log_2(p_i), \quad (5.9)$$

kde p_i představuje pravděpodobnost jednotlivých hodnot signálu.

Parametry Hjorth (aktivita, mobilita, složitost)

Parametry Hjorth jsou tři statistické ukazatele používané k charakterizaci časových řad, jako je EEG signál. Tyto parametry jsou:

Aktivita (Activity)

- Aktivita je rozptyl signálu, tedy míra intenzity signálu. Udává, jak výrazně se signál mění v průběhu času.

$$Activity = var(x), \quad (5.10)$$

kde x je signál.

Mobilita (Mobility)

- Mobilita je poměr rozptylu první derivace signálu k rozptylu samotného signálu. Udává, jak rychle se signál mění.

$$Mobility = \sqrt{\frac{var\left(\frac{dx}{dt}\right)}{var(x)}}, \quad (5.11)$$

kde $\frac{dx}{dt}$ je první derivace signálu x .

Složitost (Complexity)

- Složitost je poměr rozptylu druhé derivace signálu k rozptylu první derivace signálu, normalizovaný podle mobility. Udává, jak složitá je struktura signálu.

$$Complexity = \frac{\sqrt{var\left(\frac{d^2x}{dt^2}\right)}}{Mobility}, \quad (5.12)$$

kde $\frac{d^2x}{dt^2}$ je druhá derivace signálu x .

Energie (Energy)

- Energie udává celkové množství signálu v daném časovém intervalu.

$$Energy = \sum x^2, \quad (5.13)$$

kde x je signál.

Výkon (Power)

- Výkon signálu je průměrná energie signálu normalizovaná na jednotku času (počet vzorků).

$$Power = \frac{1}{N} \sum x^2, \quad (5.14)$$

kde x je signál a N je počet vzorků.

Kombinace těchto různých vlastností umožňuje získat komplexní pohled na charakteristiky signálů, což může vést k vyšší přesnosti a spolehlivosti detekčních systémů.

5.3.2 Discrete Wavelet Transform

DWT je jednou z dominantních technik pro analýzu časově proměnných charakteristik nestacionárních signálů. Analyzuje časově-frekvenční oblasti, což z ní činí ideální metodu pro extrakci vlastností, jelikož podstatné rysy mohou být přítomny v časové, frekvenční nebo obou doménách. DWT výrazně snižuje množství dat potřebných pro analýzu tím, že rozkládá signál na aproximační a detailní koeficienty pomocí různých filtrů [9, 49]. DWT rozděluje signál na nízkofrekvenční (dolní propust) a vysokofrekvenční (horní propust) složky do subpásem stejné šířky v první úrovni. V následujících úrovních jsou do dílčích pásem rozděleny pouze nízkofrekvenční složky [49]. Byl proveden víceúrovňový jednorozměrný vlnkový rozklad, kde na každé úrovni je rozklad charakterizován operátorem dolní propusti, poskytujícím aproximační koeficienty, a horní propusti, poskytujícím koeficienty detailů [4]. DWT tedy rozkládá signál na různá frekvenční pásma pomocí filtrů a vrací aproximační a detailní koeficienty v jednotlivých úrovních, které slouží pro výpočet vlastností určených pro vstup do klasifikátoru [9, 40].

- Koeficienty aproximace:

$$A_{j+1}[n] = \sum_k g[k] \cdot A_j[2n + k], \quad (5.15)$$

- Koeficienty detailu:

$$D_{j+1}[n] = \sum_k h[k] \cdot A_j[2n + k], \quad (5.16)$$

kde $g[k]$ a $h[k]$ jsou koeficienty nízkopropustného a vysokopropustného filtru. A_j je koeficient aproximace na úrovni j a D_j je koeficient detailu na úrovni j .

Výběr frekvenčních pásem

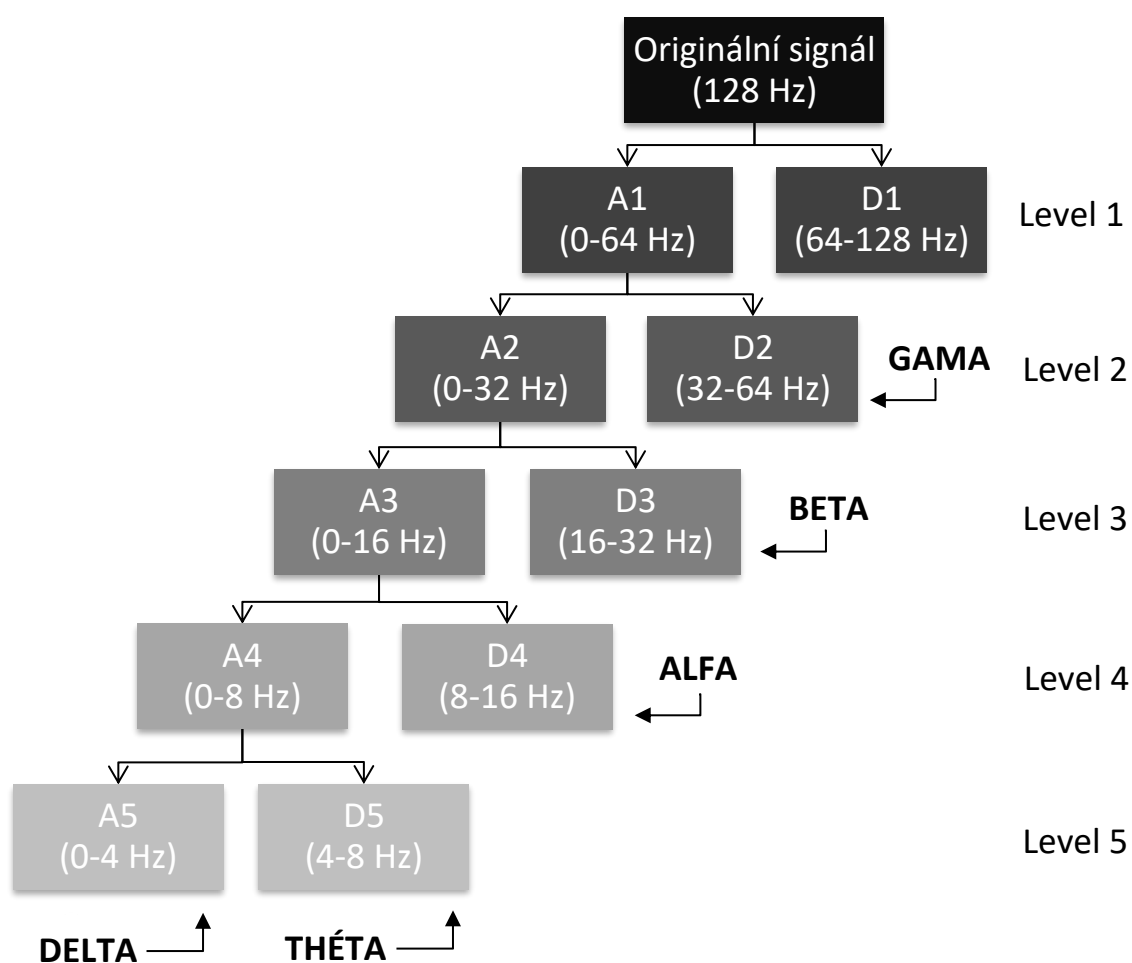
Nejinformativnější frekvenční pásma pro analýzu EEG signálů v kontextu P300 jsou alfa, beta, delta a théta. Frekvenční pásmo gama, přesahující 30 Hz, bylo eliminováno jako artefakt. Delta vlny (0,5-4 Hz) jsou klíčové, protože významně korelují s amplitudou P300, která je spojena s procesy hodnocení pozornosti a podnětů. Tato korelace činí delta vlny zvláště informativními pro úkoly zahrnující rozpoznávání známých podnětů. Théta vlny (4-8 Hz) jsou rovněž vysoce relevantní, protože jsou propojeny s procesy kódování a vyhledávání paměti, což je zásadní pro rozlišení mezi známými a neznámými tvářemi [9, 11, 46]. Alfa vlny (8-12 Hz) a beta vlny (12-30 Hz) hrají roli v kontrole pozornosti a sensoricko-motorickém zpracování. Alfa vlny se podílejí na inhibici nepodstatného zpracování pro usnadnění plnění úkolů, zatímco beta vlny jsou spojeny s bdělostí a senzomotorickými funkcemi [9, 11, 45, 46, 49].

Aplikace DWT na EEG signály

Pro efektivní analýzu EEG je nutná znalost časově-frekvenční oblasti signálu. Při detekci lži se frekvence mění, když subjekt reaguje na konkrétní podnět v daném čase. Proto byla využita DWT na odezvy ERP generované během záznamu [12]. Důležitým kritériem při použití DWT je volba počtu úrovní rozkladu signálu. Každý pokus byl rozložen do několika skupin koeficientů spojených s různými frekvenčními pásmy [11]. Každá úroveň rozkladu obsahuje aproximační a detailní koeficienty poskytující informace o frekvenčních složkách na různých úrovních. Aproximační koeficienty poskytují hrubou představu o celkovém trendu signálu na dané úrovni rozkladu, zatímco detailní koeficienty obsahují informace o jemných změnách a vysokofrekvenčních složkách signálu. Koeficienty jsou získány buď posunutím, nebo škálováním jedné funkce známé jako mateřská vlnka. Při provedení konvoluce vlnky se signálem v pravidelných intervalech výsledky napovídají, kdy byla vlnka použita. K vysokému korelačnímu koeficientu vlnky se signálem dochází v těch intervalech, kde signál obsahuje informaci o podobné frekvenci, tedy tam, kde je zvolené vlnce nejpodobnější.

DWT bylo použito k převodu získaného EEG signálu se vzorkovací frekvencí 128 Hz na reprezentaci v časově-frekvenční doméně pro zachycení hlavních frekvenčních pásem EEG relevantních pro analýzu P300. Po aplikaci DWT jsou signály rozděleny do jednotlivých frekvenčních pásem EEG (alfa, beta, delta, théta, gama + zbytek frekvence), poskytující aproximační a detailní koeficienty. V tomto případě byl proveden rozklad páté úrovně pomocí vlnkového filtru Daubechies čtvrtého řádu (db4), protože jeho vlastnosti připomínají odezvy ERP [4]. Mezní frekvence dolní a horní propusti se rovná polovině vzorkovací

frekvence. Při vzorkovací frekvenci 128 Hz jsou na první úrovni pomocí dolní propusti získávány signály ve frekvenčním rozsahu 0-64 Hz, a při aplikaci horní propusti se získá frekvenční rozsah 64-128 Hz. DWT prochází vstupní signál současně dolní a horní propustí a výstupy se na první úrovni nazývají koeficienty aproximace (A1) a detailu (D1). Stejný postup se opakuje pro získání aproximačních a detailních koeficientů další úrovně. V každém kroku rozkladu se frekvenční rozlišení zdvojnásobí filtrací a časové rozlišení se sníží na polovinu pomocí decimace. A1 až A5 představují aproximační koeficienty a D1 až D5 jsou koeficienty detailů. Tím byly získány aproximační a detailní koeficienty pro frekvenční pásma alfa (D4), beta (D3), delta (A5) a théta (D5). Z výsledných koeficientů byly poté vypočteny různé statistické vlastnosti pro podněty sond a irelevantních podnětů v jednotlivých vlnových pásmech získaných z DWT.



Obr. 35.: Rozklad signálu pomocí DWT pro 128 Hz

Tvorba vektoru

Identifikované rysy z mozkových vln generovaných v reakci na jednotlivé podněty byly primárně odvozeny z elektrody Pz umístěné v parietálním laloku. Po výpočtu průměrných ERP v časové oblasti a vlnkových koeficientů pomocí DWT v časově-frekvenční oblasti byly na každou dílčí složku aplikovány různé

statistické metody pro každé frekvenční pásmo. Tyto vlastnosti reprezentovaly hlavní charakteristiky signálu při závěrečné klasifikaci dat v kontextu detekce skrytých informací pomocí EEG, zaměřených na identifikaci složky ERP P300. Výsledné vlastnosti byly sloučeny do jediného vektoru, včetně označení tříd. Nakonec byla získána matice vlastností obsahující sedmdesát charakteristik a jeden sloupec třídy pro jednotlivé podněty u každého subjektu, která sloužila jako vstup do klasifikačního algoritmu.

Automatizace procesu

EEG data byla analyzována jak pomocí vestavěných funkcí, tak i pomocí vlastních naprogramovaných funkcí v Matlabu. Z důvodu nutnosti zpracování velkého množství dat byla použita automatizace (viz podkapitola 6.1.2).

5.4 Výběr vlastností

Výběr vhodných vlastností je klíčovým krokem pro detekci lži pomocí EEG signálů při analýze odpovědi ERP P300 v reakci na rozpoznání obličeje před provedením klasifikace. Cílem je identifikovat optimální sadu vlastností, která umožní sestavit efektivní model ze shromážděných dat [39]. Některé studie využívají ke klasifikaci EEG signálů kombinaci různých vlastností [17]. V této studii byly analýzy aspektů EEG provedeny jak v časové oblasti na datech ERP, tak i v časově-frekvenční doméně pomocí DWT. Cílem bylo porovnat EEG odpovědi na známé a neznámé tváře u jednotlivých účastníků. Výběr relevantních vlastností z těchto dat je zásadní jak pro zlepšení výpočetní efektivity, tak pro zvýšení použitelnosti systémů založených na EEG v reálných aplikacích [9]. Zkracuje dobu běhu celého procesu, protože redukuje počet prvků a následně i složitost prostoru prvků. Tím se zvyšuje efektivita výpočetních algoritmů a zlepšuje se interpretace výsledků.

Různé metody výběru vlastností nabízejí různé přístupy k identifikaci nejvíce informativních charakteristik z EEG signálů. Mezi často používané metody výběru vlastností v této oblasti patří genetický algoritmus (GA) [10, 11], Binary BAT (BBA) [4, 6] a Analysis of Variance (ANOVA) [45]. Další možností je použití Principal Component Analysis (PCA) v Classification Learneru.

Genetický algoritmus

GA, inspirované přírodním výběrem, mají schopnost prohledávat rozsáhlý a komplexní prostor vlastností pomocí evolučních principů, což může vést k nalezení optimálních vlastností. Tento algoritmus iterativně zlepšuje populaci kandidátních řešení pomocí operací, jako jsou selekce, křížení a mutace. GA jsou flexibilní a mohou být přizpůsobeny různým typům dat a kritériím optimalizace. Nicméně vyžadují značné výpočetní zdroje a čas, zejména při práci s velkými datovými sadami EEG. Výsledky mohou být závislé na zvolených parametrech algoritmu GA, což může vyžadovat experimentování a ladění [10, 11].

Analysis of Variance

ANOVA je statistická metoda, která identifikuje statisticky významné rozdíly mezi třídami v rámci jednotlivých vlastností a poskytuje kvantitativní míru vlivu každé vlastnosti na rozdělení ERP P300 signálů. Umožňuje snadno identifikovat vlastnosti s významnými rozdíly mezi skupinami, což je užitečné pro zúžení výběru vlastností. Nicméně nezahrnuje interakce mezi vlastnostmi a je omezena na lineární vztahy, což znamená, že neidentifikuje složitější, nelineární interakce. To může vést k přehlížení důležitých kombinací vlastností a následně ke snížení výkonu klasifikačního modelu [45].

Principal Component Analysis

PCA slouží k redukci dimenzionality dat před trénováním klasifikačního modelu. Transformuje původní atributy na nový soubor lineárně nekorelovaných proměnných, nazývaných hlavní komponenty, které jsou uspořádány dle klesající variance. To umožňuje zachovat co nejvíce informací obsažených v datech při použití menšího počtu komponent. Tato transformace však může vést ke ztrátě interpretovatelnosti hlavních komponent a předpokládá lineární závislost mezi proměnnými, což může být omezením v případech, kdy jsou data nelineární [47].

Shrnutí metod pro výběr vlastností

Tyto metody nabízejí různé přístupy k výběru vlastností, každá s vlastními výhodami a nevýhodami. Výběr vhodné metody závisí na specifických potřebách studie, charakteristikách dat a dostupných výpočetních zdrojích. Pro analýzu ERP P300 je vhodné zvolit specifické statistické charakteristiky a frekvenční pásma, která jsou nejvíce informativní pro tento úkol. Pásma zájmu v této studii jsou alfa, beta, delta a théta, z nichž každé odráží různé mozkové stavy subjektu. Delta pásmo je obvykle spojováno s hlubokým spánkem, ale jeho změny mohou odrážet reakce na podněty. Théta pásmo je často spojováno s kognitivními procesy a může být užitečné při analýze ERP. Alfa pásmo je spojováno s klidovým stavem mozku, ale změny mohou indikovat reakce na podněty. Beta pásmo je spojováno s aktivním myšlením a může být relevantní při analýze reakce na známé obličej.

Byly vyzkoušeny různé vlastnosti a jejich kombinace pro vstup do různých klasifikačních algoritmů, včetně různých metod pro výběr vlastností (viz podkapitola 5.5.1). Zkoumány byly různé statistické vlastnosti vypočítané na datech ERP a DWT v jednotlivých pásmech. Tyto vlastnosti byly kombinovány a byla ověřována funkčnost těchto kombinací na jednotlivých algoritmech. Pro vyhodnocení výkonu vlastností byly použity tři metody: GA, ANOVA a PCA. Nejlepších výsledků dosáhl GA, který byl poté použit pro tvorbu finálního modelu.

5.4.1 Genetický algoritmus

GA byl zvolen pro výběr optimální sady vlastností vzhledem k jeho vyšší efektivitě oproti tradičním metodám. Díky své schopnosti prozkoumávat velký prostor možností a nalézat nečekané kombinace vlastností se ukázal jako velmi efektivní. GA nejenže zohledňuje složité interakce mezi vlastnostmi, ale také zvyšuje celkový výkon a stabilitu modelu, což je klíčové pro aplikace, jako je detekce lži pomocí EEG signálů [10, 11].

Funkčnost GA

GA je účinná metoda pro nalezení optimálních řešení v prostoru možných kombinací. Jeho funkčnost spočívá v několika klíčových krocích:

- **Inicializace populace:** GA vytvoří počáteční soubor náhodných binárních vektorů, přičemž každý vektor reprezentuje kombinaci vybraných vlastností.
- **Hodnocení jedinců:** Jedinci jsou hodnoceni pomocí fitness funkce.
- **Výběr rodičů:** Rodiče jsou vybíráni z populace na základě jejich fitness hodnot, přičemž jedinci s lepšími hodnotami mají vyšší pravděpodobnost být vybráni pro vytvoření další generace.
- **Křížení:** Kombinace vlastností rodičů k vytvoření nových potomků.
- **Mutace:** Náhodná změna některých vlastností potomků pro prozkoumání nových částí prostoru řešení a zvýšení diverzity populace.
- **Tvorba nové populace:** Nová populace je vytvořena z potomků a některých nejlepších jedinců z předchozí generace.

Tento proces se opakuje iterativně, dokud není dosaženo stanoveného kritéria, jako je maximální počet generací nebo dokud není nalezeno optimální řešení.

Optimalizace nastavení GA

Pro optimalizaci výběru vlastností byl GA nastaven s následujícími parametry:

- **Maximální počet generací** (MaxGenerations): Určuje, kolikrát se populace aktualizuje. Vyšší počet generací umožňuje algoritmu provést více iterací optimalizace, což zvyšuje pravděpodobnost nalezení globálního optima a může vést k lepším výsledkům. V tomto případě byl maximální počet generací nastaven na 100 pro důkladnější průzkum prostoru řešení.
- **Velikost populace** (PopulationSize): Určuje počet jedinců (kandidátních řešení) v každé generaci. Větší populace zvyšuje genetickou diverzitu jedinců, snižuje riziko uvíznutí v lokálních optimech a může vést k lepšímu pokrytí prostoru řešení. Velikost populace byla nastavena na 200 pro zajištění dostatečné diverzity populace.
- **Křížení** (CrossoverFraction): Umožňuje kombinovat vlastnosti dvou rodičovských jedinců a vytvářet nové jedince. Vysoký podíl křížení podporuje kombinování vlastností z různých jedinců, což může vést

k nalezení lepších řešení. Pravděpodobnost křížení byla nastavena na 0,8, což znamená, že 80 % populace se bude podílet na křížení.

- **Mutate** (MutationFcn): Přidává náhodné změny do populace, které pomáhají algoritmu uniknout z lokálních optim a prozkoumat více prostoru řešení. Funkce pro mutaci jedinců byla nastavena na 0,05.

Využití Gaussian SVM v GA

Využití Gaussian Support Vector Machine (SVM) v kontextu GA pro výběr optimálních vlastností je efektivním řešením pro klasifikaci EEG dat. SVM je široce používaný klasifikační algoritmus, známý svou schopností efektivně řešit problémy s vysokou dimenzionalitou a identifikovat rozhodovací hranice mezi třídami. SVM s Gaussian jádrovou funkcí je zvláště vhodný pro zpracování komplexních a nelineárních dat, což je klíčové pro analýzu EEG dat. Tento klasifikátor dokáže lépe zachytit složité vztahy v datech díky své schopnosti transformovat vstupní prostor do vyšších dimenzí a vytvářet nelineární rozhodovací hranice.

GA je použit pro výběr optimální podmnožiny vlastností, které maximalizují klasifikační schopnosti modelu SVM a zlepšují jeho výkon. Tímto způsobem GA optimalizuje vstupní prostor pro SVM, což umožňuje lepší přizpůsobení modelu konkrétním charakteristikám EEG dat. Tato kombinace umožňuje dosažení vyšší přesnosti klasifikace a zároveň zajišťuje robustnost a schopnost generalizace modelu na nová data. Gaussian SVM je tedy klíčovou součástí GA v procesu výběru vlastností pro detekci lži z EEG dat, přičemž GA optimalizuje prostor vlastností pro maximální efektivitu tohoto klasifikátoru.

Implementace GA pro výběr relevantních vlastností:

```
% Načtení dat
data_table = readtable('features.csv');

% Identifikace sloupce s třídou
class_column = 'Class';
class_idx = strcmp(data_table.Properties.VariableNames, class_column);

% Extrakce vlastností a tříd
X = data_table(:, ~class_idx); % Všechny sloupce kromě sloupce s třídou
y = data_table(:, class_idx); % Sloupec s třídou

% Definování názvů vlastností
featureNames = data_table.Properties.VariableNames(~class_idx);
rate = 0.05;

% Nastavení možností GA
opts = optimoptions('ga', ...
    'MaxGenerations', 100, ... % Maximální počet generací
    'PopulationSize', 200, ... % Velikost populace
    'CrossoverFraction', 0.8, ... % Nastavení podílu křížení
    'Display', 'iter', ... % Zobrazení průběhu
```

```

'UseParallel', false, ... % Paralelní výpočet zrušen
'MutationFcn', {@mutationuniform, rate}); % Nastavení míry mutace

% Spuštění GA pro výběr vlastností
nFeatures = size(X, 2);
[selectedFeatures, fval] = ga(@(selectedFeatures)
evaluateFitness(selectedFeatures, X, y), ...
    nFeatures, [], [], [], [], zeros(1, nFeatures), ones(1, nFeatures),
[], 1:nFeatures, opts);

% Zobrazení vybraných vlastností
selectedFeatureNames = featureNames(selectedFeatures > 0.5);
disp('Vybrané vlastnosti:');
disp(selectedFeatureNames);

% Definování fitness funkce pro GA
function score = evaluateFitness(selectedFeatures, X, y)
    selectedIdx = find(selectedFeatures > 0.5); % Práh pro výběr více
vlastností

    minFeatures = floor(size(X, 2) / 2); % Minimálně polovina vlastností

    if length(selectedIdx) < minFeatures
        score = inf; % Penalizace pro méně než polovinu vlastností
        return;
    end

    % Provádění 10-násobné křížové validace
    cv = cvpartition(y, 'KFold', 10, 'Stratify', true);

    % Trénování a hodnocení modelu
    cvLoss = crossval(@(Xtrain, ytrain, Xtest, ytest) ...
        sum(~strcmp(ytest, predict(fitcsvm(Xtrain, ytrain,
'KernelFunction', 'gaussian', 'Standardize', true), Xtest))), ...
        X(:, selectedIdx), y, 'partition', cv);

    % Snížení penalizace pro více vlastností
    score = mean(cvLoss) + 0.001 * length(selectedIdx);
end

```

Zdrojový kód 1: Implementace GA

Tento kód implementuje GA pro výběr relevantních vlastností z datové sady, což je klíčový krok v analýze dat pro detekci lži pomocí EEG. Kód byl vytvořen v Matlabu pomocí vestavěných funkcí pro GA a SVM. GA vybírá podmnožiny vlastností pro klasifikaci a optimalizuje výběr vlastností pomocí fitness funkce, která měří výkonnost modelu SVM s Gaussian jádrem na základě křížové validace. Tento kód načte data, identifikuje vlastnosti a třídy, nastaví parametry GA k optimalizaci procesu výběru vlastností a poté pomocí fitness funkce hodnotí kvalitu vybraných vlastností. Fitness funkce *evaluateFitness* provádí 10-násobnou křížovou validaci, trénuje Gaussian SVM model na trénovací sadě dat a vrací průměrnou chybovost predikce. Tento postup umožňuje najít optimální kombinaci vlastností, která minimalizuje chybovost modelu.

GA byl použit k výběru podmnožiny nejlepších vlastností z celkového počtu 70 vlastností. GA je heuristická metoda optimalizace a jeho výstup může být ovlivněn náhodnými počátečními podmínkami a dalšími faktory. Proto bylo provedeno několik pokusů s různými parametry GA k dosažení nejlepších výsledků. Nakonec byla vybrána sada 35 vlastností, která poskytovala nejvyšší přesnost klasifikace. To snížilo dobu běhu celého procesu a složitost prostoru vlastností. Zvolené vlastnosti jsou následující: Max_ERP, P2P_ERP, Mean_ERP, Skew_ERP, Kurt_ERP, Hjorth_Activity_ERP, Hjorth_Mobility_ERP, Hjorth_Complexity_ERP, Mean_Delta, Entropy_Delta, Hjorth_Activity_Delta, Hjorth_Mobility_Delta, P2P_Theta, Mean_Theta, Skew_Theta, Kurt_Theta, Hjorth_Activity_Theta, Energy_Theta, Power_Theta, Max_Alfa, Min_Alfa, Skew_Alfa, Hjorth_Activity_Alfa, Hjorth_Mobility_Alfa, Max_Beta, Std_Beta, Median_Beta, Skew_Beta, Kurt_Beta, Entropy_Beta, Hjorth_Activity_Beta, Hjorth_Mobility_Beta, Hjorth_Complexity_Beta, Energy_Beta a Power_Beta.

Tato sada vlastností byla následně použita pro klasifikaci dat pomocí SVM. Stabilita modelu a průměrná přesnost klasifikace byla dosažena na úrovni 95,09 %, což potvrzuje efektivitu a výhody použití GA v této oblasti. Jelikož je tento proces náročný jak časově, tak výpočetně, byl v programu pro predikci zvolen hotový model s těmito vybranými vlastnostmi (viz podkapitola 6.1.3).

5.5 Klasifikace

Po extrakci a výběru vhodných vlastností EEG signálu následuje klíčový krok klasifikace dat, jehož cílem je určit přítomnost cílové informace u subjektu. Analýza reakcí na sondy a irelevantní podněty pomocí složky P300 pro odlišení vinného od nevinného lze provádět různými metodami [16]. Ve fázi klasifikace se shromážděné vlastnosti porovnávají pomocí klasifikačních algoritmů nebo statistických metod. Klasifikační algoritmy jsou v tomto typu výzkumu nejčastěji používané, kdy se výsledná data řadí do klasifikačních tříd na základě parametrů získaných v předchozí části a testuje se účinnost klasifikátorů [47].

Klasifikace kontinuálních EEG dat je náročná a závisí na typu a velikosti použité datové sady, což ovlivňuje výběr vhodného klasifikátoru a jeho výkon [10]. Různé algoritmy mají odlišné výpočetní nároky a dobu zpracování dat, což může vést k rozdílným výsledkům při aplikaci na různé datové sady. Proto je nezbytné provést předběžnou analýzu k predikci výkonu klasifikátoru na konkrétní množině dat. Výběr klasifikačního algoritmu je klíčovou součástí procesu detekce lži z EEG signálů pro přesnou identifikaci subjektů skrývajících informace [8].

Algoritmy pro automatickou detekci lži z analýzy mozkových vln jsou předmětem intenzivního zájmu výzkumné komunity. V literatuře existuje mnoho různých typů klasifikátorů, které zlepšují výkon systému při binární klasifikaci tříd vinných a nevinných pro identifikaci podvodu. Pro odhalení skrytých

informací z EEG dat se nejčastěji využívají algoritmy strojového učení, které se staly dalším nástrojem pro CIT [11]. Důležitou vlastností těchto metod je jejich nezávislost na subjektivním hodnocení, interpretaci či vyhodnocování dat. Předchozí studie využívaly různé typy algoritmů k efektivnímu oddělení výstupních tříd v datech EEG [P.1]. Mezi nejčastěji využívané algoritmy patří SVM [4, 7, 10, 17, 47], LDA [4, 7, 9-11], MLFFNN [4, 7, 10, 12], KNN [4, 7, 8, 10] a NB [4, 7, 10]. Další možností je využití Decision Trees (DT).

Support Vector Machine

SVM je jedním z nejpobulárnějších supervizovaných algoritmů pro řešení klasifikačních problémů. Jedná se o nepravděpodobnostní lineární klasifikační přístup, který volí optimální oddělovací nadrovinu, čímž rozděluje klasifikační třídy tak, aby maximalizoval vzdálenost mezi datovými body dvou tříd [10]. Datové body používané pro nalezení optimálního řešení se nazývají podpůrné vektory, které jsou zodpovědné za určení šířky nadroviny. SVM dosahuje optimálního řešení pro lineární i nelineární data a může využívat různé jádrové funkce k zajištění flexibility při oddělování nelineárních dat [4, 10].

Decision Trees

DT využívají strukturu stromu, kde uzly představují rozhodnutí na základě hodnot vstupních vlastností a listy reprezentují konečné třídy. Hlavní výhodou DT je jejich schopnost pracovat s nelineárními vztahy a interakcemi mezi vlastnostmi. Algoritmus iterativně vybírá vlastnosti, které nejlépe rozdělují data do tříd, a vytváří binární nebo vícecestné rozdělení [10].

Naive Bayes

NB je pravděpodobnostní klasifikační přístup využívající koncepty maximální pravděpodobnosti a je založen na Bayesově teorému. Bere v úvahu podmíněný pravděpodobnostní model pro klasifikaci dat do různých tříd a je považován za vhodný pro manipulaci s vysokou dimenzionalitou dat [4, 10].

K-Nearest Neighbors

KNN je neparametrický klasifikační přístup, který klasifikuje testovací data do konkrétní třídy na základě třídy většiny jejich sousedů. K identifikaci třídy daného datového bodu se vybírá k počet sousedů pomocí různých přístupů založených na vzdálenosti. Konstantu k lze vybrat na základě počtu datových bodů, které mají být klasifikovány. Pro výpočet hodnoty k neexistuje žádná předem definovaná technika a vybírá se heuristicky. KNN je nejvhodnějším přístupem pro nízkorozměrná data [4, 10].

5.5.1 Porovnání metod pro klasifikaci EEG dat

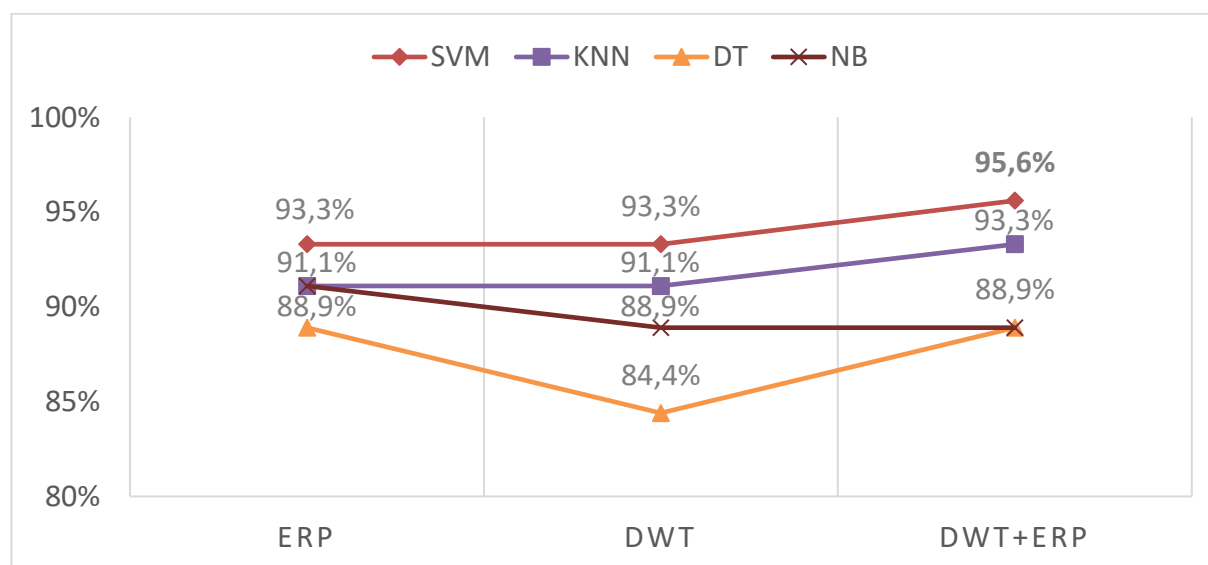
V této disertační práci byly analyzovány různé metody extrakce a výběru vlastností a klasifikační algoritmy s cílem maximalizovat přesnost detekce lži pomocí EEG signálů. Hlavním cílem bylo identifikovat kombinace metod a algoritmů, které poskytují nejvyšší přesnost klasifikace na základě analýzy ERP P300 za použití nízkonákladového zařízení.

Porovnání metod extrakce vlastností na jednotlivých elektrodách

Testovány byly různé vlastnosti signálu a jejich kombinace, jako amplitudy, parametry Hjorth, koeficienty DWT a další. Tyto metody byly použity jako vstupní data pro jednotlivé klasifikační algoritmy v nástroji Classification Learner. Samostatné použití těchto metod vedlo k nižší přesnosti klasifikace, a proto byly kombinovány do finálního vektoru zahrnujícího všechny vlastnosti z časové a časově-frekvenční domény, což umožnilo komplexnější analýzu. Úspěšnost vybraných vlastností byla porovnána pomocí různých metod a klasifikačních algoritmů na jednotlivých elektrodách a jejich kombinacích. Pro jednotlivé elektrody byly testovány různé metody extrakce vlastností (ERP, DWT a kombinace DWT+ERP) pomocí různých klasifikačních algoritmů (SVM, KNN, DT a NB) s cílem identifikovat nejúčinnější kombinace.

Elektroda Pz

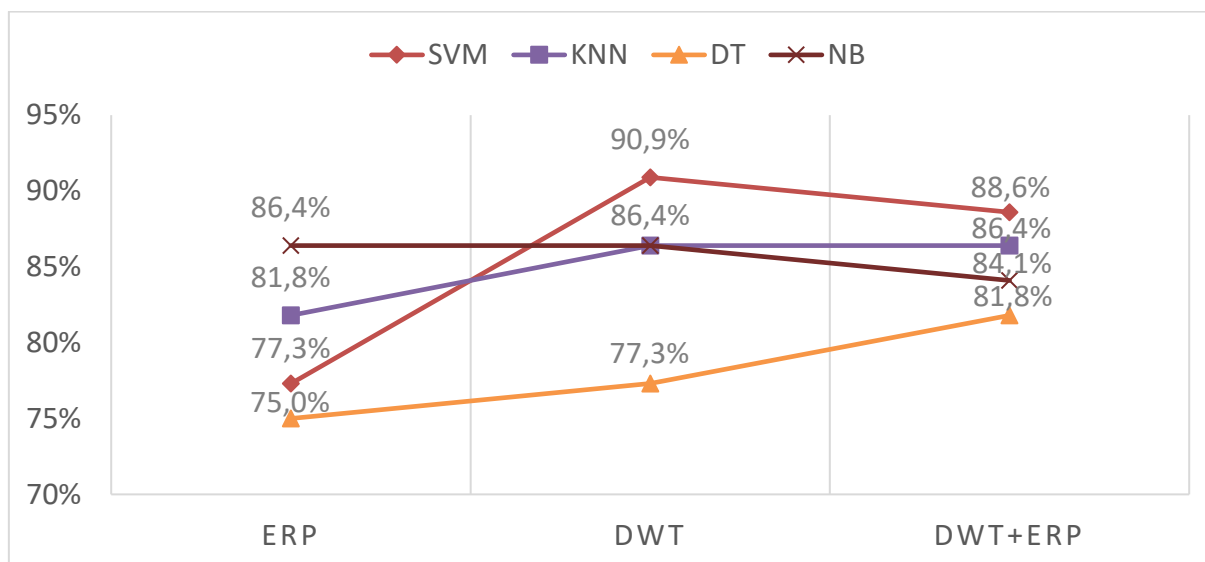
Výsledky ukazují, že nejvyšší přesnosti klasifikace na elektrodě Pz bylo dosaženo použitím kombinace DWT+ERP s algoritmem SVM, kde přesnost dosáhla až 95,6 %. Algoritmus KNN také vykazoval vysokou přesnost při kombinaci DWT+ERP, konkrétně 93,3 %. Samotné ERP nebo DWT vedly k nižší přesnosti, přičemž SVM dosahoval 93,3 % pro obě metody.



Obr. 36.: Přesnost klasifikace vlastností pomocí různých algoritmů u elektrody Pz

Elektroda T7

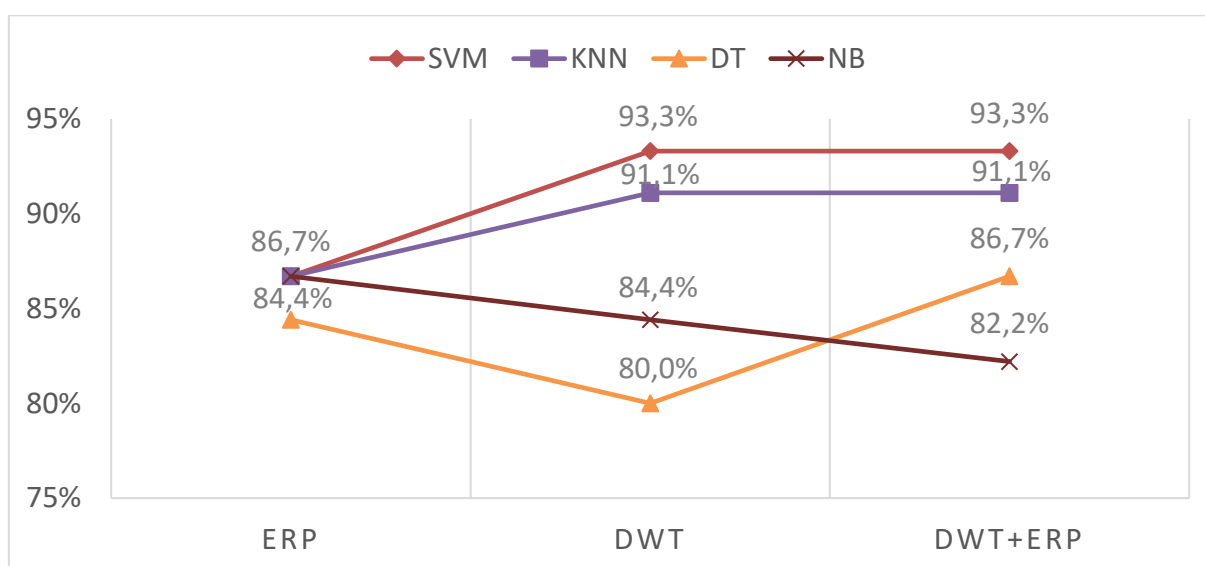
Na elektrodě T7 byla nejvyšší přesnost klasifikace dosažena pomocí algoritmu SVM při použití DWT, konkrétně 90,9 %. Kombinace DWT+ERP s algoritmem SVM vedla k mírně nižší přesnosti 88,6 %. Samotné ERP vykazovalo nejnižší přesnost mezi testovanými metodami u SVM s hodnotou 77,3 %. Algoritmy KNN a NB dosáhly přesnosti 86,4 % při použití různých metod.



Obr. 37.: Přesnost klasifikace vlastností pomocí různých algoritmů u elektrody T7

Elektroda T8

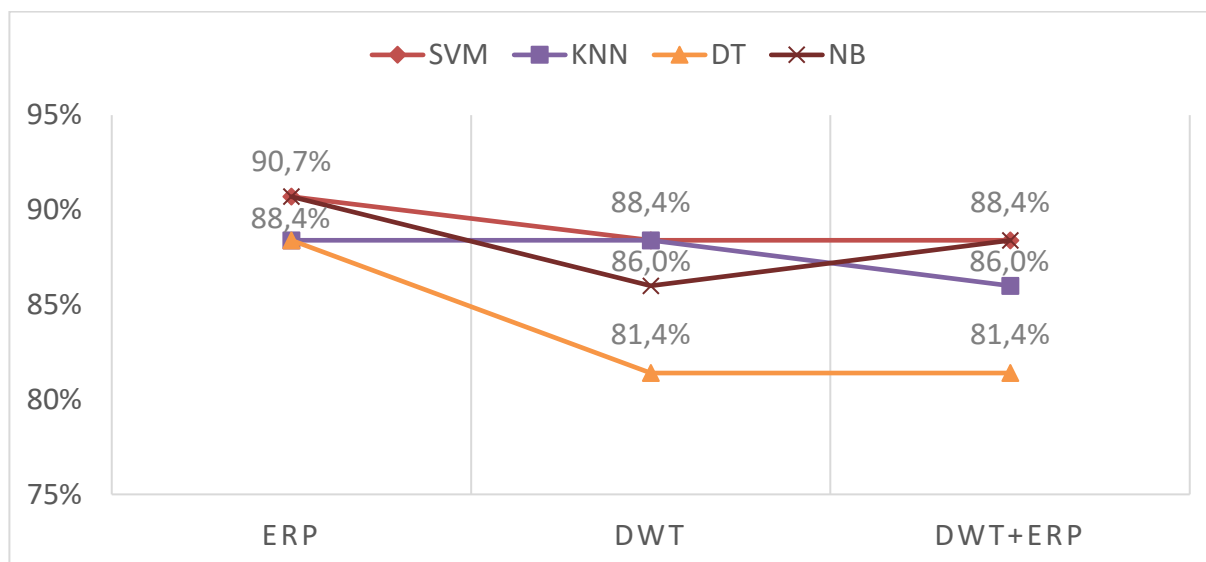
U elektrody T8 byla nejvyšší přesnost 93,3 % dosažena opět pomocí algoritmu SVM jak pro DWT, tak pro kombinaci DWT+ERP. Samotné ERP vykazovalo nižší přesnost 86,7 % při použití algoritmu SVM. Algoritmus KNN měl podobně vysokou přesnost při použití kombinace DWT+ERP a samotné DWT, konkrétně 91,1 %.



Obr. 38.: Přesnost klasifikace vlastností pomocí různých algoritmů u elektrody T8

Elektroda AF3

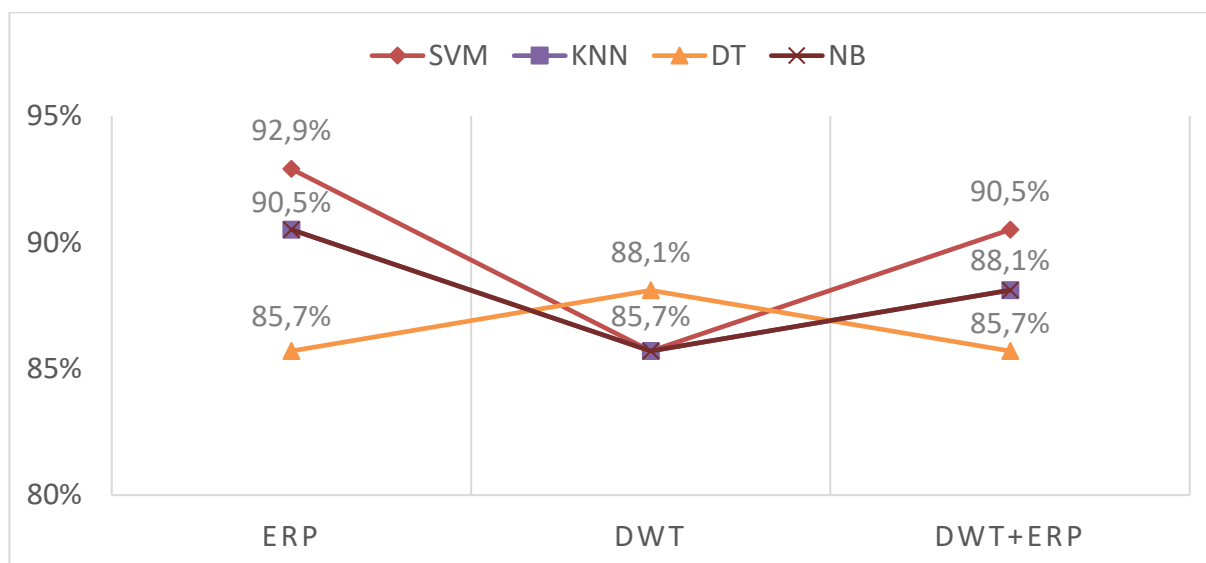
Pro elektrodu AF3 byla nejvyšší přesnost klasifikace dosažena použitím algoritmů SVM a NB na datech ERP, kde přesnost dosáhla 90,7 %. Kombinace DWT+ERP a samotné DWT vykazovaly nižší přesnost, přičemž SVM dosahoval 88,4 % u obou metod. Algoritmy KNN, NB a DT měly nižší přesnost při použití kombinace DWT+ERP a samotné DWT.



Obr. 39.: Přesnost klasifikace vlastností pomocí různých algoritmů u elektrody AF3

Elektroda AF4

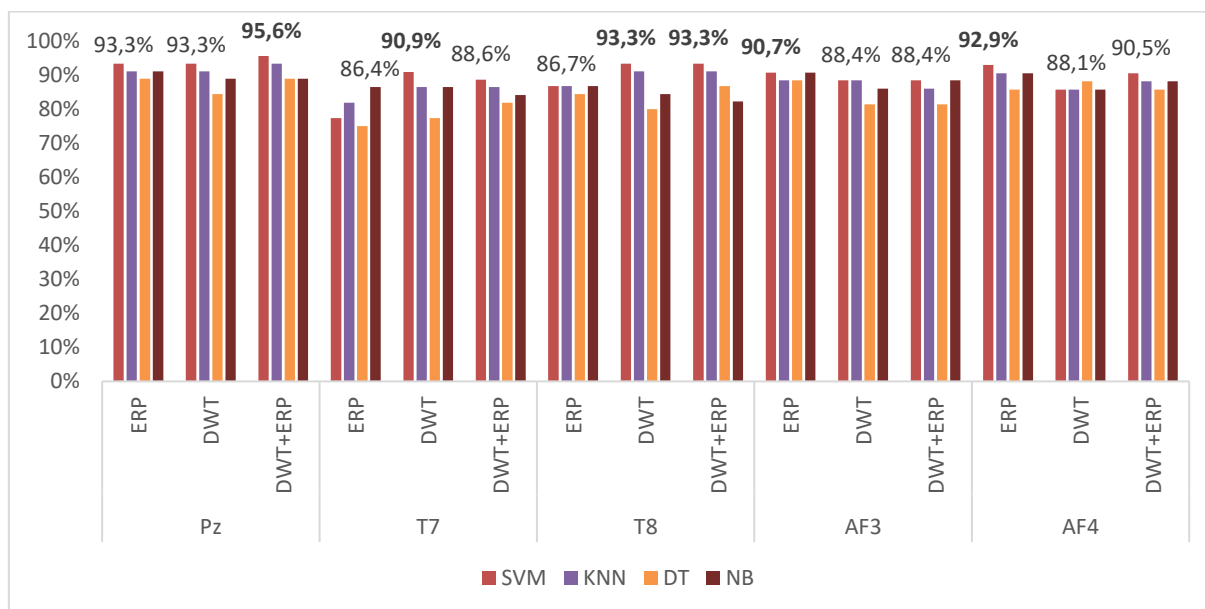
U elektrody AF4 byla nejvyšší přesnost dosažena použitím algoritmu SVM na datech ERP, konkrétně 92,9 %. Kombinace DWT+ERP a samotné DWT vykazovaly nižší přesnost, přičemž SVM dosahoval 90,5 % u kombinace DWT+ERP a 85,7 % u samotné DWT. Algoritmy KNN a NB dosáhly přesnosti 90,5 % při použití ERP.



