

## Posudek vedoucího bakalářské práce

### **Jiří Jurík: Samoučící algoritmy strojového učení pro analýzu průmyslových dat**

V bakalářské práci (programu TZSI, tedy ve spíše obecnějším typu práce řešeršního charakteru), se měl pan Jurík seznámit s pro něj zcela novou problematikou analýzy průmyslových dat a strojového učení, které přesahují výuku základního bakalářského studia na FS. Nové pro něj byla i nutnost znalosti programování, aby si mohl tyto algoritmy vyzkoušet (alespoň demonstračně, i když pokud možno na relevantních datech)

Student se o problematiku zajímal již během základního studia na FS a jeho zájem zřejmě pocházel i z působení jako technická podpora projektu v CERN na Ústavu Fyziky (naší Fakulty).

Cílem zadání této práce bylo, aby se student seznámil s aktuální problematikou analýzy průmyslových dat (data z linek, výrobní data, data pro prediktivní údržbu, detekce poruch vyjma obrazu) a s principy a využitím třídících („klastrovacích“) a samoučících algoritmů. Toto bylo v zadání jasně specifikováno.

Odevzdaná práce je pěkně a domnívám se i pečlivě zpracována, obzvlášť když vím, že student její sepisování přece jen nechal relativně dost na poslední chvíli.

Práci rozdělil na teoretickou část, ve které se s algoritmy seznámil a celkem zdařile je zdokumentoval, v praktické části se je potom snažil použít na umělá a reálná data. Sice celkem nelogicky nejprve analyzoval reálná data a potom až demonstroval, že algoritmy fungují na umělých datech, ale srovnání metody PCA s autoenkodérem pro odstranění šumu považuji za pěkný a užitečný příklad.

Pokus o obtížnou úlohu klasifikovat kompozitní struktury skenované laserovým vibrometrem (v rámci projektu řešeného na našem Ústavu) zatím výsledek bohužel nepřinesl, ale student se o to pokusil a v tomto výzkumu může nadále pokračovat.

Slabinou teoretické části práce není kvalita zpracování nebo pochopení algoritmů, to autor zvládl podle mne velmi dobře, slabinou je, že teoretická část je napsána příliš obecně o algoritmech jako takových a není řešeršně zaměřena na využití pro průmyslová data, jak jsem doufal. Také kapitola o průmyslových datech a big datech je velmi povrchní a je škoda, že student se tomuto důsledněji nevěnoval.

Nicméně délka 50 stran dobře překračuje standardní rozsah TZSI prací a práci považuji za zdařilou a může dobře posloužit i dalším zájemcům pro které je tato problematika nová.

Dotazy:

- 1) V kapitole 2 se snažíte pojednávat o průmyslových datech, uvádíte normu ISO 13374. Co na Obr.2 znamená (UI) a proč tam je? Jak to souvisí s normou?
- 2) Proč jste nezmínil i jiné normy nebo standardy týkající se průmyslových dat? Znáte ještě nějaké?

Navrhuji hodnocení známkou **C-dobře**, a souhlasím s obhajobou.