

Posudek oponenta bakalářské práce

Jméno studenta	Andrea Šíroká
Téma práce	Zobecněný lineární model a aplikace logistické regrese ve financích
Cíl práce	Tato práce bude obsahovat popis Zobecněného lineárního modelu, především z matematického hlediska, ale i z hlediska použití. Součástí práce bude podrobný popis jednoho z modelů, tedy konkrétně klasické Logistické regrese. Logistická regrese se využívá i v oblasti financí, a proto cílem této práce bude aplikace modelu na reálných datech z oblasti investování na finančním trhu podle vhodně zvolených aspektů.
Vedoucí bakalářské práce	Mgr. Jitka Kühnová, Ph.D.

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

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

Odpovídající hodnocení jednotlivých hledisek označte:

x

Připomínky a otázky k obhajobě:

Práce je rozdělena na tři kapitoly. V první kapitole se autorka věnuje matematickému popisu zobecněného lineárního modelu a logistické regrese. Náplní druhé kapitoly je investování do akcií, jsou zde zavedena měřítka pro ocenění akcií důležitá ve třetí kapitole. V ní je popsáno využití logistické regrese při rozhodování o případné koupi akcií z předem definované skupiny.

Ve třetí kapitole autorka prokázala, že je schopna logistickou regresi použít v konkrétním aplikačním problému. Velkou oporu našla ve statistickém software (RStudio), který uměla vhodně využít. K této části nemám žádnou zásadní připomínku. Pouze bych autorce doporučil, aby u obhajoby co nejlépe vysvětlila podstatu rozhodování pro koupi akcií – objasnění

v podkapitole 3.3 je místy méně srozumitelné (např. co je myšleno modelem a co predikcí v tabulce 3.3, viz formulace „Podle modelu i predikce bychom ji koupili ...“).

Na první kapitole oceňuji, že se autorka poměrně dobře vypořádala s matematickým popisem zobecněného lineárního modelu a lineární regrese v přiměřené hloubce. Přesto mám k některým částem této kapitoly konkrétní připomínky. Některé z nich, společně s výhradami ke druhé kapitole, uvádím níže:

1. Na str. 11 je zavedena matice plánu \mathbf{X} tvořená pouze naměřenými hodnotami k nezávislých proměnných a následně vztahem (1.1) lineární závislost $\mathbf{Y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Při tomto pojetí není vůbec uvažován „absolutní“ člen. Jedná se o chybné propojení dvou různých přístupů. Pokud uvažujeme matici plánu ve tvaru uvedeném v práci, musíme uvažovat lineární závislost ve tvaru $\mathbf{Y} = \boldsymbol{\alpha} + \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Avšak v následném textu uváděné odvození odhadu vektoru parametrů předpokládá platnost vztahu (1.1), pak by ovšem musela být matice plánu typu $n \times (k + 1)$ – oproti matici plánu z práce je doplněna nulovým sloupcem tvořeným jedničkami. Vektor parametrů pak musí mít $k + 1$ složek $\beta_0, \beta_1, \dots, \beta_k$. Zde má β_0 význam „absolutního“ členu podobně jako α při prvním přístupu.
2. Chybné spojení dvou přístupů k matici plánu vede k paradoxním tvrzením, jejichž příkladem je poznámka 2 na str. 11: „Pokud existuje pouze jedna nezávisle proměnná \mathbf{X} mluvíme o jednoduché lineární regresi – jedná se o klasickou regresní přímku s dvěma parametry $\rightarrow k = 2$.“
3. Na str. 11 je chybně uvedeno $\hat{\mathbf{Y}} = \mathbf{Y} \cdot \mathbf{b}$, správně má být $\hat{\mathbf{Y}} = \mathbf{X} \cdot \mathbf{b}$.
4. Na str. 12 je uvedeno: „Snaha aproximovat proměnné co nejjednodušší funkcí je z důvodu toho, abychom mohli lépe popsat jak se data chovají a určit jejich průběh, vlastnosti a vzájemné vztahy. K tomu nám poslouží nejlépe lineární funkce, jejímž grafem je přímka.“ S touto myšlenkou lze velmi polemizovat, pokud by tomu tak bylo, postrádala by význam jakákoliv jiná než lineární aproximace.
5. Na str. 13 jsou chyby ve vztazích z věty 1.1 (jeden ze sumačních indexů má být místo i správně označen j a měl by probíhat hodnoty od 0 do k ; vektor parametrů $\boldsymbol{\beta}$ nenáleží \mathbb{R}^2 , ale \mathbb{R}^{k+1}).
6. Na str. 19 je možná zbytečně oproti předchozí části textu změněno označení počtu naměřených hodnot veličiny Y z n na N . Na několika místech je však dále chybně používán stále symbol n .
7. V kapitole 1.3 se objevuje podobně chybné propojení dvou přístupů k matici plánu a vektoru parametrů, jaké je popsáno v první připomínce. Definiční vztah pro lineární prediktor na str. 22 a vztahy (1.11), (1.12) a (1.13) na str. 23 využívají výše popsaný první přístup. Ten je v rozporu se vztahy používanými v části 1.3.2 (např. (1.15) a (1.16) na str. 25 a další na str. 26), které spíše svědčí o využití druhého přístupu. Současně je v části 1.3.2 v některých vztazích uvažováno $j = 1, \dots, k$ a v některých vztazích $j = 0, \dots, k$ (což je správně při druhém přístupu, kde $\beta_0 \equiv \alpha$).
8. Na str. 31 je uvedeno: „Podíl na akciové společnosti je dán nominální hodnotou akcie, což je hodnota, za kterou akcii koupil.“ Toto tvrzení je pravdivé pouze v případě prvního majitele, a to ještě v případě prodeje za nominální hodnotu. Při koupi akcie za tržní hodnotu (např. na burze) to však neplatí.
9. EPS není hlavní složkou používanou k výpočtu P/E ratio, jak je vedeno na str. 35, ale k výpočtu PEG ratio.

Přes výše uvedené věcné připomínky se jedná o práci na dobré úrovni, naneštěstí ji však sráží poměrně velké množství pravopisných chyb (zejména chyby v psaní čárek; viz např. citace ve čtvrté připomínce), občas se objevují gramatické chyby (především nesprávná vazba) a překlepy. Poslední připomínka směřuje k práci s použitou literaturou, která ne zcela respektuje rozhodnutí děkana PŘF UHK č. 4/2020 zejména při práci s díly, která nejsou explicitně citována, ale text bakalářské práce z nich volně vychází (pokyn zmínit odkaz na ně v úvodu části textu, který z nich volně vychází). Z hlediska formální úpravy textu se jedná o práci velmi dobře zvládnutou.

K obhajobě mám tyto otázky a úkoly:

1. V práci se používá pojem „podmodel“, případně také formulace „dva do sebe vnořené modely“. Nejsou však nijak vysvětleny. Učiňte tak prosím u obhajoby.
2. Na str. 23 uvádíte v souvislosti s funkcí logit: „... vidíme, že jde dokonce o kanonickou transformační funkci.“ Vysvětlete, proč tomu tak v případě logistické regrese je.

Bakalářská práce splňuje požadavky na ni kladené, doporučuji ji k obhajobě a hodnotím ji známkou **C**.

Oponent bakalářské práce:

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

Podpis:

V Hradci Králové dne 5. 6. 2020