

## РЕЦЕНЗИЯ

От: **проф. д-р Стефан Христов Петранов**, заемащ длъжността “професор” в Стопанския факултет на СУ “Св. Климент Охридски”

Относно: дисертационен труд за присъждане на научната степен **„доктор на науките“** по професионално направление 3.8 Икономика; Икономика и управление (Индустрия) в СУ Св. Климент Охридски.

Автор на дисертационния труд: *Антон Антонов Герунов*

Тема на дисертационния труд: *Автоматизирани подходи за управление на операционния риск*

**Основание** за представяне на рецензията: участие в състава на научното жури по защита на дисертационния труд съгласно Заповед РД 38-232/24.06. 2020г. на Ректора на СУ Св. Климент Охридски.

### 1. Информация за дисертанта

Антон Герунов е доцент по управление на риска и количествени методи за анализ на публични политики в Софийския университет „Св. Климент Охридски”, където чете лекции по анализ на данни, дигитални бизнес стратегии и управление в дигитална среда. Той има научната степен „доктор по икономика“ от СУ Св. Климент Охридски и магистърски степени по икономика от СУ Св. Климент Охридски по икономика и от Стокхолмския Университет (Швеция) по компютърни науки. Бакалавърската му степен е от Университет Якобс (Бремен, Германия).

Извън академичната си дейност, авторът на предложения дисертационен труд има също така богат опит в държавната администрация като Началник на кабинета на заместник-министър председателя на Република България с основни ресори: електронно управление, административна реформа, стратегическа координация на ОП „Добро управление“, а също така и като консултант и оперативен директор на фирма в технологичния сектор. Тази комбинация от академичен фундамент и експертиза от администрацията и от технологичния бизнес, създава отлична основа за представената разработка.

## **2. Обща характеристика и актуалност на представения дисертационен труд**

Представената дисертация се състои от 364 страници, от които 332 страници основен текст, библиография и три приложения. Основният текст е разделен в уводна част, пет глави, заключение и три приложения, представящи апробираните в изследването класификационни и регресионни модели и разпределението на извадките за тези модели. Текстът е илюстриран с помощта на 141 графики и 79 таблици. Използваната литература е цитирана в библиографската справка, която включва 410 източника от български и чуждестранни автори.

Актуалността на темата на дисертационния труд не буди съмнение. Основание за това е фактът, че съвременната икономическа среда се характеризира с международна интеграция и глобализация, с висока динамика на пазарите, както и с дигитализация и автоматизация, които създават условия за проявата на редица рискови фактори. В същото време напредъкът на информационните технологии и развитието на аналитичните методи вече дават възможност с относително малки ресурси да се обработват големи масиви с данни, те да се анализират и да се извлича информация за взимане на решения. С оглед на тези условия авторът се е фокусирал върху автоматизирани подходи за управление на операционния риск чрез изследвания на приложението на методи от сферата на машинното самообучение и тяхната интеграция в цялостния управленски процес.

Отделни опити в това направление са налични и в международната и в българската специализирана литература, но в предлагания труд се постига съществено обобщение, като се извежда общ алгоритъм за управление на операционните рискове на основата на авторски разширения по метода на науката за проектиране.

## **3. Структура**

Представеният труд има структура и съдържание, отговарящи на изискванията за дисертация за научната степен «доктор на науките». Той е в достатъчно голям обем и съдържа оригинални разработки на автора. Личи си, че авторът е обвързан с практиката на разглежданата проблематика, което е предимство предвид естеството на разглежданите въпроси. Темата е разгледана изчерпателно, с широк обхват и с подходяща дълбочина и детайлност.

Като цяло работата е логично структурирана, подходът е унифициран. Преминава се от общата теория и типология на рисковете към конкретизиране на оперативния риск и приложението на класификационните и регресионните модели за управление на този риск. Използване са методи на анализ и синтез в контекста на системен интердисциплинарен подход. Изложението е ясно, последователно и логически свързано между отделните части.

#### **4. Съдържание на представения дисертационен труд**

**Увод.** Преди да пристъпи към изложението на своето изследване по същество, в Увода авторът представя основните характеристики на своята дисертация - научно-изследователска теза, цел и задачи на изследването, обхват и методология, информационно обезпечаване и др.

Научно-изследователската теза е формулирана от автора чрез четири съставни части, а именно:

1. Всички дейности, извършвани от експерти в процеса по управление на операционните рискове могат да бъдат автоматизирани.

