

Hodnocení vedoucího diplomové práce

Autor/autorka práce: **Bc. David Pivovar**

Název práce: **Detekce vybraných aktivit diabetického pacienta 1. typu**

Aktivita studenta

Student měl stejné téma DP zapsáno již podruhé. Do 23. 11. 2021 si zprovoznil frameworky TensorFlow a Frugally-Deep. Do té doby se také seznámil se prvními články současného state-of-the-art v detekci příjmu karbohydrátů. Teprve poté následovalo období aktivní práce na DP, včetně konzultací s vedoucím DP.

Spolupráce s vedoucím

Student spolupracoval s vedoucím práce při konzultacích, které sám navrhoval, přes e-mail a virtuální hovory kvůli vládním nařízením v souvislosti s vládou vyhlášenými nařízeními kvůli onemocnění Covid.

Původnost práce a práce související

Diplomantem vytvořený programový kód, a jeho dokumentace, je původní.

Kvalita řešení

Ve zdrojovém kódu C++ diplomant použil nestandardní rozšíření C++/CLI a používal konverze jazyka C. Výkonost Savitzky-Golay filtru je velmi malá a hodila by se pro jednorázové vyhlazení. Pro opakované vyhlazování signálu je její implementace nevhodná. Kód C++ tříd a jejich metod mohl diplomant strukturovat o něco lépe.

V medicínském úvodu diplomant zbytečně míchá chronické a akutní procesy v těle diabetika, takže to místy vyznívá až příliš katastroficky. Jednotlivá tvrzení, postupy a uváděné hodnoty nejsou dostatečně citovány. Ačkoliv zřejmě pochází z webu cukrovka.cz, který má odbornou redakční radu, měl si tyto informace vyhledat v odborných publikacích. Pak by mohl čtenáři poskytnout i ucelený popis homeostázy cukrů. Nicméně popsal tu část, která je pro téma diplomové práce relevantní.

V celém textu práce opakovaně používá termíny, aniž by je vysvětlil. Např. v sekci 3 píše „Kvantitativní metoda je například analýza hlavních komponent.“, aniž by uvedl, co je hlavní a co vedlejší komponenta. V sekci 3.3.1 píše „Experiment ukázal, že výsledky jsou statisticky ekvivalentní.“, ale vůbec nevíme, s čím jsou ekvivalentní. V sekci 6 se pak čtenář ztrácí v tom, co je zevrubný popis neuronových sítí, a co je vlastní diplomová práce. U obrázku 6.3 je např. zdroj URL na jakési fórum, odkud se ale daný obrázek nedá dohledat. V sekci 6.1.3 popsal architekturu vlastní neuronové sítě, ale vůbec nezdůvodnil, proč ji takto navrhnul. Pouze v sekci 6.2 se dozvíme, proč použil diskriminační analýzu. Na stránce 52 se pak objeví fragment detekčního algoritmu, ve kterém je použita nevysvětlená proměnná „activation“ a magické konstanty, které také nejsou nijak zdůvodněny. Vlastní algoritmus detekce příjmu karbohydrátů popsal velmi málo a v sekci o detekci fyzické aktivity se pak na něj ještě odkázal. Algoritmy jsou tak vlastně popsány jenom svým zdrojovým kódem.

V práci není popsán programátorský model SmartCGMS, takže sekce 8.3 nedává mnoho smyslu. V sekcích 8.2.1 a 8.3.1 chce interpolací dopočítávat chybějící hodnoty, ačkoliv toto by dělat neměl. Namísto toho měl pracovat s konceptem segmentace kontinuálně měřených hodnot.

Tabulky v textu práce jsou ve skutečnosti vloženy bitmapy, které podle některých artefaktů vypadají jako snímky obrazovky z programu MS Excel. To by např. vysvětlovalo i to, proč třeba v tabulce 9.1 chybí jednotky u absolutních hodnot.

Jako vedoucí práce jsem část textu DP viděl pouze jednou, celý text až po odevzdání DP, které se mnou diplomant nekonzultoval. Pouze několik hodin před mezním termínem odevzdání mi poslal zdrojové soubory, které jsem mu ještě poslal k opravě. Výsledná kvalita práce tomu, bohužel, odpovídá.

Využitelnost dosažených výsledků

Diplomant si správně uvědomil, že výsledky v citovaných publikacích nemusí být vzájemně porovnatelné, protože byly získány na jiných datech. Také si správně uvědomil riziko, že pacient nemusel zaznamenat všechna jídla a ani všechny fyzické aktivity. Jenomže problém je v tom, že s výchozími „thresholdy“ filtry generují příliš mnoho falešně pozitivních výsledků. To není nutně vina diplomantovy práce, ale složitosti tematiky. Bylo by tedy třeba, aby filtr evaluation generoval signál „error metric“, takže by SmartCGMS mohlo optimalizovat hodnoty „thresholdů“. Tato funkcionality však již v diplomové práci není. Detekci pomocí neuronové sítě se vedoucímu DP nepodařilo zprovoznit. Ačkoliv diplomantem vytvořený filtr úspěšně načtl natrénovaný model, výstupem byla téměř vždy nulová pravděpodobnost příjmu karbohydrátů.

K vyhlazování dat je použit filtr Savitzky-Golay, který však neumožňuje on-line vyhlazování. Diplomant měl vytvořit samostatný skript, který by data vyhladil jednou, a poté by byla analyzována. Namísto toho však diplomant upravil Savitzky-Golay filtr tak, prováděl on-line vyhlazování signálu. Z principu fungování Savitzky-Golay filtru to však může vést k nerealistickému signálu, který pak může být příčinou dalších chyb. On-line vyhlazování, o které se diplomant pokusil, je žádoucí, a je správné, že se o něj pokusil. Měl však zvolit jinou metodu, např. exponenciální klouzavý průměr.

V rámci obhajoby navrhuji, aby diplomant zodpověděl následující otázky:

- Čím by bylo vhodné nahradit filtr Savitzky-Golay, aby umožnil výpočetně únosné on-line vyhlazování signálu, a nebyl by to ani exponenciální ani vážený klouzavý průměr?
- Jak by bylo vhodné implementovat realistický „error metric“ signál, aby SmartCGMS mohl sám optimalizovat parametry pro jednotlivé pacienty?

Splnění zadání

Všechny body zadání byly splněny a práce měla velký teoretický potenciál. Bohužel ale na ní diplomant začal reálně pracovat příliš pozdě. Takže ačkoliv měl dobrý start, který prezentoval na FAV SVOČ, byl pozdě. To se pak projevilo výše uvedenými nedostatky. Jinak slibnou práci mi tak ve finále nezbyvá než hodnotit známkou **dobře**. Tímto prací doporučuji k obhajobě.