

## Posudek oponenta bakalářské práce

Jméno studenta	Monika Sochorová
Téma práce	Aplikace lineární a logistické regrese ve finančnictví
Cíl práce	Cílem je použití modelů lineární, logistické regrese a jim podobných pro predikování a klasifikaci dat z finanční oblasti s využitím knihoven jazyka Python. Práce by měla ukázat, jak lze efektivně pracovat se soubory dat s použitím knihoven numpy, pandas resp. scikit-learn a vyladit vhodný model strojového učení.
Vedoucí bakalářské práce	doc. Mgr. Dušan Bednařík, Ph.D.

náročnost tématu na	úroveň		
	nadprůměrná	průměrná	podprůměrná
teoretické znalosti		<b>x</b>	
praktické zkušenosti	<b>x</b>		
vstupní data a jejich zpracování		<b>x</b>	

kriteria hodnocení práce	úroveň			
	nadprůměrná	průměrná	podprůměrná	nelze hodnotit
stupeň splnění cíle práce	<b>x</b>			
samostatnost při zpracování tématu				<b>x</b>
logická stavba práce	<b>x</b>			
práce s českou literaturou včetně citací		<b>x</b>		
práce se zahraniční literaturou včetně citací				<b>x</b>
adekvátnost použitých metod	<b>x</b>			
hloubka provedené analýzy		<b>x</b>		
formální úprava práce (text, grafy, tabulky)		<b>x</b>		
jazyková úroveň		<b>x</b>		
nároky BP na podkladové materiály, konzultace, průzkumy ...	vysoké	průměrné	nižší	nejsou
použití analýz, matem. statistických a jiných metod, komparací apod.	ve velké míře <b>x</b>	přiměřené	částečné	absentuje
obsah a relevantnost příloh v textu či příl. části BP (tabulky, grafy, propočty apod.)	vysoce funkční	funkční <b>x</b>	méně funkční	neuspokojivé

Odpovídající hodnocení jednotlivých hledisek označte:

**x**

### Připomínky a otázky k obhajobě:

Práce je rozdělena do tří kapitol. V první kapitole studentka zavádí dvojici regresních modelů (lineární a logistický). Ve druhé kapitole je stručně popsána tvorba regresních modelů v prostředí Pythonu, včetně charakteristik knihoven NumPy, Pandas, Matplotlib, SciPy, scikit-learn a statsmodels. Konečně ve třetí kapitole jsou s využitím kódů v Pythonu vytvořeny a vyladěny dva regresní modely z oblasti finančnictví – predikce ceny nemovitosti a predikce schopnosti splácet úvěr.

Na bakalářské práci oceňuji v první řadě její ucelenou logickou výstavbu. Autorka se velmi dobře vypořádala s očekávaným rozsahem práce – uvádí vše podstatné, a přitom není výsledný text zbytečně rozsáhlý. Autorka prokázala, že je schopna v prostředí Pythonu navrhnout, vyladit a zhodnotit regresní model jak pro spojitou, tak diskrétní závislou veličinu. Ukázala též, že v tomto prostředí též ovládá zpracování a úpravu rozsáhlých podkladových datových souborů.

K práci mám pouze tyto méně zásadní připomínky. V tabulce 3.5 (str. 37) by bylo vhodné uvést rozdíl včetně znaménka (např. kladný, pokud odhad cenu nadhodnocuje, a záporný, pokud odhad cenu podhodnocuje). Práci se prolíná nejednotný způsob zápisu odrážek, bylo by vhodné ho sjednotit a dát do souladu s pravidly pravopisu. V textu se též objevují ne zcela vhodné formulace, např.:

str. 11 ... „*Jeho vzorec je ...*“, lépe: „*Vzorec pro jeho výpočet má tvar ...*“,

str. 13 ... „*... stačí, že se bude  $E(y_i)$  rovnat nějaké lineární kombinaci této funkce ...*“, lépe: „*... stačí, že se bude  $E(y_i)$  rovnat funkční hodnotě pro jistou lineární kombinaci nezávislých proměnných ...*“,

str. 40 ... „*... mezi proměnnými se nachází velké množství korelace ...*“.

Po stránce jazykové jsem pouze výjimečně narazil na pravopisné chyby nebo překlepy. Z hlediska formální úpravy textu (včetně obrázků, tabulek a grafů) se jedná o práci zdařilou.

K obhajobě mám tento úkol:

1. Na straně 42 je uvedena klasifikační tabulka, přesnost, senzitivita a specificita pro trénovací data. Prezentujte totéž i pro testovací data.

Studentka beze zbytku splnila zadání bakalářské práce. Bakalářská práce splňuje požadavky na ni kladené, doporučuji ji k obhajobě a hodnotím ji známkou **B**.

**Oponent bakalářské práce:**

Jméno, tituly: Mgr. Tomáš Zuščík, Ph.D.

Podpis:

V Hradci Králové dne 15. 8. 2022