Obr. 40.: Přesnost klasifikace vlastností pomocí různých algoritmů u elektrody AF4

Výsledky porovnání jednotlivých metod

Porovnání přesnosti klasifikace pro jednotlivé metody extrakce vlastností u jednotlivých elektrod aplikovaných na různých klasifikačních algoritmech lze vidět na Obr. 41. V grafu jsou zvýrazněny maximální dosažené hodnoty u jednotlivých metod. Hodnota přesnosti klasifikace byla vybrána na základě nejčastěji se vyskytující hodnoty v rámci několika pokusů.

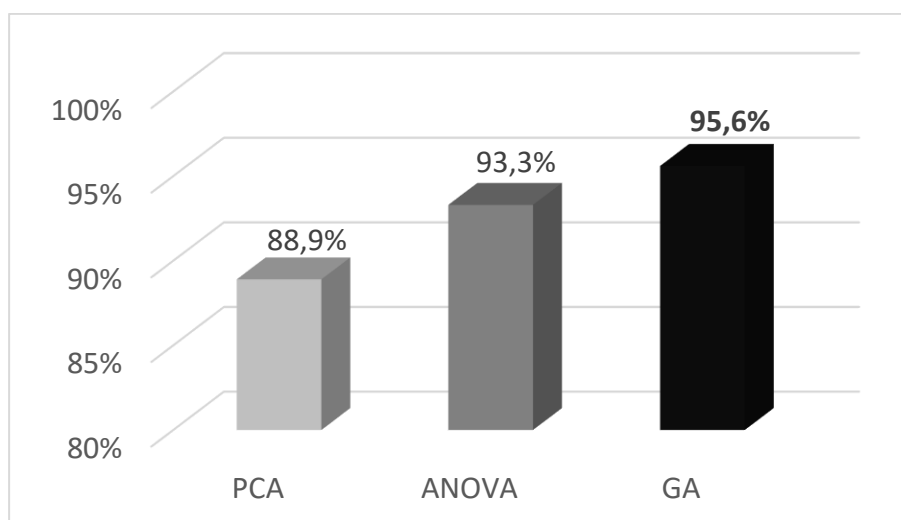


Obr. 41.: Porovnání přesnosti klasifikace na různých elektrodách

Z výsledků je zřejmé, že nejvyšší přesnosti klasifikace bylo dosaženo na elektrodě Pz při použití algoritmu SVM s kombinací metod ERP a DWT, kde přesnost dosáhla až 95,6 %. Algoritmus SVM se ukázal jako velmi vhodný pro detekci lži na základě EEG signálů. Na dalších elektrodách byla také dosažena vysoká přesnost klasifikace, zejména při použití algoritmu SVM, avšak žádná z nich nepřekonal přesnost dosaženou na elektrodě Pz.

Porovnání metod výběru vlastností

Pro optimalizaci výběru vlastností byly použity různé metody, konkrétně GA, ANOVA a PCA, s cílem získat nejvýznamnější vlastnosti z celkových 70 statistických vlastností. Testování bylo prováděno na kombinaci metod, která dosáhla nejvyšší přesnosti klasifikace, konkrétně klasifikační algoritmus SVM na elektrodě Pz za využití kombinace vlastností ERP a DWT. Výběr vlastností byl proveden také na ostatních elektrodách s využitím různých algoritmů a metod pro extrakci vlastností, nicméně žádná z metod nedosáhla vyšší přesnosti klasifikace než zvolená kombinace.



Obr. 42.: Porovnání metod výběru vlastností

Nejvyšší přesnosti klasifikace bylo dosaženo při použití GA pro výběr vlastností z ERP a DWT dat v kombinaci s klasifikačním algoritmem SVM na elektrodě Pz, což potvrzuje její vhodnost pro detekci lži pomocí EEG signálů. Analýza také ukazuje, že algoritmus SVM je obecně nejvýkonnější mezi testovanými algoritmy, zejména při použití kombinovaných metod extrakce vlastností. Metody PCA ani ANOVA neprokázaly zlepšení přesnosti klasifikace, a tudíž nejsou preferované pro tuto specifickou úlohu.

5.5.2 Gaussian Support Vector Machine

V této práci byl pro tvorbu finálního modelu použit Medium Gaussian SVM s jádrem Radial Basis Function (RBF), což je robustní algoritmus strojového učení schopný vytvářet rozhodovací hyperrovinu maximalizující vzdálenost mezi třídami dat. SVM je jedním z nejvýkonnějších a nejpopulárnějších algoritmů pro klasifikaci a je často využíván pro detekci lži, kde dosahuje vysoké přesnosti [4, 11, 17]. Gaussian SVM je zvláště vhodný pro analýzu ERP P300 signálů, které reflektují reakce mozku na známé podněty, jež mohou mít nelineární povahu. Transformuje vstupní prostor do vyšší dimenze, což umožňuje zachytit složité vztahy v datech [4]. Robustnost SVM vůči přetížení (overfitting) je zajištěna díky optimalizaci parametrů a použití jádra, které umožňuje modelu efektivně se učit i z menších množství dat.

Gaussian jádro (RBF) je definováno vzorcem:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \quad (5.17)$$

kde x_i a x_j jsou vstupní vektory dat, $\|x_i - x_j\|^2$ je Euklidovská vzdálenost mezi vektory x_i a x_j , a γ je parametr určující šířku Gaussian jádra a má zásadní vliv na schopnost modelu učit se nelineární vztahy v datech. Vyšší hodnoty γ způsobují, že jádro má menší dosah, zatímco menší hodnoty γ rozšiřují dosah jádra, což ovlivňuje přeučení a generalizaci modelu [4]. RBF jádro umožňuje

modelu efektivně rozlišovat mezi reakcemi na známé a neznámé podněty, což je klíčové pro detekci lži.

Rozhodovací funkce

SVM hledá rozhodovací hranici, která maximalizuje rozpětí mezi třídami. Rozhodovací funkce pro SVM je definována jako:

$$f(x) = \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b, \quad (5.18)$$

kde α_i jsou Lagrangeovy multiplikátory, které určují váhu každého vzoru v tréninkové množině, y_i jsou hodnoty tříd pro tréninkové vzory x_i , $K(x_i, x)$ je jádrová funkce (v tomto případě Gaussian jádro) a b je bias term, který posouvá rozhodovací hranici. Rozhodovací funkce $f(x)$ slouží k určení, na které straně rozhodovací hranice se nový vstupní vzor x nachází.

Optimalizační problém

Trénink SVM zahrnuje řešení následujícího optimalizačního problému:

$$\min_{\alpha} \left(\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^n \alpha_i \right) \quad (5.19)$$

podléhající podmínkám:

$$\sum_{i=1}^n \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq C,$$

kde α_i jsou optimalizované hodnoty Lagrangeových multiplikátorů a C je regulační parametr, který určuje penalizaci za nesprávně klasifikované vzory. Vyšší hodnota C snižuje počet chyb v tréninkové množině, ale může vést k přeučení. Nižší hodnota C může zvýšit počet chyb v tréninkové množině, ale zlepšit generalizaci.

Parametry modelu Gaussian SVM

- **Kernel Scale ($\gamma = 5,3$):** Tento parametr určuje šířku (rozptyl) Gaussian jádra. Hodnota 5,3 byla vybrána na základě optimalizace, která zajistila nejlepší výkon modelu poskytnutím rovnováhy mezi složitostí modelu a jeho schopností generalizovat.
- **Box Constraint Level ($C = 1$):** Tento parametr kontroluje regularizaci modelu, což znamená, jak moc se penalizují chyby v trénovacích datech. Hodnota 1 poskytuje rovnováhu mezi přesností modelu a jeho generalizační schopností.

- **Multiclass Method (One-vs-One):** Metoda One-vs-One rozděluje vícetřídní klasifikační úlohu na několik dvojtřídních problémů. U každé dvojice tříd je natrénován jeden SVM model a konečné rozhodnutí se provádí na základě většinového hlasování.
- **Standardize Data (True):** Data byla standardizována, aby všechny vlastnosti měly stejný vliv na rozhodovací hranici modelu, což zlepšuje stabilitu a výkon modelu.

5.5.3 Postup klasifikace dat a tvorba modelu

Tato práce představuje nový přístup ke klasifikaci EEG signálů za využití modifikovaného testu CIT založeného na ERP P300 s využitím pokročilých metod výběru vlastností a strojového učení. Klasifikace dat a tvorba modelu je klíčovým krokem pro úspěšné rozlišení mezi vinnými a nevinnými subjekty. Cílem je vytvořit přesný a robustní model, který dokáže spolehlivě predikovat třídy subjektů na základě analýzy jejich mozkových vln. Tento přístup je zásadní pro vývoj efektivního systému detekce lži využívajícího EEG signály.

Implementace klasifikačního modelu

Klasifikace dat zahrnuje vytvoření klasifikačního modelu obsahujícího pravidla pro rozdělení dat do klasifikačních tříd. Klasifikační model byl implementován s využitím binárního klasifikátoru SVM, který je široce používán pro analýzu mozkových vln. Pro trénink modelu byla data připravena ve formě matice obsahující 70 atributů (70 statistických vlastností a klasifikační třídu). Atributy představují jednotlivé statistické vlastnosti signálu vypočítané z dat ERP a DWT. Součástí matice je také hodnota třídy, přičemž každému prvku je přiřazena odpovídající klasifikační třída [63].

Extrakce a výběr vlastností

Klasifikace EEG signálů je významně ovlivněna výběrem vlastností. Nejprve byly extrahovány specifické vlastnosti EEG signálu, které poskytují důležité informace o charakteristikách ERP P300 a slouží jako vstupy pro klasifikátor. V této práci byly statistické vlastnosti extrahovány z aproximačních a detailních koeficientů získaných z DWT v každém vlnovém pásmu a z ERP dat. Výsledný vektor se skládal ze 70 statistických vlastností pro podněty sond a irelevantní podněty. Výběr vhodných vlastností byl proveden pomocí GA, který vybral 35 nejrelevantnějších vlastností z celkových 70. To vedlo k tvorbě finální matice obsahující vybrané statistické vlastnosti aplikované na hodnoty ERP a DWT pro různé podněty na elektrodě Pz pro všechny subjekty. Každý vektor v matici byl rozdělen na vinné a nevinné pomocí přiřazení jednotlivých tříd.

Křížová validace

Před trénováním byla použita křížová validace k vyhodnocení výkonnosti modelu na celkovém datasetu. Data byla rozdělena do k částí a v každé iteraci

byla jedna část použita jako testovací data, zatímco zbývající části jako trénovací data. Tento proces se opakuje k -krát, přičemž pokaždé se použije jiná část jako testovací sada. Výsledná metrika, například přesnost, je zprůměrována přes všechny iterace [8, 63]. Validace se provádí pomocí x -násobné křížové validace, nejčastěji 5-násobné [8, 10] či 10-násobné [4, 6, 9, 12]. Křížová validace se provádí za účelem ochrany proti přetrénování dat, při kterém je klasifikátor natrénován na trénovací data, ale není schopen úspěšně klasifikovat jiná data [17].

Trénink a testování modelu

Po výběru vhodných vlastností signálu byla data rozdělena na trénovací a testovací množiny s použitím desetinásobné křížové validace. Pro trénink modelu bylo využito prvních 30 subjektů (15 vinných, 15 nevinných) z celkových 50, zatímco zbývající subjekty (9 vinných, 7 nevinných) byly vyhrazeny pro finální testování modelu na nových datech k ověření jeho robustnosti a výkonu. Čtyři subjekty byly odstraněny kvůli nadměrnému množství artefaktů nebo neplatnému měření. Klasifikační algoritmus SVM s vlastnostmi generovanými pomocí DWT a ERP, vybranými pomocí GA, se ukázal jako nejúspěšnější.

Predikce a klasifikace

Konečným cílem metody CIT je určit, zda je subjekt vinný nebo nevinný. Predikce modelu byla založena na přítomnosti ERP komponenty P300 v reakci na podnět sondy v EEG signálech. Klasifikátor určuje, zda se tato komponenta vyskytuje, a následně se volí metoda pro určení označení subjektu jako vinného nebo nevinného na základě odpovědí na podněty sondy. Pokud model predikoval třídu jako 1 (vinný), subjekt byl klasifikován jako vinný. Pokud model predikoval třídu jako 0 (nevinný), subjekt byl označen jako nevinný. Pro nevinné subjekty je sonda irelevantní a pro vinné má vyšší hodnotu, což určuje, zda je subjekt vinný nebo nevinný.

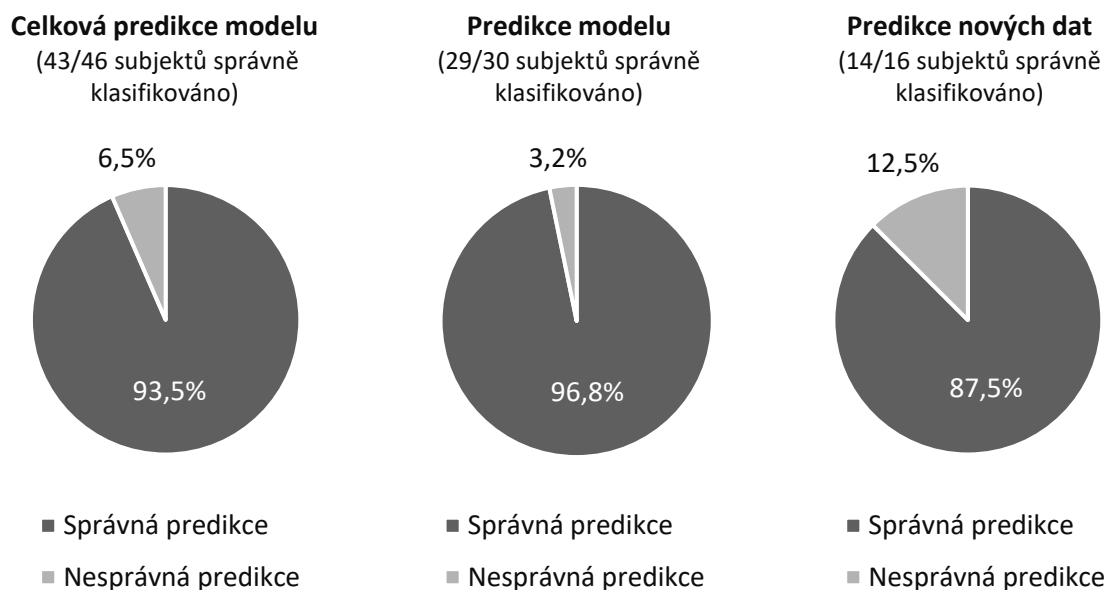
Kontrola falešných pozitiv

Během analýzy bylo důležité kontrolovat falešně pozitivní výsledky, tedy situace, kdy model nesprávně označí nevinné subjekty jako vinné. Pro zvýšení preference klasifikace subjektů jako nevinných byl zvolen parametr s menší hodnotou, který zvyšuje citlivost modelu vůči falešným pozitivům. To bylo dosaženo jemnějším nastavením jádra na 5,3 v Gaussian SVM. Menší hodnota jádra vytváří složitější rozhodovací hranice, což může minimalizovat počet falešně pozitivních výsledků, lépe generalizovat na nová data a zvýšit citlivost modelu, což je důležité pro klasifikaci v kontextu detekce lži pomocí EEG.

Optimalizace modelu a parametrů

Byly vyzkoušeny různé vlastnosti a klasifikátory s různými parametry, přičemž nejvyšší přesnosti dosáhl algoritmus SVM s vlastnostmi kombinujícími ERP a DWT. Přestože model dosáhl přesnosti klasifikace až 95,6 %, nedokázal správně

predikovat některá nová data, což je stěžejní pro detekci lži pomocí EEG. Po výběru vlastností pomocí GA a optimalizaci parametrů SVM byl model schopen správně predikovat 43 z 46 subjektů, zatímco předtím správně predikoval 40 z 46 subjektů. Jako finální model byl vybrán Medium Gaussian SVM s parametrem jádra 5,3 a pro GA byly nastaveny hodnoty mutace 0,05 a křížení 0,8, což vedlo ke stabilní přesnosti klasifikace 95,09 % a správné predikci 14 z 16 nových subjektů. Tento přístup vedl k výběru parametrů s nejlepšími výsledky, které zajistily, že vybrané vlastnosti jsou robustní a dobře generalizují na nová data, což je kritické pro praktické aplikace, jako je detekce lži pomocí EEG signálů.



Obr. 43.: Přesnost predikce finálního modelu

Automatizace procesu

Vzhledem k velkému množství dat byl použit skript v prostředí Matlab, který byl aplikován na všechny datové soubory. Tento skript automatizoval proces predikce, čímž se zjednodušil a urychlil celý postup.

Dílčí závěr

V rámci disertační práce zaměřené na detekci lži pomocí analýzy ERP P300 z EEG signálů na elektrodě Pz dosáhl finální model průměrné přesnosti klasifikace 95,09 %. Použití GA pro výběr relevantních vlastností z ERP a DWT dat v různých frekvenčních pásmech v kombinaci s SVM a optimalizací parametrů vedlo k vysoké přesnosti klasifikace. GA umožňuje prozkoumávat velký prostor možností a nalézat nečekané kombinace vlastností, čímž zvyšuje celkový výkon a stabilitu modelu. Model byl natrénován pomocí Medium Gaussian SVM s desetinasobnou křížovou validací pro minimalizaci přetížení a byl poté použit k predikci tříd (vinní a nevinní) na nezávislých datech. Tento algoritmus byl zvolen pro svou schopnost efektivně řešit nelineární klasifikační problémy s vysokou přesností. Finální model správně predikoval 43 z 46

celkových subjektů a 14 z 16 nových subjektů. Dále během experimentu šest subjektů z vinné skupiny odpovědělo, že sondu znají, i když měli tuto skutečnost utajit.

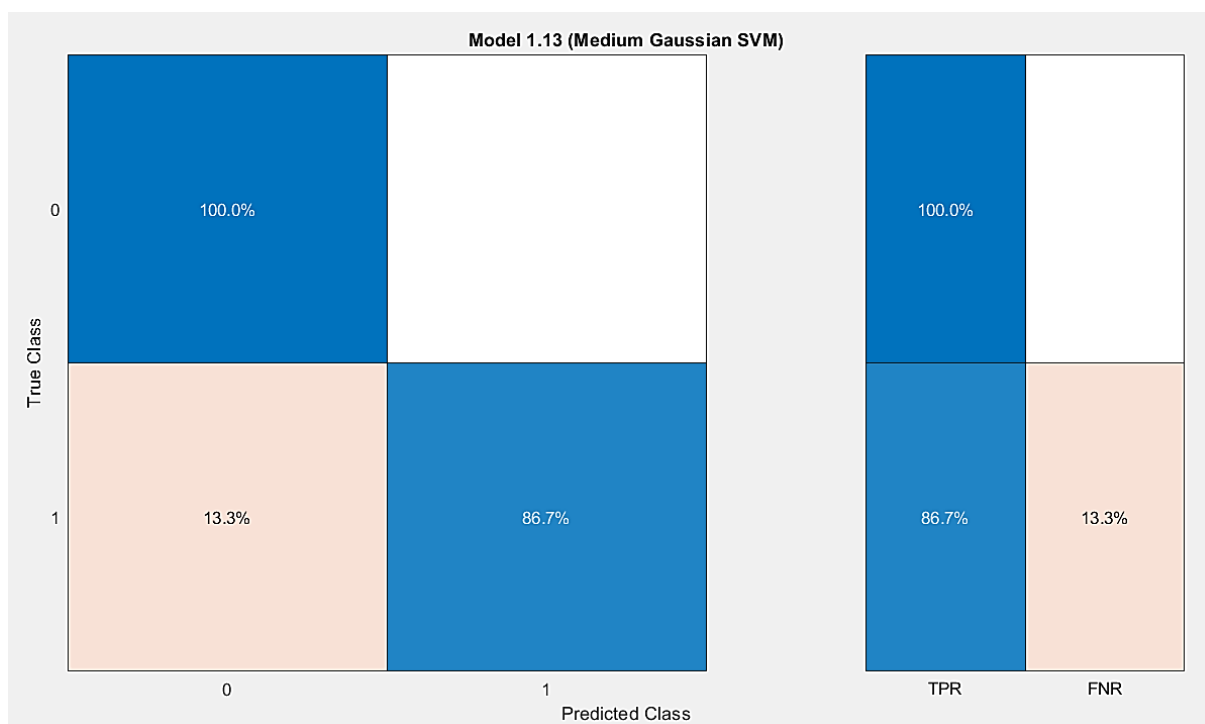
Tento výsledek potvrzuje vhodnost kombinace zvolených metod jako účinné strategie pro detekci lži pomocí EEG signálů. Finální model byl vytvořen v prostředí Classification Learner v Matlabu a poté použit ve výsledném programu pro detekci lži (viz podkapitola 6.1.3). Výkonnost modelu byla následně hodnocena pomocí různých metrik výkonu, které hodnotí správnost predikce jednotlivých tříd.

5.5.4 Vyhodnocení modelu

Vyhodnocení modelu je klíčovým krokem v procesu klasifikace, protože poskytuje informace o jeho spolehlivosti a účinnosti.

Confusion Matrix

Confusion Matrix je nástroj pro hodnocení výkonu klasifikačních modelů, který poskytuje podrobný pohled na to, jak dobře model klasifikuje různé třídy, a to tím, že ukazuje počty správných a nesprávných klasifikací pro každou třídu.



Obr. 44.: Confusion Matrix

Confusion Matrix pro binární klasifikaci je čtvercová matice 2x2, která obsahuje následující položky:

- **Skutečně negativní** (TN – true negatives): Počet negativních případů správně klasifikovaných jako negativní (nevinní správně klasifikováni jako nevinní).
- **Falešně pozitivní** (FP – false positives): Počet negativních případů nesprávně klasifikovaných jako pozitivní (nevinní chybně klasifikováni jako vinní).
- **Falešně negativní** (FN – false negatives): Počet pozitivních případů nesprávně klasifikovaných jako negativní (vinní chybně klasifikováni jako nevinní).
- **Skutečně pozitivní** (TP – true positives): Počet pozitivních případů správně klasifikovaných jako pozitivní (vinní správně klasifikováni jako vinní).

Metriky výkonu

Pro výpočet výkonu klasifikátoru se berou v úvahu různé ukazatele získané z Confusion Matrix, jako je přesnost, citlivost, specifčnost, preciznost a F-skóre [7-10]. Tyto metriky poskytují komplexní přehled o tom, jak model funguje při klasifikaci dat do jednotlivých tříd.

- **Přesnost (Accuracy):** Měří celkový podíl správných klasifikací (pozitivních i negativních). Jedná se o celkovou přesnost modelu.

$$Přesnost = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.20)$$

- **Citlivost (Sensitivity):** Měří podíl správně identifikovaných pozitivních tříd (vinní).

$$Citlivost = \frac{TP}{TP + FN} \quad (5.21)$$

- **Specifčnost (Specificity):** Měří podíl správně identifikovaných negativních tříd (nevinní).

$$Specifčnost = \frac{TN}{TN + FP} \quad (5.22)$$

- **Preciznost (Precision):** Měří podíl správně identifikovaných pozitivních příkladů mezi všemi příklady označenými jako pozitivní.

$$Preciznost = \frac{TP}{TP + FP} \quad (5.23)$$

- **F-skóre (F-measure):** Harmonický průměr citlivosti a preciznosti, poskytující vyvážené hodnocení klasifikačního modelu.

