

УДК 004.8:65.012.23

**В.І. Дубровін**, канд. техн. наук, проф.,  
**В.М. Льовкін**, інженер,  
Запорізь. нац. техн. ун-т

## МЕТОД ПРОГНОЗУВАННЯ РИЗИКУ НЕУСПІШНОСТІ ПРОЕКТІВ НА СТАДІЇ ПЕРЕДПРОЕКТНОГО ПЛАНУВАННЯ

*В.І. Дубровін, В.М. Льовкін. Метод прогнозування ризику неспішності проектів на стадії передпроектного планування.* Розглянуто підходи до розв'язання проблеми прогнозування ризиків проекту на стадії передпроектного планування на основі ансамблів нейронних мереж. Запропоновано метод прогнозування ризику неспішності проектів на стадії передпроектного планування на основі ансамблів нейронних мереж з кластеризацією та досліджено ефективність його застосування порівняно з існуючими.

*В.И. Дубровин, В.Н. Левкин. Метод прогнозирования риска неспеха проектов на стадии предпроектного планирования.* Рассмотрены подходы к решению проблемы прогнозирования рисков проекта на стадии предпроектного планирования на основе ансамблей нейронных сетей. Предложен метод прогнозирования рисков неспеха проектов на стадии предпроектного планирования на основе ансамблей нейронных сетей с кластеризацией и исследована эффективность его применения относительно существующих методов.

*V.I. Dubrovin, V.M. Liovkin. Method of project failure risk prognostication in the preproject planning process.* Project risk prognostication approaches of the preproject planning stage, based on neural networks ensembles are considered. The method of project failure risk prognostication in the pre-project planning process, based on neural network ensemble concepts with clustering, is proposed. The efficiency of its application is examined in comparison with the existing methods.

Проект — комплекс зусиль, які застосовуються з метою отримання конкретних унікальних результатів у рамках відведеного часу в межах затвердженого бюджету, що виділяється на сплату ресурсів, які використовуються або споживаються в ході проекту [1]. Перевищення запланованих значень витрат проекту або тривалості його виконання відбувається як результат впливу багатьох чинників, але всі вони пов'язані з різними формами ризиків. Аналіз цих ризиків — необхідний крок для покращення системи оцінювання й може бути використаний для діагностування слабких місць та виділення ділянок, на яких можна покращити точність оцінок проекту, або відповідного реагування на ризики за допомогою наявних фінансових ресурсів.

Управління проектами — методологія організації, планування, керівництва, координації трудових, фінансових та матеріально-технічних ресурсів протягом проектного циклу, направлена на ефективне досягнення його цілей шляхом застосування сучасних методів, технік та технологій управління для досягнення визначених у проекті результатів за складом та обсягом робіт, вартістю, часом, якістю та задоволенням вимог учасників проекту [2].

Практикуючі спеціалісти вважають, що зусилля на етапах раннього планування життєвого циклу проекту мають більший вплив на успіх проекту ніж зусилля, прийняті на пізніших етапах процесу [3]. Це видно на представленому співвідношенні кривих впливу та витрат відносно життєвого циклу проекту (рис. 1). Крива впливу відображає здатність компанії впливати на результат проекту протягом різних його стадій. Діаграма відбиває те, що на результат проекту значно легше впливати протягом стадії планування, коли витрати відносно мінімальні, ніж протягом виконання проекту або експлуатації продукту в умовах більш значних витрат. Окрім того й імовірність події, пов'язаної з ризиком, найбільша на стадії планування перед початком виконання проекту.

Агрегований показник завершеності визначення обсягу й складу робіт із проекту (Project

Definition Rating Index, PDRI) дозволяє обчислити завершеність визначення масштабу проекту перед його виконанням, на стадії передпроектного планування [4]. Для прогнозування відхилень витрат та тривалості проектів застосовувались нейронні мережі та лінійна регресійна модель [5]. Досліджувалось застосування нейронних мереж (були розглянуті тільки перцептрони) та їх ансамблів для моделювання успішності проектів [6].