2. Приложението на класификационни алгоритми от сферата на машинното самообучение за количествена оценка на рисковете може да подобри прогностичната точност спрямо традиционно използваните иконометрични методи и оттам да доведе до по-високо качество на резултатите и икономическа стойност.

3. Приложението на регресионни алгоритми от сферата на машинното самообучение за количествена оценка на рисковете може да подобри прогностичната точност спрямо традиционно използваните иконометрични методи и оттам да доведе до по-високо качество на резултатите и икономическа стойност.

4. Използването на няколко критерия за оценка на нивото на риск в рамките на автоматизиран алгоритъм ще доведе до групиране (клъстериране), а не до разнопосочност на резултатите. По този начин се гарантира надеждността на предложената система.

Въз основа на така формулираната теза, целта на дисертацията е «...да се изгради напълно автоматизиран процес за управление на операционните рискове, който да може да се възползва от широк набор от количествени методи за оценка и да бъде подкрепен от съответна специализирана информационна система“.

Към Увода могат да бъдат отправени следните бележки. На стр. 7 в раздела за изследователските задачи т.5 и 6. се говори за „най-оптимални

алгоритми“. Това е нелогично, защото няма по- или най-оптимални, а просто оптимални алгоритми. Оптималността съдържа най-доброто, според даден критерий и затова няма най-оптимално.

На стр. 12, при описанието на структурата на изследването (глава 3 и 4) целевата променлива се определя като „продължителна“, като тази терминология се среща впоследствие в текста многократно. Вероятно става дума за превод от английски език, но терминът е неудачен – в българската научна литература е трайно възприет термина „непрекъсната“ за такъв тип променливи.

**Първа глава** логично представя преглед на литературата в областта на управлението на риска. Разгледани са основните типологии на рисковете като цяло и на операционните рискове в частност. Разгледани са количествените и качествените методи за управление на операционния риск, както и общите стратегии и добри практики за този процес. В тази глава е изведена тезата, че има нужда от пълна или частична автоматизация на процеса по управление на операционния риск, защото макар количествените методи да предлагат висока прецизност и ефективност, то в настоящия си вариант те са прекомерно зависими от наличието на сравнително голям брой квалифицирани експерти, за да бъдат успешно приложени.

Имам една бележка към Първа глава. На стр. 14, при дефиниране на понятието риск, във формулата е допусната грешка – вместо в нея да участва сумата от вероятностите, погрешно е изписана очакваната стойност на случайната величина.

**Втора глава** дискутира новите развития в областта на управлението на операционните рискове и очертава основните тенденции в актуалните изследвания. В контекста на тези развития се извежда и общият алгоритъм за автоматизирано управление на операционните рискове. Той се състои от осем стъпки, отделени в пет основни управленски етапа - дефиниция на проблема, информационно обезпечаване, обучение на модел за управление на операционния риск, приложение на модела и действия за управление. Така представените етапи отчитат основните фази в управленския процес, но и позволяват алгоритмизиране на действията, така че те да могат да бъдат приложени в рамките на специализирана информационна система.

Нямам конкретни бележки към тази глава.

В **Трета глава** се идентифицират подходящи алгоритми за количествена оценка на операционните рискове в случаите на дискретна целева променлива. За тази цел се тестват 136 различни алгоритъма от сферата на статистиката и машинното самообучение, като те се прилагат за решаване на пет класификационни задачи от сферата на операционния риск.

Сред различните задачи се наблюдава преобладаващо добро представяне на методите от семейството на случайната гора, а от класическите иконометрични методи висока точност се постига от линейния дискриминантен анализ. Резултатите от изследването в тази глава дават основание да се изкаже тезата, че изборът на оптимален алгоритъм следва да бъде извършен чрез постигане на баланс между разходите за него, измерени чрез необходимите изчислителни ресурси и ползите от него, измерени чрез неговата прогностична точност.

Към главата могат да бъдат отправени няколко бележки. Текстът на стр. 144 се нуждае от по-голяма яснота. По-точно, не е изяснено какво означава отрицателен и положителен клас и как даден алгоритъм определя един клас като положителен или отрицателен.

На стр. 148 се казва, че пълната корелационна матрица на данните е представена в приложенията, но такова приложение няма. Също така, на стр. 147 една от дескриптивните характеристики на променливите е наречена „изместване“. Впоследствие този термин е използван многократно във всички таблици с дескриптивни статистики. Предполагам, че става дума за отношението между третия централен момент на променливата и стандартното отклонение, повдигнато на трета степен. Ако е така, по-добре би било да се използва термина „коэффициент на асиметрия“, който е трайно възприет в българската научна литература.

