

7. *Кривенко С. С.* Фильтрация изображений с использованием преобразования Хаара в блоках / С. С. Кривенко, Е. О. Колганова, В. В. Лукин. // Системи обробки інформації. – 2013. – №8. – С. 63–69.
8. *Healey G.* Radiometric CCD camera calibration and noise estimation / G. Healey, R. Kondepudy. // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1994. – №16(3). – С. 267–276.
9. *Еремеев В. В.* Статистическая оценка степени зашумленности космических изображений земной поверхности [Электронный ресурс] / В. В. Еремеев, В. А. Зенин, П. А. Князьков // Вестник РГРТУ № 2 (выпуск 24). – 2008. – Режим доступа до ресурсу: http://www.rsreu.ru/en/component/docman/doc_download/545-3-7.
10. *Gonzalez, R. C., Woods, R. E.* (2002). Digital image processing, (2nd ed.). Upper Saddle River, N.-J.: Prentice Hall. 793 p.

<http://doi.org/10.5281/zenodo.3610669>

Поступила 5.08.2019р.

УДК 004.942

В.Р. Вергун, аспірант, кафедра АСУ, ІКНІ, НУ “Львівська політехніка”,
І.М. Дронюк, к.ф.м.н., кафедра АСУ, ІКНІ, НУ “Львівська політехніка”,

МАТЕМАТИЧНА МОДЕЛЬ РАНЖУВАННЯ СЛУХАЧІВ ОНЛАЙН-КУРСІВ З ПРОГРАМУВАННЯ

Abstract. The article describes the mathematical models for ranking students of online IT courses. The four types of factors are investigated. Nonacademic factors are taking into account. For constructing criteria, Boolean functions are used. Criteria of success studying are extended to the studying of a group of students. On the base of mathematical model, the information technology is constructed. The schemes of information technology and construction are presented. Creating information technology prognoses the successes of online IT course students studying. The result of the application is time and others resources for accompaniment of studying minimizing.

Анотація. У статті розроблено математичну модель ранжування слухачів онлайн-курсів з програмування. Досліджено чотири типи факторів. Вперше запропоновано враховувати неакадемічні фактори, такі як креативність та зацікавленість до навчання. Для побудови критеріїв використовуються булеві функції. Представлено критерії успішності. Критерії успішності навчання поширюються на вивчення групи студентів. На основі математичної моделі побудовані представлені схеми побудови інформаційної технології. Створення інформаційних технологій прогнозує успіх студентів, які навчаються на онлайн курсах з програмування. Результатом заявки є час та інші ресурси для супроводу вивчення мінімізації.

Вступ

Індустрія інформаційних технологій відіграє важливу роль в зростанні економіки в Україні в цілому. Приблизна кількість спеціалістів залучених до індустрії перевищила 90 000 і експерти очікують зростання вдвічі впродовж наступних 3 років [1].

Таким чином, в зв'язку з популярністю і великими можливостями, кожного року збільшується інтерес до ІТ індустрії з боку вже працевлаштованого населення в інших галузях. А для того, щоби отримати відповідну кваліфікацію, потрібно приділяти велику увагу освіті, і це є найкращим способом отримати необхідні навички. Це призводить до ряду проблем. Першою проблемою є процес створення навчального плану та оптимальних графіків навчання. Через швидку зміну технологій та підходів до розробки програмного забезпечення навчальні програми не повністю покривають всі потреби індустрії. Індустрія не має можливості чекати довгий час на підготовку та перекваліфікацію працівників. Тому, здатність до швидкої та автономної розробки офлайн чи онлайн програм та висока адаптованість є критичними для будь яких навчальних закладів.

Іншою проблемою є початковий рівень кандидатів, які хочуть розпочати навчання на освітніх програмах, та коректне визначення цього рівня. Знання на початковому етапі навчання є критично важливими для того, щоби приймати правильні рішення щодо участі в навчальних програмах певних кандидатів, та для створення більш персоналізованих програм, а також онлайн курсів, запровадження яких призводить до значної економії матеріальних ресурсів. В свою чергу, такі онлайн програми дозволяють зменшити час навчання, з більшою ймовірністю завершити весь цикл і отримати необхідні знання та навички. Також окрім певного рівня знань кандидати та студенти повинні володіти навичками до швидкого навчання та здатності адаптовуватись до різноманітного формату та типу матеріалів при навчанні онлайн. А визначення таких навичок, та створення моделей для прогнозування успішності, на основі факторів, які не мають стосунку до академічних успіхів є актуальною науковою задачею.

1. Постановка задачі

Прогнозування успішності студентів онлайн-курсів - одне з найпоширеніших завдань у сучасних дослідженнях в області аналізу даних освітніх програм. Переважно дослідження зосереджуються на прогнозуванні кінцевих результатів та прогнозуванні кількості відрахувань з курсу. Дослідження проводяться на різних наборах даних з різними факторами. Слід зазначити, що вибірка необхідних та правильних даних є однією з найбільших проблем у цих дослідженнях. Для прогнозування автори розглядають набори академічних, неакадемічних та соціальних факторів. Відповідно до процесу інтелектуального аналізу даних - інформація попередньо обробляється, а інструменти машинного навчання використовуються для вирішення завдань класифікації чи кластеризації.

Інструменти машинного навчання та алгоритми є найпоширенішим інструментом досліджень способів прогнозування успішності. У дослідженні [2] автори запропонували модель нейронної мережі, щоб передбачити середній бал успішності після завершення навчання студентів-інженерів. Для дослідження були розглянуті дані 1470 студентів першого курсу. Модель включає сім факторів неакадемічних показників та одинадцять факторів історії середньої школи. Автори запропонували включити до моделі такі неакадемічні фактори, як лідерство, здатність до навчання, мотивація, тощо. Під час моделювання успішності автори досягли точності прогнозування як 71,3%. Штучні нейронні мережі широко використовуються для такого роду досліджень. Автори дослідження [3] використовували звичайні статистичні оцінки для виявлення факторів, які, ймовірно, покращують результативність студентів. Нейронна мережа була змодельована з 11 вхідних змінних. Модель нейронної мережі досягла хорошої точності прогнозування 84,8%. Використовуючи методи прогнозування, можна виявити студентів, які перебувають під ризиком відрахування, та вчасно інформувати як викладачів, так і студентів. Автори досліджень [4] проаналізували та виміряли співвідношення між демографічними показниками та академічними характеристиками для прогнозування кількості студентів, що були відраховані з навчальних програм. Такі неакадемічні фактори, як дохід батьків, рівень освіти батьків, стать та вік, були включені в модель прогнозування. Автори поєднали алгоритми з методами класифікацій, використовуючи Gradient Boosting і отримали близько 79,12% точності прогнозування із запропонованим набором даних.

Існує багато підходів прогнозування успішності студентів та виявлення тих, хто перебуває під ризиком. За даними огляду літератури [5] така кількість досліджень постійно збільшується. І основною сферою досліджень є виявлення ключових факторів що мають вплив на академічну успішність слухачів як офлайн так і онлайн курсів.

Метою даного дослідження є розробка математичної моделі ранжування слухачів онлайн-курсів з програмування з використанням неакадемічних факторів.

Для виконання даної задачі потрібно виконати наступні кроки:

- Дослідити та визначити основні неакадемічні фактори.
- Визначити фактори, які мають вплив на успішність слухачів курсів з програмування.
- Визначити функції критеріїв успішності.
- Побудувати математичну модель на основі факторів та критеріїв, та визначити на її основі схему інформаційної технології.

2. Розв'язання задачі

2.1 Опис факторів

У нашому дослідженні було виділено 4 типи факторів, що мають вплив на успішність навчання слухачів онлайн курсів з програмування, а саме:

соціальні фактори, стилі навчання, фактор зацікавленості та індекс креативності.

Розглянемо детальніше виділені нами фактори, зокрема, соціальні фактори, які були вивчені у нашому дослідженні. Стать враховувалась як фактор з описом (M or F). Освіта також була визначена булівським описом (Y or N). Фактор віку було вирішено розділити на 3 підфактори, вік менше рівне 22 років, вік у проміжку 22-29 років, не враховуючи початкове і кінцеве значення та вік більше рівне 29 років. Фактори досвіду та наявність додаткової освіти було взято з булівським описом (Y or N). Дані для визначення соціальних факторів визначались на основі даних резюме слухачів курсів, які вони подають при поступленні на навчання.

Наступна група факторів, яка має вплив на успішність навчання - це стилі навчання. Стилi навчання слухачів є особливо важливими в онлайн курсах. Оскільки різноманітний медіа-контент вимагає певних навиків до самоосвіти, та сприйняття інформації. Визначення стилів навчання відбувається за допомогою опитування. Згідно з методикою цього опитування було виділено 4 якісні ознаки стилю навчання: activist, reflector, theorist, pragmatist.



Рис. 1. Метод визначення індексу зацікавленості.

Наступні фактори, що розглядались у даному дослідженні, це фактори зацікавленості. Цей фактор впливу на успішність навчання розглядається вперше. Основою для формування цих факторів були комунікаційні потоки під час навчання, зокрема, чат і форум. Усі виділені фактори, що характеризують зацікавленість, кількісні та формуються на основі аналізу повідомлень слухачів курсів у комунікативних потоках. Було розглянуто кількість повідомлень (цифра), кількість цитувань (цифра), кількість запитань (цифра), кількість технічних запитань (цифра). На рис.1 показано схему методу визначення індексу зацікавленості. Спочатку у комунікаційних потоках вибираються повідомлення, що пов'язані з певним слухачем. На наступному кроці проводиться інтелектуальний аналіз цих даних,

повідомлення класифікуються, визначається їхня вага та формуються на цій основі цифрові значення кількісних характеристик факторів зацікавленості.

Останньою групою факторів, що була розглянута, є група факторів, що пов'язана з креативністю. Відомо, що креативність є визначальним фактором інтелектуальної економіки. Тому цей фактор є дуже важливим для слухачів online IT курсів. Для оцінки цього фактора виділено один параметр, який названий як індекс креативності. Індекс креативності у дослідженні є булівського типу з характеристикою (Y or N). Визначення даного індексу відбувається за допомогою опитування слухачів.

2.2 Математична модель

Для побудови математичної моделі введемо у розгляд булівський простір як множину, яка складається з одиниці та нуля $B^1 = \{0,1\}$, а також багатовимірний булівський простір як декартів добуток одновимірного $B^n = \prod_{i=1}^n B^1$. Операціями «додавання» та «множення» слугуватимуть звичайне булівське додавання та множення.

Для опису моделі було виділено 4 групи факторів, а саме: соціальні, навчальні, фактори зацікавленості та креативності. Для характеристики першої групи факторів об'єкта дослідження введемо у розгляд множину соціальних факторів $S = \{\vec{S}_1; \vec{S}_2\}$, де $\vec{S}_1 \in B^p$, а $\vec{S}_2 \in R^m$, де $p + m \leq n$. Будемо вважати, що фактори вектора \vec{S}_1 відповідають характеристикам, що мають бінарне визначення, а фактори вектора \vec{S}_2 описують характеристики, що мають числові значення.

Другу групу факторів, що характеризують навчання, опишемо вектором $\vec{L} \in R^k$.

Для опису третьої групи факторів, що характеризують зацікавленість слухачів курсів у навчанні, розглянемо вектор $\vec{Z} \in R^q$, а для опису четвертої групи факторів, що характеризують креативність слухачів курсів, використаємо вектор $\vec{C} \in B^r$.

