
DOI <https://doi.org/10.15407/usim.2019.02.032>
УДК 004.89

О.В. ЧОРНА, магістрантка, ПЗКС ФПМ НТУУ «КПІ ім. І. Сікорського»,
Нац. техн. ун-т України «Київський політехнічний інститут імені Ігоря Сікорського»
(НТУУ «КПІ ім. І. Сікорського»), 03056, Київ, просп. Перемоги, 37, Україна,
chorna.oksana.v@gmail.com

Л.А. ЛЮШЕНКО, канд. техн. наук, ст. викладач, кафедра ПЗКС ФПМ НТУУ,
Нац. техн. ун-т України «Київський політехнічний інститут імені Ігоря Сікорського»
(НТУУ «КПІ ім. І. Сікорського»), 03056, Київ, просп. Перемоги, 37, Україна,
yushenkol@gmail.com

Н.А. РИБАЧОК, канд. техн. наук, ст. викладач, кафедра ПЗКС ФПМ НТУУ,
Нац. техн. ун-т України «Київський політехнічний інститут імені Ігоря Сікорського»
(НТУУ «КПІ ім. І. Сікорського»), 03056, Київ, просп. Перемоги, 37, Україна,
rybachok@pzks.fpm.kpi.ua

МОДИФІКОВАНИЙ МЕТОД АВТОМАТИЗАЦІЇ ПРИЙНЯТТЯ УПРАВЛІНСЬКИХ РІШЕНЬ ДЛЯ СТВОРЕННЯ КОМАНДИ УПРАВЛІННЯ ПРОЕКТАМИ

Запропоновано модифікацію «наївного» байєсівського алгоритму з використанням байєсівської мережі для автоматизації прийняття управлінських рішень, а саме для автоматизованого підбору команди проекту. Модифікація методу полягає у використанні обмеження на структуру графа, де кожен вузол додатково може мати не більше одного з батьків серед інших вхідних змінних. Застосування модифікованого методу надає кращий результат за критеріями точності та швидкодії у підборі учасників команди.

Ключові слова: інтелектуальний аналіз даних, теорема Байєса, байєсівська мережа, класифікація, ймовірність, управління людськими ресурсами.

Вступ

Одним із найважливіших факторів успіху в проектах розробки та впровадження програмного забезпечення є ефективне управління людськими ресурсами проекту. Створення команди фахівців (дизайнерів, розробників, архітекторів, тестувальників) є досить складною задачею. Це пов'язано з необхідністю враховувати, як професійні фахові компетенції, так і поведінкові компетенції кожного члена команди проекту. Відбір ускладнюється рольовою моделлю побудови команди проекту [1]. Загалом кожен фахівець в проекті може мати декілька ролей, які потребують різних фахових компетенцій.

Підбір команди або конкретного фахівця може зробити керівник проекту самостійно, проте немає гарантії, що такий вибір буде найкращим та в подальшій роботі не призведе до виникнення конфліктів в проекті, пов'язаних як з кількістю людських ресурсів, так і з фаховою якістю проектної команди. Для зменшення ризиків, пов'язаних з людськими ресурсами в досить великих проектах, використовують програмні засоби. Найбільш ефективними програмними системами для розв'язання таких задач є автоматизовані системи підтримки прийняття рішень (АСППР).

Системи підтримки прийняття рішень використовуються в різних галузях людської діяльності, їх базовими складовими є набутий

досвід та знання, які організовані в бази даних та бази знань. Задачі аналізу даних стають дедалі актуальнішими, особливо в сучасних умовах, коли накопичено значні обсяги інформації практично з усіх напрямів людської діяльності. Автоматизовані системи підтримки прийняття рішень — це програмні комплекси, які аналізують великий обсяг інформації та допомагають приймати управлінські рішення [2]. АСППР вирішує два основних завдання:

- вибір найкращого розв'язку задачі з множини можливих (оптимізація);
- впорядкування можливих рішень (ранжування).

