

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Interpretace pozitivní detekce statisticky nezvládnutého stavu při statistické regulaci procesů
Jméno autora:	Ondřej Misař
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Vedoucí práce:	Ing. Martin Macaš
Pracoviště vedoucího práce:	CIIRC

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
<p>Detekci statisticky nezvládnutého stavu jsme měli poměrně důkladně prozkoumanou. Aby mohla být nasazena v reálném světě, bylo nutné zvýšit její důvěryhodnost a akceptovatelnost z pohledu uživatelů. K tomu významně přispívá interpretovatelnost výsledků detekce, která zde znamená odhad toho, které vstupní proměnné nejvíce způsobují pozitivní detekci podezřelého stavu procesu. Interpretace je problémem nejen metod strojového učení, ale většiny vícerozměrných metod statistické ho řízení procesů. Zadání bylo pro bakalářskou práci náročnější. Aby student mohl pracovat na interpretaci detekce statisticky nezvládnutého stavu, musel nejprve implementovat a pochopit samotnou detekci. Složitější byla i experimentální činnost, hlavně návrh evaluace a porovnání metod.</p>	

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>	
<p>Student provedl poměrně důkladnou rešerši, na jejímž základě vybral a implementoval některé metody. Dále vygeneroval syntetická data a na nich metody porovnal pomocí vhodně navržených kritérií. Tím byly všechny body zadání splněny.</p>	

Zvolený postup řešení	vynikající
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
<p>Pro porovnání byly správně vybrány tři velmi různorodé metody interpretace. První byla analytická metoda, která je původně navržena pro jeden konkrétní interpretovaný model detekce podezřelého stavu a není nezávislá na interpretovaném modelu. Další dvě vybrané metody se interpretovat učí. Jedna globálně pomocí učení neuronové sítě, jedna učí lokální vysvětlitelný model.</p> <p>Student velmi záhy zjistil, že porovnání metod není jednoduché. Nelze hodnotit pouze to, zda metody vrátí nebo nevrátí přesnou množinu posunutých znaků kvality, ale je nutné vzít v úvahu, jak hodně se odhadnutá množina posunutých znaků liší od skutečnosti. Interpretace tedy predikuje podmnožinu posunutých znaků, čímž se evaluace poměrně komplikuje. Student se zde inspiroval konceptem sensitivity a specificity, který na tento problém aplikoval, čímž byl schopen vyhodnotit, zda daná metoda má spíše tendenci některé skutečně posunuté znaky opomínat anebo naopak vrací znaky, které posunuté nejsou. Tato evaluace je podle mě správná.</p>	

Odborná úroveň**A - výborně**

Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.

Interpretace je něčím, s čím se student během svého studia nesetkal. To kompenzoval důkladnou rešerší a nalezením a pochopením několika metod. Model detekce stavů procesu založený na klasifikaci do jedné třídy student také nastudoval a velmi dobře pochopil i problematiku reálného řízení procesu obrábění motorových hlav, které posloužilo jako motivační use-case, v kterém je interpretace nutná. K odborné stránce nemám jako vedoucí žádné výhrady.

Formální a jazyková úroveň, rozsah práce**B - velmi dobře**

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Práce je psána velmi dobrou angličtinou. Prostor pro zlepšení je vždy, nicméně jazykově je práce na nadprůměrné úrovni. Musím zde také vyzdvihnout poměrně velkou náročnost popisu některých částí (např. evaluačních kritérií nebo popis generování experimentálních dat) pro studenta bakalářského studia. Pro nezainteresovaného čtenáře, popis někdy nemusí být zcela jasný a čtenářsky „přátelský“, ale jen málo lidí by ho podle mě zvládlo lépe. Rozsah práce je naprosto vyhovující. Popis je nicméně úplný. Popis metody MYT v sekci 3.2 by podle mě měl být jasnější. Celá kapitola 5 popisující hlavní závěry a budoucí práci by mohla být delší a propracovanější.

Výběr zdrojů, korektnost citací**A - výborně**

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.

Student provedl důkladnou rešerší a vybrané články korektně citoval. Převzaté obrázky doplnil komentářem „převzato z ...“ a citací. Citační zvyklosti a normy byly podle mě dodrženy.

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.

Součástí práce bylo generování syntetických dat, tedy hodnot znaků kvality, ze kterých jsou detekovány podezřelé stavy procesu. Zde jsem si vědom toho, že generovaná data jsou statická, jsou značně zjednodušena, neobsahují časový kontext a autokorelované proměnné, které se v reálu většinou vyskytují. Generování realističtějších dat je ale poměrně složité a přesahovalo by předpokládanou obtížnost bakalářské práce.

Nejzajímavějším přínosem práce je použití metody LIME. Student použil metodu LIME na interpretaci pozitivní detekce statisticky nezvládnutého stavu, což je stanovení podmnožiny znaků, které jsou za pozitivní detekci zodpovědné. V literatuře jsem tuto aplikaci LIME zatím nenalezl. Podstatnou nevýhodou metody oproti interpretaci pomocí neuronové sítě se ukázala být nutnost předem znát velikost takové podmnožiny, tedy u kolika znaků došlo ke změně. Tuto znalost ovšem v našem případě nemáme. Proto student navrhl jednoduchou heuristiku, která ale dle experimentální analýzy opomíjí některé znaky kvality, které jsou posunuté. V nerealistickém případě, že metoda LIME disponuje znalostí počtu posunutých znaků, je tato metoda srovnatelná s neuronovou sítí nebo ji dokonce překonává. Stanovení počtu znaků kvality, které mohou za pozitivní detekci podezřelého výrobního procesu student tedy správně navrhl jako téma pro budoucí zkoumání. Interpretace pomocí neuronové sítě se zdá být nejlepší a nejrobustnější. Výsledek práce je celkově velmi přínosný.

Student prokázal velkou samostatnost a schopnost rychle pochopit zadání a aktivně hledat řešení. Přestože práce je dle

mého názoru publikovatelná například v konferenčním sborníku, student ji zatím nepublikoval, což je škoda.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Student prokázal schopnost rychle se zorientovat v zadané problematice, aktivně a efektivně komunikovat řešené problémy a pracovat v týmu. Jeho práce přinesla cenné poznatky, ale upozornila na další otázky, které je potřeba v budoucnu zodpovědět. Někdy až přílišná stručnost na úkor srozumitelnosti je kompenzována dobrými citacemi zdrojů, v kterých jsou detailní popisy. Zadání práce a popis celého problému byly pro bakalářskou práci poměrně náročné a student se toho zhostil nadprůměrně dobře.

Proto předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Datum: 5.6.2022

Podpis: