

*Свиридова Л.Е.,
Аспирант 1 курс, кафедра
«Системного анализа и исследования операций»
Сибирский государственный университет науки и технологий имени
академика М.Ф. Решетнева
Россия, г. Красноярск*

ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ DATA MINING

***Аннотация:** применение data mining очень обширно и может быть использовано практически где угодно, где требуется машинный анализ текста. В статье проанализировано его применение в таких сферах как: интернет-технологии, торговля, телекоммуникации, медицина, банковское дело.*

***Ключевые слова:** data mining, статистика, анализ данных, web mining, data analysis, базы данных, искусственный интеллект.*

***Annotation:** the using of data mining is very extensive and can be used almost anywhere where computer-aided text analysis is required. The article analyzes its use in such areas as: Internet technology, trade, telecommunications, medicine, banking.*

***Key words:** data mining, statistics, data analysis, web mining, data analysis, databases, artificial intelligence.*

Data mining (Интеллектуальный анализ данных) – исследование данных, использующее методы искусственного интеллекта и ориентированное на придание системе свойств искусственного интеллекта.

Применение средств искусственного интеллекта для изучения информации, а также стремление придать системе свойства интеллектуальных систем называется интеллектуальным анализом данных, другими словами data mining [1, с. 8].

Базовые задачи, используемые в data mining:

- выявление отклонений – поиск информации, выделяющейся из основного массива данных;
- ассоциация - распознавание соотношений среди данных;
- кластерный анализ – группировка похожих событий;
- классификация – это соотнесение событий к известным шаблонам;
- регрессия – анализ функции, показывающей группу данных с минимальным выбросом;
- подведение итогов – преобразование первоначальной информации в переработанном виде, а также вывод отчетов [2, с. 69].

Применение интеллектуального анализа данных возможно во многих сферах, где требуется анализ больших массивов информации.

Интернет-технологии

Web mining — технология, использующая методы data mining для исследования и извлечения информации из web-документов и сервисов [3]. Особенностью web mining является использование типичных инструментов data mining для изучения весьма разрозненных, направленных и обширных по охвату данных, встречающихся в интернете. В данной сфере особенно заметны два направления: web content mining и web usage mining. Первое направление также называется text mining, оно занимается самостоятельным отбором и поиском подходящих данных из «захлавленных» источников. Web usage mining находит зависимости в действиях пользователя определенного веб-ресурса (или нескольких), в том числе его поведение внутри сайта [2, с. 69].

Торговля

В торговле технология интеллектуального анализа данных способна решить серьезные задачи. Например, благодаря анализу временных рядов, планирование поставок становится эффективнее, с помощью изучения прогноза продаж на основе тенденции прошлых периодов. Также изучение продаж и формирования ассортимента продукции помогает найти связи среди тех или иных товаров, и осознать, почему потребитель выбирает именно такую продукцию. Благодаря анализу данных появляется возможность понимать целевую

аудиторию, и, соответственно, использовать более тонкие инструменты маркетинга. Можно вычислить модель поведения клиента, на которого стоит ориентироваться в разных случаях. Формирование профиля покупателей дает возможность для анализа взаимосвязей, благодаря которым можно выявить данные, которые трудно получить при изучении каждого клиента по отдельности. Полученную информацию можно использовать для продвижения выгодных предложений и предложения скидок на товары, дополняющие те, что они приобрели. Данная технология позволяет организовать базу данных о клиентах, в том числе потенциальных, снизить траты на продвижение и планировать работу с поставщиками более эффективно [4].

Телекоммуникации

Современные информационные и коммуникационные системы позволяют хранить большое количество данных из транзакций. С помощью анализа данных из транзакций можно получить новые знания о пользователях продуктов / услуг / решений компании. Появляется возможность применять эти знания для определения привычек пользователя и формирования эффективных сегментов рынка, которые будут характеризоваться схожими потребительскими привычками. Данные могут быть использованы для интеллектуального анализа данных, что позволит получать новые знания о привычках и склонностях пользователей. Целенаправленный подход к отдельным сегментам рынка дает значительные конкурентные преимущества. Используя кластерный анализ в качестве предлагаемой модели сегментации рынка, можно создать привлекательные сегменты рынка. Это позволяет компаниям управлять прибыльностью и лояльностью каждого пользователя [5].

Медицина

Документация и архивы данных, содержащиеся в картотеках медицинских учреждений включают в себя огромные массивы информации о всевозможных фактах всех отдельных заболеваний. Вполне разумно предположить, что внутри этих данных существуют закономерности, которые можно извлечь для медицинских исследований. В этих случаях применяются методы

автоматической обработки данных, благодаря которым происходит та самая «добыча данных» - data mining.

Существует много систем, использующих медицинские базы для определения диагнозов. Они базируются на основе правил, описывающих сочетания симптомов различных заболеваний. Благодаря им, можно узнать не только болезнь пациента, но и как его нужно лечить [6]. Data mining помогает находить в данных шаблоны, которые оказывают помощь при выборе средств лечения, определении показаний, создании более действенного лечения, изучении факторов появления различных патологий, изучении нужности оперативного вмешательства [7].

Банковское дело

В банковском деле технология интеллектуального анализа используется в решении различных задач. С помощью изучения предыдущих операций по кредитным карточкам, проведенных мошенниками, банк может выявлять стереотипы такого поведения для обнаружения случаев мошенничества. Также для эффективной реализации маркетинговой политики можно сегментировать клиентов, реализовывая специфичные услуги каждой группе клиентов. Анализ данных позволяет прогнозировать ценностные модели своих посетителей и индивидуально подходящим способом их обслуживать [8].

Можно привести еще много примеров различных областей знания, где методы Data Mining играют ведущую роль.

ИСПОЛЬЗОВАННЫЕ ИСТОЧНИКИ

1. Мусаев, А.А. Интеллектуальный анализ данных: учебное пособие // СПб.: СПбГТИ(ТУ), 2018. – 56 с.
2. Барсегян А.А., Куприянов М. С., Холод И. И., Тесс М. Д., Елизаров С. И.. Анализ данных и процессов: учеб. Пособие// — 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009. —512 с.: ил. + CD-ROM — (Учебная литература для вузов).

3. Etzioni O. The world wide web: Quagmire or gold mine. // Communications of the ACM, 39(11):65—69, 1996.
4. Самбудагва Д.Б. Big data analysis в сфере продуктового ритейла. // Бизнес-образование в экономике знаний. – 2015. – №. 1 (1).
5. Bach, Mirjana Pejic; Simicevic, Vanja; Leskovic, Darko. Data mining in telecommunications: case study of cluster analysis. // The Free Library // URL: [https://www.thefreelibrary.com/Data mining in telecommunications: case study of cluster analysis.-a0224712437](https://www.thefreelibrary.com/Data+mining+in+telecommunications%3A+case+study+of+cluster+analysis.-a0224712437).
6. Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Классификация и снижение размерности. // М.: Финансы и статистика, 1989. – 471 с.
7. Марухина О.В., Мокина Е.Е., Берестнева Е.В. применение методов data mining для выявления скрытых закономерностей в задачах анализа медицинских данных. // Фундаментальные исследования. – 2015. – № 4. – С. 107-113.
8. Дюк В.А., Флегонтов А.В., Фомина И.К. Применение технологий интеллектуального анализа данных в естественнонаучных, технических и гуманитарных областях // Известия РГПУ им. А.И. Герцена. 2011. №138. URL: <https://cyberleninka.ru/article/n/primenenie-tehnologiy-intellektualnogo-analiza-dannyh-v-estestvennonauchnyh-tehnicheskikh-i-gumanitarnyh-oblastyah>.