

5. Лелюк Ю.М. Побудова синтетичної економічної теорії як передумова формування методологічної бази антициклічної політики / Ю.М. Лелюк // Економічний простір. – 2012. – № 58. – С. 74-79.
6. Сазонець І.Л. Сутність та основні концепції діяльності ТНК в умовах глобалізації / І.Л. Сазонець, Ю.М. Варич. Донецьк: Юго-Восток, 2009. 275 с.
7. Тарасевич В.Н. О триединой субстанции и тайне институтов / В.Н. Тарасевич // Наукові праці ДонНТУ. Серія: економічна. – 2011. – Вип. 40-2. – С. 74-77.
8. Фомина Ю.А. Инновации, организация, устойчивость и кризисы с позиций «Международного общества Й. Шумпетера» / Ю.А. Фомина // TERRA ECONOMICUS. – 2011. – Т. 9. – № 1. – С. 146-156.
9. Чухно А. Институционализм: теория, методология, значения / А. Чухно // Экономика Украины. – 2008. – № 6. – С. 4-13.
10. Шубалий О.М. Институциональна трансформація економічних відносин у лісовому комплексі: теоретичний аналіз напрямків реалізації / О.М. Шубалий // Економіст. – 2010. – № 8. – С. 29-32.
11. Regionalization in A Globalizing World: A Comparative Perspective on Forms, Actors and Processes / Michael Schulz (Editor), Fredrik Soderbaum (Editor), Joakim Ojendal (Editor) // (2001)
12. Report on progress in creating the internal gas and electricity market, Commission of the European Communities, 15.11.2010. [Електронний ресурс] – Режим доступу: // [http://ec.europa.eu/energy/gas\\_electricity/doc/2010/com\\_2010\\_0084\\_f\\_en.pdf](http://ec.europa.eu/energy/gas_electricity/doc/2010/com_2010_0084_f_en.pdf)
13. The Origins of Mercosur: Democracy and Regionalization in South America (Studies of the Americas) / Gian Luca Gardini (Author) // algrave Macmillan (February 16 2010)
14. North, Douglass C. and Weingast, Barry W. 1989. «The Evolution of Institutions Governing Public Choice in 17th Century England». Journal of Economic History, 49: 803-832.
15. North, Douglass C. Forthcoming. «Institutions, Transaction Costs, and the Rise of Merchant Empires». In Tracy, James D.

519.865

*Дюличева Ю.Ю., к.ф.-м.н., доцент, Шаранов О.Д., магистрант,  
ТНУ ии имени В.И. Вернадского*

#### **МАТЕМАТИЧЕСКИЕ МОДЕЛИ В СИСТЕМАХ КРЕДИТНОГО СКОРИНГА**

Оптимизация банковской деятельности в сфере кредитования связана с решением задачи минимизации кредитного риска. Активный спрос на розничные и потребительские кредиты, несформированный рынок кредитования, завышенные проценты по кредитам приводят к тому, что украинские банки, запуская свои розничные программы с нуля, уже через полгода могут занять долю не менее 5% рынка [1, с.58]. Несмотря на доходность рынка кредитования в Украине, наблюдается рост необоснованных отказов заемщикам, что практически вызвано отсутствием в банках полноценных систем кредитного скоринга. В настоящее время украинские банки продолжают ограничиваться скоринговыми системами, реализованными в формате MS Excel. Такие системы не защищены от несанкционированного доступа, реализуют примитивный механизм принятия решений относительно выдачи кредитов, не способны к быстрому обновлению данных [2, с.31]. Таким образом, разработка и внедрение эффективных научно обоснованных методов минимизации кредитного риска, позволяющих автоматизировать процесс принятия решений относительно платежеспособности потенциальных заемщиков - аппликантов, является актуальной задачей для банков.

Исследованию различных видов кредитного скоринга и вопросов, связанных с разработкой и внедрением систем кредитного скоринга, посвящены многочисленные работы отечественных и зарубежных авторов, среди которых следует отметить работу Дюранда (David Durand), впервые рассмотревшего задачу кредитного скоринга как задачу бинарной классификации, а также работу Д. Веста (David West) по применению нейронных сетей в кредитном скоринге; работу Мира (Mira) и Тенкони (Tenconi) по применению модели байесовской иерархической логистической регрессии для прогнозирования кредитного риска компаний.

Разработка новых алгоритмов классификации и возможность их применения для решения задачи кредитного скоринга всегда будет представлять интерес для банков как эффективный аналитический инструмент для снижения кредитных рисков.

