

Oponentní posudek dizertační práce

Jméno doktoranda: Ing. Mgr. Michal Rychnovský, MSc

Název práce: **New Methods in Credit Underwriting**

Oponent: doc. Ing. Iva Pecáková, CSc.

Předložená práce je věnována aplikaci statistiky v oblasti finančního modelování. Autor usiluje o využití statistických a matematických metod pro zkvalitnění procesu schvalování spotřebitelských úvěrů. Problematikou se zabýval a zabývá dlouhodobě z různých stránek a na různých stupních svého studia i ve své odborné praxi a jeho znalosti tomu podle mého názoru odpovídají. Dizertační práce je napsána v dobře čitelné angličtině bez zjevných chyb, má odpovídající rozsah a splňuje formální požadavky kladené na dizertační práce.

Cíle práce jsou jasně zformulovány v jejím úvodu. Autor hodlá porovnat stávající a nově navrhované metody při modelování jednak pravděpodobnosti selhání dlužníka (defaultu), jednak ztráty z tohoto selhání plynoucí. Inovativní přístupu k modelování je přitom založen na využití analýzy přežití a Coxova modelu. Třetím cílem práce je na základě výsledků prvních dvou částí promítnout rizikovost klienta do očekávané profitability úvěru v rámci modelu, který by mohl být využit ve schvalovacím procesu. V neposlední řadě pak chce autor využít statistické metody k detekci větších koncentrací selhání a podvodů v úvěrovém portfoliu. V souladu s těmito čtyřmi cíli je práce rozdělena do čtyř kapitol, dále obsahuje Úvod, Závěr a seznam použitých a citovaných zdrojů (cca 80 položek). Text je provázen správně psanými a očíslovanými vzorci, doplněn přehlednými tabulkami a grafy.

K obsahu práce:

První kapitola práce je věnována modelování pravděpodobnosti selhání dlužníka. Dnes obvykle pro tento účel používaná metoda logistické regrese ke konfrontování s Coxovým modelem založeným na konceptu analýzy přežití. Principy použitých přístupů i hodnotících nástrojů jsou rekapitulovány velmi úsporně; je to proto, jak jsem uvedla výše, že autor se tématem zabývá dlouhodobě a může se tak odvolávat na své předchozí práce. Jsou zde používány standardní pojmy i symboly, které jsou v textu (až na nepatrné výjimky) definovány. Pro analýzy jsou používána reálná data. Drobnou výhradu zde snad lze mít jen k argumentaci založené na korelačních maticích počítaných pro všechny proměnné – ne všechny jsou totiž měřitelné. Zajímalo by mě také, proč jsou v tabulce 1.9 (na str. 38) a dále 1.17 na (str. 43) či 1.21 (na str. 45) obsaženy jako vysvětlující proměnné indikátory pro „distribution channel“, když p -hodnoty testů jsou zde nad 0,05?

V závěru první kapitoly autor konstatuje, že modely jsou z hlediska náhodně zvolené validační části souboru srovnatelné, avšak z pohledu nesporně užitečného kritéria, a sice validační části tvořené nejnovějšími údaji, Coxův model logistickou regresi mírně překonává. První

zjištění je zasazeno do kontextu stejných závěrů u jiných autorů, druhé je patrně originálním výsledkem autora práce - ? Zde možná mohla být příslušná formulace jasnější.

Ve druhé kapitole jsou opět konfrontovány různé přístupy (lineární regrese, logistická regrese a model přežití) při modelování ztrátovosti úvěru prostřednictvím míry jeho návratnosti (*recovery rate*) se závěrem, že analýza přežití poskytuje poněkud lepší výsledky. Pokud jde o použítá (opět reálná) data, na rozdíl od předchozí kapitoly je zde menší prostor pro případné námitky. Autor konstatuje, že jedna vysvětlující proměnná je kategoriální a ostatní číselné a že další informace nelze s ohledem na poskytovatele dat uvést. *Věcné* informace nikoliv, podle mého názoru však podrobnější informace o *charakteru* číselných proměnných uvedeny být mohly. Vyvstává totiž otázka, zda bylo nutné volitelně nastavovat hodnoty binární vysvětlované proměnné. Podle zařazených grafů to *patrně* nutné bylo. Podrobnější vyjádření v tomto směru v textu by nicméně bylo pro čtenáře užitečné.

S tím totiž souvisí i otázka konfrontace použitých modelů, pro niž autor navrhl vlastní modifikaci koeficientu determinace s odůvodněním, že vysvětlovaná proměnná jako taková vlastně není binární. Vzhledem k tomu, že zobecněné lineární modely jsou odhadovány metodou maximální věrohodnosti a nikoliv metodou nejmenších čtverců, byly pro ně navrženy různé varianty „pseudo-R²“ statistik. Pro řešenou úlohu nebyly použitelné?