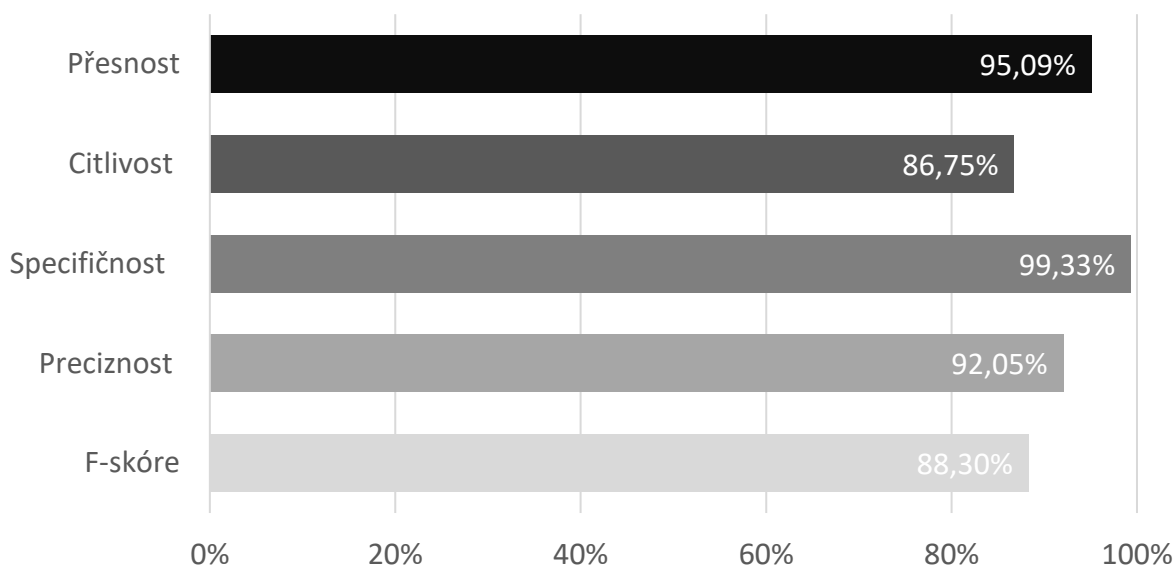
$$F - skóre = 2 \cdot \frac{Preciznost \cdot Citlivost}{Preciznost + Citlivost} \quad (5.24)$$

Tyto metriky společně poskytují ucelený pohled na výkonnost klasifikátoru a pomáhají při identifikaci jeho silných a slabých stránek [9]. Pro porovnání výsledků se současnými studii byla využita přesnost klasifikace, která je jednou z nejčastěji využívaných metrik pro hodnocení klasifikačního výkonu.

Průměrné hodnoty metrik výkonu

Finální model byl hodnocen pomocí 10-násobné křížové validace, opakované 100krát, aby se zajistila robustnost a spolehlivost výsledků. Průměrné hodnoty jednotlivých metrik výkonu jsou následující:

- **Přesnost (95,09 %):** Tento výsledek ukazuje na vysokou schopnost modelu správně klasifikovat jak lživé, tak pravdivé odpovědi. Průměrná přesnost klasifikace přesahující 95 % svědčí o tom, že model je velmi přesný při rozlišování mezi pravdou a lží na základě EEG signálů.
- **Citlivost (86,75 %):** Citlivost modelu představuje jeho schopnost správně identifikovat lživé odpovědi (pozitivní třída). Hodnota 86,75 % znamená, že model dokáže správně detekovat většinu lží, i když někteří lháři mohou zůstat neodhaleni.
- **Specifičnost (99,33 %):** Specifičnost ukazuje na schopnost modelu správně identifikovat pravdivé odpovědi (negativní třída). Hodnota 99,33 % naznačuje, že model má výjimečnou schopnost správně identifikovat pravdu s velmi malým počtem falešně pozitivních výsledků.
- **Preciznost (92,05 %):** Preciznost modelu odráží poměr správně identifikovaných lží mezi všemi odpověďmi klasifikovanými jako lži. Hodnota 92,05 % ukazuje, že většina identifikovaných lží jsou skutečně lži, což snižuje počet falešně pozitivních výsledků.
- **F-skóre (88,30 %):** F-skóre představuje harmonický průměr citlivosti a preciznosti a poskytuje vyváženou metriku pro hodnocení výkonu modelu. Hodnota 88,30 % naznačuje, že model má celkově velmi dobrý výkon při detekci lží, přičemž zohledňuje, jak schopnost správně detekovat lži (citlivost), tak minimalizovat falešně pozitivní výsledky (preciznost).



Obr. 45.: Průměrné hodnoty jednotlivých metrik výkonu

Dílčí závěr

Výsledky této studie potvrzují, že kombinace ERP a DWT s metodou výběru vlastností pomocí GA a algoritmem SVM poskytuje efektivní nástroj pro detekci lži na základě EEG signálů. Prezentovaný model dosáhl průměrné přesnosti klasifikace 95,09 %, což potvrzuje jeho potenciál v oblasti neurověd a detekce lži. Vysoká přesnost klasifikace, spolu s vynikající specifičností a solidní citlivostí, naznačuje, že tento model by mohl být velmi užitečný pro praktické aplikace, kde je důležité spolehlivě rozlišovat mezi pravdou a lží. Studie také zdůrazňuje význam kontroly falešných pozitivních výsledků a důležitost optimalizace modelu pro dosažení co nejvyšší přesnosti. Navržený model úspěšně klasifikoval 43 z 46 subjektů správně, což potvrzuje jeho robustnost a schopnost generalizace na nová data. Chybná klasifikace tří subjektů může být přičítána faktorům jako individuální variabilita EEG signálů, přítomnost artefaktů, stres a úzkost. Navržená metodika se ukázala jako nejúčinnější z hlediska přesnosti, složitosti a kompromisu mezi rychlostí a výkonem.

Celkově tento přístup demonstruje, že správný výběr vlastností signálu a použití vhodného klasifikačního algoritmu s optimalizovanými parametry jsou klíčové pro dosažení vysoké přesnosti a spolehlivosti detekce lži z EEG signálů. Navržený přístup byl úspěšně aplikován na data EEG získaná během modifikovaného testu CIT a poskytl slibné výsledky při identifikaci subjektů skrývajících informace. Výsledky této studie naznačují, že kombinace ERP P300 a pokročilých metod strojového učení pro klasifikaci může být účinným nástrojem a přispět tak k dalšímu výzkumu a vývoji spolehlivých a přesných systémů pro detekci lži pomocí analýzy mozkových vln.

5.6 Srovnání se stávajícími studii

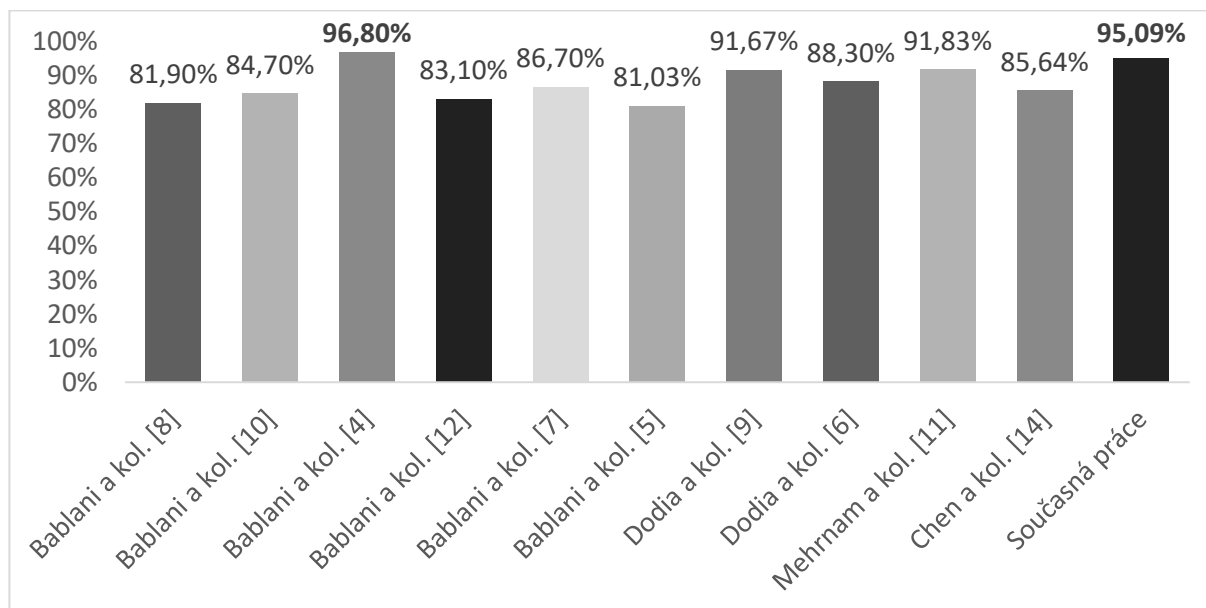
Srovnání současného přístupu se stávajícími studii je klíčovým krokem k posouzení účinnosti navržené metodiky. Tabulka 5.1 poskytuje přehled relevantních studií, včetně použitých metod extrakce vlastností, klasifikačních algoritmů, počtu subjektů a průměrné přesnosti klasifikace, která patří mezi nejčastěji používané metriky pro hodnocení výkonu klasifikace. Pro porovnání byly vybrány studie zaměřené na detekci lži pomocí EEG signálů při zkoumání reakce ERP P300 na známé tváře. Cílem je nejen porovnat jednotlivé studie, ale také zdůraznit jedinečné aspekty současného přístupu ve srovnání s metodami použitými v literatuře. Vzhledem k tomu, že se většina těchto studií zaměřila pouze na vyhodnocení výkonu modelu, bez vyhodnocení na nezávislých datech, byla pro účely porovnání použita přesnost modelu na trénovacích datech, aby se zachovala konzistence s ostatními studii.

Tabulka 5.1 Srovnání současného přístupu se stávajícími studii

Publikace	Extrakce vlastností	Klasifikační algoritmy	Počet subjektů	Průměrná přesnost klasifikace
Bablani a kol. [8]	Parametry Hjorth	KNN	10	81,9 %
Bablani a kol. [10]	DWT	LDA + SVM + MLFFNN	10	84,7 %
Bablani a kol. [4]	WT	SVM	10	96,8 %
Bablani a kol. [12]	WT	k-means + MLFFNN	10	83,1 %
Bablani a kol. [7]	CSP	fuzzy	10	86,7 %
Bablani a kol. [5]	WT	DBN	10	81,03 %
Dodia a kol. [9]	WPT	LDA	20	91,67 %
Dodia a kol. [6]	STFT	ELM	20	88,3 %
Mehrnami a kol. [11]	zákl. a nelineární vlastnosti	LDA	49	91,83 %
Chen a kol. [14]	MSCNN	CTSAM + kontrastivní učení	147	85,64 %
Současná práce	DWT + ERP	SVM	50	95,09 %

Porovnání výsledků

Výsledky dosažené v této disertační práci významně převyšují většinu stávajících studií zaměřených na detekci lži pomocí EEG. Použitím kombinace metod časové a časově-frekvenční analýzy, výběru vlastností pomocí GA a algoritmu SVM pro klasifikaci se podařilo dosáhnout průměrné přesnosti 95,09 %. Tato hodnota je pozoruhodná v kontextu aktuální literatury, kde se přesnosti dosahované různými metodami extrakce a klasifikace pohybují v rozmezí 81,03 % až 96,8 %, přičemž pouze jedna studie dosáhla vyšší přesnosti než současná práce.



Obr. 46.: Srovnání současného přístupu se stávajícími studiemi

Použité metody extrakce vlastností

V této studii byla využita kombinace metod DWT a ERP pro extrakci vlastností, což se ukázalo jako velmi efektivní. Jiná studie, Bablani a kol. [4], dosáhla srovnatelné přesnosti 96,8 % pomocí samotné metody WT, což potvrzuje, že časově-frekvenční analýza je velmi vhodná pro tento typ aplikací.

Použitý klasifikační algoritmus

Klasifikační algoritmus SVM použitý v této studii se ukázal jako nejúčinnější a nejvýkonnější mezi testovanými algoritmy, zejména při použití kombinovaných metod extrakce vlastností. Výzkum Bablani a kol. [4] rovněž prokázal vysokou přesnost při použití SVM, čímž se potvrzuje vhodnost tohoto algoritmu pro detekci lži pomocí EEG.

Počet subjektů

Tato studie dosáhla průměrné přesnosti klasifikace 95,09 %, což ji řadí mezi nejúspěšnější metody v této oblasti. Přestože jedna studie, Bablani a kol. [4], dosáhla o něco vyšší přesnosti (96,8 %), je důležité poznamenat, že současná

studie pracovala s větším vzorkem (50 subjektů) než většina ostatních studií, které často pracovaly s relativně malým počtem subjektů (10 subjektů). Toto rozšíření počtu subjektů přispívá k vyšší důvěryhodnosti, diverzitě a generalizovatelnosti výsledků. Chen a kol. [14] sice pracovali s větším vzorkem (147 subjektů), ale dosažená přesnost 85,64 % je výrazně nižší než přesnost v současné studii.

Vliv počtu kanálů

Současná studie také ukazuje, že detekce lži pomocí EEG může být účinná i s menším počtem kanálů. Výsledky pomocí vytvořeného datového souboru s využitím jednoho kanálu EEG byly lepší než výsledky získané z jiných vícekanálových datových souborů. To naznačuje, že správně vybrané a zpracované signály z jednoho kanálu mohou poskytnout dostatečné informace pro přesnou klasifikaci lži. Tento přístup navíc snižuje složitost a náklady spojené s měřením, což z něj činí atraktivní řešení pro praktické aplikace.

Unikátní aspekty současné studie

Jedním z významných přínosů této práce je použití nízkonákladového zařízení EEG, což demonstruje, že vysoké přesnosti lze dosáhnout i s méně nákladným vybavením. To je zvláště důležité pro praktické aplikace, kde náklady na vybavení mohou být limitujícím faktorem. Další důležitou inovací je použití GA pro výběr vlastností v kombinaci s klasifikátorem SVM, což poskytuje vysokou přesnost klasifikace. Metoda GA umožnila efektivní identifikaci relevantních vlastností z vysokodimenzionálních dat, což vedlo k lepšímu rozpoznání ERP komponenty P300. Použití SVM s RBF jádrem bylo klíčové díky jeho schopnosti provádět nelineární mapování dat do vyšší dimenze, což vedlo k optimálním výsledkům a vyšší úspěšnosti modelu.

Dílčí závěr

Výsledky této studie jsou výjimečné, s průměrnou přesností klasifikace 95,09 %, což převyšuje výsledky většiny zmíněných studií a přináší významné přínosy pro oblast detekce lži pomocí EEG v reakci na známé a neznámé tváře. Vysoká přesnost dosažená pomocí nízkonákladového vybavení, pokročilých metod strojového učení a výběru vlastností ukazuje na potenciál pro praktické aplikace a další výzkum. Použití nízkonákladového zařízení je významné, neboť ukazuje, že i s méně nákladným vybavením při použití pouze jednoho kanálu je možno dosáhnout vysoké úrovně přesnosti. To ukazuje na proveditelnost detekce lži z EEG signálů s menším počtem kanálů, což umožňuje nákladově efektivní implementaci. Větší počet subjektů ve srovnání s většinou ostatních studií zvyšuje validitu a generalizovatelnost výsledků, což představuje významný přínos k dalšímu rozvoji této oblasti výzkumu.

6. VÝVOJ APLIKACE PRO DETEKCI LŽI

Vytvoření aplikace pro detekci lži pomocí EEG signálů bylo motivováno potřebou zjednodušit a urychlit proces identifikace klamání. Tradiční metody analýzy EEG dat a následná klasifikace vyžadují značné odborné znalosti a jsou časově náročné. V současné době existuje snaha o vývoj nástrojů, které tento proces automatizují a zpřístupní širší veřejnosti.

EEG signály, konkrétně ERP komponenta P300, jsou klíčovým nástrojem pro analýzu mozkové aktivity při detekci lži. V této práci byly analyzovány EEG signály získané během experimentálních sezení za účelem binární klasifikace subjektů na vinné a nevinné. Tento proces zahrnuje komplexní kroky, jako je extrakce a výběr užitečných vlastností z naměřených amplitud P300 a následné použití metod strojového učení. Navržený rámec pro detekci lži využívá analýzu ERP komponenty P300 k rozhodnutí o vině nebo nevině subjektů. Z důvodu vysoké náročnosti tohoto procesu byla vyvinuta speciální aplikace, která umožňuje automatickou a rychlou detekci lži pomocí EEG signálů. Cílem této kapitoly je představit vývoj a funkčnost navržené aplikace, jejíž implementace usnadňuje a zrychluje proces detekce lži, čímž zvyšuje její dostupnost a efektivitu pro praktické využití v různých oblastech, včetně bezpečnosti a právního systému.

Zjednodušení detekčního postupu

Pro zajištění přístupnosti a aplikovatelnosti detekce lži pomocí EEG byl navržen proces zjednodušení detekčního postupu pro běžné uživatele. Byla vyvinuta aplikace, která umožňuje rychlou a jednoduchou detekci lži bez potřeby hlubších znalostí v dané problematice. Aplikace poskytuje výstup ve formě zprávy indikující, zda subjekt lže či nelže, což výrazně zjednodušuje celý proces. Dále byla vytvořena dokumentace s podrobným návodem pro proces detekce lži pomocí EEG, zahrnující pokyny pro snímání dat během navrženého experimentu a použití aplikace.

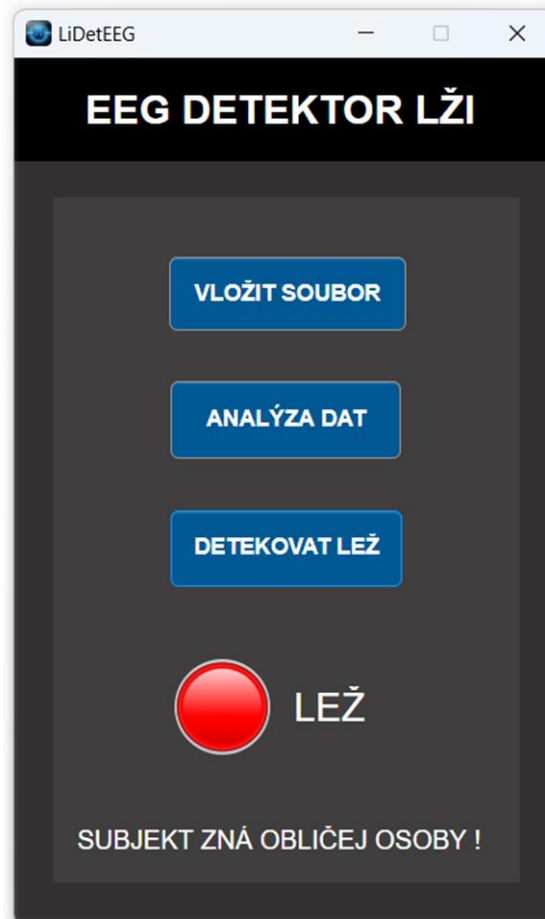
Ošetření výjimek a chyb

Aby byla aplikace uživatelsky přívětivější, byla implementována opatření pro ošetření různých výjimek a možných chyb. To zahrnuje odstranění artefaktů, správné načtení souboru, ošetření chyb v předzpracování dat a zajištění správného provedení všech kroků nezbytných pro funkčnost aplikace.

Vývoj aplikace

Aplikace byla vytvořena pomocí nástroje App Designer v Matlabu, což zaručuje moderní design a funkčnost GUI. Tento systém detekuje lež prostřednictvím analýzy ERP komponenty P300 získané z EEG signálů. Klíčové vlastnosti byly vybrány pomocí GA z ERP a DWT dat a následně zpracovány pomocí algoritmu SVM. Jednotlivé metody byly zvoleny na základě výsledků

provedených experimentů a vyvinuté metodiky v předchozích částech práce. Tento přístup zajišťuje vysokou přesnost a spolehlivost klasifikace, čímž přispívá k vývoji detekce lži za použití pokročilých technologií a strojového učení. Výsledná aplikace je připravena k distribuci jako samostatný exe soubor bez nutnosti vlastnění licence k Matlabu.



Obr. 47.: Aplikace pro detekci lži

Funkčnost aplikace

Aplikace umožňuje rozlišení subjektů na vinné a nevinné. Uživatel nejprve načte data z provedeného experimentu pomocí tlačítka „Vložit soubor“. To umožní načtení a úpravu dat do vhodného formátu pro následnou analýzu, která je provedena po kliknutí na tlačítko „Analýza dat“. Poté uživatel stiskne tlačítko "Detekce lži", načež program načte analyzovaná data z předchozího kroku, provede predikci a zobrazí výsledek ve formě zprávy "LEŽ" nebo "PRAVDA". Výsledek „LEŽ“ naznačuje, že daný subjekt zná obličej konkrétní osoby, což může indikovat zatajení informací souvisejících s trestným činem.

Uživatelské rozhraní a interakce

Aplikace je postavena na GUI, které umožňuje uživatelům jednoduše načítat a analyzovat data pomocí interaktivních prvků. GUI poskytuje intuitivní tlačítka, která vedou uživatele k jednotlivým krokům procesu detekce lži.

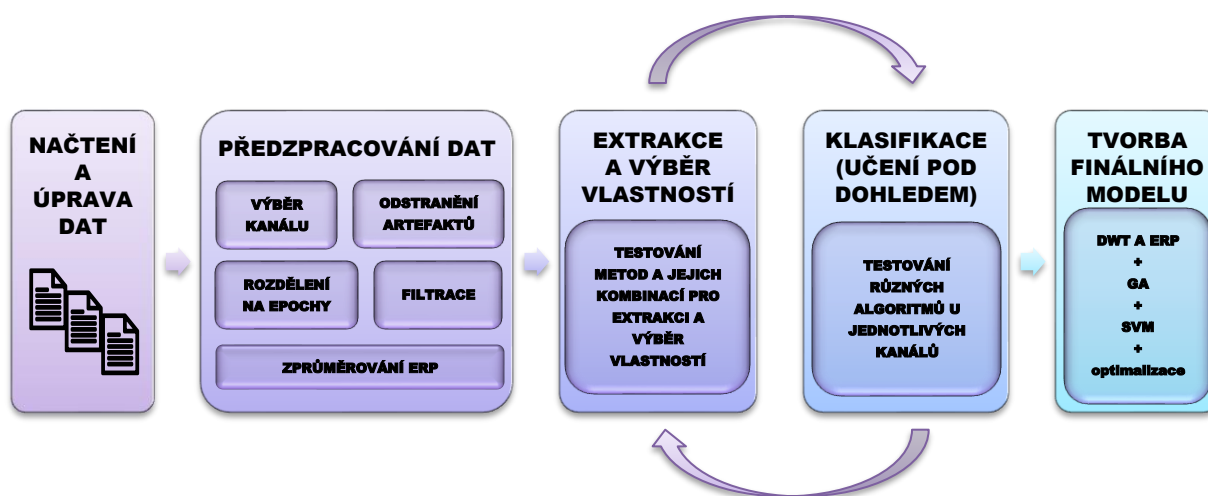
GUI komponenty

- **Tlačítka:** Pro načítání dat, spuštění analýzy a zobrazení výsledků.
- **Dialogová okna:** Pro výběr souboru a zobrazení chybových zpráv.
- **Zprávy:** Informují uživatele o úspěšnosti nebo neúspěšnosti jednotlivých kroků.

Trénink a predikce modelu

Proces trénování modelu se opakuje, dokud není dosaženo optimálního modelu. Následně je natrénovaný model integrován do aplikace pro predikci.