Інвестора цікавить не тільки бюджет проекту, але й ризик того, що проект виявиться неуспішним і зокрема те, на скільки відхилиться його витрати та тривалість реалізації. Адже окрім того, що такі відхилення найчастіше призводять до підвищених фінансових вимог до самого інвестора, вони ще й можуть призвести до санкцій компаній-партнерів. Якщо інвестор перед початком реалізації отримає прогнозоване значення ризику неуспішності проекту, то він зможе або змінити своє рішення на користь іншого проекту (якщо така можливість є і за сукупністю критеріїв інший проект має більш привабливий вигляд) або, спираючись на отриману оцінку, ініціювати перегляд проекту. Таким чином, прогнозоване значення ризику неуспішності проекту може стати ключовим чинником вибору проекту серед альтернативних або його доопрацювання для отримання більш надійного плану, а це в свою чергу приведе до більш ефективного використання фінансових ресурсів, чого безпосередньо й прагне інвестор. З цього й випливає актуальність завдання прогнозування ризику неуспішності проекту.

Виходячи з наведених аргументів, на етапі передпроектного планування необхідно надати стратегічну інформацію, за допомогою якої інвестор (або власник) міг би дослідити ризик та прийняти рішення щодо направлення ресурсів, максимізуючи ймовірність досягнення успіху проекту. Таким чином, у процесі дослідження необхідно дослідити ефективність застосування різних методів (у тому числі нейронних мереж та їх ансамблів) для прогнозування значення ризику неуспішності проекту, виходячи з його стану під час передпроектного планування.

Початкові значення ваг у нейронній мережі зазвичай встановлюються випадковим чином. Це призводить до того, що результати прогнозування мереж, які навчалися на одній і тій самій вибірці

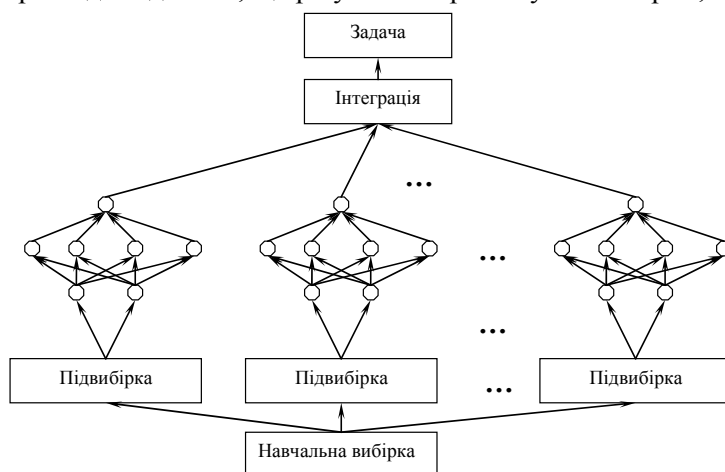


Рис. 2. Ансамбль нейронних мереж на основі методу bootstrap aggregation

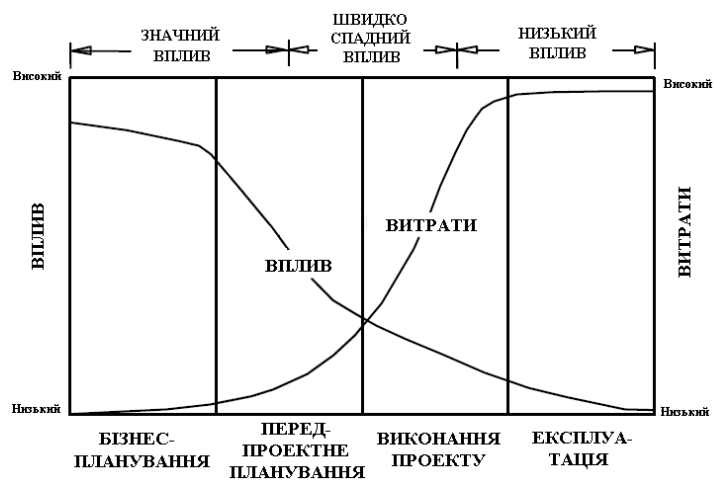


Рис. 1. Криві впливу та витрат відносно життєвого циклу проекту

даних, можуть різнитися. З метою отримання більш точних і стабільних результатів були запропоновані ансамблі нейронних мереж, які складаються з декількох мереж, а результат прогнозування ансамблю формується з результатів прогнозування окремих мереж шляхом усереднення або іншим чином.

