

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ЮРИДИЧНА АКАДЕМІЯ»
ФАКУЛЬТЕТ КІБЕРБЕЗПЕКИ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра інформаційних технологій

КВАЛІФІКАЦІЙНА РОБОТА

на тему:

**«МЕТОДИ ТА АЛГОРИТМИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ
ДЛЯ АНАЛІЗУ ПОВЕДІНКИ СТУДЕНТА
В СИСТЕМІ ДИСТАНЦІЙНОГО НАВЧАННЯ»**

Виконав: здобувач 2 курсу

Рівень магістр

Галузь знань 12 «Інформаційні
технології»,

спеціальність 124 «Системний аналіз»

Божко Павло Володимирович

Керівник: к.т.н., доцент

Манаків Сергій Юрійович

Рецензент: к.пед.н., доцент

Логінова Наталія Іванівна

Одеса – 2024

АНОТАЦІЯ

Кваліфікаційна робота на тему “Методи та алгоритми інтелектуального аналізу даних для аналізу поведінки студента в системі дистанційного навчання” на здобуття другого (магістерського) рівня вищої освіти за спеціальністю 124 «Системний аналіз», освітньою програмою «Аналіз та безпека даних», містить 11 рисунків, 1 таблицю, 29 літературних джерел за переліком посилань. Робота виконана на 67 сторінках загального тексту і 56 сторінках основного тексту.

Метою роботи є застосування методів та алгоритмів інтелектуального аналізу даних для дослідження поведінки студента у системі електронної освіти.

Методи досліджень базуються на використанні положень теорії алгоритмів, інтелектуального аналізу даних, лінійної алгебри та теорії ймовірностей.

Розроблено математичну модель інтелектуального аналізу поведінки студента при в системі електронного навчання. Проведено математичне та комп'ютерне моделювання алгоритму CFSFDP-HD, представлено експеримент та порівняння результатів аналізу роботи різних методів.

Основний науковий результат магістерської дисертації полягає в тому, що застосування методів та алгоритмів інтелектуального аналізу даних для дослідження поведінки студента у системі електронної освіти може бути покращено за рахунок застосування алгоритмів CFSFDP та CFSFDP-HD. Застосування в системах електронної освіти алгоритмів таких алгоритмів є ефективним та доцільним.

ДОСЛІДЖЕННЯ ПОВЕДІНКИ СТУДЕНТА, СИСТЕМА ЕЛЕКТРОННОГО НАВЧАННЯ, ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ, CFSFDP, CFSFDP-HD.

ABSTRACT

The qualification work on the topic " Data mining methods and algorithms for analyzing student behavior in a distance learning system" for obtaining the second (master's) level of higher education in the specialty 124 "System Analysis", educational program "Data Analysis and Security", contains 11 figures, 1 table, 29 literary sources according to the list of references. The work was completed on 67 pages of the general text and 56 pages of the main text.

The purpose of the work is the application of methods and algorithms of intellectual data analysis for the study of student behavior in the electronic education system.

Research methods are based on the use of algorithm theory, intelligent data analysis, linear algebra and probability theory.

A mathematical model of intellectual analysis of student behavior in the electronic learning system has been developed. Mathematical and computer modeling of the CFSFDP-HD algorithm was carried out, an experiment and a comparison of the results of the analysis of the work of different methods were presented.

The main scientific result of the master's thesis is that the application of methods and algorithms of intelligent data analysis for the study of student behavior in the electronic education system can be improved due to the use of CFSFDP and CFSFDP-HD algorithms. Application of such algorithms in electronic education systems is effective and expedient.

STUDENT BEHAVIOR RESEARCH, E-LEARNING SYSTEM, INTELLIGENT DATA ANALYSIS, CFSFDP, CFSFDP-HD.

ЗМІСТ

ВСТУП.....	7
1 АНАЛІЗ ПОВЕДІНКИ СТУДЕНТІВ У СИСТЕМІ ЕЛЕКТРОННОГО НАВЧАННЯ.....	10
1.1 Системи електронного навчання	10
1.2 Моделі аналізу процесу онлайн-навчання та постановка задачі.....	14
2 BIG DATA В ЕЛЕКТРОННОМУ НАВЧАННІ	23
2.1 Big Data у сучасних системах електронного навчання	23
2.2 Форма подання даних про студентів.....	24
2.3 Алгоритми вивантаження великих даних.....	29
2.4 База даних у системі електронної освіти	31
3 АЛГОРИТМИ АНАЛІЗУ BIG DATA	34
3.1 Інтелектуальний аналіз даних для Big Data.....	34
3.2 Методики експертної оцінки поведінки студента у системах електронної освіти	38
3.3 Методи інтелектуального аналізу даних у системі електронного навчання ...	40
3.3.1 Прогнозування	40
3.3.2 Виявлення структури	42
3.3.3 Виявлення взаємозв'язків	42
3.4 Алгоритми інтелектуального аналізу даних у системі електронного навчання	43
3.4.1 Зсув середнього значення.....	44
3.4.2 K-means.....	45
3.4.3 K-medoids	46
3.4.4 DBSCAN.....	47
3.4.5 Ієрархічна кластеризація	48
3.4.6 CFSFDP та CFSFDP-HD	49
4 КОМП'ЮТЕРНА МОДЕЛЬ АЛГОРИТМА CFSFDP- HD ДЛЯ ОЦІНКИ ПОВЕДІНКИ СТУДЕНТА В СИСТЕМІ ЕЛЕКТРОННОГО НАВЧАННЯ.....	53
4.1 Комп'ютерна модель алгоритму CFSFDP-HD.....	53

4.2 Аналіз даних про студентів у системі електронного навчання за допомогою алгоритмів K-means та CFSFDP-HD.....	55
4.2.1 Алгоритм CFSFDP-HD	55
4.2.2 Порівняння результатів аналізу даних різними методами інтелектуального аналізу	57
4.2.3 Результати дослідження	60
ВИСНОВКИ.....	62
ПЕРЕЛІК ПОСИЛАНЬ	64

ВСТУП

В даний час електронне навчання – важливий елемент навчальної та викладацької діяльності у переважній кількості коледжів та університетів по всьому світу. У сучасних вищих навчальних закладах основні бібліотечні служби, ІТ та служби підтримки підтримують послуги з електронного навчання. Це багато у чому зумовлено сучасною епідеміологічною ситуацією у світі та сьогоdnішніми бойовими діями в Україні.

Електронна освіта містить у собі електронні підручники, навчальні посібники, освітні послуги та технології. Сучасні студенти та школярі – це покоління, для яких електронна інформація та способи її отримання є повсякденною складовою життя. Загалом, сучасні технології в освіті сприймаються студентами позитивно, – знання, вміння та навички, які вони здобувають, використовуються надалі не тільки в кар'єрному зростанні, а й у повсякденному житті.

В огляді українського та світового ринку електронного навчання [1] виділяється те, що застосування технологій у навчанні дає безмежні можливості для розвитку та отримання корисних знань, навичок. Здобувати знання у XXI столітті стало швидко, захоплююче та доступно. З погляду інтересів держави, електронна освіта надає можливість здобуття однакового рівня освіти людям, які живуть у віддалених точках країни, тобто. електронна освіта дозволяє усунути освітню нерівність.

Актуальність дослідження обумовлена необхідністю застосувати методи інтелектуального аналізу даних для оцінки поведінки студента, що дозволить з більшою ймовірністю прогнозувати можливі помилки та підміну даних у системі електронного навчання.

В умовах формування системи безперервної освіти та підвищення кваліфікації, дистанційне навчання стало оптимальною формою випереджувального навчання та професійної перепідготовки. Розвиток дистанційної освіти потребує спеціальних інформаційних можливостей для забезпечення безперебійної роботи мережі, а також для полегшення викладачів та

студентів. Проте, питання якості залишається відкритим.

Завдання, які вирішуються за допомогою методів інтелектуального аналізу даних, можна розділити на описові (англ. descriptive) та передбачувальні (англ. predictive) [2].

У описових завданнях дається наочний опис наявних прихованих закономірностей, тоді як у передбачуваних вирішується питання про передбачення тих випадків, котрим даних ще немає. Описові завдання:

- пошук правил чи патернів;
- кластерний аналіз;
- знаходження регресійної моделі.

Передбачувальні завдання:

- класифікація об'єктів;
- регресійний аналіз.

В даний час, інтелектуальний аналіз даних є напрямком, що розвивається, за допомогою якого можна поліпшити якість електронної освіти. Це підкреслює актуальність даної досліджень у даному напрямку.

Метою роботи є застосування методів та алгоритмів інтелектуального аналізу даних для дослідження поведінки студента у системі електронної освіти.

Об'єкт дослідження – аналіз поведінки студента під час проходження курсу електронного навчання.

Предмет дослідження – методи та алгоритми інтелектуального аналізу даних у системі електронного навчання.

Гіпотеза дослідження полягає в тому, що можна застосовувати методи та алгоритми інтелектуального аналізу для ухвалення рішення про коректність поведінки студента в системі електронного навчання, якщо:

- позначені методи оцінки поведінки студента;
- реалізовано технологію, яка забезпечує достовірний аналіз поведінки студента під час проходження курсу електронного навчання;
- проведено дослідження поведінки студента за допомогою методів та алгоритмів інтелектуального аналізу даних.

Виходячи з мети дослідження та для перевірки висунутої гіпотези, необхідно вирішити такі завдання:

- провести порівняльний аналіз існуючих способів та засобів аналізу поведінки студента під час проходження курсу електронного навчання;
- дослідити можливі критерії порівняння поведінки студентів;
- дослідити методи та алгоритми інтелектуального аналізу даних;
- розробити математичну модель оцінювання студента за даними його поведінки;
- спроектувати та реалізувати математичну модель інтелектуального аналізу поведінки студента під час проходження курсу електронного навчання;
- перевірити ефективність розробленої моделі та визначити результативність її впровадження.

Наукова новизна дослідження полягає у використанні сучасних методів та алгоритмів інтелектуального аналізу даних для оцінки поведінки студента у системі електронної освіти.

Практична значимість полягає в аналізі поведінки студента при проходженні курсу електронного навчання з використанням методів та алгоритмів інтелектуального аналізу даних, що дозволить з більшою ймовірністю прогнозувати можливі помилки та підміну даних у системі дистанційного навчання.

ВИСНОВКИ

У ході виконання магістерської роботи було сформульовано цілі та завдання, визначено об'єкт та предмет, висунуто гіпотезу, позначено актуальність теми дослідження.

Виконано завдання магістерської роботи:

- проведено порівняльний аналіз існуючих способів та засобів аналізу поведінки студента під час проходження курсу електронного навчання;
- досліджено можливі критерії порівняння поведінки студентів;
- досліджено методи та алгоритми інтелектуального аналізу даних;
- розроблено математичну модель оцінювання студента за даними його поведінки;
- спроектовано та реалізовано математичну модель інтелектуального аналізу поведінки студента при проходженні курсу електронного навчання;
- перевірено ефективність розробленої моделі.

Грунтуючись на аналізі робіт з теми кваліфікаційної роботи, було описано теоретико-методологічні засади інтелектуального аналізу даних поведінки студентів у системі електронного навчання. Визначено основні пункти аналізу поведінки студентів у системі електронного навчання. Розглянуто методи інтелектуального аналізу даних у системі електронного навчання. Виділено головні переваги та недоліки існуючих програмних пакетів щодо оцінки поведінки студентів у системі електронного навчання. Описано існуючі програмні пакети, які реалізують методи інтелектуального аналізу даних у системі електронного навчання.

Було проведено аналіз існуючих розробок на тему «Методи та алгоритми інтелектуального аналізу даних для оцінки поведінки студента в системі електронного навчання», розглянуто існуючі системи електронного навчання та їх особливості. Досліджено моделі аналізу процесу онлайн-навчання, а також Big Data у системах електронної освіти. Описано форму подання даних про студентів, алгоритми вивантаження великих даних та подано логічну модель

даних у системах електронної освіти.

Досліджено методи інтелектуального аналізу даних:

- прогнозування;
- виявлення структури;
- виявлення взаємозв'язків.

Досліджено алгоритми інтелектуального аналізу даних, позначено їх переваги та недоліки:

- зсув середнього значення;
- K-means;
- K-medoids;
- DBSCAN;
- ієрархічна кластеризація;
- CFSFDP та CFSFDP-HD.

Проведено математичне та комп'ютерне моделювання алгоритму CFSFDP-HD, представлено експеримент та порівняння результатів аналізу роботи різних методів.

Основний науковий результат магістерської дисертації полягає в тому, що застосування методів та алгоритмів інтелектуального аналізу даних для дослідження поведінки студента у системі електронної освіти може бути покращено за рахунок застосування алгоритмів CFSFDP та CFSFDP-HD. Застосування в системах електронної освіти алгоритмів CFSFDP та CFSFDP-HD є ефективним та доцільним.