Для аналізу та надання варіантів рішень в АСППР використовують багато різних підходів та їх комбінацій: інформаційний пошук, управління на основі прецедентів, імітаційне моделювання, еволюційне обчислення, генетичні алгоритми, ситуаційний аналіз, нейронні мережі та інтелектуальний аналіз даних.

Інтелектуальний аналіз даних (ІАД) дозволяє знаходити не тільки статистичні закономірності, а й створювати моделі, які описують та прогнозують явища, що є нерегулярними. ІАД базується на таких дисциплінах як: теорія баз даних, машинне навчання, штучний інтелект, статистика та ін. Управління проектами відбувається в значній мірі в області невизначеності, тому ІАД для управлінських рішень в проектах є досить ефективним, оскільки існує необхідність доповнити дослідження середніх тенденцій індивідуальними особливостями.

Аналіз існуючих рішень

Одним з базових методів ІАД є алгоритми перебору. Простий перебір всіх досліджуваних об'єктів вимагає $O(2^N)$ операцій, де N — кількість об'єктів. Зі збільшенням кількості даних обсяг обчислень зростає експоненційно.

Отже, розв'язок будь-якого складного завдання таким методом є практично неможливим.

Іншим прикладом застосування ІАД для проектування АСППР є метод пошуку асоціативних правил, що є окремим випадком задачі класифікації. Суть методу полягає в аналі-

зі послідовності подій, що відбуваються. Якщо будуть виявлені закономірності таких подій, то можна з певною ймовірністю передбачити появу подій у майбутньому [3]. Недоліком цього методу може стати надмірна кількість вхідної інформації, що може ускладнити аналіз закономірностей та спричинити зайву генерацію та підрахунок асоціативних правил.

Системи міркувань на основі прецедентів (аналогічних рішень) використовують для прогнозування майбутніх рішень або для вибору правильних рішень. Системи *Case Base Reasoning (CBR)* приймають рішення на підставі інформації про минулі аналоги конкретної ситуації та обирають відповідь, що була правильною.

Недоліком даних систем є вибір рішення тільки з використанням масивів попередніх даних. За їх допомогою відбувається тільки узагальнення попереднього досвіду, *CBR* не створюють нові правила та моделі.

Для розв'язання задачі підбору фахової команди проекту у статті розглядається один з методів глибокого аналізу даних — найвний байєсівський алгоритм. Перевагою цього методу для розв'язання задачі прийняття управлінських рішень при виборі персоналу проекту є швидке виконання багатоскладової класифікації.

Постановка задачі

Для розв'язання задачі підбору ефективної фахової команди проекту пропонується використання модифікованої байєсівської мережі. Найвний байєсівський алгоритм (НБА) — це алгоритм класифікації, заснований на теоремі Байєса з припущенням про незалежність ознак. Іншими словами, НБА передбачає, що наявність якої-небудь ознаки в класі не пов'язана з наявністю будь-якої іншої ознаки. Навіть якщо ці ознаки залежать одна від одної або від інших ознак, в будь-якому випадку вони вносять незалежний вклад у ймовірність того, що об'єкт, який розглядається належить до класу даних ознак. У зв'язку з таким припущенням алгоритм називається «найвним» [4].

Моделі на основі НБА досить прості і вкрай корисні при роботі з дуже великими наборами даних. При своїй простоті НБА здатний перевершити навіть деякі складні алгоритми класифікації [5]. Однак, недоліком НБА при обробці даних для класифікації, згідно поставленої задачі, є умовна незалежність атрибутів, тобто їх незалежність від класів. Тому пропонується використання модифікованої байєсівської мережі. Модифікація полягатиме у використанні обмеження на структуру графа і буде розглядатися таке розширення «наївної» моделі Байєса, де кожен вузол додатково може мати не більше одного з батьків серед інших вхідних змінних.

