



# Posudek oponenta závěrečné práce

**Oponent práce:** Ing. Magda Friedjungová, Ph.D.  
**Student:** Jan Jeníček  
**Název práce:** Generativní metody vhodné pro anonymizaci a tvorbu testovacích dat v bankovníctví  
**Obor / specializace:** Znalostní inženýrství  
**Vytvořeno dne:** 7. 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

Zadání považuji za splněné.

### 2. Písemná část práce

88 /100 (B)

Odevzdaná práce je logicky členěna a text v anglickém jazyce je dobře srozumitelný. Po písemné a informační stránce jsem s kvalitou práce spokojená, vytkla bych pouze několik následujících drobností.

Místy se vyskytují překlepy, drobné chyby ve formátování a v matematických zápisech (např. chyba ve vzorci 3.1). Místy se vyskytují nepřesné formulace jako např. F-Score se nepoužívá pouze pro binární klasifikaci. Obrázky by mohly být vektorové. V sekci 5.1 mi chybí citace. Dále student teoretickou část věnuje dvěma generativním modelům - VAE a GAN. Je škoda, že v praktické části však VAE nepoužívá, když s jeho popisem strávil tolik času. Argumentaci na straně 18 rozumím, jen by mi srovnání TVAE a CTGANu přišlo zajímavé. Dále bych vytkla popis sekce 5.2, kdy bych uvítala detailní popis výpočtu jednotlivých metrik. Navazující "SDV Evaluation" (str. 33) mi tak nepřijde srozumitelná, obzvláště pak "The average score aggregated from all metrics..." a porovnatelnost jednotlivých metrik.

### 3. Nepísemná část, přílohy

85 /100 (B)

Odevzdaný kód sice nelze vyzkoušet (data jsou KB), nicméně byl odevzdán i s výsledky jednotlivých buněk, samotný kód je komentovaný a přehledný. Vytkla bych snad jen to, že

student nepoužívá vlastní implementaci CTGANu, a tak většina praktické práce spočívá v přípravě datasetu, pipeline experimentů a vyhodnocení.

#### **4. Hodnocení výsledků, jejich využitelnost**

90 /100 (A)

Jedná se o komplexní práci s dobrými výsledky, která pravděpodobně poslouží dalším účelům zadavatele - Komerční banka.

#### **Celkové hodnocení**

85 /100 (B)

Práci navrhuji hodnotit stupněm B a to z důvodu, že praktická část má potenciál pro větší úsilí, přestože věřím, že práce dobře poslouží v praxi.

#### **Otázky k obhajobě**

Přesně vysvětlete použité metriky v "SDV Evaluation" (str. 33) a popsané v sekci 5.2, včetně výpočtu zmíněného průměru.

## **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. 26/2017, 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.