Trénink (opakuje se, dokud se nezíská nejlepší model)



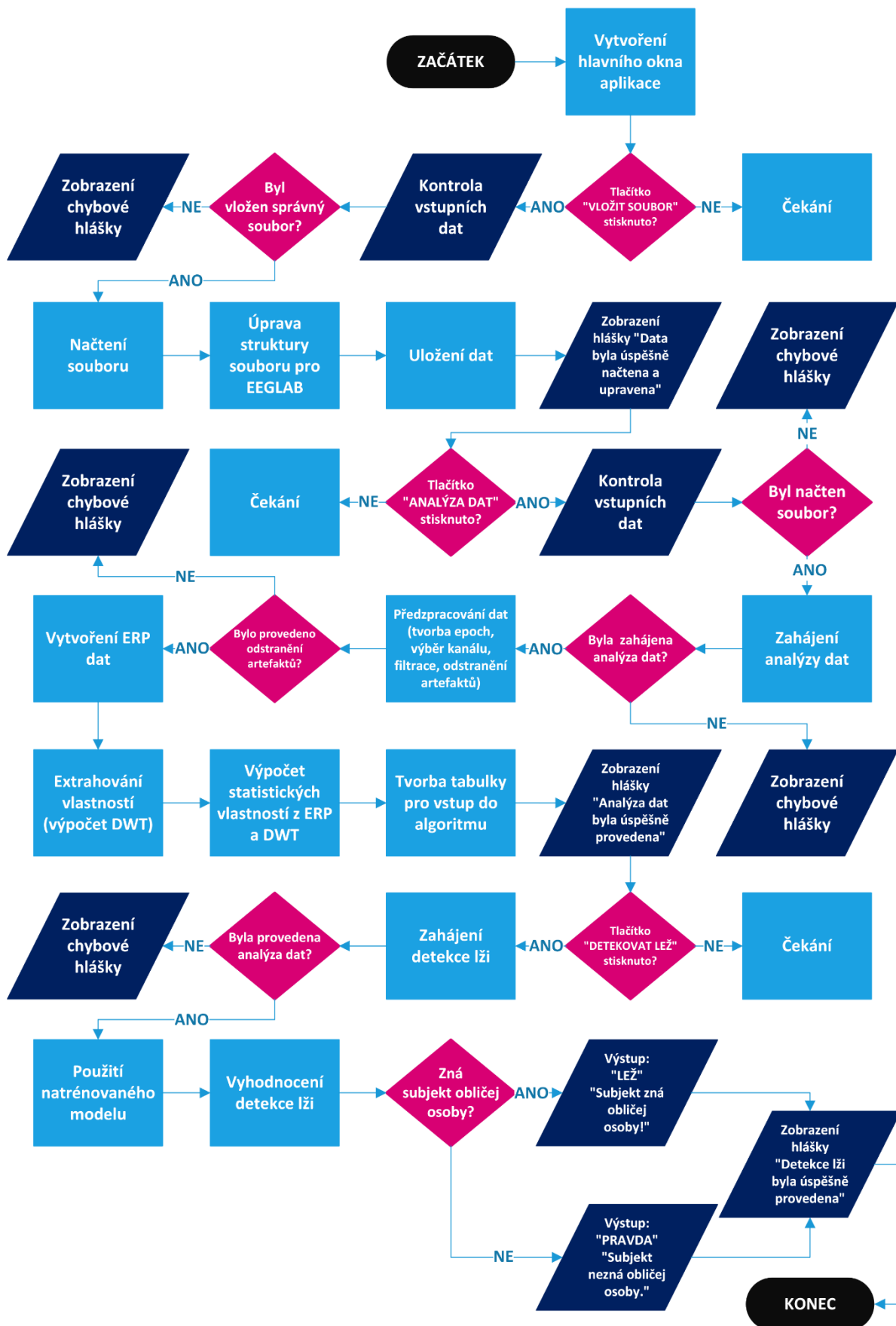
Predikce (integrace natrénovaného modelu do aplikace)



Obr. 48.: Pracovní postup aplikace

Vývojový diagram aplikace

Vývojový diagram ilustruje krok za krokem celý proces detekce lži pomocí EEG signálů, zahrnující načítání souborů, předzpracování dat, extrakci vlastností a klasifikaci.



Obr. 49.: Vývojový diagram aplikace pro detekci lži

6.1 Popis hlavních funkcí aplikace

Hlavní funkce spolupracují na automatizaci procesu detekce lži pomocí EEG, od načtení souboru a extrakce vlastností až po identifikaci možných lží. Důležitou částí je robustní ochrana proti chybám a důkladné předzpracování dat, což zajišťuje spolehlivost a přesnost výsledků.

6.1.1 Načtení dat

Tato část programu obsahuje specifické funkce, které načítají a upravují data do vhodného formátu pro další zpracování z CSV souboru vybraného uživatelem, přičemž každá část obsahuje ochranu proti chybám a kontrolu kvality dat.

Hlavní funkce *LoadDataPushed*

- **Výběr souboru:** Funkce se spustí po kliknutí na tlačítko „Vložit soubor“. Umožňuje uživateli vybrat soubor s daty ve formátu CSV a provádí jejich načtení a úpravu.
- **Import dat:** Funkce *importAndSaveData* volána pro načtení a úpravu dat.
- **Zobrazení chybových zpráv:** Pokud dojde k chybě během načítání souboru, zobrazí se chybová zpráva.

```
% Funkce pro načtení a úpravu dat
function LoadDataPushed(app, event)

    try
        % Import dat ze souboru
        [fname, fdir] = uigetfile('*.csv*', 'Text Files (*.csv)');
        loadedFile = fullfile(fdir, fname);

        % Ošetření situace, kdy uživatel zruší výběr souboru
        if isequal(fname, 0)
            msgbox('Výběr souboru byl zrušen', ' ', 'modal');
            return;
        end

        % Import a úprava dat
        importAndSaveData(loadedFile);

        if isempty(app.editData)
            else
                msgbox('Data byla úspěšně načtena a upravena', ' ');
            end

    catch ME
        msgbox(['Soubor nelze načíst, vyberte prosím podporovaný formát
souboru (.csv). Chyba: ', ME.message], 'Chyba', 'error');
    end

end
```

Zdrojový kód 2: Implementace funkce pro načtení a úpravu dat

Funkce *importAndSaveData*

- **Načítání dat:** Načítá data ze zvoleného CSV souboru a upravuje je.
- **Výběr relevantních sloupců:** Vybere sloupce obsahující EEG data na základě názvů sloupců.
- **Konverze na pole:** Vybraná data jsou konvertována na pole pro další zpracování.
- **Uložení dat:** Data jsou uložena do nového CSV souboru ve formátu ASCII.

```
% Funkce pro import dat ze souboru a uložení do proměnné
function importAndSaveData(loadedFile)

    try
        % Import dat ze souboru
        data = readtable(loadedFile);

        % Výběr sloupců podle názvů
        selectedData = data(:, {'EEG_AF3', 'EEG_T7', 'EEG_Pz', 'EEG_T8',
'EEG_AF4'});

        % Konverze vybraných dat na pole (odstranění názvů proměnných)
        eegdata = table2array(selectedData);

        % Uložení upraveného souboru a převod do ascii formátu
        save EEG.csv eegdata -ascii;

        % Uložení dat do globální proměnné
        app.editData = 'EEG.csv';

    catch ME
        % Zobrazení chybové zprávy
        disp(['Došlo k chybě: ', ME.message]);
        msgbox('Soubor nelze načíst, vyberte soubor se správnou strukturou
ve formátu .csv', 'Chyba načtení souboru', 'error');

        % Návrat prázdné proměnné v případě chyby
        app.editData = [];
    end
end
```

Zdrojový kód 3: Implementace funkce pro import dat

6.1.2 Analýza dat

Tato část programu se zaměřuje na analýzu EEG dat. Nejprve je provedeno předzpracování dat, včetně rozdělení dat na epochy, výběru kanálu a aplikace různých filtrů pro odstranění artefaktů, což je klíčové pro získání kvalitních a použitelných dat pro následnou analýzu. Data ERP jsou poté zprůměrována a je aplikována DWT, pro kterou byl vytvořen vlastní algoritmus. Následně je provedena extrakce jednotlivých statistických vlastností z ERP a DWT dat. Celkově tato funkce umožňuje komplexní automatizovanou analýzu EEG dat

zahrnující předzpracování signálů s cílem extrahovat relevantní vlastnosti pro následnou klasifikaci, přičemž ošetřuje možné chyby a kontroluje provedení předchozího kroku.

Hlavní funkce *AnalysisPushed*

- **Analýza dat:** Funkce se spustí při kliknutí na tlačítko „Analýza dat“ a kontroluje, zda byl soubor načten.
- **Ošetření výjimek:** Kontroluje existenci potřebných nástrojů, validuje vstupní soubor a podobně.
- **Hlavní kroky analýzy dat:** Filtrace, import eventů a event listu, rozdělení na epochy, výběr kanálu, odstranění artefaktů, převzorkování, průměrování ERP, uložení dat do proměnných pro jednotlivé podněty, implementace DWT, výpočet statistických vlastností na datech ERP a DWT a tvorba tabulky obsahující jednotlivé statistické vlastnosti vstupující do algoritmu.

Funkce *removeArtifacts*

- **Odstranění artefaktů:** Automatická detekce a odstranění artefaktů z EEG dat.
- **Kontrola odstraněných dat:** Zajišťuje, že zůstanou jen dostatečně kvalitní data pro další analýzu. Pokud je odstraněno více než 80 % dat nebo všechna data, je vyvolána chyba upozorňující na potřebu nového měření.

```
% Funkce pro odstranění artefaktů
function EEG = removeArtifacts(EEG)

    % Počet epoch před odstraněním artefaktů
    numEpochsBefore = size(EEG.data, 3);

    % Odstranění artefaktů na základě automatické detekce
    EEG = pop_artextval(EEG, 'Channel', 1, 'Flag', 1, 'LowPass', -1,
        'Threshold', [-100 100], 'Twindow', [-50 950]);

    % Získání indexů epoch označených jako artefakty
    rejectedEpochs = find(EEG.reject.rejmanual);

    % Odstranění označených epoch
    if ~isempty(rejectedEpochs) % ošetření, pokud není nalezen žádný
    artefakt
        EEG = pop_rejepoch(EEG, rejectedEpochs, 0);
    end

    % Počet epoch po odstranění artefaktů
    numEpochsAfter = size(EEG.data, 3);

    % Výpočet procentuálního podílu odstraněných dat
    removedPercentage = (numEpochsBefore - numEpochsAfter) /
    numEpochsBefore * 100;
```

```

% Kontrola, zda bylo odstraněno více než 80 % dat jako artefakty
if removedPercentage > 80
    error('Data obsahují mnoho artefaktů. Provedte nové měření pro
dosažení kvalitnějších výsledků.');
```

end

```

if isempty(EEG.data)
    error('Všechna data byla odstraněna jako artefakty. Provedte nové
měření.');
```

end

end

Zdrojový kód 4: Implementace funkce pro odstranění artefaktů

Funkce *my_dwt_multilevel*

- **Provádí víceúrovňovou DWT:** Každá úroveň transformuje signál pomocí funkce *my_dwt* a výstupní aproximace slouží jako vstup pro další úroveň.
- **Výstup:** Výstupem jsou jednotlivá pole obsahující aproximační a detailní koeficienty pro každou úroveň.

```

% Funkce pro výpočet víceúrovňové DWT
function [approximations, details] = my_dwt_multilevel(x, level, wavelet)

% Inicializace buněk pro aproximace a detaily
approximations = cell(1, level);
details = cell(1, level);

% DWT
for i = 1:level
    [approximations{i}, details{i}] = my_dwt(x, wavelet);
    x = approximations{i}; % Přepíše data pro další úroveň transformace
end
end
```

Zdrojový kód 5: Implementace funkce pro víceúrovňovou DWT

Funkce *my_dwt*

- **Provádí jednoúrovňovou DWT:** Používá předdefinované filtrační koeficienty *dec_lo* a *dec_hi* pro vlnku db4.
- **Aplikace koeficientů filtru:** Pro každý prvek výsledného signálu prochází příslušné části vstupního signálu a aplikuje koeficienty filtru.

```

% Funkce pro výpočet DWT
function [approximation, detail] = my_dwt(x, wavelet)

% Filtrační koeficienty pro DWT - vlnka db4
dec_lo = [0.4829629131445341, 0.8365163037378079, 0.2241438680420134, -
0.1294095225512604];
```

```

    dec_hi = [-0.1294095225512604, -0.2241438680420134, 0.8365163037378079,
-0.4829629131445341];

    % Inicializace aproximace a detailu
    approximation = zeros(1, length(x)/2);
    detail = zeros(1, length(x)/2);

    % Provedení DWT
    for i = 1:2:length(x)
        for j = 1:length(dec_lo)
            k = i + j - 1;
            if k <= length(x)
                approximation((i+1)/2) = approximation((i+1)/2) + dec_lo(j)
* x(k);
                detail((i+1)/2) = detail((i+1)/2) + dec_hi(j) * x(k);
            end
        end
    end
end
end

```

Zdrojový kód 6: Implementace funkce pro jednoúrovňovou DWT

Funkce *calculate_statistics*

- **Počítá statistické vlastnosti:** Vybrané statistické vlastnosti zahrnují maximum, minimum, peak-to-peak, průměr, směrodatnou odchylku, medián, šikmost, špičatost, entropii, aktivitu, mobilitu, složitost, energii a výkon. Tyto vlastnosti jsou vypočteny v jednotlivých frekvenčních pásmech.
- **Výstup:** Výstupem je struktura obsahující všechny vypočítané vlastnosti.

```

% Výpočet statistických vlastností
function stats = calculate_statistics(details, approximations)

    % Výpočet statistických vlastností pro jednotlivá pásma
    delta = approximations{5}';
    theta = details{5}';
    alfa = details{4}';
    beta = details{3}';

    % Výpočet Hjorth parametrů
    hjorth_params_delta = hjorth_parameters(delta);
    hjorth_params_theta = hjorth_parameters(theta);
    hjorth_params_alfa = hjorth_parameters(alfa);
    hjorth_params_beta = hjorth_parameters(beta);

    % Výpočet energie
    energy_delta = sum(delta .^ 2);
    energy_theta = sum(theta .^ 2);
    energy_alfa = sum(alfa .^ 2);
    energy_beta = sum(beta .^ 2);

```

```

% Výpočet výkonu
power_delta = mean(delta .^ 2);
power_theta = mean(theta .^ 2);
power_alfa = mean(alfa .^ 2);
power_beta = mean(beta .^ 2);

stats = [max(delta), min(delta), max(delta) - min(delta), mean(delta),
std(delta), median(delta), skewness(delta), kurtosis(delta),
entropy(delta), hjorth_params_delta, energy_delta, power_delta, ...
max(theta), min(theta), max(theta) - min(theta), mean(theta),
std(theta), median(theta), skewness(theta), kurtosis(theta),
entropy(theta), hjorth_params_theta, energy_theta, power_theta, ...
max(alfa), min(alfa), max(alfa) - min(alfa), mean(alfa),
std(alfa), median(alfa), skewness(alfa), kurtosis(alfa), entropy(alfa),
hjorth_params_alfa, energy_alfa, power_alfa, ...
max(beta), min(beta), max(beta) - min(beta), mean(beta),
std(beta), median(beta), skewness(beta), kurtosis(beta), entropy(beta),
hjorth_params_beta, energy_beta, power_beta];

end

```

Zdrojový kód 7: Implementace funkce výpočtu statistických vlastností

6.1.3 Detekce lži

Tato část programu je klíčovým krokem aplikace pro detekci lži pomocí EEG a je navržena tak, aby využívala analyzovaná data z předchozí části a prováděla predikci pomocí předem natrénovaného modelu strojového učení. Funkce nejprve ověřuje, zda byla provedena analýza dat, a pokud ne, upozorní uživatele. Výsledkem predikce je určení, zda jde o lež nebo pravdu, což se vizuálně zobrazuje uživateli prostřednictvím změny textových popisků a barevných indikátorů. Pokud dojde k chybě, uživatel je o této skutečnosti informován pomocí chybové zprávy.

Hlavní funkce *LieDetectionPushed*

- **Spuštění funkce:** Spustí se, když uživatel klikne na tlačítko „Detekovat lež“.
- **Kontrola:** Zjišťuje, zda byla provedena analýza dat.
- **Příprava dat:** Přípravuje data pro následnou predikci.
- **Načtení modelu:** Načítá předtrénovaný model strojového učení.
- **Provádění predikce:** Používá model k provedení predikce na vstupních datech.
- **Zobrazení výsledku:** Zobrazuje výsledek predikce uživateli.

```

function LieDetectionPushed(app, event)

% Tělo funkce pro detekci lži
Try
    % Kontrola, zda byla provedena analýza dat

```

```

    if isempty(app.finalni_tabulka_variable)
        msgbox('Analýza dat nebyla provedena. Proved'te ji pomocí
tlačítka "Analýza dat"', ' ', 'modal');
        return;
    end

    data_for_prediction = app.finalni_tabulka_variable;

    % Načtení modelu pro predikci
    loadedModel = load('FinalModelSVM.mat');

    % Přístup k natrénovanému modelu
    trainedModel = loadedModel.FinalModelSVM;

    % Predikce pomocí natrénovaného modelu
    [yfit, ~] = trainedModel.predictFcn(data_for_prediction);

    % Predikce na základě modelu
    if any(yfit == 1)
        app.NEZNAOBLICEJOSOBYLabel.Visible = 'off';
        app.ZNAOBLICEJOSOBYLabel.Visible = 'off';
        app.PRAVDALamp.Visible = 'off';
        app.Label.Text = " ";
        app.Label_2.Text = " ";

        app.Label.Text = "LEŽ";
        app.PRAVDALamp.Visible = 'on';
        app.PRAVDALamp.Color = 'r';
        app.ZNAOBLICEJOSOBYLabel.Visible = 'on';
    else
        app.NEZNAOBLICEJOSOBYLabel.Visible = 'off';
        app.ZNAOBLICEJOSOBYLabel.Visible = 'off';
        app.PRAVDALamp.Visible = 'off';
        app.Label.Text = " ";
        app.Label_2.Text = " ";

        app.Label_2.Text = "PRAVDA";
        app.PRAVDALamp.Visible = 'on';
        app.PRAVDALamp.Color = 'g';
        app.NEZNAOBLICEJOSOBYLabel.Visible = 'on';
    end

    % Zobrazení hlášky
    msgbox('Detekce lži byla úspěšně provedena', ' ', 'modal');

    catch ME
        msgbox(['Došlo k chybě při detekci lži: ', ME.message], 'Chyba',
'error');
    end
end

```

Zdrojový kód 8: Implementace funkce pro detekci lži

6.2 Distribuce aplikace pro detekci lži pomocí EEG

Detekce lži pomocí EEG signálů prostřednictvím zařízení Emotiv Insight představuje klíčový prvek této studie. Tento nízkonákladový přístup umožňuje snímání mozkových aktivit s využitím dostupné technologie, která je k dispozici široké veřejnosti.

Automatizovaný proces detekce lži zahrnuje snímání dat pomocí navrženého experimentu a následnou analýzu EEG dat pomocí aplikace. Získávání dat probíhá pomocí zařízení Emotiv Insight a placeného softwaru EmotivPro, který je nezbytný pro sběr dat z EEG zařízení. Pro analýzu naměřených dat je použita aplikace LiDetEEG.exe, která vyžaduje instalaci bezplatného Matlab Runtime R2020b (9.9) nebo placené verze Matlab R2020b.

V této disertační práci byl sestaven podrobný manuál pro proces detekce lži pomocí EEG, který zahrnuje vzorovou prezentaci pro experiment, vyvinutou aplikaci a kompletní dokumentaci s podrobnými instrukcemi správného postupu pro získávání dat a použití aplikace. Veškeré nezbytné materiály pro detekci lži pomocí EEG jsou k dispozici jako příloha disertační práce (viz PŘÍLOHA P IX: DOKUMENTY A APLIKACE PRO DETEKCI LŽI POMOCÍ EEG).

7. VÝSLEDKY A DISKUSE

Disertační práce se zaměřila na vývoj systému pro automatickou, rychlou a efektivní detekci intencionální lži pomocí analýzy EEG signálů snímaných nízkonákladovým zařízením, s využitím pokročilých metod výběru vlastností a strojového učení. Cílem bylo identifikovat nejvhodnější kombinaci metod pro extrakci a výběr vlastností a klasifikaci, která zajistí co nejvyšší přesnost detekce lži pomocí EEG.

Metodika

Byla navržena a ověřena metodika pro celý proces detekce lži pomocí EEG, která zahrnovala podrobný popis jednotlivých fází analýzy dat. Navrhovaný přístup poskytuje nový pohled na využití nízkonákladových zařízení pro detekci lži s vysokou mírou přesnosti a spolehlivosti.

Experiment

Experimentální návrh se zaměřil na detekci skrytých informací pro identifikaci podvodu pomocí protokolu CIT založeného na paradigmatu ERP P300, sledující mozkovou aktivitu reagující na vizuální podněty známých a neznámých tváří. Byl vytvořen scénář falešného zločinu, kde se zjišťovalo, zda je subjekt obeznámen s danou tváří či nikoli. EEG signály byly získány pomocí zařízení Emotiv Insight a softwaru EmotivPro.

Předzpracování dat a výběr elektrod

Předzpracování zahrnovalo filtrování, odstranění artefaktů, rozdělení signálů na jednotlivé epochy a výpočty ERP průměrů pro sondy a irelevantní podněty. Pro analýzu byly testovány elektrody Pz, AF3, AF4, T7 a T8, přičemž nejvyšší přesnost byla dosažena na elektrodě Pz. Data byla zpracována pomocí softwaru Matlab a jeho nástrojů EEGLAB a ERPLAB.

Extrakce vlastností

Byly použity různé metody extrakce vlastností v časové i časově-frekvenční doméně. Nakonec bylo vybráno 70 statistických vlastností, které byly aplikovány na data ERP a DWT. Kombinace těchto metod zvyšuje klasifikační přesnost, poskytuje komplexnější pohled na signál EEG a zlepšuje detekci lži.

Výběr vlastností

Před přistoupením ke klasifikaci byl proveden výběr optimální sady vlastností pomocí metod GA, ANOVA a PCA za účelem zlepšení úspěšnosti klasifikace. Optimalizace sady vlastností byla provedena pomocí GA, který se osvědčil jako účinný pro identifikaci relevantních vlastností, což vedlo ke zlepšení generalizace na nová data a snížení výpočetní náročnosti ze 70 vlastností na 35. Vybrané vlastnosti byly následně zadány jako vstup do klasifikačního algoritmu.

Klasifikace

Pro klasifikaci byly využity různé metody strojového učení, konkrétně SVM, KNN, DT a NB, aplikované na data z jednotlivých elektrod s různými kombinacemi metod pro extrakci a výběr vlastností. Nejvyšší přesnosti klasifikace bylo dosaženo pomocí algoritmu SVM na elektrodě Pz s výběrem vlastností pomocí GA z dat ERP a DWT, který dosáhl přesnosti až 95,6 %. Použití SVM s Gaussian jádrem na EEG signály zaměřené na analýzu složky P300 vykazuje vysokou přesnost klasifikace, což demonstruje jeho efektivitu a potenciál pro aplikace v reálném světě, jako je detekce lži. SVM byl vybrán pro svou schopnost udržet rovnováhu mezi přesností na trénovací sadě dat a schopností generalizace na testovací data. Rozdělení do dvou tříd na vinné a nevinné bylo provedeno pomocí klasifikátoru SVM založeného na analýze komponenty ERP P300.

Validace a robustnost modelu

Pro validní statistické vyhodnocení výsledků bylo do studie zařazeno 50 subjektů. Model pro trénování algoritmu byl hodnocen na 30 subjektech s desetinásobnou křížovou validací a dosáhl průměrné klasifikační přesnosti 95,09 %, což potvrzuje jeho účinnost v detekci lži pomocí EEG signálů. Zbylé subjekty byly ponechány jako validační pro ověření robustnosti modelu a schopnosti generalizovat na nová data.

Optimalizace modelu

Optimalizace parametrů SVM umožnila přesnější predikci modelu na nových datech. Po výběru nejrelevantnějších vlastností pomocí GA a optimalizaci parametrů SVM byl vytvořen finální model, který dokázal správně klasifikovat 43 z 46 celkových subjektů a 14 z 16 nových subjektů, čímž vykázal robustnost při testování na nových datech.

Interpretace výsledků

Výsledky ukazují, že navržená metodika byla schopna správně detekovat většinu subjektů díky kombinaci ERP a DWT, na které byl aplikován výběr vlastností pomocí GA pro vstup do klasifikátoru SVM. Dále bylo prokázáno, že známé podněty vyvolaly u vinných subjektů P300 s vysokou amplitudou, což podporuje hypotézu, že P300 může být využita k detekci skrytých informací zaměřených na rozpoznávání známých tváří.

Srovnání se stávajícími studii

Výsledky této práce výrazně předčily přesnost dosaženou ve většině současných studií, což potvrzuje efektivitu navržené metodiky a významný přínos pro oblast detekce lži pomocí EEG.

Praktická uplatnitelnost a nákladová efektivita

Pro zjednodušení a zefektivnění sběru dat lze využít uživatelsky přívětivá EEG zařízení. Použitím nízkonákladového zařízení, pokročilých metod výběru vlastností a strojového učení bylo dosaženo vysoké přesnosti klasifikace, naznačující proveditelnost a nákladovou efektivitu navrženého postupu, což je klíčové pro praktické využití v oblasti bezpečnosti a kriminalistiky.

Vývoj aplikace pro detekci lži

Proces detekce lži pomocí EEG je náročný jak časově, tak znalostně. Proto byla vyvinuta aplikace včetně návodu pro snímání mozkové aktivity, která usnadňuje automatickou detekci lži pomocí EEG s využitím nízkonákladové náhlavní soupravy Emotiv Insight. Aplikace umožňuje analýzu EEG dat bez potřeby hlubších znalostí v oblasti neurovědy či programování, čímž zjednodušuje a urychluje celý proces detekce lži a zvyšuje praktickou aplikovatelnost výsledků.

Diskuse

Výsledky naznačují, že kombinace metod výběru vlastností z časové a časově-frekvenční domény na elektrodě Pz a optimalizace parametrů klasifikačních algoritmů se ukazuje jako klíčová pro dosažení vysoké přesnosti klasifikace. Navrhovaný přístup vykazuje slibné výsledky pro detekci lži pomocí EEG signálů se zaměřením na ERP komponentu P300.

Výzvy a omezení

Implementace této metody do praxe představuje dlouhodobější výzvu. Vyžaduje zvýšení povědomí o této technice detekce lži, interdisciplinární přístup a další výzkum k vylepšení současných postupů a rozšíření aplikací v praktických scénářích. Tato metoda je komplexní a vyžaduje rozsáhlé znalosti v oblasti neurovědy, zpracování signálů a strojového učení. Pro zlepšení přesnosti klasifikace by bylo vhodné aplikovat pokročilejší metody pro klasifikaci, optimalizaci, extrakci a výběr vlastností. Dále je důležité provádět validace na větších a různorodějších souborech dat, aby bylo možné lépe ověřit robustnost a generalizovatelnost navržených metod.

Etické aspekty

Snímání mozkové aktivity může vyvolávat etické otázky a obavy o soukromí, zejména v kontextu detekce lži. Účastníci mohou cítit nepohodlí nebo úzkost z možnosti odhalení skrytých informací. Je nezbytné, aby integrace EEG technologií v reálných scénářích respektovala práva jednotlivců a byla prováděna s vysokým standardem etické odpovědnosti. Tyto aspekty musí být pečlivě zváženy při návrhu a implementaci systémů pro detekci lži pomocí EEG. Zajištění soukromí a ochrany osobních údajů subjektů je klíčové. Účastníci musí být plně informováni o účelu studie a způsobu použití jejich dat, a musí být přijata opatření

ke snížení nepohodlí a obav účastníků. Tato technologie může být užitečná v bezpečnostních kontrolách či psychologických testech, vyžaduje však rigorózní testování a přísná etická pravidla pro ochranu soukromí jednotlivců [64].

Možnosti zavedení do praxe

Zavedení detekce lži pomocí EEG do praxe vyžaduje zajištění vysoké úrovně přesnosti v různých reálných scénářích, standardizaci metodiky a technologií a řešení právních, etických a regulačních otázek. Přestože detekce lži pomocí EEG ukazuje slibné výsledky v laboratorních podmínkách, její zavedení do praxe čelí několika výzvám. Jednou z hlavních výzev je zajištění vysoké úrovně přesnosti a spolehlivosti v různých reálných scénářích, jako jsou policejní vyšetřování nebo soudní procesy. Dalším faktorem je potřeba standardizace metodiky a technologií používaných pro detekci lži pomocí EEG. Současná literatura obsahuje různé přístupy k předzpracování signálu, extrakci vlastností a klasifikaci, což ztěžuje porovnání výsledků mezi různými studiemi. Pro zavedení do praxe je tedy nezbytné vyvinout standardizované protokoly a metodiky, které zajistí konzistentní a spolehlivé výsledky. Implementace detekce lži pomocí EEG v praxi by také vyžadovala řešení právních a regulačních otázek. Různé jurisdikce mohou mít odlišné právní rámce pro používání těchto technologií, proto je nutné zajistit dodržování všech právních a etických norem, včetně ochrany osobních údajů, práv na soukromí a zajištění, že technologie nebude použita k porušování lidských práv. Detekce lži pomocí EEG, zejména s využitím ERP P300, představuje inovativní přístup s potenciálem výrazně zlepšit schopnost odhalovat skryté informace. Avšak zavedení této technologie do praxe vyžaduje pečlivé zvážení mnoha aspektů. Klíčové kroky zahrnují zajištění informovaného souhlasu, ochranu proti zneužití technologie a vývoj standardizovaných metodik, aby byla technologie nejen účinná, ale i eticky a právně přijatelná [3, 65].

Závěrečné shrnutí

Tato disertační práce přináší nový přístup k rozpoznávání skrytých informací s podrobným popisem celého procesu odhalování lži pomocí EEG. Byla navržena aplikace pro automatickou a rychlou detekci intencionální lži, se zaměřením na analýzu ERP komponenty P300, využívající pokročilé metody strojového učení a výběru vlastností. Výsledky ukázaly, že kombinace metod ERP a DWT s výběrem vlastností pomocí GA a optimalizovaným SVM modelem poskytuje vysokou přesnost klasifikace, přičemž elektroda Pz se ukázala jako nejrelevantnější. Navržený postup je proveditelný s nízkonákladovým EEG zařízením, což zvyšuje jeho praktickou aplikovatelnost v oblasti bezpečnosti a kriminalistiky. Vyvinutá aplikace dále usnadňuje celý proces detekce lži pomocí EEG, čímž zvyšuje přístupnost a využitelnost této metody pro široké spektrum uživatelů.

8. VYHODNOCENÍ CÍLŮ PRÁCE

Hlavním cílem této disertační práce byl vývoj systému pro efektivní a rychlou detekci lži pomocí EEG signálů snímaných nízkonákladovým zařízením, s využitím pokročilých metod výběru vlastností a strojového učení. Práce se zaměřuje na detekci intencionální lži prostřednictvím analýzy ERP komponenty P300. Měření byla provedena pomocí vizuálního ERP experimentu, který zkoumá neurofyziologické změny při pokusu o lhaní. Následně byly porovnány a vyhodnoceny rozdíly v průběhu vln P300 účastníků v reakcích na vizuální podněty známých a neznámých tváří pomocí GA pro výběr vlastností z DWT a ERP dat a algoritmu SVM pro klasifikaci. Navržený přístup dosáhl vysoké přesnosti klasifikace, což je důležité v mnoha aplikacích výzkumu, jako je bezpečnost. Níže je uvedeno bodové vyhodnocení jednotlivých dílčích cílů:

- **Vývoj a validace metodiky pro detekci lži pomocí EEG:** Byla vytvořena a ověřena komplexní metodika pro identifikaci intencionální lži pomocí analýzy EEG signálů.
- **Návrh a realizace vizuálního ERP experimentu:** Byl navržen a realizován experiment s použitím nízkonákladového EEG zařízení Emotiv Insight, který měřil ERP komponenty P300 při vystavení účastníků vizuálním podnětům známých a neznámých tváří.
- **Analýza dat EEG a volba vhodných kritérií:** Byly identifikovány a aplikovány optimální metody, parametry a kritéria pro zpracování a analýzu EEG dat zaměřených na detekci lži.
- **Klasifikace dat EEG pomocí strojového učení:** Byly využity různé pokročilé metody strojového učení, přičemž nejlepší výsledky byly dosaženy pomocí algoritmu SVM pro klasifikaci, v kombinaci s metodami DWT a ERP pro extrakci vlastností a GA pro výběr vlastností. Finální model dosáhl průměrné přesnosti klasifikace 95,09 %, což potvrzuje jeho vysokou spolehlivost při detekci lži.
- **Vývoj uživatelsky přívětivého systému:** Byl vyvinut systém umožňující rychlou a efektivní detekci lži pomocí nízkonákladového zařízení. Aplikace je navržena tak, aby byla uživatelsky přívětivá a dostupná i pro uživatele bez odborných znalostí v oblasti EEG. Díky navrženému paradigmatu a aplikaci lze detekovat lež přibližně za 10-15 minut, včetně měření.
- **Ověření a validace navrženého systému:** Výsledky práce dosáhly lepší přesnosti klasifikace než většina studií v této oblasti. Účinnost a přesnost navrženého systému byla dále testována a ověřena na nezávislých datových sadách, přičemž systém dosáhl vysoké úspěšnosti při predikci nových dat.

Dosažené výsledky potvrzují, že navržená aplikace pro zjednodušení procesu detekce lži pomocí nízkonákladového EEG zařízení je efektivní, rychlá a snadno použitelná, což je významné pro právní a bezpečnostní kontexty.

9. PŘÍNOS PRÁCE PRO VĚDU A PRAXI

Disertační práce přináší významné přínosy pro vědeckou komunitu i praktické aplikace. Následující body shrnují hlavní přínosy této práce:

Vědecký přínos

- **Inovativní použití nízkonákladového zařízení pro detekci lži pomocí EEG:** Studie dokládá, že i s využitím cenově dostupného EEG vybavení lze dosáhnout spolehlivých výsledků v detekci lži. Tento přístup rozšiřuje možnosti EEG výzkumu na širší škálu aplikací a může inspirovat další výzkumníky k využití nízkonákladových zařízení ve svých studiích.
- **Hlubší analýza a porozumění mozkovým procesům při lhaní:** Výzkum přispívá k hlubšímu pochopení vztahů mezi mozkovými vlnami, zejména ERP komponentou P300, a intencionální lží. Tento přístup zlepšuje chápání neurofyzologie lhaní a poskytuje základ pro budoucí výzkum v neurověděch a psychologii.
- **Integrace genetických algoritmů a strojového učení:** Využití GA a SVM pro optimalizaci a klasifikaci zlepšuje přesnost detekce lži a může být inspirací pro další výzkum v oblastech strojového učení a neurověd.
- **Analýza skrytých informací z mozku pomocí EEG:** Výzkum přináší nové možnosti pro analýzu mozkové aktivity a identifikaci skrytých informací, což otevírá nové směry pro budoucí neurovědecké studie.
- **Podrobný návod včetně komplexní metodiky pro detekci lži pomocí EEG:** Práce poskytuje komplexní metodiku a návod pro celý proces detekce lži pomocí EEG, který může sloužit jako referenční příručka.
- **Výsledky výzkumu a získané poznatky pro další experimenty:** Výsledky a poznatky získané v této disertační práci mohou sloužit jako inspirace pro další výzkumné a experimentální činnosti v akademickém kontextu.

Praktický přínos

- **Využití EEG technologie pro detekci lži:** Studie ukazuje, že EEG technologie má potenciál stát se spolehlivým nástrojem pro detekci lži, což může být využito v oblasti bezpečnosti a soudnictví.
- **Dostupnost technologie:** Vyvinutá aplikace je navržena tak, aby byla snadno použitelná i pro osoby bez hlubokých znalostí EEG, což usnadňuje její implementaci v různých oblastech.
- **Zlepšení detekce lží:** Přesnost a spolehlivost detekce lži pomocí EEG mohou přispět k vytvoření efektivnějších a spravedlivějších nástrojů pro detekci lží, čímž podporují lepší vymáhání práva a prevenci podvodů.
- **Omezení manipulace:** EEG metoda je méně náchylná k vědomému ovlivnění subjektem ve srovnání s tradičními metodami, jako jsou polygrafy, čímž zvyšuje její důvěryhodnost.

- **Adaptace systému pro identifikaci podvodů v reálných scénářích:** Vyvinutý systém detekce lži má potenciál pro aplikaci v různých kontextech, jako jsou policejní vyšetřování, soudní procesy a bezpečnostní kontroly. Tato adaptabilita zvyšuje praktickou hodnotu technologie a možnosti jejího využití v reálných scénářích.
- **Snížení nákladů prostřednictvím efektivních metod a minimalizace počtu elektrod:** Zavedení nízkonákladových technologií pro detekci lži a použití menšího počtu elektrod přispívá k výraznému snížení finančních nákladů spojených s těmito procesy a zjednodušuje experimentální protokoly.
- **Efektivní a rychlá detekce lži prostřednictvím vyvinuté aplikace:** Vyvinutá aplikace umožňuje automatickou, efektivní a rychlou detekci intencionální lži pomocí EEG, včetně podrobných instrukcí pro provedení experimentu. Tento přístup zkracuje čas potřebný pro vyhodnocení a zvyšuje efektivitu procesu, čímž zlepšuje její použitelnost v praxi pro širokou škálu uživatelů.

Tato disertační práce přináší významné přínosy jak pro vědeckou sféru, tak pro praktické aplikace. Vyvinutá aplikace a metodika pro detekci lži pomocí EEG signálů mají potenciál změnit přístupy k odhalování lží a podvodů v akademických výzkumech i v reálných scénářích. Budoucí výzkum by měl být zaměřen na optimalizaci metod a ověření jejich efektivity na větších a diverzifikovanějších populacích v různých praktických scénářích, což umožní širší implementaci této inovativní technologie.

ZÁVĚR

Identifikace lži je klíčová pro zajištění bezpečnosti, neboť správné odhalení skrytého chování jedince má kritický význam. Tato disertační práce se zaměřila na rozpoznávání skrytých informací za účelem odhalení intencionální lži pomocí EEG signálů získaných neinvazivním nízkonákladovým nositelným zařízením, které snímá elektrickou aktivitu mozku při prezentaci vizuálních podnětů.

Proces detekce lži pomocí EEG zahrnoval získávání dat, jejich předzpracování, extrakci a výběr vlastností a následnou klasifikaci. Data byla snímána pomocí nízkonákladového náhlavního zařízení Emotiv Insight, které nabízí nositelné řešení s přijatelným počtem EEG kanálů, představující rychlou a levnou alternativu k tradičním polygrafickým technikám a má různé výhody oproti klasickým náhlavním soupravám používaných ve výzkumech detekce lži. Nejlepších výsledků dosáhla kombinace vlastností získaných z DWT a ERP dat, přičemž elektroda Pz vykazovala nejvyšší přesnost klasifikace. Nejúčinnější metodou pro výběr relevantních vlastností byl GA, jelikož umožňuje prozkoumávat velký prostor a nalézat nečekané kombinace vlastností, čímž zvyšuje stabilitu a celkový výkon modelu. Klasifikace byla provedena pomocí algoritmu SVM, který dosáhl nejlepší přesnosti mezi různými klasifikačními algoritmy, což potvrzuje jeho vhodnost pro tento typ analýzy. EEG data jsou nelineárně oddělitelná, proto byl použit klasifikátor SVM s RBF jádrem pro nelineární mapování dat do vyšší dimenze k dosažení optimálního řešení.

Výzkum se zaměřil na využití ERP komponenty P300 v CIT k rozpoznávání známých a neznámých tváří prostřednictvím nízkonákladového EEG zařízení pro detekci lži. Navržené paradigma CIT simulující falešný zločin využívá EEG signály k analýze chování při lhaní, přičemž byly zkoumány vlny P300 v reakci na vizuální stimuly. Tento scénář poskytuje účinný způsob, jak zjistit, zda jedinec nese nějaké skryté informace. Použité zařízení lze rychle nasadit a využít s řádně navrženým experimentem k odhalení utajování informací. Reakce subjektů byly klasifikovány do dvou tříd, vinných a nevinných, pomocí klasifikačního algoritmu SVM s vlastnostmi vybranými pomocí GA z ERP a DWT dat. Navržený přístup dosáhl průměrné klasifikační přesnosti 95,09 % a správně predikoval 43 z 46 subjektů.

Tato práce představuje komplexní metodiku pro detekci intencionální lži pomocí EEG signálů, se zaměřením na ERP P300 a zdůrazňuje úspěšnost navržených přístupů a jejich potenciál pro praktické využití v různých aplikacích, včetně bezpečnosti a forenzních věd. Experimenty ukázaly vysokou přesnost klasifikace při využití nízkonákladového zařízení jako rychlého, snadného a účinného detektoru lži. Studie dále demonstruje, že kombinace vlastností z různých domén a integrace pokročilých metod výběru vlastností spolu s optimalizací algoritmů strojového učení může výrazně zlepšit přesnost detekce

lži, což otevírá široké možnosti aplikací. Tato zjištění podtrhují důležitost komplexního přístupu k analýze EEG signálů pro účinné rozpoznání skrytých informací za účelem odhalení podvodů.

Pro zajištění využitelnosti této metody byla vyvinuta aplikace s podrobným návodem pro celý proces detekce lži pomocí EEG se zaměřením na rozpoznání obličeje. Aplikace umožňuje jednoduchou, rychlou a efektivní detekci lži prostřednictvím analýzy ERP komponenty P300, což přispívá k rozšíření povědomí o možnostech využití technologie EEG pro identifikaci podvodu. Navržená aplikace je přístupná i uživatelům bez hlubokých odborných znalostí, čímž podporuje praktičnost a snadnou implementaci této technologie do běžné praxe. Přílohou práce je složka se všemi potřebnými materiály a návodem pro proces detekce lži pomocí EEG, včetně navržené aplikace.

Výsledky této práce potvrzují možnost využití ERP P300 a vizuálních podnětů v CIT pomocí EEG k identifikaci, zda podezřelý zná konkrétní obličej. Navržený přístup může sloužit jako základ pro budoucí studie zaměřené na zlepšení robustnosti detekčních modelů vůči individuálním variacím v EEG signálech a ověření přesnosti metody na větším vzorku subjektů v různých reálných scénářích. Celkově lze konstatovat, že cíle této disertační práce byly úspěšně splněny. Vyvinutá aplikace a metodika pro detekci lži pomocí EEG signálů představují významný přínos pro neurofyziologický výzkum a nabízejí potenciál pro praktické aplikace v různých oblastech.

SEZNAM POUŽITÉ LITERATURY

- [1] ABOOTALEBI, Vahid, Mohammad Hassan MORADI a Mohammad Ali KHALILZADEH. A new approach for EEG feature extraction in P300-based lie detection. *Computer methods and programs in biomedicine*. 2009, **94**(1), 48-57. Dostupné z: <https://doi.org/10.1016/j.cmpb.2008.10.001>
- [2] LIN, Xiaohong, Liyang SAI a Zhen YUAN. Detecting Concealed Information with Fused Electroencephalography and Functional Near-infrared Spectroscopy. *Neuroscience*. 2018, **386**, 284-294. Dostupné z: <https://doi.org/10.1016/j.neuroscience.2018.06.049>
- [3] FARWELL, Lawrence A. a Sharon S. SMITH. Using brain MERMER testing to detect knowledge despite efforts to conceal. *Journal of Forensic Sciences*. 2001, **46**(1), 135-143. PMID: 11210899.
- [4] BABLANI, Annushree, Damodar Reddy EDLA, Diwakar TRIPATHI, Shubham DODIA a Sridhar CHINTALA. A Synergistic Concealed Information Test With Novel Approach for EEG Channel Selection and SVM Parameter Optimization. *IEEE Transactions on Information Forensics and Security*. 2019, **14**(11), 3057-3068. Dostupné z: <https://doi.org/10.1109/TIFS.2019.2913798>
- [5] BABLANI, Annushree, Damodar Reddy EDLA a Venkatanareashbabu KUPPILI. Deceit Identification Test on EEG Data Using Deep Belief Network. *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. 2018, 1-6. Dostupné z: <https://doi.org/10.1109/ICCCNT.2018.8494124>
- [6] DODIA, Shubham, Damodar Reddy EDLA, Annushree BABLANI a Ramalingaswamy CHERUKU. Lie detection using extreme learning machine: A concealed information test based on short-time Fourier transform and binary bat optimization using a novel fitness function.

- Computational Intelligence*. 2019, **36**(2), 637-658. Dostupné z: <https://doi.org/10.1111/coin.12256>
- [7] BABLANI, Annushree, Damodar Reddy EDLA, Venkatanareshbabu KUPPILI a Ramesh DHARAVATH. Lie Detection Using Fuzzy Ensemble Approach With Novel Defuzzification Method for Classification of EEG Signals. *IEEE Transactions on Instrumentation and Measurement*. 2021, **70**, 2509413, 1-13. Dostupné z: <https://doi.org/10.1109/TIM.2021.3082985>
- [8] BABLANI, Annushree, Damodar Reddy EDLA a Shubham DODIA. Classification of EEG Data using k-Nearest Neighbor approach for Concealed Information Test. *Procedia Computer Science*. 2018, **143**, 242-249. Dostupné z: <https://doi.org/10.1016/j.procs.2018.10.392>
- [9] DODIA, Shubham, Damodar Reddy EDLA, Annushree BABLANI, Dharavath RAMESH a Venkatanareshbabu KUPPILI. An efficient EEG based deceit identification test using wavelet packet transform and linear discriminant analysis. *Journal of Neuroscience Methods*. 2019, **314**, 31-40. Dostupné z: <https://doi.org/10.1016/j.jneumeth.2019.01.007>
- [10] BABLANI, Annushree, Damodar Reddy EDLA, Diwakar TRIPATHI a Venkatanareshbabu KUPPILI. An efficient Concealed Information Test: EEG feature extraction and ensemble classification for lie identification. *Machine Vision and Applications*. 2019, **30**, 813-832. Dostupné z: <https://doi.org/10.1007/s00138-018-0950-y>
- [11] MEHRNAM, A. H., A. M. NASRABADI, Mahrads GHODOUSI, A. MOHAMMADIAN a S. TORABI. A new approach to analyze data from EEG-based concealed face recognition system. *International journal of psychophysiology*. 2017, **116**, 1-8. Dostupné z: <https://doi.org/10.1016/j.ijpsycho.2017.02.005>

- [12] BABLANI, Annushree, Damodar Reddy EDLA, Venkatanaresbhabu KUPPILI a Dharavath RAMESH. A multi stage EEG data classification using k-means and feed forward neural network. *Clinical Epidemiology and Global Health*. 2020, **8**(3), 718-724. Dostupné z: <https://doi.org/10.1016/j.cegh.2020.01.008>
- [13] ALSUFYANI, Abdulmajeed, Omid HAJILOU, Alexia ZOUMPOULAKI, Marco FILETTI, Hamed ALSUFYANI, Christopher J. SOLOMON, Stuart J. GIBSON, Roobaea ALROOBAEA a Howard BOWMAN. Breakthrough percepts of famous faces. *Psychophysiology*. 2019, **56**(1), e13279. Dostupné z: <https://doi.org/10.1111/psyp.13279>
- [14] CHEN, Chao, Lingfeng FAN, Ying GAO, Shuang QIU, Wei WEI a Huiguang HE. EEG-FRM: a neural network based familiar and unfamiliar face EEG recognition method. *Cognitive Neurodynamics*. 2024, **18**. Dostupné z: <https://doi.org/10.1007/s11571-024-10073-5>
- [15] CHANG, Wenwen, Hong WANG, Chengcheng HUA, Qiaoxiu WANG a Yue YUAN. Comparison of different functional connectives based on EEG during concealed information test. *Biomedical Signal Processing and Control*. 2019, **49**, 149-159. Dostupné z: <https://doi.org/10.1016/j.bspc.2018.12.008>
- [16] AKHAVAN, Amir, Mohammad Hassan MORADI a Safa Rafiei VAND. Subject-based discriminative sparse representation model for detection of concealed information. *Computer Methods and Programs in Biomedicine*. 2017, **143**, 25-33. Dostupné z: <https://doi.org/10.1016/j.cmpb.2017.02.007>
- [17] SAINI, Navjot, Saurabh BHARDWAJ a Ravinder AGARWAL. Classification of EEG signals using hybrid combination of features for lie detection. *Neural Computing and Applications*. 2020, **32**(8), 3777-3787. Dostupné z: <https://doi.org/10.1007/s00521-019-04078-z>

- [18] AKHAVAN, Amir a Mohammad Hassan MORADI. Detection of Concealed Information Using Multichannel Discriminative Dictionary and Spatial Filter Learning. *IEEE Transactions on Information Forensics and Security*. 2018, **13**(10), 2616-2627. Dostupné z: <https://doi.org/10.1109/TIFS.2018.2825940>
- [19] AMBER, Faryal, Adeel YOUSAF, Muhammad IMRAN a Khurram KHURSHID. P300 Based Deception Detection Using Convolutional Neural Network. In: *2019 2nd International Conference on Communication, Computing and Digital systems (C-CODE), Islamabad, Pakistan*. IEEE, 2019, 201-204. Dostupné z: <https://doi.org/10.1109/C-CODE.2019.8681025>
- [20] LIU, Qi, Hong-guang LIU a Lei ZHANG. Awareness of Crime-related Information and Concealed Information Detection method. In: *2019 IEEE 18th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), Milan, Italy*. IEEE, 2019, 372-378. Dostupné z: <https://doi.org/10.1109/ICCICC46617.2019.9146100>
- [21] ASLAN, Musa, Muhammet BAYKARA a Talha Burak ALAKUS. LieWaves: dataset for lie detection based on EEG signals and wavelets. *Medical & Biological Engineering & Computing*. 2024, **62**. Dostupné z: <https://doi.org/10.1007/s11517-024-03021-2>
- [22] XIONG, Yijun, Junfeng GAO a Ran CHEN. Connectivity network analysis of EEG signals for detecting deception. *Journal of Physics: Conference Series*. 2019, **1176**(3), 032051. Dostupné z: <https://doi.org/10.1088/1742-6596/1176/3/032051>
- [23] WANG, Hanwen, Yu QI, Hang YU, Yueming WANG, Cong LIU, Guoping HU a Gang PAN. RCIT: An RSVP-Based Concealed Information Test Framework Using EEG Signals. *IEEE Transactions on*

