

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Prediction of Epileptic Seizures from Intracranial EEG
Jméno autora:	Bc. Lenka Zoulová
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Oponent práce:	Ing. Eduard Bakštein, Ph.D.
Pracoviště oponenta práce:	Katedra kybernetiky, FEL, ČVUT

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání bylo zaměřeno na vývoj klasifikátoru intrakraniálních EEG dat z otevřené databáze, použitý přístup (příznaky, model atd.) i rozsah použitých analýz byl ponechán na studentce. Za zmínku stojí rozsah datasetu, který obsahoval až téměř dvoutýdenní záznamy o 80 i více kanálech.	

Splnění zadání	splněno
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Podle mého názoru splnila studentka zadání v plné míře a kvalitně, výhrady mám pouze k samotné textové části práce.	

Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Studentka provedla kvalitní exploraci dat a vyhodnotila možný vliv chybějících dat, outlierů a možné závislosti trénovačích a testovacích dat na přesnost odhadu výkonu klasifikátorů a vyzkoušela několik přístupů k práci s nevyváženým datasetem. Porovnávala velké množství příznaků z publikovaných studií i klasifikačních algoritmů včetně optimalizace nastavení jejich parametrů. Více přístupů (dopředná selekce a evoluční algoritmy) byly testovány i pro redukci příznakového vektoru. Z mého pohledu se tedy jedná o uváženou a podle všeho dobře odvedenou práci, ač bez velkých inovací oproti stávajícímu stavu vědění.	

Odborná úroveň	B - velmi dobře
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Práce využívá standardní klasifikační schéma: rozdělení signálu na segmenty, ze kterých jsou vypočteny příznaky a na ně jsou aplikovány standardní algoritmy z toolboxu PRTTools. Výběr předzpracování i klasifikačních metod je proveden uváženě a je kladen důraz na nezávislé zhodnocení kvality dosažených výsledků. Na druhou stranu se výsledky klasifikace zdají s ohledem na provedenou rešerši méně přesné než nejlepší stávající metody, což není v práci téměř diskutováno. Samotný text práce je poměrně strohý a je obtížné přesně rekonstruovat některé části použité metodiky: např. postup při aplikaci navržených klasifikátorů na testovací data atp.	

Formální a jazyková úroveň, rozsah práce	C - dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je psána v angličtině s poměrně velkým množstvím gramatických chyb, text je nicméně srozumitelný. Na několika místech v práci jsem postrádal podrobnější popis, který by přesně popisoval použitou metodiku (např. str. 16 – jak přesně jsou příznaky extrahovány z vícekanalového signálu, popř. str. 29, 4.1. validation methodology – aplikace klasifikátorů na trénovací/testovací data apod.), což částečně kompenzuje obsah příloženého DVD, obsahujícího zdrojové kódy v Matlabu. Některé grafy jsou nedostatečně popsány (např. ROC křivky ve Fig. 18 – není jasné, co znamenají jednotlivé barvy apod, Fig. 12 obsahuje 10 metod předzpracování, ale přidružená tabulka 9 vysvětluje pouze 5 z nich atd).	

Výběr zdrojů, korektnost citací

C - dobře

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Studentka nastudovala a vhodně citovala přiměřené množství odborné literatury včetně primárních zdrojů, a výsledky aplikovala do vlastního řešení. V zadání uvedená literatura však nebyla v práci citována, přestože by se nabízela jako vhodná např. do části 1.3. „State of the art“ – jeden ze článků je odborná review, která by mohla snadno přispět k lepšímu přehledu o stavu vědění v dané problematice.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Celkově dosáhla studentka poměrně dobrých výsledků v klasifikaci vzorků, předcházejících epileptickému záchvatu. Vyřešení této úlohy, která je v současnosti horkým výzkumným tématem, by mohlo přinést výrazné zlepšení života pacientů s epilepsií. Výsledky se zdají horší v porovnání s citovanou literaturou, na druhou stranu patrně neexistuje konsenzus, jak dané algoritmy vyhodnotit - u naměřených signálů v čase není a-priori patrné, kde začínají vzorky „bezprostředně předcházející záchvatu“. Studentka musela tedy sama navrhnout postup hodnocení přesnosti klasifikace a výsledky nejsou tedy s ostatními pracemi přímo porovnatelné, což bohužel není v práci příliš diskutováno.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Studentka prokázala schopnost analyzovat předložený problém, navrhnout a vyhodnotit velké množství možností metod pro předzpracování, provést adekvátní automatickou selekci příznaků a výběr vhodného klasifikátoru. Celkově považuji navržené řešení za kvalitní a dobře provedené, práci však sráží strohý text, který nedokumentuje zcela přesně navržený postup, ani dostatečně nediskutuje dosažené výsledky. K obhajobě předkládám následující otázky:

1. Objasněte použité klasifikační schéma, zejm. s ohledem na testovací data. Které parametry klasifikátoru (příznakový vektor, model a jeho parametry) byly zachovány z trénovacích dat, a které byly znovu naučeny na testovacích pacientech?
2. Na str. 29 uvádíte, že data z jedné hodiny mohou být závislá a není vhodné je rozdělovat do trénovací/testovací množiny. Pokud se jedná o „celé hodiny“: jak se liší závislost vzorků získaných ve 11:50 a 12:00 od závislosti vzorků získaných ve 12:00 a 12:10? Objasněte a zdůvodněte Vaše rozhodnutí.
3. Na str. 13 uvádíte, že (na základě stat. testu) „we can not accept nor reject dependency“ – prosím, vysvětlíte, popř. korigujete.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře**.

Datum: 5.6.2017

Podpis: