

## Posudek vedoucího diplomové práce

Student: Nika Beranová  
Číslo studenta: E15680  
Název diplomové práce: Detekce finančních podvodů metodami strojového učení  
Cíl práce: Cílem práce pro detekci finančních podvodů metodami strojového učení bylo zabezpečit sběr dat, popsat použitá data, charakterizovat vybrané metody strojového učení, navrhnout vhodný predikční model, verifikovat jej na předzpracovaných datech a výsledky statisticky porovnat.  
Vedoucí práce: doc. Ing. Petr Hájek, Ph.D.  
Studijní program: Systémové inženýrství a informatika  
Akademický rok: 2017/2018

### Náročnost tématu

	výborně	velmi dobře	vyhovující	nevyhovující	nelze hodnotit
Teoretické znalosti	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vstupní údaje a jejich zpracování	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Použité metody	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### Kritéria hodnocení práce

	výborně	velmi dobře	vyhovující	nevyhovující	nelze hodnotit
Stupeň splnění cíle práce	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Původnost zpracování tématu	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka provedené analýzy (ve vztahu k tématu)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba práce a rozsah	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s českou a zahraniční literaturou včetně citací	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Formální úprava práce (text, grafy, tabulky)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková úroveň (styl, gramatika, terminologie)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### Využitelnost výsledků práce

	vysoká	střední	nízká	nelze hodnotit
Pro teorii	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pro praxi	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Identifikace záznamu:  
Agenda závěrečných prací

.....

## Ostatní připomínky k práci

Detekce finančních podvodů metodami strojového učení lze považovat za vysoce aktuální téma, neboť strojové učení je stále více používáno jako nástroj na podporu rozhodování nejen regulátory finančních trhů, ale také ostatními zainteresovanými osobami. Jak se ale ukazuje i v této diplomové práci, odlišit od sebe podvodné a nepodvodné subjekty je stále složitější, což vede k použití komplexních metod strojového učení. Ke splnění cíle práce tedy byly potřeba nejen rozsáhlé teoretické poznatky z oblasti strojového učení, ale také z oblasti financí.

Diplomantka ve své práci postupuje logicky. Nejprve definuje a klasifikuje finanční podvody. Dále představuje současné přístupy k jejich včasné detekci. Zde oceňuji zejména pečlivou práci se zahraniční literaturou. V dalších kapitolách jsou definovány použité metody a popsán sběr dat. Použité atributy jsou opět podepřeny relevantní literaturou. Samotný predikční model zahrnuje tři základní klasifikátory a tři meta-klasifikátory v současnosti používané v souvisejících studiích. Použitá měřítka kvality predikce jsou dostačující. V současnosti se k tomuto účelu používají také náklady klasifikace, ale ty se v dostupné literatuře dosti liší (od 1:2 do 1:100). Navíc diplomantka pracuje s různými typy finančních podvodů, takže stanovení těchto nákladů by bylo složité. Z výsledků je zřejmé, že diplomantka provedla velké množství experimentů s cílem maximalizovat použitá měřítka kvality predikce na testovacích datech. Výsledky jsou doplněny vhodně zvolenými neparametrickými statistickými testy. I když se výsledná kvalita predikce zdá být nižší než u obdobných studií, autorka dosažené výsledky v závěru dostatečně vysvětluje.

Formální úprava odpovídá požadavkům na tento typ prací. Přílohy jsou funkční, obsahují ukázkou dat, nastavení experimentů i výsledky jednotlivých experimentů.

## Vyjádření k výstupům ze systému Theses

Kontrola plagiátorství: nejvyšší míra podobnosti 0 %, práce není plagiát.

## Otázky a náměty k obhajobě

Vzhledem k tomu, že diplomantka práci pravidelně s vedoucím práce konzultovala, nemám do diskuze žádné další otázky.

## Závěrečné hodnocení

Práci **doporučuji** k obhajobě.

Tuto diplomovou práci navrhuji hodnotit známkou: **A**

V Pardubicích 22.5.2018

Podpis .....  .....