- Cognitive and Developmental Systems*. 2022, **14**(2), 541-551. Dostupné z: <https://doi.org/10.1109/TCDS.2021.3053455>
- [24] BAGHEL, Neeraj, Divyanshu SINGH, Malay Kishore DUTTA, Radim BURGET a Vojtech MYŠKA. Truth Identification from EEG Signal by using Convolution neural network: Lie Detection. In: *2020 43rd International Conference on Telecommunications and Signal Processing (TSP), Milan, Italy*. IEEE, 2020, 550-553. Dostupné z: <https://doi.org/10.1109/TSP49548.2020.9163497>
- [25] LUKACS, Gaspar, Alicja GRZADZIEL, Marleen KEMPKES a Ulrich ANSORGE. Item Roles Explored in a Modified P300-Based CTP Concealed Information Test. *Applied Psychophysiology and Biofeedback*. 2019, **44**(3), 195-209. Dostupné z: <https://doi.org/10.1007/s10484-019-09430-6>
- [26] ZHANG, Pengcheng, Lingyue YAN a Qiangqiang WANG. Neural processes underlying faking and concealing a personal identity: An electroencephalogram study. *Social Behavior and Personality*. 2023, **51**(2), 1-12. Dostupné z: <https://doi.org/10.2224/sbp.12119>
- [27] LIU, Qi, Xiao-Guang ZHAO, Zeng-Guang HOU a Hong-Guang LIU. Deep Belief Networks for EEG-Based Concealed Information Test. In: CONG, F., LEUNG, A., WEI, Q. (eds) *Advances in Neural Networks. Lecture Notes in Computer Science*. 2017, **10262**, 498-506. Dostupné z: https://doi.org/10.1007/978-3-319-59081-3_58
- [28] FUNICELLI, Michel, Lauren WHITE, Sabina UNGUREANU a Jean-Roch LAURENCE. An Independent Validation of the EEG-Based Complex Trial Protocol with Autobiographical Data and Corroboration of its Resistance to a Cognitively Charged Countermeasure. *Applied Psychophysiology and Biofeedback*. 2021, **46**(3), 287-299. Dostupné z: <https://doi.org/10.1007/s10484-021-09506-2>

- [29] YOUSSEF, Amira E., Hebatalla T. OUDA a Mohamed AZAB. MUSE: A Portable Cost-efficient Lie Detector. *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Kanada*. 2018, 242-246. Dostupné z: <https://doi.org/10.1109/IEMCON.2018.8614795>
- [30] KOHAN, Marzieh Daneshi, Ali Motie NASRABADI, Ali SHARIFI a Mohammad Bagher SHAMSOLLAHI. Interview based connectivity analysis of EEG in order to detect deception. *Medical Hypotheses*. 2020, **136**, 109517. Dostupné z: <https://doi.org/10.1016/j.mehy.2019.109517>
- [31] WAGNER-ALTENDORF, Tobias A., Arie H. van der LUGT, Jane F. BANFIELD, Carsten MEYER, Caterina ROHRBACH, Marcus HELDMANN a Thomas F. MÜNTE. The Electro cortical Signature of Successful and Unsuccessful Deception in a Face-to-Face Social Interaction. *Frontiers in Human Neuroscience*. 2020, **14**, 00277. Dostupné z: <https://doi.org/10.3389/fnhum.2020.00277>
- [32] KOHAN, Marzieh Daneshi, Ali Motie NASRABADI, Mohammad Bagher SHAMSOLLAHI a Ali SHARIFI. EEG/PPG effective connectivity fusion for analyzing deception in interview. *Signal, Image and Video Processing*. 2020, **14**(5), 907-914. Dostupné z: <https://doi.org/10.1007/s11760-019-01622-1>
- [33] THAKUR, Santosh, Ramesh DHARAVATH a Damodar Reddy EDLA. Spark and Rule-KNN based scalable machine learning framework for EEG deceit identification. *Biomedical Signal Processing and Control*. 2020, **58**, 101886. ISSN 1746-8094. Dostupné z: <https://doi.org/10.1016/j.bspc.2020.101886>
- [34] ALSUFYANI, Abdulmajeed, Kathryn HARRIS, Alexia ZOUMPOULAKI, Marco FILETTI a Howard BOWMAN. Breakthrough

- percepts of famous names. *Cortex*. 2021, **139**, 267-281. ISSN 0010-9452. Dostupné z: <https://doi.org/10.1016/j.cortex.2021.02.030>
- [35] YUAN, Zhen a Xiaohong LIN. Mapping of the brain activation associated with deception using fused EEG and fNIRS. *Proceedings SPIE 10865, Neural Imaging and Sensing*. 2019, 108650D. Dostupné z: <https://doi.org/10.1117/12.2508257>
- [36] CHANG, Wenwen, Hong WANG, Zhiguo LU a Chong LIU. A Concealed Information Test System Based on Functional Brain Connectivity and Signal Entropy of Audio-Visual ERP. *IEEE Transactions on Cognitive and Developmental Systems*. 2020, **12**(2), 361-370. Dostupné z: <https://doi.org/10.1109/TCDS.2020.2991359>
- [37] SAINI, Navjot, Saurabh BHARDWAJ a Ravinder AGARWAL. An Intelligent Approach of Measurement and Uncertainty Estimation for Hidden Information Detection Using Brain Signals. *MAPAN - Journal of Metrology Society of India*. 2022, **37**(1), 81-95. Dostupné z: <https://doi.org/10.1007/s12647-021-00493-7>
- [38] TATUM, William O., Aatif M. HUSAIN, Selim R. BENBADIS a Peter W. KAPLAN. *Handbook of EEG interpretation*. New York: Demos Medical Publishing, 2008. ISBN 978-1-933864-11-2.
- [39] KRAJČA, Vladimír a Jitka MOHYLOVÁ. *Číslicové zpracování neurofyziologických signálů*. Praha: České vysoké učení technické, 2011. ISBN 978-80-01-04721-7.
- [40] MARCELO, Cirach Anthony G., Zenina Rose B. PASQUIN, Arvin Dale T. PICHAY, Miguel Louis D. TAN, Ma. Fia Kristine N. SIMON, Seigfred V. PRADO, Charlene V. SAN a Marissa S. NICASIO. Characterization and comparison of brain wave signals during deception. *2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management*

- (*HNICEM*). 2017, 1-6. Dostupné z: <https://doi.org/10.1109/HNICEM.2017.8269508>
- [41] PURVES, Dale, George J. AUGUSTINE, David FITZPATRICK, William C. HALL, Anthony-Samuel LAMANTIA, Richard D. MOONEY, Michael L. PLATT a Leonard E. WHITE. *Neuroscience*. Sixth edition. Sunderland, Massachusetts: Sinauer Associates, 2018. ISBN 9781605353807.
- [42] PÁNEK, David. *Elektroencefalografické koreláty pohybového chování a výkonnostní zátěže*. Praha: Univerzita Karlova, Nakladatelství Karolinum, 2016. ISBN 978-80-246-3459-3.
- [43] Brain lobes. In: *Pixabay* [online]. [cit. 2022-04-19]. Dostupné z: <https://pixabay.com/illustrations/brain-lobes-neurology-human-body-1007686>
- [44] MECARELLI, Oriano. *Clinical Electroencephalography*. 1. vyd. Cham: Springer International Publishing, 2019. ISBN 978-3-030-04573-9.
- [45] LAI, Ying-Fang, Mu-Yen CHEN a Hsiu-Sen CHIANG. Constructing the lie detection system with fuzzy reasoning approach. *Granular Computing*. 2018, **3**, 169-176. Dostupné z: <https://doi.org/10.1007/s41066-017-0064-3>
- [46] LIU, Peng, Hongkui SHEN a Shumei JI. Functional Connectivity Pattern Analysis Underlying Neural Oscillation Synchronization during Deception. *Neural Plasticity*. 2019, 2684821. Dostupné z: <https://doi.org/10.1155/2019/2684821>
- [47] ANWAR, Syed, Tahira BATOOL a Muhammad MAJID. Event Related Potential (ERP) based Lie Detection using a Wearable EEG headset. In: *2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*. IEEE, 2019, 543-547. Dostupné z: <https://doi.org/10.1109/IBCAST.2019.8667131>

- [48] GAO, Junfeng, Jian SONG, Yong YANG, Shun YAO, Jinan GUAN, Huifang SI, Hui ZHOU, Sheng GE a Pan LIN. Deception Decreases Brain Complexity. *IEEE Journal of Biomedical and Health Informatics*. 2019, **23**(1), 164-174. Dostupné z: <https://doi.org/10.1109/JBHI.2018.2842104>
- [49] JAGANNATH, Mohan, K. ADALARASU a Keerthik R. DHIVYA. Neural correlates of guilt knowledge analysis using electroencephalographic technique. *The Journal of Forensic Psychiatry & Psychology*. 2022, **33**(2), 231-251. Dostupné z: <https://doi.org/10.1080/14789949.2022.2042360>
- [50] FABER, Josef. *Elektroencefalografie a psychofyziologie*. 1. vyd. Praha: ISV, 2001. Lékařství. ISBN 80-85866-74-9.
- [51] Emotiv Insight 2.0: 5 Channel Mobile Brainwear [online]. *Emotiv*. © 2024 [cit. 2024-04-29]. Dostupné z: <https://www.emotiv.com/insight/>
- [52] Headset Comparison Chart [online]. *Emotiv*. © 2024 [cit. 2024-07-18]. Dostupné z: <https://www.emotiv.com/pages/comparison>
- [53] Brain Products EasyCap [online]. *EasyCap*. © 2024 [cit. 2024-07-18]. Dostupné z: <https://www.easycap.de/products/>
- [54] Biosemi ActiveTwo [online]. *Biosemi*. © 2024 [cit. 2024-07-18]. Dostupné z: <https://www.biosemi.com/>
- [55] Neuroscan EEG Systems [online]. *Compumedics Neuroscan*. © 2024 [cit. 2024-07-18]. Dostupné z: <https://compumedicsneuroscan.com/>
- [56] Lafayette Instrument Polygraph Systems [online]. *Lafayette Instrument*. © 2024 [cit. 2024-07-18]. Dostupné z: <https://lafayetteinstrument.com/>
- [57] EmotivPro: The Most Advanced EEG Software [online]. *Emotiv*. © 2024 [cit. 2024-04-29]. Dostupné z: <https://www.emotiv.com/emotivpro/>
- [58] DELORME, Arnaud a Scott MAKEIG. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent

- component analysis. *Journal of Neuroscience Methods*. 2004, **134**, 9-21.
Dostupné z: <http://dx.doi.org/10.1016/j.jneumeth.2003.10.009>
- [59] EEGLAB Wiki [online]. *EEGLAB*. [cit. 2024-05-06]. Dostupné z: <https://eeglab.org/>
- [60] LOPEZ-CALDERON, Javier a Steven J. LUCK. ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in human neuroscience*. 2014, **8**, 213. Dostupné z: <https://doi.org/10.3389/fnhum.2014.00213>
- [61] ERPLAB Tutorial [online]. *ERPLAB*. [cit. 2024-05-06]. Dostupné z: <https://github.com/ucdavis/erplab/wiki/ERPLAB-Tutorial>
- [62] MA, Debbie S., Joshua CORRELL a Bernd WITTENBRINK. The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*. 2015, **47**(4), 1122-1135. Dostupné z: <https://doi.org/10.3758/s13428-014-0532-5>
- [63] PÁTEK, Adam. *Analýza emocí při čtení mapy a řešení prostorových úloh*. Olomouc, 2017. Diplomová práce. Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra geoinformatiky. Vedoucí práce Mgr. Stanislav Popelka, Ph.D. Dostupné z: http://www.geoinformatics.upol.cz/dprace/magisterske/patek17/data/DP_Adam_Patek.pdf
- [64] BUILES, Juan Camilo Carvajal a Carolina Gutierrez de PINERES. Instruments, measures, procedures and background of studies produced for deceit detection by using psycho-physiological actions. *Revista Criminalidad*. 2018, **60**(2), 75-88.
- [65] FARWELL, Lawrence A. Brain fingerprinting: a comprehensive tutorial review of detection of concealed information with event-related brain potentials. *Cognitive Neurodynamics*. 2012, **6**(2), 115-154. Dostupné z: <https://doi.org/10.1007/s11571-012-9192-2>

SEZNAM OBRÁZKŮ

Obr. 1.: Mezinárodní systém 10-20 [39]	19
Obr. 2.: Anatomické rozčlenění jednotlivých částí mozkové kůry [43]	20
Obr. 3.: EEG signály	21
Obr. 4.: Frekvenční pásma EEG	23
Obr. 5.: Odezva P300	24
Obr. 6.: Zařízení Emotiv Insight [51]	30
Obr. 7.: Umístění elektrod u zařízení Emotiv Insight [51]	30
Obr. 8.: Schéma navržené metodiky pro detekci lži pomocí EEG	32
Obr. 9.: Experimentální postup	36
Obr. 10.: Časová osa experimentu pro jeden blok	38
Obr. 11.: Správné umístění zařízení Emotiv Insight [51]	39
Obr. 12.: Kontrola kvality EEG signálů – nízká (vlevo) vs vysoká (vpravo) kvalita	40
Obr. 13.: Spuštění nahrávání záznamu v EmotivPro	40
Obr. 14.: Průběh měření v EmotivPro	41
Obr. 15.: Export dat z EmotivPro	41
Obr. 16.: Import dat v EEGLAB	42
Obr. 17.: EEG signály	42
Obr. 18.: Nastavení umístění kanálů v EEGLAB	43
Obr. 19.: Nastavení filtrace pomocí Notch filtru v ERPLAB	44
Obr. 20.: Import informací o eventech v EEGLAB	44
Obr. 21.: Rozdělení na jednotlivé eventy	45
Obr. 22.: Vytvoření seznamu událostí v ERPLAB	45
Obr. 23.: Tvorba epoch v ERPLAB	46
Obr. 24.: Rozdělení na jednotlivé epochy	46
Obr. 25.: Nastavení filtrace pomocí IIR Butterworth filtru v ERPLAB	47
Obr. 26.: Filtrace pomocí IIR Butterworth filtru	47
Obr. 27.: Nastavení odstranění artefaktů v ERPLAB	48

Obr. 28.: Odstranění artefaktů	48
Obr. 29.: Průměrování signálu v ERPLAB.....	49
Obr. 30.: Export dat z ERPLAB	49
Obr. 31.: Nastavení parametrů pro vykreslení ERP dat v ERPLAB	50
Obr. 32.: Průběh vlny P300 na elektrodě Pz pro jednotlivé podněty	50
Obr. 33.: PSD pro vinný (vlevo) vs nevinný (vpravo) subjekt.....	51
Obr. 34.: Průměr ERP na elektrodě Pz pro všechny subjekty	52
Obr. 35.: Rozklad signálu pomocí DWT pro 128 Hz	59
Obr. 36.: Přesnost klasifikace vlastností pomocí různých algoritmů u elektrody Pz.....	67
Obr. 37.: Přesnost klasifikace vlastností pomocí různých algoritmů u elektrody T7	68
Obr. 38.: Přesnost klasifikace vlastností pomocí různých algoritmů u elektrody T8	68
Obr. 39.: Přesnost klasifikace vlastností pomocí různých algoritmů u elektrody AF3.....	69
Obr. 40.: Přesnost klasifikace vlastností pomocí různých algoritmů u elektrody AF4.....	69
Obr. 41.: Porovnání přesnosti klasifikace na různých elektrodách	70
Obr. 42.: Porovnání metod výběru vlastností	71
Obr. 43.: Přesnost predikce finálního modelu	75
Obr. 44.: Confusion Matrix.....	76
Obr. 45.: Průměrné hodnoty jednotlivých metrik výkonu.....	79
Obr. 46.: Srovnání současného přístupu se stávajícími studiemi	81
Obr. 47.: Aplikace pro detekci lži.....	84
Obr. 48.: Pracovní postup aplikace	85
Obr. 49.: Vývojový diagram aplikace pro detekci lži.....	86

SEZNAM TABULEK

Tabulka 3.1 Rozdělení frekvenčních pásem EEG [42]	22
Tabulka 5.1 Srovnání současného přístupu se stávajícími studiemi	80

SEZNAM ZDROJOVÝCH KÓDŮ

Zdrojový kód 1: Implementace GA	64
Zdrojový kód 2: Implementace funkce pro načtení a úpravu dat	87
Zdrojový kód 3: Implementace funkce pro import dat	88
Zdrojový kód 4: Implementace funkce pro odstranění artefaktů	90
Zdrojový kód 5: Implementace funkce pro víceúrovňovou DWT	90
Zdrojový kód 6: Implementace funkce pro jednoúrovňovou DWT	91
Zdrojový kód 7: Implementace funkce výpočtu statistických vlastností ...	92
Zdrojový kód 8: Implementace funkce pro detekci lži	93

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

a kol.	A kolektiv
ANOVA	Analysis of Variance
ANS	Autonomic Nervous System (autonomní nervový systém)
BBA	Binary BAT Algorithm
BPF	Band-Pass Filter
CIT	Concealed Information Test
CNS	Central Nervous System (centrální nervový systém)
CTP	Complex Trial Protocol
CTSAM	Channel and Time-Spatial Attention Module
DBN	Deep Belief Network
DIT	Deceit Identification Test
DT	Decision Trees
DWT	Discrete Wavelet Transform
EEG	Electroencephalography (Elektroencefalografie)
EKG	Electrocardiography (Elektrokardiografie)
ELM	Extreme Learning Machine
EMG	Electromyography (Elektromyografie)
EOG	Electrooculography (Elektrookulografie)
ERP	Event-Related Potential
FFT	Fast Fourier Transform
fMRI	Functional magnetic resonance imaging (funkční magnetická rezonance)

fNIRS	Functional near-infrared spectroscopy (funkční blízka infračervená spektroskopie)
GA	Genetic Algorithm (genetický algoritmus)
GKT	Guilty Knowledge Test
GUI	Graphical User Interface (grafické uživatelské rozhraní)
KNN	K-Nearest Neighbors
LDA	Linear Discriminant Analysis
MERMER	Memory and Encoding Related Multifaceted Electroencephalographic Response
MLFFNN	Multi-Layer Feed Forward Neural Network
MSCNN	Multi-Scale Convolutional Neural Network
NB	Naive Bayes
PCA	Principal Component Analysis
PNS	Peripheral Nervous System (periferní nervový systém)
RBF	Radial Basis Function
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
PSD	Power Spectral Density
WPT	Wavelet Packet Transform
WT	Wavelet Transform
WV	Weighted Voting

SEZNAM PŘÍLOH

- P I Schválení experimentu Etickou komisí UTB
- P II Informovaný souhlas
- P III Účel a princip experimentu – vinní
- P IV Účel a princip experimentu – nevinní
- P V Umístění kanálů
- P VI Seznam eventů
- P VII Seznam událostí
- P VIII Dokumentace s návodem pro detekci lži pomocí EEG
- P IX Dokumenty a aplikace pro detekci lži pomocí EEG

PŘÍLOHA P I: SCHVÁLENÍ EXPERIMENTU ETICKOU KOMISÍ UTB

Zápis ze zasedání Etické komise pro biomedicínský výzkum

Předmět zasedání: Projednání žádosti o schválení experimentu k disertační práci "Rozpoznávání skrytých informací k určení intencionální lži pomocí elektroencefalografie"

Předkladatelka Ing. Martina Žabčíková

Termín zasedání: 15.1.2024

Forma zasedání: elektronická

Dne 29.1.2024 obdržel předseda komise žádost o projednání a schválení experimentu k disertační práci "Rozpoznávání skrytých informací k určení intencionální lži pomocí elektroencefalografie" Žádost byla doručena elektronicky s přílohami obsahujícími Data management plan, Informovaný souhlas; CV předkladatelky.

Dne 1.2.2024 požádal předseda členy komise o projednání předloženého materiálu per rollam.

Výsledek hlasování:

<i>Počet hlasů schvalujících předloženou žádost:</i>	<i>4</i>
<i>Počet hlasů schvalujících předloženou žádost s připomínkou:</i>	<i>0</i>
<i>Počet hlasů neschvalujících předloženou žádost:</i>	<i>1 (prof. Jašek nehlasoval z důvodu střetu zájmu)</i>

Závěr: Etická komise pro biomedicínský výzkum schvaluje předložený plán experimentu k disertační práci "Rozpoznávání skrytých informací k určení intencionální lži pomocí elektroencefalografie"

Verifikoval:

15.2.2024

prof. Ing. Petr Humpolčák, Ph.D.



Přílohou zápisu jsou dokumenty dodané společně se žádostí.

PŘÍLOHA P II: INFORMOVANÝ SOUHLAS

Informovaný souhlas s účastí ve výzkumu

Studentka doktorského studia Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně, Martina Žabčíková v rámci své disertační práce s názvem *Rozpoznávání skrytých informací k určení intencionální lži pomocí elektroencefalografie* hodlá realizovat níže uvedený výzkum.

Cílem výzkumu je rozpoznávání skrytých informací pro odhalení lži pomocí signálů EEG (elektroencefalografie) získaných neinvazivním nositelným zařízením snímající elektrickou aktivitu mozku při prezentaci vizuálních stimulů známých a neznámých tváří. Měření je nebolestivé, nemá žádnou kontraindikaci a není k němu nutná žádná speciální předchozí příprava.

Data budou zpracována anonymně. Vstupními informacemi o účastníkovi/účastnici výzkumu (dále pro zjednodušení bude uváděno pouze „účastník“) jsou:

- rok narození,
- pohlaví,
- datum experimentu,

kteří uvede účastník na konci tohoto Informovaného souhlasu. Poté mu bude přiřazeno číslo, pod kterým bude experiment realizován. Informace o věku a pohlaví budou využity pro vyhodnocení, zda existují rozdíly v reakcích na jednotlivé vizuální stimuly mezi pohlavími a různými věkovými kategoriemi. Údaje, které budou získány v rámci realizace experimentu jsou anonymní a není možné na jejich základě identifikovat osobu zúčastněnou v experimentu. Všechna získaná data budou využita výhradně pro účely této disertační práce. Tento experiment nebude zahrnovat osoby v závislém postavení.

Před experimentem bude účastníkovi nasazeno bezdrátové neinvazivní zařízení pro měření mozkové aktivity. Hlavním cílem experimentu je získání mozkových signálů účastníka v reakci na vizuální podněty.

Role účastníka spočívá ve sledování obrazovky, na které budou prezentovány obrázky všeobecně známých i neznámých tváří osob, na něž bude vzápětí účastník odpovídat, zda danou osobu zná či ne. Délka experimentu by měla být přibližně 15 minut, včetně instalace zařízení a samotné prezentace. Z experimentu je možno kdykoli odstoupit, a to bez udání důvodu. Pokud se rozhodnu odstoupit, informuji vedení experimentu o svém rozhodnutí písemně.

Byl/a jsem informován/a o cíli studie, o jejich průběhu a postupech, a o tom, co se ode mě očekává. Byl/a jsem informován/a o účelu experimentu, kterým je sběr dat mozkových signálů, pro potřeby výzkumu v disertační práci. Zapojením se do výzkumu souhlasím se svou účastí v něm. Tento projekt byl schválen Etickou komisí UTB pro výzkum.

Děkuji za Vaši ochotu a účast na tomto výzkumu. Vaše spolupráce je pro mě velmi cenná.

Martina Žabčíková

Informace o účastníkovi/ci výzkumu:

- účastník/ce byl/a do experimentu zařazen/a pod číslem:
- rok narození:
- pohlaví:
- datum experimentu:

PŘÍLOHA P III: ÚČEL A PRINCIP EXPERIMENTU – VINNÍ

Účel a princip experimentu

Jedná se o experiment založený na falešném zločinu, kde se zjišťuje, zda je subjekt obeznámen s tváří konkrétní osoby či nikoliv. Experimentální scénář zahrnuje sérii snímků obličejů osob prezentovaných na obrazovce. Tyto obrázky musíte rozpoznat a po každém obrázku **rychle a jednoznačně odpovědět** ANO (znám osobu na obrázku) či NE (neznám osobu na obrázku). Je důležité, aby vaše odpovědi byly spontánní a co nejrychlejší. Nepřemýšlejte příliš dlouho nad každým obrázkem. Obrázek černého fixačního kříže na bílém pozadí slouží pro zaměření se zpět na střed obrazovky (**příležitost k mrkání**).

