

Posudek oponenta diplomové práce

Autor práce: **Filip Jani**

Název práce: **Workflows for electrophysiological data processing**

Obsah práce

Úvod práce je spíše obecný a cílům DP se příliš nevěnuje. V další kapitole již autor popisuje konkrétní potřeby neuroinformatické laboratoře KIV a postupně ozřejmuje, že cílem je poskytnout open source řešení pro zpracování elektroencefalografických (EEG) dat. Na základě poměrně podrobného srovnání systémů pro vytváření a správu workflow byl zvolen systém Orange3. Realizační část práce začíná kapitolou 6 Implementace. Zde se domnívám, že autor trochu zbytečně uvádí princip fungování uvedených metod (patří spíše do teorie) a uživatelské rozhraní (patří spíše do uživatelského manuálu).

Poměrně rozsáhlou práci (78 stránek) doplňuje přes 30 str. příloh obsahující zejména uživatelský manuál pro jednotlivé bloky.

Kvalita řešení a dosažených výsledků

Implementační část práce spočívá ve vytvoření potřebných pluginů (metod, bloků) do Orange3 v jazyce Python, které umožní zpracování EEG signálu a klasifikaci. Vytvořené metody jsou funkční a kódy celkem přehledné. Některé bloky jsou příliš specializované na jeden typ experimentu (např. Prepare Vectors, Create Classification Struct, Concatenate Classification Structs), což souvisí i s absencí parametrů. Kvůli obtížnější implementaci grafického rozhraní byla omezena konfigurace konvolučních neuronových sítí.

Ověření funkčnosti bloků autor provedl na základě tří workflow, které byly v minulosti použity a publikovány v rámci výzkumu v neuroinformatické výzkumné skupině. Vzhledem k občas složitě reprodukovatelným krokům v citovaných studiích (ruční odstraňování artefaktů, náhodná inicializace klasifikátorů) autor požaduje vizuální podobnost výsledků (ke srovnání nepoužívá statistické metody).

Formální úroveň

Práce je celkem vhodně členěna do kapitol a podkapitol. Je psaná dobrou angličtinou s ojediněle se vyskytujícími chybami (např.: "signal filtration" místo filtering v úvodu), nebo překlepy. Práci doplňuje dostatečné množství obrázků a tabulek. Škoda jen, že v případě prvního workflow na ověření práce nejsou grafy/tabulky původních prací a práce autora blízko sebe, ani se stejnými měřítky, což trochu komplikuje jejich srovnání.

Práce s literaturou

Práce se odkazuje na 56 literárních zdrojů. Kromě on-line zdrojů a různých tutoriálů se vyskytují i vědecké články. Citace jsou vhodně umístěny v textu.

Splnění zadání

Zadání považuji za splněné bez výhrad.

Dotazy k práci

1. V průběhu hodnocení nástrojů pro správu workflow jste odmítnul NeuroPype kvůli licenčním podmínkám navzdory rozsáhlé knihovně metod pro EEG. Domníváte se, že v případě změny licenční politiky by byl tento nástroj vhodnější, zejména s ohledem na rozšiřitelnost?
2. Jaké vidíte možnosti praktického nasazení popsaného řešení v dalších typických workflow neuroinformatiky a v jakých případech se domníváte, že šetří čas ve srovnání se standardním skriptováním?

I přes uvedené výhrady lze konstatovat, že autor úspěšně zvládl složitou analýzu EEG dat, časově velmi náročné sestavování workflow a představil uvěřitelné výsledky experimentů. Navrhuji hodnocení známkou **velmi dobře** a práci doporučuji k obhajobě.

V Plzni 19. 8. 2020

Ing. Lukáš Vařeka, Ph.D.