Серед методів, які використовуються для навчання нейронних мереж у ансамблі, можна виділити методи беггінг (bagging) та посилення (boosting).

У методі беггінг (bagging скорочено від "bootstrap aggregation") відбувається зважене голосування

базових мереж, навчання яких проводиться на різних підвбірках даних [7]. Кожна така підвбірка формується з загальної вибірки даних для навчання випадковим чином, зокрема пропонується обирати розмір підвбірки, що приблизно дорівнює 60 % від загальної вибірки (рис. 2).

Метод посилення (boosting) відрізняється від попереднього методу тим, що нейронні мережі навчаються послідовно. Кожній парі даних встановлюється вага, яка залежить від того, наскільки точно класифікували даний випадок попередні мережі. Таким чином, більш складні для класифікації завдання отримують більшу вагу. Це приводить до того, що такі завдання будуть обрані для навчання наступної мережі. Схема даного методу представлена (рис. 3).

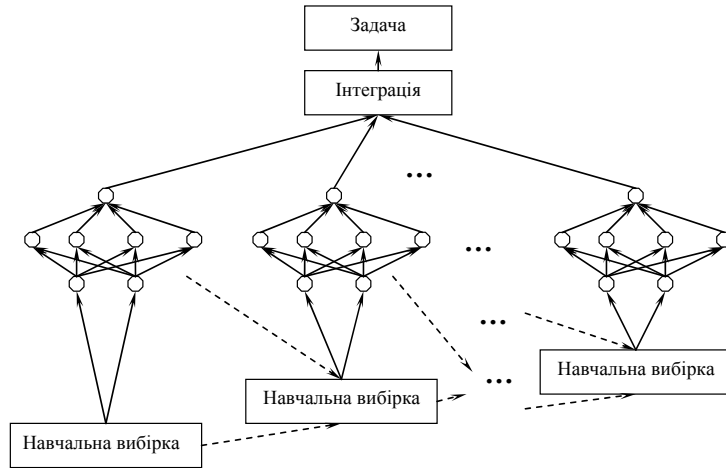


Рис. 3. Ансамбль нейронних мереж на основі методу boosting

Метод адаптивного посилення (AdaBoost, скорочено від “Adaptive Boosting”) може бути представлений таким чином [8].

Вибірка даних представлена у вигляді:  $(x_1, y_1), \dots, (x_m, y_m)$ , де  $x_i \in X, y_i \in Y$ , а кожній парі даних відповідає вага, представлена елементом матриці  $\mathbf{D}_t$ . Елементи матриці пропонується ініціювати значеннями  $D_1(i) = 1/m, i = 1, \dots, m$ .

Для кожного  $t = 1, \dots, T$ , де  $T$  — кількість ітерацій (класифікаторів), виконати:

- провести навчання нейронної мережі  $t$  за допомогою навчальної вибірки  $\mathbf{D}_t$ ;
- класифікувати за допомогою мережі  $t$  набір даних  $D_t$ , отримавши похибку  $h_t: X \rightarrow Y$ ;
- обчислити похибку  $h_t: e_t = \sum_{i: h_t(x_i) \neq y_i} D_t(i)$  (якщо  $e_t > 0,5$ , встановити  $T = t - 1$  і завершити

навчання);

- встановити  $\beta_t = e_t / (1 - e_t)$ ;

- оновити ваги:  $D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} \beta_t, & \text{якщо } h_t(x_i) = y_i; \\ 1, & \text{інакше,} \end{cases}$

де  $Z_t$  — нормалізуюча константа.

Визначити остаточне значення  $h_{fin}(x) = \arg \max_{y \in Y} \sum_{t: h_t(x)=y} \log \frac{1}{\beta_t}$ .