Модифікований метод автоматизації прийняття управлінських рішень на основі інтелектуального аналізу даних

Метод автоматизації прийняття управлінських рішень на основі Теорема Байєса. Теорема Байєса використовується для отримання максимально точної оцінки ймовірності події, яка уточнюється з отриманням великої кількості додаткової інформації. Принцип розв'язання задачі з підбору команди проекту полягає саме в оцінці ймовірності події, що деякий співробітник стане учасником команди. Велика кількість додаткових даних, таких як досвід роботи, кількість виконаних проектів конкретного типу, швидкість виконання задач, знання певних технологій допомагає максимально уточнити оцінку вибору учасника команди.

Теорема Байєса дозволяє розрахувати апостеріорну ймовірність $P(C|X)$ на основі $P(C)$, $P(X)$, $P(X|C)$ і за формулою:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)},$$

де $P(C|X)$ — апостеріорна ймовірність даного класу C (тобто даного значення цільової змінної) при даному значенні ознаки X ;

$P(C)$ — апіорна ймовірність даного класу;

$P(X|C)$ — правдоподібність, тобто ймовірність даного значення ознаки при даному класі;

$P(X)$ — апіорна ймовірність даного значення ознаки.

Переваги наївного байєсівського алгоритму:

- Класифікація, в тому числі багатоскладова, виконується легко і швидко.

- Коли припущення про незалежність виконується, НБА перевершує інші алгоритми, такі, наприклад, як логістична регресія (*logistic regression*), і при цьому вимагає менший обсяг навчальних даних.

- НБА краще працює з категорійними ознаками ніж з безперервними. Для безперервних ознак передбачається нормальний розподіл, що є досить сильним припущенням.

Для візуалізації отриманих даних використовуємо рекурсивний алгоритм побудови дерева рішень *ID3*.

Вхід: S — навчальна вибірка, B — множина базових предикатів.

Вихід: коренева верхівка дерева, побудованого за вибіркою S .

Опис алгоритму, реалізованого процедурою *LearnID3(S)*:

- Якщо всі об'єкти з вибірки S належать одному класу, то утворюємо новий лист v і повертаємо (v);

- Розбиваємо вибірку на дві частини за предикатом $S = S_0 \cup S_1$;

- Створюємо новий лист, якщо він належить класу, в якому знаходиться більшість об'єктів вибірки S ;

- Створюємо новий лист, якщо він не належить класу, то створюємо нову внутрішню верхівку v ;

- Будуємо ліву гілку $Lv = \text{LearnID3}(S_0)$ та будуємо праву гілку $Rv = \text{LearnID3}(S_1)$, повертаємо (v).

В результаті роботи алгоритму отримаємо дерево рішень (рис. 1). Дерево складається з листків, що вказують на клас та вузол. Воно може використовуватися для класифікації об'єктів, що не увійшли в навчальну множину. Пошук починається з кореня, поки не буде виявлений клас, що відповідає об'єкту.

Однак, недоліком НБА при обробці даних для класифікації згідно поставленої задачі, є умовна незалежність атрибутів, тобто їх неза-

лежність від класів. Тому пропонується використання модифікованої байєсівської мережі.

Метод автоматизації прийняття управлінських рішень з використанням Байєсівської мережі. Байєсівською мережею називається спрямований граф без циклів, який дозволяє представляти спільний розподіл випадкових змінних. Кожен вузол графа являє випадкову змінну, а дуги — прямі залежності між ними [6, 7]. Мережа описує, як кожна змінна залежить тільки від безпосередніх батьків. Таким чином, граф описує обмеження на залежність змінних одна від одної, що зменшує кількість параметрів спільного розподілу. Параметри спільного розподілу кодуються набором таблиць для кожної змінної в формі умовних розподілів за умови на значення змінних-батьків. Структура графа і умовні розподіли вузлів при значеннях їх батьків однозначно описують спільний розподіл всіх змінних, що дозволяє розв'язувати задачу класифікації як визначення значення змінної класу, максимізуючи її умовну ймовірність при заданих значеннях вхідних змінних. Схематичне представлення байєсівської мережі зображено на рис. 2.

Очевидно, що «наївний» алгоритм Байєса є окремим випадком байєсівської мережі, де кожна вхідна змінна залежить тільки від змінної класу, яка є єдиним коренем графа (рис.3).