В Таблица 32 променливата съотношение общо пасиви/общо активи е представена като величина, която има средно 0.45. В същото време по дефиниция в счетоводството общо активите на фирмата са равни на общо пасивите и това съотношение би трябвало да е винаги 1-ца. Така че явно в таблицата става дума за друго съотношение.

**Четвърта глава** следва логиката на Трета глава, но разглежда случаи, когато целевата променлива на задачата е непрекъсната. Анализирани са както традиционни иконометрични методи, така и методи от сферата на машинното самообучение. Общо се тестват 109 различни алгоритъма, като те се прилагат за решаване на пет ситуации от сферата на операционния риск с регресионен характер.

Авторът представя резултати, които показват и в този случай много добро представяне на методите от семейството на случайната гора, но така също и на невронните мрежи. Представя и резултати, според които

класическите иконометрични методи на линейна регресия имат чувствително по-слабо представяне спрямо методи от областта на машинното самообучение и установява отново сравнително слаба връзка между прогностична точност и ресурсна интензивност.

Бележките ми към тази глава са следните. Има объркване на номерацията на разделите – от стр. 208 до стр. 221 подразделите (изписани с „италик“) започват с цифрата 5, докато основният раздел е с цифра 4. Освен това, заглавието на Таблица 45 и Таблица 46 е сбъркано, те не се отнасят към данните, показани в таблиците.

На стр. 230 авторът стига до заключението, че класическият модел на линейна регресия има най-слаби резултати. Това твърдение е пресилено, защото не се потвърждава от данните от Таблица 52. От тях се вижда, че от пет разглеждани модела, този модел има 2-ри резултат по критерия ME, 3-ти резултат по критерия RMSE (който в цялото изследване авторът възприема като основен), 2-ри резултат по критерия MAE и 5-ти по критериите MPE и MAPE. Т.е. резултатите са по-скоро смесени. В следствие този извод се обобщава на стр. 266, но е необходима по-убедителна аргументация.

На стр. 217 авторът отбелязва, че алгоритъмът на случайната гора, който дава най-добър резултат, стига до нелогичната зависимост, че рискът от прекомерно отсъствие на служители зависи статистически значимо от височината на служителите. В този случай коментарът е ограничен до едно изречение в смисъл, че това надделяване на формалните критерии над структурната логиката на модела е един от основните проблеми на алгоритмите за машинно самообучение. Моето мнение е, че би било от полза за работата, ако този проблем е разгледан по-детайлно. Защото цялостната теза на предложениия труд е за предимствата и силните страни на алгоритмите за самообучение и в този контекст, особено в дисертация за „доктор на науките“ би било добре да се коментират и техните слаби страни. По своята същина смисълът на такъв модел е да послужи за управление на рисковете, което в случая с примера на ръста на служителите не може да има практическа стойност.

Данните от Таблица 55 показват дескриптивни статистики на величината „цена на единица площ“ за имоти в Тайван. Като максимална такава цена (от 414 наблюдения) е показана 117.5 тайвански долара за 1 пинг. В същото време оцененият линеен иконометричен модел, според написаното в текста, показва, че увеличаването на възрастта на сградата с една година води до намаляване на цената с 2740 тайвански долара. Увеличаването на броя на близките магазини, от своя страна, води до увеличаване на цената с 1208 тайвански долара. Тези резултати са

невъзможни, при максимална цена от 117.5 долара и тук вероятно има някаква грешка. В текста на същата страница е сбъркана и таблицата – реферира се към Таблица 19, а трябва да е Таблица 56.

В Таблица 61 има разместване на редовете и данните от 2-ра до 6-та колона не съответстват на променливите от 1-ва колона. Ако се отчете това разместване, се вижда, че многомерният линеен регресионен модел постига коефициент на детерминация 1, регресионни коефициенти 1-ца за променливите Поръчки от вид А, В, С и 0-леви регресионни коефициенти за всички други променливи. Това е възможно да се получи, ако броят на общите поръчки (прогнозираната променлива) е равен на броя на поръчките от вид А, В и С, т.е., ако всяка поръчка в общия брой е от един от трите вида. От показаните данни в текста не може да се съди дали това е така, но ако е така, то чрез тези променливи не е правилно да се прогнозира общото търсене, защото зависимостта между тях и прогнозираната променлива не е регресионна, а е равенство по дефиниция. Би трябвало да се моделира търсенето на самите поръчки от вид А, В, С.