Целью статьи является исследование особенностей некоторых математических моделей, основанных на построении композиций обучающих алгоритмов и применяемых для разработки зарубежных и отечественных систем кредитного скоринга.

Одним из формализованных подходов для оценки платежеспособности заемщиков (application scoring) является использование модели кредитного скоринга, которая позволяет на основании анкетных данных аппликанта и кредитной истории вычислить балл или оценку (score), характеризующую вероятность невозврата кредита. Простейшая модель кредитного скоринга заключается в построении линейной скоринговой функции вида  $s(X) = x_1\omega_1 + x_2\omega_2 + \dots + x_n\omega_n$ , которая по заданному набору признаков или атрибутов  $x_1, x_2, \dots, x_n$ , характеризующих аппликанта, определяет весовые коэффициенты  $\omega_i$  атрибутов  $x_i$ , что позволяет вычислить величину или балл  $s(X)$ , по которому аппликант  $X$  будет отнесен либо к классу «благонадежных», либо к классу «неблагонадежных» заемщиков.

Исследуем особенности математических моделей, применяемых для разработки систем кредитного скоринга.

1) Байесовские сети (Bayesian Network) в кредитном скоринге. Байесовская сеть  $B = \langle G, \theta \rangle$  задается ориентированным ациклическим графом  $G$  и таблицей условных вероятностей  $\theta$ . Вершинам байесовской сети сопоставляются случайные переменные (для задачи кредитного скоринга – атрибуты), а ориентированным дугам – отношения условной независимости между этими переменными. При обучении байесовской сети сначала определяется структура графа  $G$ , а затем оцениваются параметры  $\theta$ .

Особенность применения байесовских сетей для решения задачи кредитного скоринга заключается в использовании понятия марковского покрытия (Markov Blanket) и методов Монте Карло по схеме марковской цепи (Markov Chain Monte Carlo) для построения байесовской сети на основе выбора из совокупности байесовских сетей, каждая из которых порождается марковской цепью с простым распределением вероятностей [3, с.49]. Результаты, полученные на основе байесовских сетей и деревьев решений легко интерпретировать в отличие от нейронных сетей. Фрагмент байесовской сети для решения задачи кредитного скоринга представлен на рис. 1.

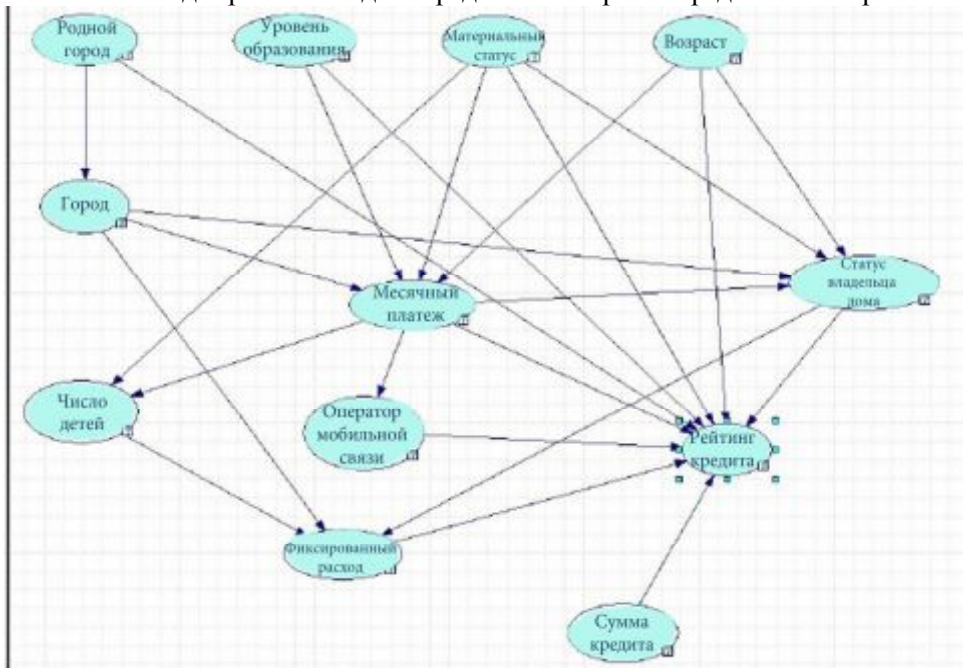


Рис.1. Фрагмент байесовской сети для решения задачи кредитного скоринга, источник [4]

2) Метод опорных векторов (Support Vector Machine (SVM)) и гибридная модель на основе ансамбля SVM в кредитном скоринге.