Навчання байєсівських мереж стало одним з актуальних напрямів обчислювальної математики і сьогодні є предметом активних досліджень. Однак, до цього часу визначення структури байєсівської мережі в загальному вигляді є складним завданням як з теоретичної, так і з обчислювальної точки зору. Підхід в загальному вигляді має такі недоліки:

- обчислювальну складність;
- при спробі врахувати велику кількість залежностей між змінними, оцінки умовних ймовірностей набувають великі значення дисперсії, так як їх спільна поява в даних є мало-ймовірною подією.

Оцінки параметрів можуть стати недостовірними, що може призвести до погіршення якості класифікації навіть у порівнянні з «наївним» алгоритмом Байєса;

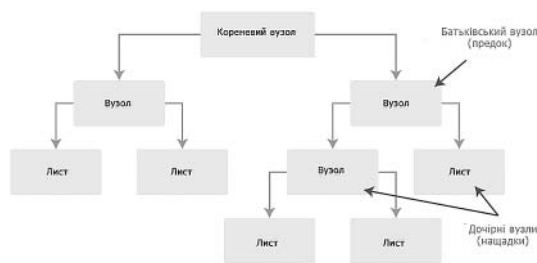


Рис. 1. Дерево рішень за алгоритмом ID3

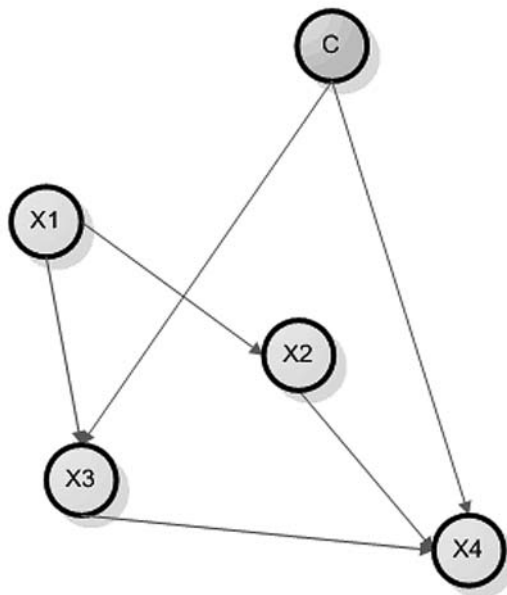


Рис. 2. Байєсівська мережа

▪ через велику кількість параметрів, модель виходить занадто орієнтованою на навчальні дані. Це призводить до дуже гарних результатів класифікації на навчальних даних і незадовільних — на тестових. Тобто модель описує не загальні закономірності в структурі даних, а скоріше набір окремих випадків в навчальній вибірці.

Метод автоматизації прийняття управлінських рішень з використанням модифікованої моделі Байєса. Для вирішення цих проблем запропоновано використовувати обмеження на структуру графа і розглядати таке розширення «наївної» моделі Байєса, де кожен вузол додатково може мати не більше одного з батьків серед інших вхідних змінних (рис. 4).

За допомогою такої модифікації буде досягнута більша точність оцінки підбору кожного

