

# DOPORUČENÍ ŠKOLITELE K VYKONÁNÍ OBHAJOBY DISERTAČNÍ PRÁCE

Doktorand: Ing. Monika Papoušková  
Fakulta ekonomicko-správní  
Univerzita Pardubice

Téma: Modelování skórovacích karet pro rozhodování v nebankovních finančních institucích

Studijní program: Aplikovaná informatika

Školitel: doc. Ing. Petr Hájek, Ph.D.  
Ústav systémového inženýrství a informatiky  
Fakulta ekonomicko-správní, Univerzita Pardubice

Předložená disertační práce Ing. Moniky Papouškové se zabývá problematikou modelování skórovacích karet (úvěrového rizika) pomocí metod strojového učení, zejména pomocí souborů klasifikačních a regresních metod, včetně jejich aplikace na reálných datech nebankovní finanční instituce. Práce je rozdělena do 8 kapitol a má 127 stran.

## *Aktuálnost disertační práce*

Modelování skórovacích karet (úvěrového rizika) patří mezi nejčastěji řešené klasifikační a regresní ekonomické úlohy. To je dáno nejen řadou problémů spojených s charakteristikou dat (nevyváženost tříd, odlišné náklady spojené s chybnou klasifikací, atd.). Problém tohoto modelování dále není záležitostí pouze bankovních institucí, ale také v poslední době také nebankovních finančních institucí a peer-to-peer platform. To umožňuje modelování skórovacích karet na stále větších datových souborech, což vede ke zpřesňování modelů založených na strojovém učení. Ty se v současnosti jeví pro uvedené úlohy jako nejefektivnější. Téma disertační práce proto považuji za vysoce aktuální. Přínosy této disertační práce navíc umožňují využití v mnoha dalších aplikačních oblastech s podobnou charakteristikou dat.

## *Dosažení cílů disertační práce*

Cílem disertační práce je navrhnout model skórovacích karet, který umožní modelování pravděpodobnosti defaultu a expozice při defaultu pomocí heterogenních kombinací metod strojového učení. Takto stanovený cíl je dostatečně náročný a inovativní, neboť výše uvedená charakteristika úlohy vyžaduje integraci klasifikačních a regresních modelů. V případě klasifikačního modelu autorka kombinuje algoritmus pro vybalancování tříd s heterogenní kombinací klasifikátorů. Pro minimalizaci nákladů chybné klasifikace navíc autorka navrhuje původní metriku relativních nákladů. Jedna z klasifikovaných tříd je navíc spojena s různými náklady, což autorka řeší druhým, regresním modelem. Tím je oproti současným přístupům dosaženo vyšší přesnosti. Vzhledem k tomu, že navržený model je empiricky otestován na reálných datech a dostatečně porovnán s dalšími v současnosti používanými modely, považuji cíl práce za splněný.

### *Postup řešení, výsledky a přínosy disertační práce*

Autorka ke splnění cíle rozdělila práci na 8 kapitol. Postup řešení problému je logický a je založený na běžně používané metodologii specifikované v kapitole 3. Současně používané metody strojového učení používané pro modelování skórovacích karet jsou přehledně představeny v kapitole 4. Klíčovou částí práce je kapitola 6, kde je navržen dvoufázový model a uvedeny výsledky modelování na v tomto kontextu poměrně rozsáhlém datovém souboru české nebankovní finanční instituce. Efektivnost navrženého přístupu je z výsledků zřejmá a podpořená statistickým porovnáním se současně používanými modely. Kromě původního přístupu k modelování úloh s nevybalancovanými třídami s různým rozdělením nákladů autorka ukazuje také ekonomické přínosy pro vybranou instituci. Jak je naznačeno v kapitole 8, navržený model může mít daleko širší využití, než je pouze modelování skórovacích karet. Jedná se o řadu dalších klasifikačních / regresních úloh s podobnou charakteristikou dat.

### *Formální úprava disertační práce*

Disertační práce je na dobré technické úrovni, je napsaná jasně a srozumitelně. Všechny obrázky jsou dobře čitelné. V práci je také velmi málo překlepů. V textu se objevuje pro skupinu autorů „et al.“ i „a kol.“. Autoři v tabulce 2 by neměli být uváděni v závorce. Obojí je ale způsobeno použitým citačním software, takže se nejedná o chybu doktorandky. Celkově je práce napsána velmi pečlivě.

### *Publikační aktivita autora*

Přehled vlastních publikací doktorandky je uveden na straně 112. Doktorandka je autorkou nebo spoluautorkou 3 článků v časopisech, z toho jeden s impaktním faktorem a jeden v časopise s SJR. Dále je autorkou nebo spoluautorkou sedmi příspěvků ve sborníku z konference. Nejvýznamnější publikace má vysoký impaktní faktor a reprezentuje výsledky dosažené v rámci řešení disertační práce. Je také v souladu se studijním programem. Publikační aktivita doktorandky je zejména v posledním roce studia na velmi dobré úrovni.

### *Závěrečné vyjádření školitele*

Doktorandka začala svou disertační práci zpracovávat pod vedením prof. RNDr. Viery Pacákové, Ph.D. Jejím školitelem jsem se stal v roce 2017, neboť se jako nejefektivnější při realizaci práce ukázaly modely založené na strojovém učení, zejména v porovnání s tradičními statistickými metodami. Díky tomu se doktorandka mohla zaměřit na návrh kombinací modelů strojového učení a dosáhnout v porovnání se současně používanými modely unikátních výsledků. Oceňuji také její schopnost samostatné vědecké práce a aplikaci navržených modelů na realizaci reálného ekonomického problému.

Disertační práce Moniky Papouškové „Modelování skórovacích karet pro rozhodování v nebankovních finančních institucích“ dle mého názoru splňuje požadavky kladené na doktorské disertační práce v oboru Aplikovaná informatika, a proto ji **doporučuji** k obhajobě.

V Pardubicích dne 5. dubna 2019

  
doc. Ing. Petr Hájek, Ph.D.