В Пета глава разглеждат възможностите за интегриране на предложени общ алгоритъм за управление на операционните рискове и съответните методи за количествена оценка на риска в рамките на специализирана информационна система. Извежда се референтна архитектура на информационна система, която съдържа четири основни подсистеми - подсистемата за съхранение и обработка, подсистемата за моделиране, подсистемата за анализи и подсистемата за управление. Извежда се консенсусен критерий за аномалност на базата на четири алгоритъма, като наблюденията, класифицирани като аномални от три от тях се определят за рискови, а тези класифицирани от четири – за високорискови. Направена е и оценка на съответствието с етичните изисквания за системи с автономни агенти.

Нямам конкретни бележки към тази глава.

## **5. Оценка на научните и научно-приложни приноси**

Дисертационният труд има научни и научно-приложни приноси, като от първите бих открил предложеният алгоритъм за автоматизирано управление на операционните рискове и извършените многобройни тестове с широк кръг модели от сферата на машинното обучение. От вторите откроявам предложени консенсусен критерий за определяне на рискови наблюдения и етичните критерии, разработени за анализ на

системи с автономно взимане на решения. Като цяло приемам изброените приноси на стр. 332-333.

## **6. Оценка на публикациите по дисертацията**

Основните резултати от дисертационния труд са апробирани в научната общност чрез публикации в авторитетни икономически издания. В материалите по процедурата за защита са представени общо десет публикации на Антон Герунов. Осем от публикациите са в реферирани научни списания в България и чужбина. Сред публикациите са монографичен учебник, три студии, както и пет статии, една от които е в индексирано в Scopus списание. От публикациите пет са на английски език, а останалите – на български език.

И в количествено и в качествено отношение, представените публикации отговарят на възприетите критерии за докторска дисертация, като особено тежат монографичният учебник Gerunov, A. (2017). *Notes on Risk Management*. Sofia University “St. Kliment Ohridski”, Faculty of Economics and Business Administration. ISBN: 978-954-9399-45-5; студията Герунов, А. (2020). Приложение на класификационни алгоритми за моделиране на икономически избори. *Икономическа мисъл*, 2, 45-67, както и статията Gerunov, A. (2019). Modelling economic choice under radical uncertainty: machine learning approaches. *International Journal of Business Intelligence and Data Mining*, 14(1-2), 238-253.

Дисертационният труд е оригинален и не повтаря темата или значителна част от съдържанието на труда, за който е получена образователната и научна степен „доктор”. С това той отговаря на изискването на чл. 76 (2) от Правилника за условията и реда за придобиване на научни степени и заемане на академични длъжности в СУ „Св. Климент Охридски”.

## **7. Оценка на автореферата**

Авторефератът представя пълно и обективно съдържанието на дисертационния труд.

## **8. Критични бележки и препоръки**

В Раздел 4 съм представил бележки към отделните глави на дисертацията. Тук ще дам една по-обща бележка, която се състои в следното. При коментиране на предимствата на алгоритмите, базирани на



машинно обучение не винаги е удачно сравнението с представените линейни многомерни регресионни модели, тъй като не е ясно доколко адекватни са те. Илюстрация в това отношение е моделът от Таблица 46. При 740 наблюдения и 21 кандидат-обясняващи променливи, моделът постига коригиран коефициент на детерминация от само 0.17. Това навежда на мисълта, че или кандидат-обясняващите променливи не са подбрани подходящо или зависимостите между тях и обясняваната променлива не са линейни или и двете. В този смисъл този модел не е убедителна база за сравнение.

## 9. Заключение

На базата на постигнатите резултати и изложените по-горе аргументи считам, че предложеният дисертационен труд на Антон Антонов Герунов отговаря на критериите и изискванията, залегнали в ПРАВИЛНИКА ЗА УСЛОВИЯТА И РЕДА ЗА ПРИДОБИВАНЕ НА НАУЧНИ СТЕПЕНИ И ЗАЕМАНЕ НА АКАДЕМИЧНИ ДЛЪЖНОСТИ В СУ „СВ. КЛИМЕНТ ОХРИДСКИ” и в ПРАВИЛНИКА ЗА ПРИЛАГАНЕ НА ЗАКОНА ЗА РАЗВИТИЕТО НА АКАДЕМИЧНИЯ СЪСТАВ В РЕПУБЛИКА БЪЛГАРИЯ. С настоящата рецензия изразявам моето **положително мнение** за присъждането научната степен “доктор на науките” на Антон Герунов.

30.8.2020/София

Подпис: .....

*(Проф. д-р Стефан Петранов)*