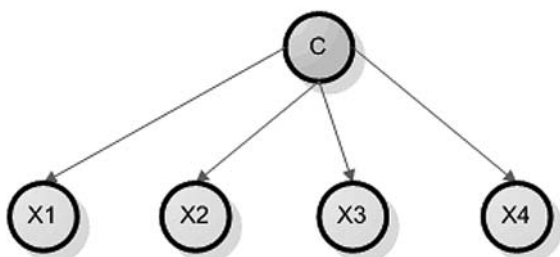


Рис. 3. «Наївний» алгоритм Байєса

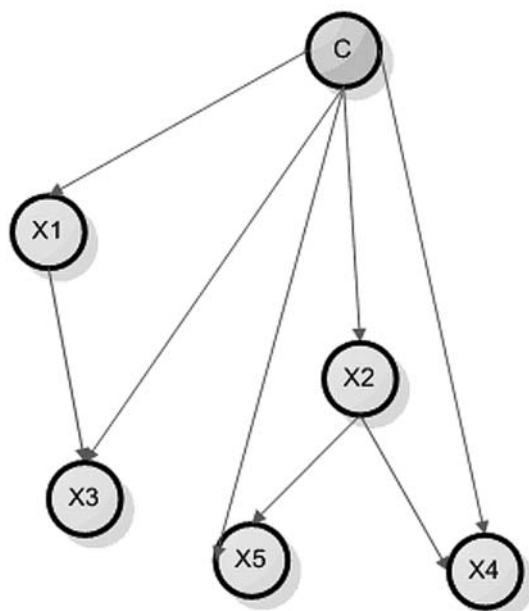


Рис. 4. Модифікований деревовидний алгоритм Байєса

з учасників команди при врахуванні великої кількості вхідних параметрів, які описують професійні навички та компетенції робітників.

Розширимо «наївну» модель Байєса, додавши взаємозв'язки між вузлами моделі. Однак обмежимося розглядом тільки таких графів, де у кожного вузла не може бути більше одного батька, за винятком вузла *C*.

Для вирішення зазначених проблем пропонується використання наступних модифікацій методу:

- визначення кореневого вузла, як вузла, що в найбільшій мірі впливає на результат класифі-

кації за змінною *C*. Як критерій ступеня впливу обрано взаємну інформацію між кожним незалежним вузлом і класифікуючим вузлом;

- введення мінімального порогу на значення взаємної умовної інформації між вузлами. Якщо взаємна інформація між вузлами виявиться менше цього порогу, будемо вважати відповідні вузли незалежними. Як такий поріг виберемо середнє значення взаємної умовної інформації за всіма вузлами.

Суть модифікованої байєсівської мережі та використання алгоритму *ID3* у площині прийняття управлінських рішень полягає у реалізації наступних етапів.

1. Обробка початкових даних проекту.

На даному етапі особа, яка приймає рішення про склад команди проекту, вводить початкові дані проекту (тематику, терміни, кількість членів команди і т. ін.) для створення профілю проекту. За введеними даними виконується обробка початкових текстових даних для кожного члена команди за наступними критеріями: досвід роботи, кількість виконаних проектів, відповідальність, дотримання термінів, розуміння бізнес-процесів; далі отримані дані класифікуються або кластеризуються і на їх основі формується база знань та будуються моделі.

2. Обробка даних для прийняття рішення з бази знань.

Далі з бази знань робиться вибірка працівників компанії та їх даних (атрибутив), які співпадають з профілем даного проекту.

3. Побудова дерева рішень для проекту.

На основі отриманої вибірки будується дерево рішень, яке відображає склад команди проекту на основі алгоритму *ID3*, описаному у даній статті.

4. Визначення оптимального складу команди проекту.

Перелік учасників команди проекту, отриманий при побудові дерева рішень перевіряється механізмом оцінки якості рішення. У разі низької якості команди, проводиться уточнення початкових даних проекту і виконується побудова дерева рішень. Ці дії виконуються до тих пір, поки не буде прийнято оптимального рішення з підбору складу команди.

Експериментальні результати

Для аналізу ефективності запропонованої модифікації методу реалізовано відповідне програмне забезпечення на мові C#, а також створено програмне забезпечення в середовищі *Matlab*, що реалізує стандартний «наївний» алгоритм Байєса.

Проведення тестування відбувалось з використанням різної кількості особових справ співробітників.

Модифікований метод автоматизованого прийняття рішення з вибору команди проекту показав кращі результати за критерієм точності та швидкодії, ніж стандартний метод у *MatLab*.

При завантаженні 30 особових справ модифікований метод показав точність підбору учасника команди на 6 відсотків краще та на 30 відсотків швидше. При завантаженні 100 особових справ модифікований метод показав точність підбору учасника команди на 11 відсотків краще та на 45 відсотків швидше.