Зазвичай на практиці метод AdaBoost дає кращі результати порівняно з методом bagging. Однак, даний метод має і недоліки, серед яких:

- схильність до перенавчання при наявності значного шуму в даних;
- вимагає достатньо великих навчальних вибірок (метод bagging дозволяє будувати ансамбль схожої якості за меншою кількістю даних);
- стратегія послідовного додавання мереж часом призводить до неоптимального набору базових алгоритмів.

На основі проаналізованих методів навчання ансамблів нейронних мереж пропонується метод, який, базуючись на методі AdaBoost для навчання ансамблів нейронних мереж, дозволяє досягти спеціалізації базових нейронних мереж за рахунок розподілення вхідних даних на кластери відповідно до значення PDRІ та запланованого обсягу витрат (тривалості) проекту. Окрім того метод дозволяє вирішити проблему, виявлену під час дослідження застосування нейронних мереж, коли деякі архітектури виявлялись більш ефективними з одними даними й

поступались іншим архітектурам на іншій навчальній вибірці. Приклад ансамблю нейронних мереж, який пропонується побудувати в даному методі, зображений на рис. 4.

Метод прогнозування ризику неуспішності проектів на стадії передпроектного планування на основі ансамблів нейронних мереж з кластеризацією складається з таких процедур:

— сформувати навчальну вибірку у вигляді  $(\{x_i^1, x_i^2\}, y_i)$ , де  $x_i^1$  — значення показника PDRI  $i$ -го проекту,  $x_i^2$  — запланований обсяг витрат (тривалості)  $i$ -го проекту,  $y_i$  — клас, до якого належить  $i$ -й проект (0 або 1);

— провести кластеризацію даних на основі карт самоорганізації Кохонена, відповідно до двох параметрів: показника PDRI та запланованого обсягу витрат (тривалості), сформувавши  $K$  кластерів;

— сформувати ансамбль нейронних мереж, де для кожного з  $K$  кластерів передбачити окремий набір нейронних мереж, кожний з яких представлений  $M \times N$  нейронними мережами, де  $M$  — кількість різних архітектур, а  $N$  — кількість нейронних мереж кожної архітектури в даному кластері (нейронна мережа в такому випадку характеризується множиною  $(k, m, n)$ , де  $n$  — порядковий номер нейронної мережі архітектури  $m$  з кластера  $k$ );

— провести навчання  $n$  нейронних мереж з архітектурою  $m$  ( $m = 1, \dots, M$ ) у середині кожного  $k$ -го кластера ( $k = 1, \dots, K$ ) за допомогою методу AdaBoost на основі частини навчальної вибірки, дані якої належать до кластера  $k$ ;

— рішення про відношення проекту, що характеризується показниками  $\{x_i^1, x_i^2\}$  та належить до кластера  $k$ , до одного з двох класів (0 або 1) приймається нейронними мережами з підансамблю  $k$ , шляхом незалежного винесення рішення, а загальне рішення ансамблю приймається рівномірним голосуванням всіх  $M \times N$  нейронних мереж кластера  $k$ ;

— розрахувати загальний ризик неуспішності проекту шляхом усереднення результатів прогнозування нейронних мереж, що брали участь у голосуванні в пункті д):

$$R = \frac{\sum_{m=1}^M \sum_{n=1}^N r_{(k,m,n)}}{MN},$$

де  $r_{(k,m,n)}$  — прогнозоване  $n$  нейронною мережею архітектури  $m$  з підансамблю  $k$  значення ризику (визначається як прогнозований ступінь належності проекту до класу неуспішних) проекту, що за характеристиками було віднесено до кластера  $k$ .

Для розв'язання задачі класифікації для наявного набору даних було проведено дослідження показників  $K, M, N$ . Пропонується обирати число кластерів  $K$ , таким, що дорівнює трьом, використовувати  $M = 2$  (нейронні мережі двох архітектур: зворотного поширення з одним шаром та чотирма нейронами на ньому та двома шарами з 12 та чотирма нейронами на них), а  $N$  приймати таким, що дорівнює п'яти.

