

УДК 004.62

ИСПОЛЬЗОВАНИЕ ПАРСЕРОВ КАК СРЕДСТВО ОТСЛЕЖИВАНИЯ АКТУАЛЬНЫХ ЦЕН НА МАРКЕТПЛЕЙСАХ

Михаил Евгеньевич Филипчик

студент

miklfil@mail.ru

Наталья Викторовна Картечина

кандидат сельскохозяйственных наук, доцент

kartechnatali@mail.ru

Лариса Ивановна Никонорова

кандидат сельскохозяйственных наук, доцент

lenaniknrva@rambler.ru

Мичуринский государственный аграрный университет

г. Мичуринск, Россия

Аннотация. Использование парсеров для сбора и анализа данных о ценах на торговых онлайн-площадках, включая маркетплейсы, позволяет автоматизировать процесс отслеживания актуальных цен, анализировать тенденции рынка и контролировать ценообразование. Для сбора и обработки данных веб-страниц использовались язык программирования Python и библиотека BeautifulSoup.

Ключевые слова: парсинг, маркетплейс, мониторинг цен, автоматизация поиска, сравнение цен, сбор данных.

Парсинг представляет собой процесс, который помогает извлекать нужные данные из структурированной информации, такой как HTML-код веб-страниц, XML-файлы или текстовые документы. Это позволяет автоматизировать сбор информации с веб-ресурсов, что может быть полезно для самых разных задач: от мониторинга цен до создания баз данных.

Когда речь идет о мониторинге цен на маркетплейсах, парсинг становится особенно актуальным. Он помогает собирать информацию о ценах на товары и услуги с различных интернет-ресурсов, таких как онлайн-магазины и маркетплейсы. Парсер анализирует HTML-код страницы, находит нужные данные, например, цены и описания товаров, и извлекает их для дальнейшего использования.

Маркетплейсы создают платформу для взаимодействия продавцов и покупателей, но при этом не всегда предоставляют открытую информацию о ценах, что усложняет их отслеживание для конкурентов и поставщиков.

Существует несколько методов парсинга:

Веб-скрейпинг — это когда парсер собирает данные с веб-страницы. Такие парсеры требуют регулярного обновления, чтобы поддерживать изменения на сайте.

API — метод, при котором парсер использует программный интерфейс маркетплейса для сбора информации. Основная проблема здесь — это отсутствие открытого доступа к API.

Браузерный парсинг — метод, при котором парсер имитирует действия браузера, чтобы собирать данные в режиме реального времени. Однако, каждый маркетплейс требует отдельной настройки парсера.

Для работы с HTML и XML файлами отлично подходит библиотека Python под названием BeautifulSoup. Она упрощает навигацию, поиск и изменение структуры данных, экономя много времени.

Из-за отсутствия открытых баз данных на маркетплейсах, парсеры становятся полезным инструментом для отслеживания цен. Они позволяют собирать информацию о ценах, анализировать и сравнивать их.

Процесс начинается с выбора маркетплейса и настройки парсера для сбора нужной информации, такой как название товара, цена и дата. Парсер имитирует действия браузера и собирает данные, которые затем можно сохранить в базе данных или текстовом файле.

Одной из проблем парсеров является то, что изменения на веб-страницах могут нарушить их работу, поэтому их нужно регулярно обновлять. Парсеры – это мощный инструмент, но при их использовании важно учитывать ограничения, риски, обеспечивать безопасность и соблюдать законодательство.

Однако, несмотря на все сложности, парсинг открывает перед бизнесом новые способы влияния на рынок. С его помощью компании могут не только следить за ценами конкурентов, но и анализировать тренды на рынке, выявлять популярные товары и предугадывать спрос. Это дает возможность оперативно реагировать на изменения и адаптировать свою стратегию продаж.

С развитием технологий и появлением новых инструментов, таких как машинное обучение и искусственный интеллект, возможности парсинга будут только расширяться. Это позволит компаниям еще более эффективно использовать данные для принятия стратегических решений.

Таким образом, парсинг – это не только технический процесс, но и мощный инструмент, который может значительно облегчить жизнь как бизнесу, так и обычным пользователям.

Список литературы:

1. Митчелл Р. Современный скрапинг веб-сайтов с помощью Python. Питер. 2021. 336 с.
2. Рашка С. Python и машинное обучение. Пер. с англ. А. В. Логунова. М.: ДМК Пресс. 2017г., 420с.

3. Шарден Б., Массарон Л., Боскетти А. Крупномасштабное машинное обучение вместе с Python / Пер. с англ. А.В. Логунова. М.: ДМК Пресс. 2018 г. 358с.

4. Документация библиотеки BeautifulSoup // Crummy – URL: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

UDC 004.62

USING PARSERS AS A MEANS OF TRACKING CURRENT PRICES ON MARKETPLACES

Mikhail Ev. Filipchik

student

miklfil@mail.ru

Natalya V. Kartechina

candidate of agricultural sciences, associate professor

kartechnatali@mail.ru

Larisa Iv. Nikonorova

candidate of agricultural sciences, associate professor

lenaniknrva@rambler.ru

Michurinsk State Agrarian University

Michurinsk, Russia

Abstract. Using parsers to collect and analyze price data on online trading platforms, including the marketplace, allows you to automate the process of tracking current prices, analyze market trends and control pricing. The Python programming language and the BeautifulSoup library were used to collect and process web page data.

Keywords: parsing, marketplace, price monitoring, search automation, price comparison, data collection.

Статья поступила в редакцию 11.11.2024; одобрена после рецензирования 20.12.2024; принята к публикации 25.12.2024.

The article was submitted 11.11.2024; approved after reviewing 20.12.2024; accepted for publication 25.12.2024.