Kapitola 3 je nejspíš zajímavým příspěvkem do oblasti finančního modelování. Směřuje k vytvoření matematického modelu, na jehož základě by mohla být odhadována ziskovost jednotlivých půjček na základě charakteristik klientů a půjček samých. Pro vyjádření očekávané ztráty, která musí nutně do modelu vstupovat, je potřebná pravděpodobnost nesplacení (selhání dlužníka) u každé platby, a také míra návratnosti (*recovery rate*) půjčky, které autor modeloval v předchozích kapitolách. Tradiční použití logistické regrese zde přináší určité problémy – autor v návaznosti na existující řešení navrhuje řešení vlastní, mj. opět s využitím analýzy přežití. Pro posouzení ziskovosti z různých hledisek definoval autor čtyři různé míry ziskovosti. Za účelem simulování různých modifikací automatizovaného schvalovacího procesu půjček na reálném datovém souboru autor model implementoval do MS Excel. V – řekněme – teoretické i praktické části této kapitoly prokázal schopnost aplikovat podle potřeby řešené úlohy statistické nástroje.

V poslední, čtvrté části práce autor hledá praxí poptávané řešení problému vyhledávání vysokých koncentrací selhání dlužníků (*defaultů*) v portfoliu, tedy vlastně odhalování jejich potenciálně podvodného chování. Zatímco u ostatních kapitol je nesporný vědecký přínos, u této kapitoly bych ho viděla spíše v rovině praktické. Pojednávané přístupy jsou dílem relativně jednoduché, dílem jde o jednu praktickou ukázkou použití (a porovnání jeho výsledku) sofistikovaných *dataminingových* nástrojů pro tento účel. S ohledem na název práce („nové metody“) možná mohl být prezentovaným jednoduchým přístupům věnován menší a novějším nástrojům poněkud větší prostor, už jenom proto, jak autor sám konstatuje, že jde o „pozorně sledované téma“ a „všechny nástroje s pozitivními výsledky v tomto směru jsou praxí dobře přijímány“.

Závěr:

Je zřejmé, že mnou vznesené připomínky k předložené práci nejsou nikterak zásadní. Celkově hodnocenou práci považuji za zajímavý příspěvek dokládající možnosti využití statistických metod v praxi úvěrových společností a motivující další výzkum v této oblasti.

Práci doporučuji k obhajobě před příslušnou komisí a v případě jejího pozitivního průběhu doporučuji udělit autorovi vědeckou hodnost Ph.D.

30. 6. 2017



Iva Pecáková

Hodnocení disertační práce v oboru Statistika (VŠE v Praze)

(vyplňuje školitel a oponenti - příloha posudku oponenta)

Jméno doktoranda: Ing. Mgr. Michal Rychnovský, MSc.....

Název práce:New Methods in Credit Underwriting

Jméno hodnotitele:doc. Ing. Iva Pecáková, CSc. jako oponent.....

Kritéria hodnocení

1. Odpovídá název práce jeho obsahu (zcela, částečně)? ...zcela.....
.....
 2. Je vymezení cílů vyhovující, je uvedeno v úvodu, abstraktu?Cíle práce jsou jasné
vymezeny v úvodu i abstraktu.
 3. Jak jsou vytčené cíle splněny, je shrnuto v závěru, abstraktu? ...ano.....
.....
 4. Jsou zřetelně odlišeny metody převzaté z literatury a vlastní přístupy?ano.....
 5. Co je vědeckým přínosem (např. z hlediska teorie statistiky či metod analýzy dat)?
.....Inovativní aplikace analýzy přežití v oblasti modelování pravděpodobnosti selhání dlužníka,
modelování ztráty plynoucí z tohoto selhání; využití těchto modelů pro modelování profitability pro
schvalování retailových úvěrů
 6. Úroveň rešerše poznatků a použité literatury ve zkoumané oblasti: ...dobrá...
.....
 7. Úroveň aplikací, hodnocení a porovnání metod (zařazení tabulek a grafů):
.....dobrá.....
 8. Způsob vyjadřování, jazyková úroveň: ...práce je napsána srozumitelnou
angličtinou.....
- Jsou řádně definovány používané pojmy, zkratky a symboly?ano.....
.....
9. Formální úroveň (matematických výrazů, tabulek, grafů): ...dobrá
.....

Datum: 30. 6. 2017.....


Doc. Ing. Iva Pecáková, CSc.