В работе [5] особенность применения метода опорных векторов к решению задачи кредитного скоринга заключается в построении гибридной двухуровневой модели, объединяющей идеи кластеризации и классификации. На первом уровне построения гибридной модели используется метод нечеткой кластеризации C-средних (fuzzy C-Means) с последующей нормализацией данных. На втором уровне строится совокупность (ансамбль) из 10 SVM. Принятие решений совокупностью алгоритмов осуществляется не по мажоритарному правилу, а с учетом весов, приписываемых каждому SVM.

3) Генетические алгоритмы и генетическое программирование в кредитном скоринге.

Генетические алгоритмы и генетическое программирование основаны на построении популяций особей и применении к ним операторов кроссинговера (скрещивания) и мутации с целью получения оптимального решения-особи.

Особенность применения генетического программирования для решения задачи кредитного скоринга в работе [6] заключается в разработке двухэтапного алгоритма: на первом этапе на основании генетического программирования выводятся закономерности вида «если...,то...», на втором этапе на основе редуцированных данных снова используется генетическое программирование для построения классифицирующей функции. На первом этапе каждая особь популяции реализована в виде дерева, внутренним вершинам которого сопоставлены только условные операторы (IF...THEN), операторы сравнения ( $\geq$ ,  $=$ ,  $\leq$ ) и булевы операторы (AND, OR, NOT). Именно эта особенность построения деревьев на первом этапе алгоритма позволила строить легкие для интерпретации решающие правила.

4) Деревья решений и леса решений в кредитном скоринге. Построение единичных деревьев решений традиционно основано на применении алгоритмов C4.5 и CART. Для увеличения обобщающей способности алгоритмов классификации на основе деревьев решений используют различные их композиции. Наиболее известными композициями деревьев решений являются леса, композиции на основе бустинга и беггинга.

В работе [7] исследуются особенности применения бустинга на деревьях решений для решения задачи кредитного скоринга. Согласно алгоритму бустинга на начальном этапе всем аппликантам приписывается один и тот же вес. После построения классификатора-дерева решений вес каждого аппликанта изменяется в соответствии с ответом, полученным классификатором. Второй классификатор строится по обучающей выборке с измененными весами аппликантов (объектов). Такая процедура повторяется многократно. Окончательная классификация аппликанта определяется значением взвешенной суммы отдельных классификаций по всем построенным деревьям решений. Существуют различные подходы к определению весов аппликантов, наиболее популярным среди них является алгоритм бустинга AdaBoost.

«Ядро» конкретной банковской скоринговой системы является коммерческой тайной. Перечислим некоторые наиболее популярные банковские скоринговые системы: система кредитного скоринга физических лиц SAS Credit Scoring Solution компании SAS, использующая комплексное аналитическое решение на основе алгоритмов Data Mining и специализированных алгоритмов для решения задачи кредитного скоринга; программный комплекс CreditAnalyst компании «Бизнес Нейро-Системы» на основе нейронных сетей и алгоритмов Data Mining; система Scorto Solutions компании Scorto; система «Финансы и аналитика» компании Франклин&Грант; система кредитного скоринга Credit4Cast компании Forecsys на основе алгоритмов модифицированной логистической регрессии, решающего списка логических правил, взвешенного голосования логических правил и т.д. [8]; система кредитного скоринга «KXEN:Scoring» компании «Ксема», математическое ядро которой реализует идеи минимизации структурного риска; скоринговое решение, реализованное в платформе Deductor компании BaseGroup Labs с использованием деревьев решений, нейронных сетей, самоорганизующихся карт, логистической регрессии и т.д.

Разработчики банковских скоринговых систем должны учитывать следующие особенности использования обучающих алгоритмов: нейронные сети и практически все статистические алгоритмы обучения не эффективны для использования в качестве «ядра» систем кредитного скоринга, в случае обучающих выборок (кредитных историй) небольшого объема; нейронные сети не эффективны в случае наличия в обучающей выборке нерелевантных признаков; для обучающих выборок небольшого объема рекомендуется использование логических алгоритмов обучения и деревьев решений; деревья решений и байесовские сети позволяют строить наглядные и легко интерпретируемые решающие правила.

В области создания и внедрения банковских скоринговых систем приоритетным для дальнейших исследований остается направление по разработке методов обнаружения мошенничества с кредитными картами.

