

Елена Г. Бутурлакина

ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТИ ВОЗНИКНОВЕНИЯ СИТУАЦИИ НЕДОСТАТОЧНОГО ФИНАНСИРОВАНИЯ ПРИ РЕАЛИЗАЦИИ ИНВЕСТИЦИОННЫХ ПРОЕКТОВ В РЕСПУБЛИКЕ КАЗАХСТАН

В статье рассмотрено прогнозирование возникновения проблемы недостаточного финансирования при реализации инвестиционных проектов. Сравнены результаты применения таких методов анализа данных как нейронные сети, карты Кохонена, деревья решений и логистическая регрессия. Полученные модели позволяют верно определить 70–80% исходов. На их основе возможно создание эффективного инструмента инвестиционного анализа, дополняющего стандартные методологии.

Ключевые слова: инвестиционный проект; прогнозирование; интеллектуальный анализ данных; недостаточное финансирование.

Табл. 1. Рис. 1. Лит. 16.

Олена Г. Бутурлакiна

ВИЗНАЧЕННЯ ЙМОВІРНОСТІ ВИНИКНЕННЯ СИТУАЦІЇ НЕДОСТАТНЬОГО ФІНАНСУВАННЯ ПРИ РЕАЛІЗАЦІЇ ІНВЕСТИЦІЙНИХ ПРОЕКТІВ У РЕСПУБЛІЦІ КАЗАХСТАН

У статті розглянуто прогнозування виникнення проблем з недостатнім фінансуванням при реалізації інвестиційних проектів. Порівняно результати застосування таких методів аналізу даних як нейронні мережі, карти Кохонена, дерева рішень та логістична регресія. Отримані моделі дозволяють вірно визначити 70–80% вихідних ситуацій. На основі цього можливе створення ефективного інструменту інвестиційного аналізу, що доповнював би стандартні методології.

Ключові слова: інвестиційний проект; прогнозування; інтелектуальний аналіз даних; недостатнє фінансування.

Yelena G. Buturlakina¹

FORECASTING THE LACK OF FUNDING PROBLEM FOR INVESTMENT PROJECTS IN KAZAKHSTAN

This paper explores the prediction of the possible lack of funding problem arising during the implementation phase of investment projects. The numerical results were obtained and compared with the help of artificial neural networks, Kohonen maps, decision trees, and logistic regression. The resulting models allow identify correctly about 70–80% of the outcomes. Based on the results obtained an effective tool can be developed for investment analysis in addition to existing techniques.

Keywords: investment project; forecasting; data mining; lack of funding.

Постановка проблеми. Оценка инвестиционных проектов — первая и важнейшая составляющая любого предпринимательского начинания. Никто не в состоянии с полной уверенностью ответить, что случится завтра, через неделю, через год, а при принятии инвестиционных решений приходится рассматривать значительно более длительные временные периоды. Каждому инвестору приходится рисковать более или менее значительными финансовыми средствами, и желание минимизировать этот риск легко объяснимо.

Анализ последних исследований и публикаций. Оценка инвестиционных проектов включает следующие задачи:

¹ Sericbayev East Kazakhstan State Technical University, Ust-Kamenogorsk, Kazakhstan.

- оцeнка фiнансової реалiзуемостi проєкта;
- оцeнка вигоди;
- виявлення граничних умов ефективної реалiзацiї проєкта;
- оцeнка ризика i стiйливостi проєкта [2].

Для рeшення цих задач розробленi стандартнi методологiї i рекомендацiї [3]. Однак, сучаснi вичислительнi засоби дозволяють застосовувати методи математичного моделювання для бiльш конкретних задач. В частинi, математичнi моделi широко використовуються для визначення ефективностi наукових розробок, странового ризика, кредитного скоринга, прогнозування банкрутства компанiй, курсу акцiй i др. i демонструють хорошу ефективностi. Тому багато вiдomoвцiв намагаються застосовувати цi методи для рeшення частинних задач, наприклад, в [4] авторами поставлена задача визначення ймовiрностi отримання акцiй, що, на їхню думку, є важливим показником активностi компанiї i дозволяє судити про перспективи банкрутства.

В вiдomoвствi використанi методи аналізу даних, широко застосовуванi в рiзноманiтних економiчних розробках. Наприклад, в згаданому проєкті [4] автори використовували в вiдomoвствi логистичну регресiю, з допомогою цього ж інструменту в [15] визначається кредитний ризик фiрми на основі даних про її активностi. Автори [13] застосовують логистичну регресiю, щоб довести наявностi зв'язку між деловою репутацiєю фiрми i її готовностiю першими впровадити новi «чисті» технології.