Висновки

Запропоновано модифікований метод автоматизації прийняття управлінських рішень на основі одного з підходів інтелектуального аналізу даних — байєсівської мережі та використання алгоритму побудови дерева рішень *ID3*.

Даний метод має наступні переваги: вирішення проблем обчислювальної складності; покращення якості класифікації навіть у порівнянні з «наївним» алгоритмом Байєса; навіть при використанні великої кількості параметрів будуть отримані результати класифікації з більшою точністю як для навчальних даних, так і для тестових даних.

Наведено детальний опис та модифікацію запропонованого методу. Метод може бути розширений додатковими критеріями щодо формування команди проекту. Модифікований метод автоматизованого прийняття рішення по вибору команди проекту показав кращі результати за критерієм точності та швидкодії ніж стандартний метод, реалізований в *MatLab*.

ЛІТЕРАТУРА

1. Шапиро В.Д., Ильин Н.И. и др. Управление проектами. СПб.: Газпром, 1996, 610 с.
2. Keen Peter. Decision support systems: a research perspective. Cambridge, Massachusetts: Center for Information Systems Research, Alfred P. Sloan School of Management. 1980.
3. John Wiley & Sons – Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data. 2014. P. 12–19.
4. Maron M.E. Automatic Indexing: An Experimental Inquiry. 1961.
5. Онишук Р.М., Терещенко І.М. Опис алгоритму для побудови наївного байєсовського класифікатора. 2015.
6. Pearl Judea. Causality: Models, Reasoning, and Inference. Cambridge University Press. 2000. ISBN 978-0-521-77362-1
7. Згуровський М.З., Бідюк П. І., Терентьев О.М., Присянкін-Жарова Т.І. Байєсівські мережі в системах підтримки прийняття рішень. К.: ТОВ «Видавниче Підприємство «Едельвейс», 2015. 300 с.

Надійшла 01.04.2019

REFERENCES

1. Shapiro, V.D., Il'in, N.I. 1996. Project Management. San-Peterburg, Gazprom, 610 p. (In Russian).
2. Keen P., 1980. Decision support systems: a research perspective. Cambridge, Massachusetts: Center for Information Systems Research, Alfred P. Sloan School of Management.
3. John Wiley & Sons – Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data. 2014, pp. 12–19.
4. Maron M.E., 1961. Automatic Indexing: An Experimental Inquiry.
5. Onyschuk, R.M., Tereschenko, I.M., 2015. Algorithm description for constructing a Naive Bayes classifier. (In Ukrainian).
6. Pearl Judea. Causality: Models, Reasoning, and Inference. Cambridge University Press. 2000. ISBN 978-0-521-77362-1
7. Zgurovsky, M.Z., Bidyuk, P.I., Terentyev, O.M., Prosyankina-Zharova, T.I., 2015. Bayesian Networks in Decision Support Systems. K.: Edelweis, 300 p. (In Ukrainian).

Received 01.04.2019

O.V. Chorna, Master student of the Applied Mathematics Faculty, National Technical University of Ukraine "Igor Sikorsky Kyiv Politechnic Institute",
03056, Peremohy Ave 37, Kyiv, Ukraine,
chorna.oksana.v@gmail.com

L.A. Lyushenko, PhD, Senior Lecturer, Computer Systems Software Department of the Applied Mathematics Faculty, National Technical University of Ukraine "Igor Sikorsky Kyiv Politechnic Institute", 03056, Peremohy Ave 37, Kyiv, Ukraine,
yushenkol@gmail.com

N.A. Rybachok, PhD, Senior Lecturer, Computer Systems Software Department of the Applied Mathematics Faculty, National Technical University of Ukraine "Igor Sikorsky Kyiv Politechnic Institute", 03056, Peremohy Ave 37, Kyiv, Ukraine,
rybachok@pzs.fpm.kpi.ua

MODIFIED METHOD OF AUTOMATION DECISION-MAKING MANAGEMENT PROCESS FOR CREATION OF THE PROJECT MANAGEMENT TEAMS

Introduction. One of the most important factors in the success of software development projects is the effective management of the human resources. For these purposes automated decision support systems are commonly used that can be based on data mining methods.

