

## ФУНКЦІОНАЛЬНІ ОСОБЛИВОСТІ GUI–ІНТЕРФЕЙСУ БІБЛІОТЕКИ АЛГОРИТМІВ XELOPES

Інтелектуальний аналіз даних (Data Mining) набуває широке поширення в науці і бізнесі [1]. Актуальною стає задача підготовки фахівців-аналітиків, які знають, як зберігаються, обробляються, фільтруються дані, як виконуються класифікація, регресія, кластеризація. У галузі освіти спостерігається процес стрибкоподібного впровадження Business Intelligence–систем як методів, технологій, засобів видобування і представлення знань [2]. Для вивчення нових технологій аналізу даних запроваджуються нові дисципліни. Мета навчання полягає у підготовці сучасного фахівця, здатного вирішувати складні аналітичні задачі і ухвалювати науково обгрунтовані рішення.

Для набуття практичних навичок аналізу даних на заняттях можна використовувати популярні програмні засоби Data Mining: SPSS, MatLab, PolyAnalyst, Statistica, Xelopes. Перелічені програми доступні, однак освоєння нових інструментів вимагає певних зусиль.

Бібліотека алгоритмів Xelopes привертає увагу своїми особливостями — підтримкою стандартів Data Mining, незалежністю від платформи, незалежністю від вихідних даних, доступністю. Бібліотека розроблена компаніями ProdSys і ZSoft [3]. Для забезпечення зручності разом з бібліотекою у вигляді окремого додатку поставляється графічний інтерфейс користувача (GUI). Функції інтерфейсу: завантаження даних; проглядання даних у вигляді таблиці; проглядання атрибутів даних; проглядання статистичної інформації про дані; побудова моделі Data Mining; візуалізація моделі; збереження моделі; застосування моделі.

Завантаження вихідних даних виконується в діалоговому режимі натисненням кнопки **Open Mining Data** на панелі інструментів. Текстовий файл формату ARFF відкривається натисненням на кнопку **Open**.

Проглядання вихідних даних виконується в табличному виді (кнопка **View Input Data**). Рядки таблиці відповідають об'єктам, які досліджуються, а стовпці – атрибутам.

GUI–інтерфейс Xelopes дозволяє одержати докладну інформацію про атрибути завантажених даних (кнопка **Display Data Description**). У діалоговому вікні **Variables** відображаються відомості про категоріальні і числові атрибути.

Настройка відображення статистичної інформації про дані виконується в діалоговому вікні **Statistics** (кнопка **Display Descriptive Statistics**). Вибираються параметри: атрибут по осі X, атрибути по осі Y, тип діаграми – плоска або просторова. На діаграмі можна відобразити наступні статистичні параметри: кількість об'єктів, мінімальне і максимальне значення, ранг, суму, середнє.

Для побудови Mining моделі застосовуються методи Data Mining (кнопка **Build Mining Model**). У діалоговому вікні пропонується перелік моделей. Доступні 8 моделей: 1)Association Rules Mining Model; 2)Decision Tree Mining Model; 3)SVM (Support Vector Machine Mining Model; 4)Sequential Mining Model; 5)Customer Sequential Mining Model; 6)Partition Clustering Mining Model; 7)CDBased Clustering Mining Model; 8)Hierarchical Clustering Mining Model.

Параметри настройки процесу побудови моделі залежать від типу моделі і виконуються на вкладці **Settings**. Вибір алгоритму виконується на вкладці **Algorithm**. Список доступних алгоритмів також залежить від типу моделі. Після виконання настройок для побудови моделі необхідно натиснути на кнопку **Build**.

Побудовану модель можна зберегти натисненням на кнопку **Save Model as PMML**. Формат PMML (predictive modeling mark-up language) є мовою розмітки, яка використовується для статистичних моделей і моделей Data Mining. Візуальне представлення (Browse Model) GUI-інтерфейс Xelopes забезпечує тільки для трьох видів моделей: асоціативних правил, дерев рішень, ієрархічної кластерної моделі — у вигляді діаграми даних (datagram).

Побудована модель може бути застосована в задачах класифікації і регресії (кнопка **Apply Model**). Вибирається файл з новими даними. Результати відображаються у стовпці з назвою: predicted\_attribute.

1. Барсегян, А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP: уч. пос. / [А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод].–2-е изд. – СПб.: БХВ-Петербург, 2007. – 384 с.
2. Перминов, Г. И. Системы интеллектуального анализа данных (Business Intelligence). Учебно-метод. комплекс. — М.: ГУ – ВШЭ, 2007.
3. Компания ZSoft. Библиотека Xelopes / <http://www.zsoft.ru/page.php?14>.