Нейроннi мережi дозволяють широко варіювати їхню структуру i алгоритми навчання, тому в вiдomoвствi зазвичай використовуються декілька рiзних моделей. Автори [8] порiвнюють ефективностi нейронних мереж з рiзною кількiстю внутрiшніх шарiв у визначенні кредитного ризика. Самоорганiзовуючися карти Кохонена – одна з рiзновидностей нейронних мереж – використанi в [9; 11] для прогнозування цiн на акцiї i формування портфеля з мінімальним ризиком, автори [14] застосовують їх для оцiнки якостi колективних інвестиційних рiшень. Вiдomoвцi [5] використовують класифікацію на основі мереж Кохонена для прогнозування поведінки фiрм за показниками їхньої фiнансової активностi. В [6] на основі моделі, побудованої з використанням самоорганiзовуючихся карт, доводиться гiпотеза про інвестиційну поведінку фiрм рiзної величини.

Дерева рiшень отримали широке поширення у аналізі інвестиційних проєктiв. В [12] з допомогою дерева рiшень визначається ефективностi проєктiв у сфері інформаційних технологiй, в [10; 16] аналізуються енергетичнi проєкти, проєкти вiтрових установок зокрема. В [7] на основі нечітких дерев рiшень авторами розроблено програмний продукт для визначення кредитного ризика банківських позичальникiв.

Цель вiдomoвства. В своєму вiдomoвстві ми розглядаємо ефективностi рiзних аналітичних методiв для визначення найбiльш ефективного інструменту рeшення поставленої задачі.

Основнi результати вiдomoвства. В вiдomoвстві була використана база даних інвестиційних проєктiв, реалiзованих у рiвних Республiки

Казахстан и финансируемых из государственного бюджета. База данных доступна в системе мониторинга по адресу <http://ip.economy.kz>. Она содержит сведения о названии проектов, сроках их реализации, ответственных органах, стоимости проектов, регионе и отрасли реализации, а также о проблемах, возникших при реализации. Всего имеется информация о 157 инвестиционных проектах. При их реализации отмечены следующие проблемы:

- недостаток финансирования;
- организационные;
- неразвитость инфраструктуры;
- выделение/выкуп земли;
- нехватка ресурсов;
- отсутствие рынка сбыта;
- другие.

Проблем может быть несколько (максимум 3). Имеется описание проблем и предполагаемые меры по их решению.

Задача исследования состоит в том, чтобы создать инструмент инвестиционного проектирования, позволяющий спрогнозировать вероятность появления проблем на основе внешних данных об инвестиционном проекте. Поскольку в базе данных доступна информация только о проектах, реализуемых в Республике Казахстан, то применять полученный финансовый инструмент возможно только для проектов, реализация которых планируется в регионах Республики Казахстан. Применяемая методика является универсальной и, при наличии соответствующих данных о проектах в регионах других стран, может быть адаптирована для любой страны.

В основе исследования лежит гипотеза о существовании взаимосвязи между такими показателями проекта, как величина финансирования, регион и отрасль реализации и возникающими при реализации проекта проблемами. Если такая взаимосвязь будет установлена, то полученная математическая модель зависимости будет полезна, во-первых, потенциальным инвесторам, так как, зная о вероятности возникновения проблемы, они смогут принять меры по ее избеганию уже на этапе планирования. Во-вторых, выявленные закономерности необходимо отметить органам высшего руководства республики, чтобы попытаться выяснить и устранить их причины.

В исследовании использовались следующие методики анализа, реализованные средствами платформы "Deductor Academic 5.2": логистическая регрессия, самоорганизующиеся карты Кохонена, деревья решений, нейронные сети.

Как правило, в подобных задачах классификации используются обширные наборы обучающих и тестовых данных. Поскольку количество доступных для анализа проектов не очень велико, большая часть данных, а именно 141 проект, использована для обучения (90%). Даже незначительный прирост обучающего набора в 10% позволил значительно улучшить результат для некоторых методов. Проекты в обучающий набор выбираются по порядку, начиная с начала таблицы. Инструментарий позволяет отбирать данные в случайном порядке, но необходимо, чтобы обучающий набор содержал все возможные варианты значений категориальных переменных, а при случайном отборе этого нельзя гарантировать. К тому же, для сравнения эффективности разных

методов обучающий и тестовый наборы должны быть идентичны. Для тестирования моделей используется оставшиеся 16 (10%) проектов.

Большая часть проектов предполагают реализацию проектирования, строительства, модернизации, реконструкции различных объектов: школ, водо- и газопроводов, производственных предприятий. Объекты расположены в регионах Республики Казахстан, суммы затрат на реализацию проектов составляют от 0,03 млрд тенге («Внедрение капельного орошения в полеводстве и строительство картофелехранилища») до 925 млрд тенге («Строительство интегрированного нефтехимического комплекса»).

