

ЗАСТОСУВАННЯ МЕТОДІВ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ ДЛЯ РОЗВ'ЯЗАННЯ ОСВІТНІХ ТА УПРАВЛІНСЬКИХ ЗАВДАНЬ

Проведення ефективної політики та реформ у сфері освіти вимагає застосування нових методів аналізу для підготовки організаційних і управлінських рішень, адекватних сучасним завданням. У даній ситуації інформаційно-аналітичне забезпечення стає одним з головних «сервісів» у вирішенні проблеми модернізації управління якістю освіти.

У зв'язку із зростаючими обсягами статистичної інформації в навчально-виховній та організаційно-управлінській діяльності ВНЗ, що накопичується в розподілених, розрізнених джерелах даних, і вимогами до аналізу інформації, які постійно змінюються актуальним стає використання методів інтелектуального аналізу даних (Data Mining) для моніторингу навчальної діяльності, аналізу стану системи освіти у ВНЗ, прогнозування її розвитку, тощо [1]. Інтелектуальний аналіз даних – це потужна технологія для аналізу важливої інформації зі сховища даних. Ця технологія аналізу даних використовується для ідентифікації прихованих закономірностей у великому наборі даних. Інтелектуальний аналіз даних успішно використовується в різних областях, включаючи й освітнє середовище.

В даній час методи Data Mining отримали широке поширення в різних сферах діяльності. Дослідженнями в цій області займаються такі вчені, як А.А. Барсегян, М.С. Купріянов, Г. Пятецькій-Шапіро, Х. Ромесбург, Дж. Хан. Проблеми аналізу даних освітнього процесу розглядалися в роботах таких вчених, як Р. Бакер, Л.І. Григор'єв та інші.

Інтелектуальний аналіз освітніх даних (Educational Data Mining) є цікавим напрямом дослідження, який витягує корисні, раніше невідомі закономірності з навчальних баз даних для кращого розуміння і поліпшення успішності та оцінки процесу навчання студентів та прийняття ефективних управлінських рішень [2]. Застосування методів інтелектуального аналізу даних до освітніх баз даних дозволить підвищити ефективність системи вищої освіти [3]. Неявна інформація, отримана від видобутку освітніх наборів даних може бути застосована наприклад, для точного передбачення підсумкової оцінки студентів, зниження відсіву студентів, для класифікації студентів за додатковими предметами, які найбільш їм підходять, і так далі. Це допоможе особам, які приймають рішення виявити асоціації, шаблони і тенденції, які можуть привести до поліпшення освітніх процесів.

Метою роботи є проведення аналізу дослідження можливостей застосування методів інтелектуального аналізу даних для розв'язання різних освітніх та управлінських завдань з використанням програмних пакетів MatLab та Weka.

Основна мета інтелектуального аналізу даних – повністю автоматичне або напівавтоматичне знаходження в зібраних даних залежностей, що представляють практичну цінність в контексті сфери застосування даної технології. Методи інтелектуального аналізу даних поділяються на три групи [4,5]: пошуковозалежні (discovery), прогнозування (predictive modelling) і аналіз аномалій (forensic analysis). Пошук залежностей полягає в перегляді бази даних з метою автоматичного виявлення залежностей. Проблема тут полягає у відборі дійсно важливих залежностей з величезного числа даних існуючих в базі даних. Прогнозування передбачає, що користувач може пред'явити системі записи з незаповненими полями і запросити відсутні значення. Система сама аналізує вміст бази і робить правдоподібне пророкування щодо цих значень. Аналіз аномалій – це процес пошуку підозрілих даних, які сильно відхиляються від стійких залежностей. Технології інтелектуального аналізу даних дозволяють вирішувати безліч завдань із залученням методів математичної статистики і теорії ймовірності, а також методів штучного інтелекту. Найбільшого поширення знайшли методи, що дозволяють вирішувати наступні завдання:

- Класифікація – віднесення об'єкта (події, предмета) до одного із задалегідь відомих класів за його характеристикам;
- Регресія – прогнозування значення будь-якого вихідного параметра об'єкта по набору вхідних параметрів;
- Кластеризація – завдання полягає в поділі об'єктів на кластери за значеннями, які притаманні об'єктам параметрів. Вирішення цього завдання допомагає краще зрозуміти дані;
- Пошук асоціативних правил – виявлення закономірностей між якими-небудь пов'язаними об'єктами. Вирішення цього завдання допомагає краще зрозуміти природу аналізованих даних і може служити для прогнозування появи подій.
- Прогнозування послідовностей – знаходження залежностей між об'єктами або подіями у формі правил, що вказують, після якої події А настає подія В;
- Аналіз відхилень – аналіз даних на предмет входження явних нехарактерних шаблонів.

Проблеми аналізу і моделювання освітнього процесу у ВНЗ формулюються схожим чином, і вирішення більшості з них зводиться до тієї чи іншої задачі інтелектуального аналізу даних або до їх комбінації.

Для розв'язання освітніх завдань з використанням методів інтелектуального аналізу даних використовуються програмні пакети MatLab (Matrix Laboratory) – сучасний інструмент аналізу даних, в якому реалізовані різні методи Data Mining та Weka (Waikato Environment for Knowledge Analysis) – вільно поширюваний програмний пакет з відкритим вихідним кодом для аналізу даних, що являє собою набір засобів візуалізації і алгоритмів для інтелектуального аналізу даних.

У доповіді пропонується аналіз дослідження можливостей застосування методів інтелектуального аналізу даних для розв'язання освітніх та управлінських завдань, зокрема аналіз стану системи освіти у ВНЗ, прогнозування її розвитку, прогнозування «траєкторії» навчальної та наукової роботи студентів різних спеціальностей, тощо. Для розв'язання поставлених завдань застосовуються різні методи Data Mining, зокрема дерева рішень, нейронні мережі, нечітка логіка. Для практичного розв'язання освітніх завдань використовуються програмні пакети MatLab та Weka.

Список використаних джерел

1. Абриков В.С., Ефремов Л.Г., Кошеев И.Г. Новые подходы к разработке моделей системы поддержки принятия решений и управления вузом // Вестник Чувашского университета. 2013. № 1. С. 224 - 229.
2. Goyal, Monika. Applications of Data Mining in Higher Education. International journal of computer science, 2012, 9 (2), p. 113.
3. Програмний інструментарій Education Intelligence як засіб підвищення ефективності креативного навчання / В. Гришачов, Д. Замятін, О. Кебало, А. Михайлюк, Л. Огнівчук, В. Тарасенко // Міжнародний науковий журнал "Комп'ютинг", 2011, Том 10, Випуск 2 - с. 114-132.
4. Барсегян А.А., Куприянов М.С. и др. Методы и модели анализа данных: OLAP и Data Mining. – СПб.: БХВ-Петербург, 2004.
5. Fayyad, U. Advances in Knowledge Discovery and DataMining / U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy.– AAAI/MIT Press, 1996.