



# Posudek oponenta závěrečné práce

**Oponent práce:** Ing. Mgr. Ladislava Smítková Janků, Ph.D.  
**Student:** Michal Šolc  
**Název práce:** Detekce podvodných finančních výkazů v KB  
**Obor / specializace:** Znalostní inženýrství  
**Vytvořeno dne:** 8. června 2021

## Hodnotící kritéria

### 1. Splnění zadání

- [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Výhrady jsou popsány níže, zahrnují primárně chybějící popis formátu vstupních dat a přesný popis transformace dat na datasety (tj. popis zajišťující reprodukovatelnost metody na jiných datech, absence originálních dat nevadí), a dále jsou některé části textové části práce dle mého názoru příliš zredukovány.

### 2. Písemná část práce

70/100 (C)

Práce tématicky odpovídá bakalářské práci a student v ní prokazuje přiměřené znalosti, nicméně zásadním nedostatkem práce je to, že text práce je nevyvážený. Některé části jsou zpracovány v dostatečném rozsahu, jiné mají naopak nedostatečný rozsah. Zásadní připomínky jsou: 1) jedním bodem zadání byla analýza existujících přístupů: kapitola 1 věnovaná tomuto bodu má 2 strany a obsahuje 3 odkazy, celkově ji lze jen obtížně nazvat analýzou, 2) v kapitole 3 chybí přesný popis formátu vstupních dat a transformace údajů z finančních výkazů na datasety, student pouze uvádí, že provedl redukci redundantních dat, a dále uvádí příklady několika sledovaných položek. Vynechání tohoto kroku de facto znemožňuje provést reprodukci postupu popsaného v práci. 3) V zadání je uveden bod, kdy má student diskutovat s vedoucími členy týmu specifikaci a význam vstupních dat, v práci není odkaz na proběhlou diskusi se členy týmu KB a výsledek této diskuse (buď ve formě rozboru vstupních dat, sledovaných položek výkazů či aspoň ve formě stručného shrnutí, co bylo v rámci tohoto bodu probráno apod.). 4) Nepřiměřeně stručný je závěr a kapitola věnovaná navrhovaným vylepšením.

Pozitivem práce je vlastní programovací část, popis aplikace vybraných metod a grafické a tabulární zpracování výsledků, které je zpracováno jasně, přehledně a v dostatečném rozsahu. Velmi pozitivně také hodnotím vyjadřovací styl studenta, z práce je patrné, že

problematiku popisuje vlastními slovy. Reference jsou až na výše zmiňovaný chybějící odkaz na ústní předání informací od členů týmu adekvátní, student v přiměřeném rozsahu cituje zdroje, též uvádí jasný výčet technologických nástrojů, které použil. Po typografické stránce je práce v pořádku. Celkově je patrné, že na poli programování a experimentů student odvedl poměrně dost práce, bodové ohodnocení snižuje nevyváženost textové části a chybějící údaje.

### 3. Nepísemná část, přílohy

70 /100 (C)

Nepísemná část obsahuje export jupyter notebooků, csv soubory s výsledky, atd. Z pochopitelných důvodů neobsahuje originální vstupní data (jednalo se o data klientů KB). Bohužel nejsou přiloženy ani vzorky virtuálních vstupních dat (nevycházejících z dat klientů), které by umožňovaly aspoň částečně si předložený software otestovat. Současně práce neobsahuje přesnou specifikaci vstupních dat a obsahuje jen velmi vágní popis transformace vstupních finančních výkazů z účetnictví na datasety (nereprodukovatelnost), takže nebylo možné si virtuální vstupní data vygenerovat. Funkčnost předloženého software tedy nemohla být testována. Nemohu též posoudit, zda předložené výsledky jsou výstupem předložených algoritmů či nikoliv. Kódy se jeví v pořádku, mohly by být lépe komentované.

### 4. Hodnocení výsledků, jejich využitelnost

75 /100 (C)

Výsledky jsou zajímavé jako studie použitelnosti vybraných metod AI k detekci podvodných výkazů. Praktickou využitelnost vzhledem k chybějícím datům (aspoň virtuálním) a chybějící specifikaci formátu vstupních dat nemohu posoudit (viz bod 3 posudku). V práci se autor odkazuje na spolupráci s KB, ale využitelnost výsledků není doložena písemným vyjádřením KB přiloženým k práci.

## Celkové hodnocení

75 /100 (C)

Po doplnění údajů uvedených v otázkách k obhajobě práci doporučuji k obhajobě.

## Otázky k obhajobě

- 1) Doplněte přesnou specifikaci formátu vstupních dat.
- 2) Doplněte postup transformace vstupních dat z výkazů na formát dat použitý v datasetech (redukci redundantních dat) a přesnou specifikaci formátu dat v datasetech. Proč přesná specifikace formátu dat není součástí práce?

## **Instrukce**

### **Splnění zadání**

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

### **Písemná část práce**

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

### **Nepísemná část, přílohy**

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

### **Hodnocení výsledků, jejich využitelnost**

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

### **Celkové hodnocení**

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.