Для анализа выбрана наиболее часто встречающаяся проблема — «недостаток финансирования». Наблюдения, соответствующие проектам, столкнувшимся с ней, помечены как «истинные», не имеющие данной проблемы — как «ложные». В качестве входных параметров для всех методов используются «сумма проекта» (непрерывный тип), «регион» (категориальный тип) и «отрасль» (категориальный тип).

Проблема «недостаток финансирования» встречается в 45% проектов: в 57 проектах она указана в качестве единственной проблемы, еще в 14 проектах фигурирует наряду с другими проблемами. Результаты приведены в табл. 1.

Таблица 1. Результаты классификации на обучающем и тестовом наборах, авторская разработка

Метод	Результаты классификации							
	Обучающий набор (141 проект)				Тестовый набор (16 проектов)			
	ТТ	FF	TF	FT	ТТ	FF	TF	FT
Логистическая регрессия	29	61	34	17	2	9	5	0
Каргы Кохонена	44	71	19	7	5	8	2	1
Нейросеть	38	70	25	8	4	8	3	1
Дерево решений	16	76	46	2	3	9	4	0

ТТ — верно классифицированные истинные значения; FF — верно классифицированные ложные значения; TF — неверно классифицированные истинные значения (ошибки 1-го рода); FT — неверно классифицированные ложные значения (ошибки 2-го рода).

На рис. 1 отмечено качество классификации тестовых наборов каждым из методов.

Видно, что все методы дают хороший результат классификации — от 70 до 80% случаев классифицированы верно.

Однако, кроме доли верно классифицированных элементов, необходимо рассмотреть избирательность и чувствительность методов. Все методы стремятся классифицировать элементы как ложные, т.е. число неверно классифицированных ложных значений значительно меньше числа неверно классифицированных истинных значений. Однако в случае прогнозирования риска такой подход опасен, поскольку гораздо важнее распознать вероятность возникновения рисков ситуации и принять меры по ее недопущению, чем справляться с последствиями возникшего риска. Поэтому важным критерием классификации в данном случае является чувствительность.

Методы, при возможности, были настроены на максимальную чувствительность, например, снижен порог, при котором элемент должен классифицироваться как истинный.

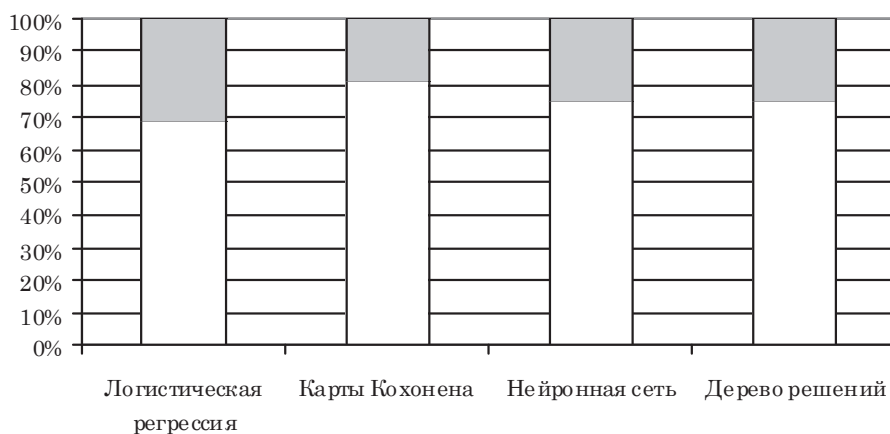


Рис. 1. Сравнение результатов классификации, %, авторская разработка

Из табл. 1 видно, что наилучшей чувствительностью, как на обучающем, так и на тестовом наборах, обладает метод «карты Кохонена».

Применение полученных результатов. Планируется использовать полученные результаты в разрабатываемой системе поддержки принятия инвестиционных решений. При этом наиболее удобным инструментом, несмотря на худший результат классификации, является метод логистической регрессии, поскольку он позволяет получить в качестве прогноза вероятность наступления ситуации. Зная эту вероятность, инвестор может самостоятельно сделать вывод и принять решение. Например, при величине вероятности в 45% ситуация будет оценена классификатором как безрисковая, однако такая вероятность все еще очень высока и требует повышенного внимания. Кроме того, логистическая регрессия – единственный метод, который дает возможность быстро перенести функции классификатора в любую среду разработки. Для этого достаточно полученные коэффициенты регрессии, доступные в соответствующей вкладке визуализации, включить в соответствующую функцию. В то же время академическая версия программы не позволяет экспортировать обученные классификаторы других методов. Их применение возможно только внутри платформы с помощью аналитического механизма «что-если».

