

ТЕОРИЯ И МЕТОДИКА ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

Жуков Николай Николаевич

ассистент кафедры

ФГБОУ ВПО «Российский государственный
педагогический университет им. А.И. Герцена»

г. Санкт-Петербург

ЭЛЕКТРОННОЕ ОБУЧЕНИЕ ИЗВЛЕЧЕНИЮ ДАННЫХ ИЗ СОЦИАЛЬНЫХ СЕТЕЙ

Аннотация: в статье дано описание элементов электронного обучения извлечению данных из социальных сетей, отдельных элементов интеллектуального анализа данных. Автор приходит к заключению о целесообразности вложения в основу подготовки студентов открытой курсовой модели электронного обучения.

Ключевые слова: электронное обучение, извлечение данных, социальные сети, интеллектуальный анализ данных, программирование, социальная сеть, стигмергия, *data mining*, *knowledge discovery in databases*.

Актуальность обучения студентов инженерных специальностей извлечению данных из социальных сетей объясняется взрывным характером распространения последних и сосредоточением в них больших объемов информации, представляющей исследовательский интерес, для обработки которой необходимы различные продвинутое алгоритмы. Обучение студентов, с одной стороны, представляет собой прикладную задачу веб-программирования, а с другой стороны, позволяет развернуть исследовательскую деятельность в электронной информационно-образовательной среде [1, с. 71], направленную на изучение свойств социальных сетей и закономерностей информационных процессов, протекающих в них.

Количество зарегистрированных пользователей одной только социальной сети Foursquare/Swarm превысило 55 миллионов пользователей, совершивших

более 6 миллиардов отметок о посещении различных мест по всему миру. В таких сетях одни пользователи разделяют с другими различную информацию: текстовые заметки, изображения и видеоролики, свое местоположение, мнение об объектах (точках) и опыт их посещения, используя для этого современные возможности мобильных устройств (например, смартфонов и планшетов).

В сектор информационных технологий требуется все больше и больше высококвалифицированных специалистов различных направлений, что, в свою очередь, делает актуальным подготовку квалифицированных кадров в рамках высшего профессионального обучения. Таким образом, современный уровень подготовки инженеров в области информационных технологий предполагает изучение таких направлений как обработка больших объёмов данных (big data), интеллектуальный анализ данных (intelligent data analysis) и извлечение данных (data mining). Это общемировые тенденции, зафиксированные в том числе и в стратегии развития информационных технологий в Российской Федерации.

Термины «извлечение данных» (data mining) и «интеллектуальный анализ данных» (intelligent data analysis), а также термин «обнаружение знаний в базах данных» (knowledge discovery in databases), являющийся близким по значению впервые появились в англоязычных научных публикациях. На сегодняшний день, их, как правило, считают синонимами как в англоязычном, так и в русскоязычном научном сообществах.

Подготовка к извлечению данных может быть организована в рамках таких курсов, как «Программирование», «Разработка мобильных приложений» или «Веб-проектирование и веб-дизайн». В основу подготовки студентов целесообразно положить открытую курсовую модель электронного обучения [2, с. 164], при которой:

- 1) основное содержание подготовки сосредоточено в открытых курсах MOOC (например, на площадке Coursera);
- 2) стигмергическое взаимодействие студентов осуществляется в локальной среде электронного обучения (Moodle);

3) демонстрационные примеры и компоненты портфолио студентов размещаются на основе онлайн-редактора кода (например, kodaktor.ru).

Процесс извлечения данных предполагает изучение множества тем, в числе которых «клиент-серверное» взаимодействие, которое в свою очередь предполагает активное использование API социальных сетей. С учетом актуальности направления «извлечение данных» и изучения студентами «клиент-серверного» взаимодействия, можно сформулировать ряд задач, основанных на использовании API популярных социальных сетей, таких как Instagram, Twitter, Foursquare/Swarm и Flickr.

Приведем примеры таких задач:

– получение информации о пользователях одной социальной сети и медиаконтенте, который они разместили, а также найти их профили в других социальных сетях;

– получение медиаконтента пользователей, опубликовавших его в определенной локации, используя социальную сеть Instagram и кластеризовать пользователей этой сети (выделить группы схожих пользователей), учитывая информацию, которую они размещают;

– получение всех записей пользователей, опубликовавших (с помощью Twitter) их на определенной территории (геолокации) с определенными ключевыми словами (хэштегами) и классификация записи (выделив записи определенной тематики), с использованием заранее заданные категории.

Решение задач, описанных выше, можно условно разделить на два этапа:

– техническое получение данных;

– обработка и анализ данных, добыча знаний (извлечение данных).

Список литературы

1. Государев И.Б. Стигмергия в образовании на платформе веб и телекоммуникационных систем // Современное образование: традиции и инновации. – 2015. – №2. – С. 69–75.

2. Государев И.Б. Электронное обучение: тенденции развития моделей и опыт применения // Известия Российского государственного педагогического университета имени А.И. Герцена. – СПб., 2013. – №162. – С. 162–166.