Введемо у розгляд сукупну множину визначення усіх факторів моделі X, яку задамо наступною формулою

$$X = \{B^p, R^m, R^k, R^q, B^r\}, \quad (1)$$

де величина $N=p+m+k+q+r$ визначає сумарну кількість виділених факторів для прогнозування успішності навчання слухача курсів. Тоді успішність навчання слухача курсів можна визначити на основі наступної булівської цільової функції

$$F: X \rightarrow B^1. \quad (2)$$

Успішне проходження навчання можна описати формулою

$$F(x) \equiv TRUE, \text{ де } x \in X, \quad (3)$$

де X та F задані співвідношеннями (1), (2).

Якщо в групі на навчання є K слухачів, то загальну успішність групи слухачів будуть характеризувати наступні функції

$$\Psi = F_1 \wedge F_2 \wedge \dots \wedge F_K, \quad (4)$$

де символ \wedge означає кон'юнкцію, а кожна з функцій F_i задана за формулою (2), отже функція (4) є кон'юнктивною, що означає її істинність тоді, коли усі слухачі курсів успішно пройшли навчання.

Інша функція, що описує навчання групи є диз'юнктивна, а саме

$$\Phi = F_1 \vee F_2 \vee \dots \vee F_K, \quad (5)$$

де символ \vee означає кон'юнкцію, а кожна з функцій F_i задана за формулою (2), отже ця функція буде істинна, якщо хоча б один слухач успішно закінчить курси.

З практики проведення навчання на курсах вдається виділити два основні критерії для визначення успішності завершення навчання слухача. Перший критерій відповідає тому, що слухач успішно пройшов всі навчальні завдання до завершення навчання і не кинув його, тобто функцію F з (3) для слухачів курсів можна конкретизувати у вигляді $F^{(1)}$

$$F^{(1)}(x) \equiv TRUE, \quad (6)$$

де $F^{(1)}$ - має зміст булівської функції, що залежить від усіх визначених факторів та вона набуде істинне значення тоді, якщо слухач успішно дійшов до кінця навчання.

Другий критерій відповідає тому, що слухач успішно здав завершальний тест, тобто функцію F з (3) для цього випадку запишемо у вигляді $F^{(2)}$

$$F^{(2)}(x) \equiv TRUE, \quad (7)$$

де $F^{(2)}$ - має зміст булівської функції, що залежить від усіх визначених факторів та вона набуде істинне значення тоді, якщо слухач успішно здав завершальний тест.

Успішне завершення навчання одного слухача тоді можна охарактеризувати наступною формулою

$$F^{(1)}(x) \wedge F^{(2)}(x) \equiv TRUE. \quad (8)$$

Таким чином формула (8) виражає чіткий математичний критерій успішного завершення одним слухачем online IT курсів та може бути використана для побудови інформаційної технології.

Для опису успішного навчання групи, що містить K слухачів, підставимо формулу (6) у вирази (4) та (5), отримаємо функції, що описують успішне навчання групи слухачів.

У кон'юнктивній формі,

$$\Psi(x) = (F_1^{(1)}(x) \wedge F_1^{(2)}(x)) \wedge (F_2^{(1)}(x) \wedge F_2^{(2)}(x)) \wedge \dots \wedge (F_K^{(1)}(x) \wedge F_K^{(2)}(x)) \quad (9)$$

У диз'юнктивній формі

$$\Phi(x) = (F_1^{(1)}(x) \wedge F_1^{(2)}(x)) \vee (F_2^{(1)}(x) \wedge F_2^{(2)}(x)) \vee \dots \vee (F_K^{(1)}(x) \wedge F_K^{(2)}(x)) \quad (10)$$

У формулах (9) та (10) змінна $x \in X$ з множини (1) та містить N різноманітних факторів, що впливають на успішність навчання. Таким чином побудовано функції (9) (10) критеріїв успішного навчання групи K студентів, що залежить від N факторів різної природи.