Розглянуті методи та моделі застосування для розв'язання проблеми класифікації проектів.

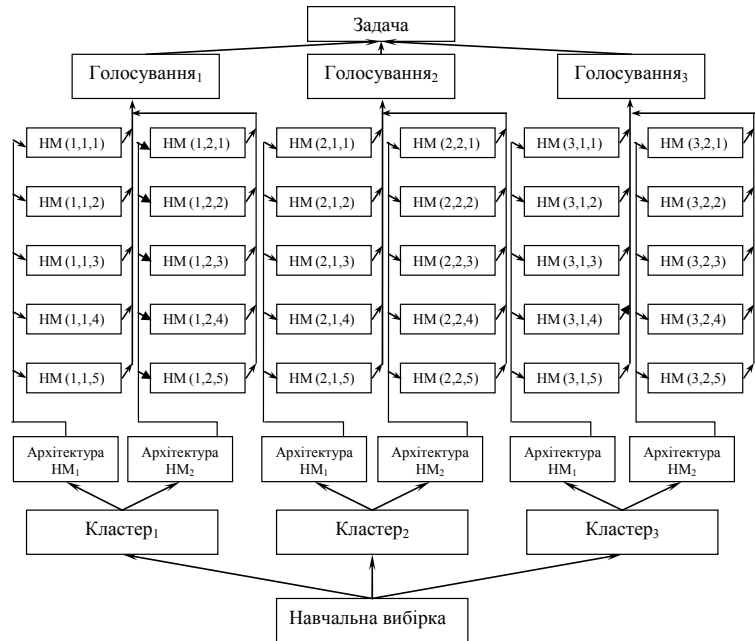


Рис. 4. Ансамбль нейронних мереж для методу AdaBoost

Для навчання кожний раз випадковим чином із загальної вибірки, що була сформована з 78 будівельних проектів, обирались 50, решта проектів формувала тестову вибірку [3].

Вважаємо успішним проект, для якого виконуються такі умови:

$$IB_B = \frac{\text{Фактичні витрати} - \text{Планові витрати}}{\text{Планові витрати}} \leq 0; \quad (1)$$

$$IB_T = \frac{\text{Фактична тривалість} - \text{Планова тривалість}}{\text{Планова тривалість}} \leq 0, \quad (2)$$

де  $IB_B$  — показник успішності проекту за бюджетом;

$IB_T$  — показник успішності проекту за часом.

Якщо хоча б одне з рівнянь (1), (2) не виконується, то проект ідентифікується як неуспішний. Тоді завдання прогнозування ризику неуспішності буде мати вигляд як класифікація проектів на два класи (успішні й неуспішні).

Прогнозування успішності проектів за бюджетом та програмою відбувалось окремо. Проаналізовано ефективність застосування для класифікації нейронних мереж, з яких були обрані ті, що показали найкращі результати класифікації: ансамблів нейронних мереж на основі методів boosting та AdaBoost, методу прогнозування ризику неуспішності проектів на стадії передпроектного планування на основі ансамблів нейронних мереж з кластеризацією. Результати наведені (див. таблицю).

*Результати класифікації нейронних мереж*

Метод (модель)	Базова нейронна мережа	Похибка класифікації на тестовій вибірці, %	Ранг
Ансамбль нейронних мереж boosting	Нейронна мережа зворотного поширення [12, 4]	10,71	6
	Нейронна мережа зворотного поширення [4]	10,71	6
Ансамбль нейронних мереж AdaBoost	Нейронна мережа зворотного поширення [12, 4]	10	5
	Нейронна мережа зворотного поширення [4]	9,29	4
Метод прогнозування ризику неуспішності проектів на стадії передпроектного планування	Нейронна мережа зворотного поширення [12, 4]	7,86	3
	Нейронна мережа зворотного поширення [4]	7,14	2
	—	6,43	1
Нейронна мережа зворотного поширення [12, 4]	—	11,43	9
Нейронна мережа зворотного поширення [4]	—	10,71	6

Для запропонованого методу досліджувались також його варіації на основі нейронної мережі зворотного поширення з одним прихованим шаром і чотирма нейронами на ньому та двома шарами й 12 та чотирма нейронами на них, а також безпосередньо сам метод. Усі три розглянуті варіанти продемонстрували кращі результати порівняно з іншими моделями на основі ансамблів нейронних мереж, а також звичайними нейронними мережами.