#### Литература

1. Румянцев А. Современное состояние скоринга в банках Украины // Банкирь. – 2007. – С.58-59.
2. Брітченко І.Г. Особливості використання скорингових систем у банківському кредитуванні фізичних осіб // Фінанси, облік і аудит [Електронний ресурс] : зб. наук. праць / М-во освіти і науки, молоді та спорту України, ДВНЗ "Київ. нац. екон. ун-т ім. В. Гетьмана" ; відп. ред. А. М. Мороз. – К.: КНЕУ, 2011. – Вип. 17. – С. 31–37.

3. Baesens B., Egmont-Petersen M., Castelo R., Vanthienen J. Learning Bayesian network classifiers for credit scoring using Markov Chain Monte Carlo search // Proceedings of International Congress on Pattern Recognition. – 2002. – P.49-52.
4. Bicer I., Sevis D., Bilgic T. Bayesian credit Scoring Model with Integration of Expert Knowledge and Customer Data // International Conference 24<sup>th</sup> in EURO Conference “Continuous Optimization and Information Technologies in the Financial Sector” (MEC EurOPT 2010). – 2010. – P.324-329.
5. Ghodselahi A.A Hybrid Support Vector Machine Ensemble Model for Credit Scoring // International Journal of Computer Applications. – 2011. – Vol.17, No.5. – P.1-5.
6. Huang J., Tzeng G., Ong C. Two-stage genetic programming (2SGP) for the credit scoring model // Applied Mathematics and Computation. – 2006. – Vol. 174. – P.1039-1053.
7. Bastos J.A. Credit Scoring with Boosted Decision Trees // City: Munich Personal RePEc Archive, P. 262-273.
8. Кочедыков Д. А., Ивахненко А. А., Воронцов К. В. Система кредитного скоринга на основе логических алгоритмов классификации // Математические методы распознавания образов-12. — М.: МАКС Пресс, 2005. — С. 349–353.

338.4

*Солдатов М.А., к.ф.-м.н., доцент,  
Горбунов О.В., магистрант,  
ТНУ имени В.И. Вернадского*

### **АНАЛИЗ ВАЖНОСТИ ИЗУЧЕНИЯ РИСКОВ ЭТАПА КОДИРОВАНИЯ ПРОЦЕССА РАЗРАБОТКИ ПРОГРАММНЫХ ПРОДУКТОВ**

В современном мире наблюдается тенденция увеличения роли информационных технологий в жизни общества. Однако информационные технологии являются сравнительно весьма молодой и динамично развивающейся сферой жизни общества, поэтому объективно существует недостаток научных исследований в данной области. Как правило, информационные технологии на современном этапе развития науки рассматриваются с точки зрения следующих аспектов:

**Физико-математического.** Информационные процессы рассматриваются в своей глубинной сути. Изучаются физические процессы, происходящие при работе компьютерных устройств, а также математические алгоритмы, которые определяют направленность упомянутых физических процессов.

**1. Программистского.** Данный аспект получил мало развития как предмет научного исследования. Работы в данной области имеют в основном сугубо практический характер и не систематизированы в научном формате.

**Управленческого.** Процесс разработки программных продуктов рассматривается с аспекта менеджмента, как науки.

Необходимость научных знаний для управления разработкой программного обеспечения стала проявляться, когда проекты стали настолько крупными, что управлять без определенного системного подхода уже стало нельзя. Таким образом, наука менеджмента разработки программного обеспечения имеет отношения по большей части только к крупным проектам. Действительно, на практике большинство методов организации разработки для небольших проектов являются избыточными и вредными. Это связано с тем, что современный менеджмент практически не уделяет внимания вопросам организации самого процесса кодирования. А так как для небольших проектов этот процесс имеет значительно большую долю от всего процесса разработки программного обеспечения, чем в больших проектах, выясняется, что современный менеджмент эффективней не применять в формате устоявшихся книжных практик, чем применять.

Публикации на тему анализ рисков разработки программных продуктов в основном принадлежат иностранным исследователям: Don Gotterbarn (2005) “Responsible Risk Analysis for Software Development: Creating the Software Development Impact Statement”[1], Ropponen J., N. Lyytinen and N. Kalle (2000) “Components of Software Development Risk: How to Address them? A Project Manager Survey”[2], Highsmith, J. (1997) “Agile Project Management”[3], Watson R. T., Turban, E., Mclean E., and Wetherbe J. (2009) “Information Technology for Management, Improving Quality and Productivity”[4], Schmidt R. Lyytinen K., Keil M., And Cule P. (2001) ”Identifying Software Project Risks: An International Delphi Study”[5].

Рассмотренные научные работы игнорируют или уделяют недостаточное внимание проблеме анализа рисков, которые возникают в процессе этапа кодирования разработки программных продуктов. Однако данные риски существуют и занимают важное место в списке рисков проектов