2.3 Інформаційна технологія

На основі розробленої математичної моделі розроблено інформаційну технологію прогнозування успішності навчання слухачів online IT курсів. Розроблена інформаційна технологія реалізує збір, опитування та збереження статистичних та динамічних даних слухачів. На основі цих даних проводить інтелектуальний аналіз зібраних даних, класифікацію та формує відповідні значення факторів, що впливають на успішність навчання. На основі цього інформаційна технологія проводить ранжування слухачів курсів. На рис.2 показана структурна схема роботи інформаційної технології ранжування слухачів курсів.

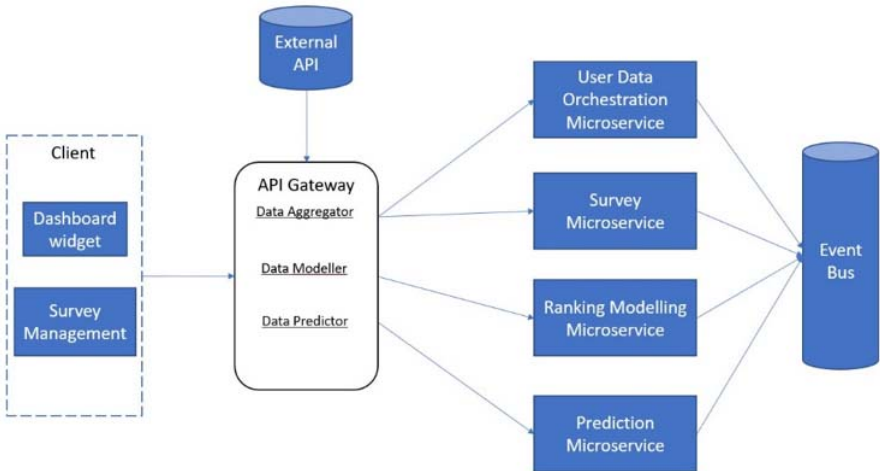


Рис. 2. Структурна схема інформаційної технології

Дана інформаційна технологія володіє можливістю інтеграції з системами підтримки онлайн навчання, для збору даних слухачів онлайн курсів. Основним елементом системи є сервіс ранжування слухачів, за допомогою якого відбувається прогнозування з випередженням заданих критеріїв.

Висновки

У даній роботі визначено основні неакадемічні фактори для побудови математичної моделі ранжування слухачів онлайн курсів з програмування. Запропоновано використання індексу зацікавленості, як один з методів аналізу комунікативних потоків онлайн курсів. Використано індекс креативності, як один з вирішальних факторів успішності слухачів саме IT курсів. Створено математичну модель ранжування групи слухачів на основі запропонованих неакадемічних факторів. Визначено структуру інформаційної технології на основі математичної моделі, яка дозволяє інтеграцію з іншими системами підтримки навчання, для збору інформації та здійснення прогнозування з випередженням.

1. <https://www.n-ix.com/software-development-in-ukraine-2019-2020-market-report/>
2. *Jin, Q., Imbrie, P. K., Chen, X. (2011, June)*. AC 2011-1608: A multi-outcome hybrid model for predicting student success in engineering. Paper presented at the 2011 American Society for Engineering Education, Vancouver, BC.
3. *Lau, E.T., Sun, L. & Yang, Q.* SN Appl. Sci. (2019) 1: 982. <https://doi.org/10.1007/s42452-019-0884-7>
4. *Hutagaol, Nindhia & Suharjito, Suharjito.* (2019). Predictive Modelling of Student Dropout Using Ensemble Classifier Method in Higher Education. *Advances in Science, Technology and Engineering Systems Journal*. 4. 10.25046/aj040425.
5. *Hellas, A., Ihantola, P., Petersen, A., Ajanovski, V. V., et al.* (2018). Predicting Academic Performance: A Systematic Literature Review. In *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*, (pp. 175–199). (ITiCSE 2018 Companion). New York, NY, USA: ACM. <https://doi.org/10.1145/3293881.3295783>

<http://doi.org/10.5281/zenodo.3610673>

Поступила 8.08.2019р.