Найменший розмір похибки в 6,43 % дозволив отримати безпосередньо метод прогнозування ризику неуспішності проектів на стадії передпроектного планування на основі ансамблів нейронних мереж з кластеризацією.

Запропонований метод дозволив покращити результати прогнозування, що дає можливість більш ефективно приймати рішення щодо управління проектами та ризиками на стадії передпроектного планування.

Окрім того, варто відзначити, що метод дозволяє отримати не тільки клас проекту, але й безпосередньо загальний рівень ризику неуспішності проекту. На основі прогнозу власник (або інвестор) може прийняти рішення щодо вкладання фінансових ресурсів у один з альтернативних проектів, вибору одного з варіантів реалізації проекту або прийняти рішення щодо тих чи інших дій (на основі окремих елементів індексу PDRI), що дозволить зменшити ризик.

Таким чином, аналіз існуючих методів навчання ансамблів нейронних мереж дозволив виявити їх недоліки. У процесі дослідження було також проаналізовано ефективність їх застосування для проблеми прогнозування ризику неуспішності проектів на основі базових нейронних мереж, які продемонстрували найбільшу точність.

Запропоновано метод прогнозування ризику неуспішності проектів на стадії передпроектного планування, який дозволив підвищити точність результатів прогнозування та розрахувати загальний ризик неуспішності проекту, що сприяє більш ефективному управлінню ризиками на ранніх стадіях, перед початком реалізації проекту, дозволяючи зменшити фактичні витрати на його реалізацію.

### Література

1. Арчибальд, Р.Д. Управление высокотехнологичными программами и проектами / Д. Арчибальд; пер. с англ. Мамонтова Е.В.; под ред. Баженова А.Д., Арефьева А.О. — 3-е изд., перераб. и доп. — М.: Компания АйТи, 2004. — 472 с.
2. Мазур, И.И. Управление проектами: учебное пособие / И.И. Мазур, В.Д. Шапиро, Н.Г. Ольдерогге; под общ. ред. И.И. Мазура. — 5-е изд., перераб. — М.: Омега-Л, 2009. — 960 с.
3. Wang, Yu.-R. Applying The PDRI in Project Risk Management: Ph.D. Thesis / Yu.-R. Wang. — Austin: TX, 2002. — 268 p.
4. Cho, C.-S. Building Project Scope Definition Using Project Definition Rating Index / Ch.-S. Cho, G.E. Gibson // *J. of Architectural Eng.* — 2001. — Vol. 7, № 4. — P. 115 — 125.
5. Wang, Yu.-R. A Study of Preproject Planning and Project Success Using ANN and Regression Models / Yu.-R. Wang, G.E. Gibson // *The 25<sup>th</sup> Intern. Symp. on Automation and Robotics in Construction.* — Vilnius: Vilnius Gediminas Technical University, 2008. — P. 688 — 695.
6. Wang, Yu.-R. Applying Neural Network Ensemble Concepts for Modelling Project Success / Yu.-R. Wang, Yi-J. Chen, C.F. Jeffrey Huang // *26<sup>th</sup> Intern. Symp. on Automation and Robotics in Construction.* — Austin, TX: University of Texas at Austin, 2009. — P. 199 — 204.
7. Breiman, L. Bagging Predictors / L. Breiman // *Machine Learning.* — 1996. — Vol. 24, № 2. — P. 123 — 140.
8. Freud, Y. Experiments with a New Boosting Algorithm / Y. Freud, R.E. Schapire // *Machine Learning: Proc. of the Thirteenth International Conf.* — San-Francisco: Morgan Kaufmann, 1996. — P. 148 — 156.

Рецензент д-р техн. наук, проф. Одес. нац. політехн. ун-ту Гогунський В.Д.

Надійшла до редакції 29 грудня 2010 р.