Выводы. В результате проведенного исследования показана высокая эффективность применения методов "data mining" для прогнозирования риска возникновения ситуации недостаточного финансирования при реализации инвестиционных проектов в Казахстане. Можно утверждать, что существуют факторы взаимосвязи между величиной суммы проектных средств, регионом и отраслью реализации проекта, влияющие на возникновение данной ситуации. Несмотря на небольшой объем обучающей выборки получены стабильные результаты для всех методов, которые могут применяться в качестве дополнительного инструмента при проведении инвестиционного анализа. Спрогнозированная высокая величина риска позволит привлечь внимание инвестора и заставит уделить больше внимания проектированию финансовых

потоків проекту. К сожалению, над базой данных, использованной в исследовании, не ведется планомерной работы. Данные не актуализируются, в [1] отмечены факты несоответствия информации, представленной в базе, реальному состоянию дел. Однако, необходимость развития подобных систем мониторинга в Казахстане и других странах не вызывает сомнений, поскольку, помимо своей непосредственной функции контроля, они предоставляют важнейшую информацию для многих исследователей и инвесторов.

1. *Бурбаева А.* Инвестиционная показуха: от количества к... проблемам // Общественная позиция: республиканский ежедневник. — 21.02.2013.
2. *Виленский П.Л., Лившиц В.Н., Смоляк С.А.* Оценка эффективности инвестиционных проектов. Теория и практика: Учеб. пособие. — 2-е изд. — М.: Дело, 2002. — 888 с.
3. Методические рекомендации по оценке эффективности инвестиционных проектов / В.В. Коссов, В.Н. Лившиц, А.Г. Шахназаров и др. — Третья ред., исправ. и доп. — М.: Экономика, 2008. — 232 с.
4. *Andriospoulos, D.G., Pasiouras, F., Zopounidis, C.* (2012). An application of multicriteria decision aid models in the prediction of open market share repurchases. *Omega*, 40: 882–890.
5. *Chen, J., Huang, C., Chen, A.* (2010). Applying Self-Organizing Mapping Neural Network for Discovery Market Behavior of Equity Fund. *WSEAS transactions on information science and applications*, 7(2): 230–240.
6. *Fioretti, G.* (2004). The investment acceleration principle revisited by means of a neural network. *Neural Computing & Applications*, 13(1): 16–23.
7. *Gonzalez-Carrasco, I., Colomo-Palacios, R., Lopez-Cuadrado, J.L., Garci'a-Crespo A., Ruiz-Mezcua, B.* (2012). PB-ADVISOR: A private banking multi-investment portfolio advisor. *Inf Sci*, 206: 63–82.
8. *Khashman, A.* (2010). Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. *Expert Syst Appl*, 37(9): 6233–6239.
9. *Lukyanitsa, A.A.* (2009). Risk minimization with self-organizing maps for mutual fund investment. *Evolutionary Computation*, 2009. CEC '09. IEEE Congress.
10. *Munoz, J.I., Contreras, J., Caamano, J., Correia, P.F.* (2011). A decision-making tool for project investments based on real options: the case of wind power generation. *Annals of Operations Research*, 186(1): 465–490.
11. *Nakaoka, I.* (2005). A portfolio selection by SOM and an asset allocation of risk/nonrisk assets by fuzzy reasoning using the selected brands. *Systems, Man and Cybernetics*, 2005 IEEE International Conference.
12. *Pendharkar, P.C.* (2010). Valuing interdependent multi-stage IT investments: A real options approach. *Eur J Oper Res*, 201(3): 847–859.
13. *Petkova, A.P., Wadhwa, A., Yao, X., Jain, S.* (2011). Being first or making a difference? A study of US venture capital investments into the "clean" energy sector. *Frontiers of Entrepreneurship Research*, 31(2): 49–61.
14. *Pham, H.V.* (2011). Human Reasoning Awareness Quantified by Self-Organizing Map Using Collaborative Decision Making for Multiple Investment Models. *Complex, Intelligent and Software Intensive Systems (CISIS)*, 2011 International Conference.
15. *Psillaki, M., Tsolas, I.E., Margaritis, D.* (2010). Evaluation of credit risk based on firm performance. *Eur J Oper Res*, 201(3): 873–881.
16. *Tan, B., Anderson, E.G., Dyer, J.S., Parker, G.G.* (2010). Evaluating system dynamics models of risky projects using decision trees: alternative energy projects as an illustrative example. *System Dynamics Review*, 26(1): 1–17.

Стаття надійшла до редакції 30.07.2014.