Pokyny

- odložte prosím svůj **mobilní telefon**
- během testu se **uvolněte**, sed'te klidně, **minimalizujte pohyby těla a omezte pohyby očí**, úst a mrkání
- **oči mějte fixované na střed** obrazovky

V tomto experimentu budete hrát roli účastníka, který byl **obviněn z trestného činu**. Vaše účast a spolupráce jsou klíčové pro úspěch této studie a umožní lepší porozumění procesům spojeným s identifikací známých a neznámých osob na základě obličeje.

Scénář (vinní): Na Fakultě aplikované informatiky UTB ve Zlíně byl ukraden mobilní telefon vysoké hodnoty této osobě:

Fotografie oběti trestného činu

Abyste skryli svoji vinu, **identifikujte prosím tuto osobu jako neznámou** (odpovězte co nejrychleji NE). Chovejte se mentálně jako nevinní a pokuste se skrýt své informace o známé tváři tím, že ji klasifikujete jako neznámou. Taktéž si dávejte pozor na mimické výrazy v obličeji, když se budete snažit zalhat, jako při klasickém výslechu. Na **ostatní obrázky osob odpovídejte čestně**, tedy pro známé osoby odpovídejte ANO a pro neznámé NE.

Děkuji Vám za váš čas a příspěvek k výzkumu.

PŘÍLOHA P IV: ÚČEL A PRINCIP EXPERIMENTU – NEVINNÍ

Účel a princip experimentu

Jedná se o experiment založený na falešném zločinu, kde se zjišťuje, zda je subjekt obeznámen s tváří konkrétní osoby či nikoliv. Experimentální scénář zahrnuje sérii snímků obličejů osob prezentovaných na obrazovce. Tyto obrázky musíte rozpoznat a po každém obrázku **rychle a jednoznačně odpovědět** ANO (znám osobu na obrázku) či NE (neznám osobu na obrázku). Je důležité, aby vaše odpovědi byly spontánní a co nejrychlejší. Nepřemýšlejte příliš dlouho nad každým obrázkem. Obrázek černého fixačního kříže na bílém pozadí slouží pro zaměření se zpět na střed obrazovky (**příležitost k mrkání**).

Pokyny

- odložte prosím svůj **mobilní telefon**
- během testu se **uvolněte**, sed'te klidně, **minimalizujte pohyby těla a omezte pohyb očí**, úst a mrkání
- **oči mějte fixované na střed** obrazovky

V tomto experimentu budete hrát roli účastníka, který je **nevinný**. Vaše účast a spolupráce jsou klíčové pro úspěch této studie a umožní lepší porozumění procesům spojeným s identifikací známých a neznámých osob na základě obličeje.

Scénář (nevinný): Na Fakultě aplikované informatiky UTB ve Zlíně byl ukraden mobilní telefon vysoké hodnoty. Jelikož budete v tomto experimentu hrát roli nevinného, budete **odpovídat čestně na všechny obrázky osob**. Pokud tvář na obrázku poznáváte, odpovězte ANO, zatímco pokud tvář neznáte, odpovězte NE.

Děkuji Vám za váš čas a příspěvek k výzkumu.

PŘÍLOHA P V: UMÍSTĚNÍ KANÁLŮ

Soubor Insight.ced se používá pro nastavení umístění jednotlivých elektrod na povrchu pokožky hlavy u zařízení Emotiv Insight v EEGLAB.

Number	labels	theta	radius	X	Y	Z	sph_theta	sph_phi	sph_radius	type
1	AF3	-23	0.411	0.885	0.376	0.276	23	16	1	1
2	T7	-90	0.511	6.12e-17	0.999	-0.0349	90	-2	2	2
3	Pz	180	0.256	-0.719	-8.81e-17	0.695	-180	44	3	3
4	T8	90	0.511	6.12e-17	0.999	-0.0349	-90	-2	4	4
5	AF4	23	0.411	0.885	-0.376	0.276	-23	16	1	5

PŘÍLOHA P VI: SEZNAM EVENTŮ

Soubor pro označení nástupu jednotlivých stimulů (eventy.txt).

Latency Type Position

2	irelevantní 1	104	cíl 35	206	sonda 69
5	sonda 2	107	irelevantní 36	209	irelevantní 70
8	irelevantní 3	110	irelevantní 37	212	irelevantní 71
11	irelevantní 4	113	irelevantní 38	215	irelevantní 72
14	irelevantní 5	116	cíl 39	218	cíl 73
17	cíl 6	119	irelevantní 40	221	irelevantní 74
20	irelevantní 7	122	cíl 41	224	irelevantní 75
23	cíl 8	125	irelevantní 42	227	irelevantní 76
26	irelevantní 9	128	cíl 43	230	sonda 77
29	irelevantní 10	131	irelevantní 44	233	irelevantní 78
32	irelevantní 11	134	irelevantní 45	236	irelevantní 79
35	irelevantní 12	137	irelevantní 46	239	cíl 80
38	cíl 13	140	sonda 47	242	irelevantní 81
41	irelevantní 14	143	irelevantní 48	245	cíl 82
44	sonda 15	146	irelevantní 49	248	irelevantní 83
47	irelevantní 16	149	irelevantní 50	251	irelevantní 84
50	irelevantní 17	152	cíl 51	254	irelevantní 85
53	irelevantní 18	155	irelevantní 52	257	sonda 86
56	cíl 19	158	irelevantní 53	260	irelevantní 87
59	irelevantní 20	161	irelevantní 54	263	irelevantní 88
62	irelevantní 21	164	sonda 55	266	irelevantní 89
65	irelevantní 22	167	irelevantní 56	269	cíl 90
68	sonda 23	170	irelevantní 57	272	irelevantní 91
71	irelevantní 24	173	irelevantní 58	275	cíl 92
74	irelevantní 25	176	irelevantní 59	278	irelevantní 93
77	irelevantní 26	179	cíl 60	281	sonda 94
80	irelevantní 27	182	cíl 61	284	irelevantní 95
83	cíl 28	185	irelevantní 62	287	irelevantní 96
86	irelevantní 29	188	irelevantní 63	290	irelevantní 97
89	cíl 30	191	irelevantní 64	293	cíl 98
92	irelevantní 31	194	cíl 65	296	irelevantní 99
95	irelevantní 32	197	irelevantní 66	299	irelevantní 100
98	irelevantní 33	200	irelevantní 67		
101	sonda 34	203	irelevantní 68		

PŘÍLOHA P VII: SEZNAM UDÁLOSTÍ

Soubor obsahující seznam jednotlivých událostí (event_list.txt).

1	"sonda"	1	"sonda"
2	"cíl"	2	"cíl"
3	"irelevantní"	3	"irelevantní"

PŘÍLOHA P VIII: DOKUMENTACE S NÁVODEM PRO DETEKCI LŽI POMOCÍ EEG

Dokumentace s návodem pro detekci lži pomocí EEG

Tato dokumentace slouží jako návod pro proces detekce lži pomocí EEG signálů s využitím nízkonákladového zařízení Emotiv Insight. Obsahuje postup snímání dat a následné použití aplikace, včetně licenčních informací.

Postup snímání dat EEG pomocí zařízení Emotiv Insight

Jedná se o experiment založený na falešném zločinu, kde se zjišťuje, zda je subjekt obeznámen s tváří konkrétní osoby či nikoliv. Experimentální scénář zahrnuje sérii snímků obličejů osob prezentovaných na obrazovce. Délka experimentu je přibližně 15 minut, přičemž příprava trvá cca 10 minut a prezentace 5 minut. Pro snímání dat je třeba zajistit klidné prostředí a postupovat dle následujících kroků:

1) Příprava prezentace

- Vložte do prezentace (vzor_prezentace_DetekceLzi.pptx) fotografii osoby postižené trestným činem do označených slidů a upravte ji dle pokynů.
- Fotografie musí být vycentrována zarovnáním oční linie do stejné horizontální polohy, ukazující čelní pohledy na tváře s neutrálním výrazem, bílým pozadím a ve stupních šedi s upraveným kontrastem (viz ostatní fotografie).

2) Předložení informovaného souhlasu

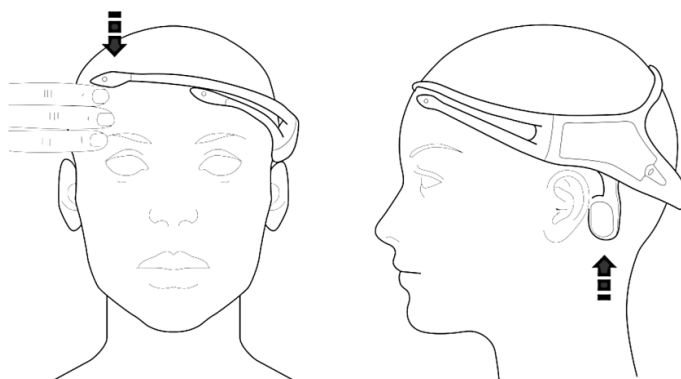
- Před zahájením experimentu předložte subjektům písemný informovaný souhlas s účastí ve výzkumu.

3) Pokyny pro subjekty

- Odložte mobilní telefon.
- Během testu se uvolněte, sedněte klidně, oči fixujte na střed obrazovky, minimalizujte pohyby těla a omezte pohyb očí, úst a mrkání.
- Odpovídejte čestně na všechny fotografie osob. Pokud tvář poznáváte, odpovězte nahlas ANO; pokud ne, odpovězte NE.
- Odpovědi musí být spontánní a co nejrychlejší.

4) Správné umístění zařízení a navlhčení senzorů

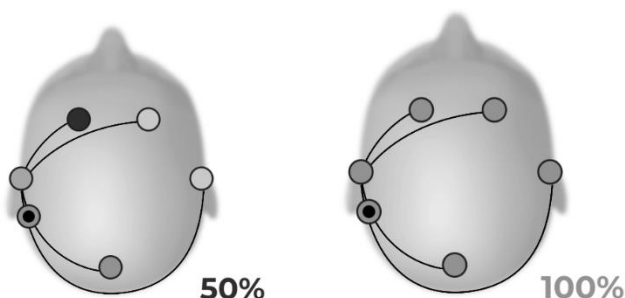
- Před měřením navlhčete senzory a nasadte náhlavní soupravu (viz Obr. 1).
- Zapněte zařízení a zkontrolujte, že všechny elektrody jsou v kontaktu s pokožkou.



Obr. 1.: Správné umístění zařízení Emotiv Insight

5) Kontrola kvality signálů

- Kvalitu signálu ověřte v softwaru EmotivPro (viz Obr. 2).
- Signály s nízkou kvalitou jsou červené, optimální signály zelené. Pokud senzory vykazují nízkou kvalitu, znovu je navlhčete.



Obr. 2.: Kontrola kvality EEG signálů – nízká (vlevo) vs vysoká (vpravo) kvalita

6) Zahájení měření

- Prezenci a měření v EmotivPro spustíte současně.
- Po dosažení optimální kvality signálu zahajte záznam signálu (viz Obr. 3).

Obr. 3.: Spuštění nahrávání záznamu v EmotivPro

7) Ukončení měření a export dat

- Po ukončení měření zastavte nahrávání a opatrně sejměte EEG zařízení.
- Proveďte export naměřených dat z EmotivPro do formátu CSV (viz Obr. 4).

Obr. 4.: Export dat z EmotivPro

Dokumentace k aplikaci LiDetEEG

Popis aplikace

Aplikace LiDetEEG.exe je určena pro detekci lži pomocí EEG signálů. Byla vytvořena v MATLAB App Designeru a používá specializované toolboxy pro analýzu a klasifikaci signálů.

Pokyny k instalaci a spuštění aplikace

- Pro spuštění aplikace je nutno mít nainstalovaný Matlab R2020b nebo bezplatný Matlab Runtime R2020b (9.9).
- Pokud vlastníte Matlab verze R2020b, není potřeba žádná další instalace.
- Pokud nevládníte Matlab, stáhněte a nainstalujte bezplatnou verzi Matlab Runtime R2020b (9.9) z oficiálních stránek MathWorks:
 - <https://www.mathworks.com/products/compiler/matlab-runtime.html>

Kroky pro použití aplikace

- 1) **Import dat:** Klikněte na tlačítko "Vložit soubor" a načtěte EEG data vyexportovaná z EmotivPro.
- 2) **Analýza dat:** Klikněte na tlačítko "Analýza dat" pro předzpracování a analýzu vložených EEG dat.
- 3) **Detekce lži:** Klikněte na tlačítko "Detekovat lež" pro spuštění detekce lži na základě vložených dat. Výsledkem je zpráva, zda subjekt zná či nezná obličej dotazované osoby.

Použité toolboxy a uznání

- **EEGLAB:** Open-source software distribuovaný pod GNU General Public License (GPL). Viz <https://sccn.ucsd.edu/eeglab/index.php>.
- **ERPLAB:** Open-source software distribuovaný pod GNU GPL. Viz <https://erpinfo.org/erplab>.
- **Signal Processing Toolbox a Statistics and Machine Learning Toolbox:** MATLAB toolboxy licencované společností MathWorks.

Licenční informace

Aplikace byla vytvořena a zkompileována pomocí MATLAB Compiler, který umožňuje distribuci aplikace bez potřeby instalace MATLABu na cílovém počítači. Pro použití MATLAB Compiler je nutná platná licence, která je součástí akademické licence naší instituce.

V souladu s licencí GNU GPL musí být aplikace, která zahrnuje EEGLAB a ERPLAB, distribuována se stejnou licencí a zdrojovým kódem. Pro přístup ke zdrojovému kódu se obraťte na oddělení tvůrčích činností prostřednictvím e-mailu prodekan-vyzkum@fai.utb.cz.

Kontakt

Pro jakékoli dotazy mě prosím kontaktujte na zabcikova@utb.cz.

PŘÍLOHA P IX: DOKUMENTY A APLIKACE PRO DETEKCI LŽI POMOCÍ EEG

Příloha zahrnuje všechny soubory nezbytné pro proces detekce lži pomocí EEG, které jsou veřejně přístupné elektronicky prostřednictvím informačního systému STAG UTB v sekci Kvalifikační práce. Tyto soubory zahrnují vzorovou prezentaci pro měření signálů EEG (vzor_prezentace_DetekceLzi.pptx), dokumentaci s návodem pro snímání dat a použití aplikace (Dokumentace.pdf), testovací data (testovaci_subjekt.csv), samotnou aplikaci (LiDetEEG.exe) a zdrojový kód (LiDetEEG.m), včetně všech souvisejících souborů potřebných pro plnou funkčnost aplikace. Zdrojový kód je k dispozici na požádání.

PUBLIKAČNÍ AKTIVITY AUTORA

Články ve sborníku z mezinárodních konferencí

- [P.1] ZABCIKOVA, Martina, Zuzana KOUDELKOVA a Roman JASEK. EEG-based lie detection using ERP P300 in response to known and unknown faces: An overview. In: *2022 26th International Conference on Circuits, Systems, Communications and Computers (CSCC)*, Crete, Greece. 2022, 11-15.
- [P.2] KOUDELKOVA, Zuzana, Martina ZABCIKOVA a Roman JASEK. Communication Tool for Disabled People Based on Surface Electromyography. In: *Proceedings of the 2020 4th International Conference on Medical and Health Informatics (ICMHI 2020)*. 2020, 86-89.
- [P.3] KOUDELKOVA, Zuzana, Martina ZABCIKOVA a Roman JASEK. Proposal of Hybrid BCI Applicable in Rehabilitation. In: *6th SWS International Scientific Conference on Social Sciences (SGEM 2019)*. 2019, **6**(3), 637-642. ISBN 978-619-7408-93-5, ISSN 2682-9959.
- [P.4] KOUDELKOVA, Zuzana, Martina ZABCIKOVA a Roman JASEK. The sEMG Biofeedback: Controlling LEGO EV3 by Arduino Uno with sEMG. In: *6th SWS International Scientific Conference on Social Sciences (SGEM 2019)*. 2019, **6**(3), 807-812. ISBN 978-619-7408-93-5, ISSN 2682-9959.
- [P.5] ZABCIKOVA, Martina. Visual and Auditory Stimuli Response, Measured by Emotiv Insight Headset. In: *23rd International Conference on Circuits, Systems, Communications and Computers (CSCC 2019)*. 2019, **292**.
- [P.6] ZABCIKOVA, Martina. Measurement of Visual and Auditory Stimuli Using EEG Headset Emotiv Epop+. In: *23rd International Conference*

on Circuits, Systems, Communications and Computers (CSCC 2019).
2019, **292**.

- [P.7] KOUDELKOVA, Zuzana, Martina ZABCIKOVA a Roman JASEK. The Possibility of Using Diagnostic Methods EEG and sEMG in Rehabilitation. In: *5th International Conference on Computing and Artificial Intelligence (ICCAI'19)*. 2019, 240-243.

Články v časopisech evidovaných v databázi Web of Science a Scopus

- [P.8] ZABCIKOVA, Martina, Zuzana KOUDELKOVA a Roman JASEK. Concealed information detection using EEG for lie recognition by ERP P300 in response to visual stimuli: A review. *WSEAS Transactions on Information Science and Applications*. 2022, **19**, 171-179. ISSN: 1790-0832.
- [P.9] ZABCIKOVA, Martina, Zuzana KOUDELKOVA, Roman JASEK a José Javier LORENZO NAVARRO. Recent advances and current trends in brain-computer interface research and their applications. *International Journal of Developmental Neuroscience*. 2022, **82**(2), 107-123.
- [P.10] ZABCIKOVA, Martina, Zuzana KOUDELKOVA a Roman JASEK. Examining the Efficiency of Emotiv Insight Headset by Measuring Different Stimuli. *WSEAS Transactions on Applied and Theoretical Mechanics*. 2019, **14**, 235-242. ISSN 1991-8747.
- [P.11] ZABCIKOVA, Martina, Zuzana KOUDELKOVA a Roman JASEK. Investigation of Emotiv Epoc+ Headset Usability by Measuring Various Stimuli. *WSEAS Transactions on Applied and Theoretical Mechanics*. 2019, **14**, 184-191. ISSN 1991-8747.
- [P.12] STRMISKA, Martin, Martina ZABCIKOVA a Zuzana KOUDELKOVA. Measuring Brain Signals Using Emotiv Devices.

WSEAS Transactions on Systems and Control. 2018, **13**. ISSN 1991-8763.

Ostatní publikace

- [P.13] ŽABČÍKOVÁ, Martina, Zuzana KOUDELKOVÁ a Roman JAŠEK. Možná využití modelování a simulace v Brain-Computer Interface systémech. *Trilobit*. 2021.
- [P.14] JAŠEK, Roman, David MALANÍK, Milan OULEHLA, Petr ŽÁČEK, Lukáš KRÁLÍK, Zuzana KOUDELKOVÁ, Vít ŠTĚPÁNEK, Roman ŠENKEŘÍK, Zuzana KOMÍNKOVÁ OPLATKOVÁ, Bronislav CHRAMCOV, Peter JANKŮ, Pavel VAŘACHA, Radek VALA a Martina ŽABČÍKOVÁ. *Řízení procesů a aplikace moderních technologií – Kybernetická bezpečnost*. 2019.
- [P.15] JAŠEK, Roman, Bronislav CHRAMCOV, Zuzana KOUDELKOVÁ, Peter JANKŮ, Zuzana KOMÍNKOVÁ OPLATKOVÁ, Pavel VAŘACHA, Vít ŠTĚPÁNEK, Jan KOLEK, Alžběta TUREČKOVÁ, Martina ŽABČÍKOVÁ, Adam VIKTORIN a Jana KORZENIOWSKA. *Řízení inteligentních systémů pomocí technologie Brain Computer Interface a Brain Machine Interface jako fyzioterapeutické metody v klinické praxi*. 2019.
- [P.16] KOUDELKOVÁ, Zuzana, Martina ŽABČÍKOVÁ, Michal FILIP, Roman JAŠEK, Martin STRMISKA, Bronislav CHRAMCOV, Peter JANKŮ, Vít ŠTĚPÁNEK, Pavel VAŘACHA, Šárka DAŇKOVÁ, Irena CHMELOVÁ a Petr LINZER. Perspektivy využití technologie BCI (Brain Computer Interface) ve fyzioterapii. *Trilobit*. 2018, **9**(2). ISSN 1804-1795.

ODBORNÝ ŽIVOTOPIS AUTORA

OSOBNÍ ÚDAJE

Ing. Martina Žabčíková

📍 Masarykovo nám. 1129, 763 12 Vizovice

✉ zabcikova@utb.cz

☎ +420 731 904 489

📅 15. 3. 1993



VZDĚLÁNÍ

2018 – SOUČASNOST

Ph.D. | UNIVERZITA TOMÁŠE BATI VE ZLÍNĚ / FAKULTA APLIKOVANÉ INFORMATIKY

Inženýrská informatika

2016 – 2018

Ing. | UNIVERZITA TOMÁŠE BATI VE ZLÍNĚ / FAKULTA APLIKOVANÉ INFORMATIKY

Učitelství informatiky pro střední školy

2012 – 2016

Bc. | UNIVERZITA TOMÁŠE BATI VE ZLÍNĚ / FAKULTA APLIKOVANÉ INFORMATIKY

Informační a řídicí technologie

2008 – 2012

Maturita | STŘEDNÍ PRŮMYSLOVÁ ŠKOLA STAVEBNÍ VALAŠSKÉ MEZIRÍČÍ

Stavební obnova

PRACOVNÍ ZKUŠENOSTI

10/2022 – SOUČASNOST

Výzkumný projektový pracovník | UNIVERZITA TOMÁŠE BATI VE ZLÍNĚ / FAKULTA APLIKOVANÉ INFORMATIKY

Člen řešitelského týmu na projektu s názvem „Rekonstrukce scénáře bezpečnostního incidentu v prostředí virtuální reality“

07/2017 – 11/2017

IT tester, IT analytik | BUSINESS LOGIC

Spolupráce na IT projektech s moderními technologiemi. Testování webů, práce s databází SQL, vytváření wireframe.

2008 – 2017

Doplňující informace o pracovních zkušenostech

Sezónní brigády

Brigády při studiu (většinou manuální zaměření)

Praxe – stavebnictví, pedagogika

AKADEMICKÁ STÁŽ _____

09/2019 – 12/2019

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

PROJEKTY _____

2022 – 2025

VJ02010043 (MVČR)

Rekonstrukce scénáře bezpečnostního incidentu v prostředí virtuální reality – v průběhu řešení

2023 – 2025

INTERNÍ GRANTOVÁ AGENTURA (IGA)

Inteligentní datová analýza s využitím metod umělé inteligence – v průběhu řešení

2022

INTERNÍ GRANTOVÁ AGENTURA (IGA)

Výzkum v oblasti zpracování fyziologických a neurofyziologických signálů – hlavní řešitel projektu

2021

INTERNÍ GRANTOVÁ AGENTURA (IGA)

Výzkum Brain-Computer Interface pro řízení inteligentních systémů – hlavní řešitel projektu

2020

INTERNÍ GRANTOVÁ AGENTURA (IGA)

Využití diagnostických metod EEG a EMG pro sběr dat a řízení inteligentních systémů

2019

INTERNÍ GRANTOVÁ AGENTURA (IGA)

Aplikační výzkum Brain Computer Interface pro sběr dat a řízení inteligentních systémů – hlavní řešitel projektu

CERTIFIKÁTY

2022, 2019

INTERNATIONAL CONFERENCE OF WSEAS – INASE – CSCC

Certificate of Appreciation

2024, 2022

THE WORKSHOP ON APPLIED INFORMATICS

Certificate of Attendance

2018

UNIVERZITA TOMÁŠE BATI VE ZLÍNĚ

Kurz anglického jazyka se spol. R. Bosch České Budějovice

JAZYKOVÉ ZNALOSTI

Anglický jazyk

středně pokročilý (B2)

POČÍTAČOVÉ ZNALOSTI

Microsoft Office	pokročilý
Microsoft Windows	pokročilý
Adobe Premiere Pro	základy
Adobe After Effects	základy
Wolfram Alpha	základy
HTML	základy
SQL	základy
CSS	základy
C#	základy

Ing. Martina Žabčíková

**Rozpoznávání skrytých informací k určení intencionální lži
pomocí elektroencefalografie**

Detecting Concealed Information to Identify Intentional Deception Using
Electroencephalography

Disertační práce

Vydala Univerzita Tomáše Bati ve Zlíně,
nám. T. G. Masaryka 5555, 760 01 Zlín.

Náklad: vyšlo elektronicky

Sazba: autor

Publikace neprošla jazykovou ani redakční úpravou.

Rok vydání 2024