Purpose. The purpose of this research is to develop a modified method of automation decision-making management process based on data mining.

Methods. Modification of the method consists in using the restriction on the structure of the graph and considering such an extension of the Naive Bayes classifier, where each node can additionally have no more than one parent among other input variables. Using this modification, better accuracy will be achieved in assessing the selection of each team member, taking into account a large number of input parameters that describe the professional skills and competencies of the workers.

Results. To analyze the effectiveness of the proposed method modification, appropriate software was implemented using C# language. There was also implemented Naive Bayes algorithm in Matlab for comparison. Testing was carried out using a variety of personal cases. The modified method of automation decision-making management process of choosing a project team showed better results for accuracy criteria than the standard MatLab method, as well as a higher performance.

Conclusion. The presented modified method of automation decision-making management process based on data mining can be used for an efficient selection of the project team.

Keywords: *Data mining, Bayes' theorem, Bayes' network, classification, probability, human resource management.*

О.В. Черная, магістрантка, кафедра ПОКС ФПМ НТУУ «КПИ им. И. Сикорского»,
Нац. техн. ун-т Украины «Киевский политехнический институт имени Игоря Сикорского»
НТУУ «КПИ им. И. Сикорского», просп. Победы, 37, Киев, 03056, Украина,
chorna.oksana.v@gmail.com

Л.А. Люшенко, кандидат технических наук, ст. преподаватель, кафедра ПЗКС ФПМ НТУУ
Нац. техн. ун-т Украины «Киевский политехнический институт имени Игоря Сикорского»
НТУУ «КПИ им. И. Сикорского», просп. Победы, 37, Киев, 03056, Украина,
yushenkol@gmail.com

Н.А. Рыбачок, кандидат технических наук, ст. преподаватель, кафедра ПЗКС ФПМ НТУУ
Нац. техн. ун-т Украины «Киевский политехнический институт имени Игоря Сикорского»
НТУУ «КПИ им. И. Сикорского», просп. Победы, 37, Киев, 03056, Украина,
rybachok@pzks.fpm.kpi.ua

МОДИФИЦИРОВАННЫЙ МЕТОД АВТОМАТИЗАЦИИ ПРИНЯТИЯ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ ДЛЯ СОЗДАНИЯ КОМАНДЫ УПРАВЛЕНИЯ ПРОЕКТАМИ

Вступление. Одним из наиболее важных факторов успеха проектов разработки программного обеспечения является эффективное управление человеческими ресурсами проекта. Создание команды специалистов (дизайнеров, разработчиков, архитекторов, тестировщиков) является достаточно сложной задачей. Это связано с необходимостью учитывать, как профессиональные, так и поведенческие компетенции каждого члена команды проекта. Отбор усложняется ролевой моделью построения команды проекта.

Для решения задачи подбора эффективной команды для создания проекта обычно используются автоматизированные системы поддержки принятия решений, которые могут основываться на методах интеллектуального анализа данных.

Целью данного исследования является разработка модифицированного метода автоматизации процесса принятия решений на основе интеллектуального анализа данных.

Методы. Используя эту модификацию, будет достигнута лучшая точность при оценке выбора каждого члена команды с учетом большого количества входных параметров, которые описывают профессиональные навыки и компетенции работников.

Результаты. Для анализа эффективности предложенного способа модификации было реализовано соответствующее программное обеспечение с использованием языка C#. Для сравнения был также реализован алгоритм Наивного Байеса в *Matlab*. Тестирование проводилось с использованием различных дел. Модифицированный метод автоматизации процесса принятия решений по выбору проектной команды показал лучшие результаты по критериям точности, чем стандартный метод *MatLab*, а также более высокую производительность.

Заключение. Представленный модифицированный метод автоматизации процесса принятия решений на основе интеллектуального анализа данных может быть использован для эффективного отбора проектной команды.

Ключевые слова: интеллектуальный анализ данных, теорема Байеса, байесовская сеть, классификация, вероятность, управление человеческими ресурсами.