

23. Дослідження можливостей використання методів Web Mining для оцінки ефективності функціонування web-ресурсу

Анастасія Малоок

Національний університет харчових технологій

Вступ. Метою дослідження є аналіз даних веб-ресурсу (спільноти соціальної мережі «ВКонтакте») для виявлення недоліків, підвищення відвідуваності і, як наслідок, збільшення кількості підписантів групи та клієнтів салону краси.

Матеріали і методи. У соціальній мережі «ВКонтакте» існує спільнота, яка рекламує послуги салону краси. Адміністратору цієї групи доступна інформація про різноманітні характеристики підписантів спільноти, їх кількість, реакція на запропоновану рекламу салону, статистика відвідувань спільноти тощо.

Дана інформація структурована і є основою для проведення аналізу з метою підвищення відвідуваності спільноти користувачами мережі «ВКонтакте» та для збільшення потоку клієнтів салону краси. Аналіз отриманих даних проводився за допомогою інструментальної платформи Microsoft SQL Server Analysis Services та Microsoft Office Excel.

Дані з соціальної мережі вивантажуються в Excel-документ за допомогою вбудованої функції вивантаження даних. Отриманою інформацією на основі розроблених процедур SQL Server Integration Services (SSIS) заповнюється сховище даних. SSIS включає в себе набір потужних засобів очищення, перетворення та

завантаження, які дозволяють здійснювати з даними всі маніпуляції, необхідні при заповненні сховища. Аналіз накопиченої інформації проводиться за допомогою засобів інтелектуального аналізу даних, включених до складу Microsoft SQL Server служби Analysis Services.

Метод кластеризації з набору методів Data Mining дозволяє розподіляти відвідувачів спільноти за характеристиками на групи «схожих» об'єктів, що називаються кластерами. За замовчуванням моделі кластеризації Microsoft SQL Server Analysis Services використовують алгоритм масштабованої максимізації очікування, який спочатку визначає зв'язки в наборі даних та формує ряд кластерів на основі цих зв'язків. Після першого визначення кластерів алгоритм обчислює, як кластери представляють групування точок, а потім намагається повторно визначити групування, щоб створити кластери, які краще представляють дані. Алгоритм послідовно виконує цей процес доти, доки поліпшити результати, визначаючи кластери, буде неможливо.

Таким чином, підписанти були розбиті на групи, що дозволило дізнатися про те, люди яких вікових категорій, статі, місця проживання тощо реагують на різні типи та види реклами.

Результати. Отже, були застосовані служби Microsoft SQL Server та Microsoft Office Excel. Ці програмні компоненти є зручними у використанні та доступні користувачам Інтернет, тому розв'язання поставленої задачі не вимагає великих витрат.

Аналіз отриманих даних дозволив розробити рекомендації щодо збільшення кількості підписантів спільноти. Визначивши спільні за характеристиками та вподобаннями групи людей на сайті, підбирається цікавий саме для цих користувачів контент, розробляються системи акцій та знижок.

Розуміння потреб цільової аудиторії дозволяє налаштовувати рекламу таким чином, що під час її запуску спостерігається підвищення відвідуваності спільноти, що свідчить про доцільність та ефективність її використання.

Висновки. Web Mining є новим перспективним напрямком аналізу інтернет-ресурсів для оптимізації структури веб-сайтів, дослідження вподобань відвідувачів спільноти, що дозволяє підібрати необхідний для успішного функціонування та «розкрутки» сайту контент.

Література

1. Барсегян А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP [Текст] / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод. — 2-е изд., перераб. и доп. — СПб. : БХВ-Петербург, 2009. — 384 с. : ил. — CD-ROM.
2. Markov Z. Data Mining the Web: uncovering patterns in Web content, structure, and usage / Markov Z., Larose D. T. — John Wiley & Sons Inc., 2007. — 218 p.

Науковий керівник: к. т. н., доц. Харкянен О. В.