

Збір та обробка історичних даних для системи прийняття рішень

Вінницький національний технічний університет

Анотація

В даному документі представлено реалізацію збору та обробки даних для прийняття рішень. Обґрунтовується необхідність використання веб-скрапінгу для завантаження історичних даних.

Ключові слова: трейдинг, системи прийняття рішень, аналіз, історичні дані, задачі.

Abstract

This paper presents the implementation of data collection and processing for decision-making. The need to use web scrapping to download historical data is substantiated.

Keywords: trading, decision-making systems, analysis, historical data, tasks.

Актуальність

Прогнозування фінансових часових рядів - необхідний елемент будь-якої інвестиційної діяльності. Сама ідея інвестицій - вкладення грошей зараз з метою одержання доходу в майбутньому - ґрунтується на ідеї прогнозування майбутнього. Відповідно, прогнозування фінансових історичних даних лежить в основі діяльності всієї індустрії інвестицій - усіх бірж і небіржових систем торгівлі цінними паперами [1].

Проблемами аналізу та прогнозування займалися декілька століть назад, їм будуть приділяти увагу і надалі, оскільки хто володіє кращим інструментом аналізу, той має перевагу над іншими, а отже, буде мати й очікуваний прибуток [2].

З появою комп'ютерів і сучасних інформаційних систем стали бурхливо розвиватися методи технічного аналізу, що використовують фільтрацію чи математичну апроксимацію [3]. Через відносну простоту їхнього обчислення і легкість інтеграції в інформаційні системи було створено сотні технічних індикаторів. І усі вони, починаючи з Moving Average, RSI, MACD, Momentum, Stochastics, є цифровими фільтрами, тому що змінюють частотний спектр сигналу в деякому напрямку [1].

Реалізація пошуку історичних даних з допомогою веб-скрапінгу

Оскільки об'єм даних постійно збільшується, доцільним стає використання програмного забезпечення, яке буде автоматизовувати та покращувати процес збору даних, використовуючи технічний та фундаментальний аналіз, для їх обробки та фільтрації. Виходячи з цього перевагою такого програмного забезпечення буде швидкість, зручність та точність обробки даних, а також можливість для коригування вхідних та вихідних даних, для більш детального огляду отриманих даних. Одним із популярних методів збору інформації є веб-скрапінг.

Веб скрапінг - це процес програмного збору інформації з інтернету. Використовується для отримання доступу до World Wide Web через використання протоколу передачі гіпертексту (http), або через веб-переглядач. З допомогою веб скрапінгу можна провести послідовний синтаксичний аналіз інформації, розміщеної на інтернет сторінках.

Основні мови програмування, що допомагають в розробці програмного забезпечення для вебскрапінгу:

- Python
- Php
- Java
- C#

Для розробки програмного забезпечення для веб-скрапінгу часто використовують Python, так як він має високу швидкість та підтримується широким набором бібліотек для обробки та запису даних : - BeautifulSoup

- mySql.connector
- requests

Beautiful Soup - бібліотека Python яка обробляє дані з HTML і XML файлів. Вона працює з парсером, щоб забезпечити ідіоматичні способи навігації, пошуку та модифікації дерева синтаксичного розбору.

MySql.connector – бібліотека Python, яка дозволяє підключатись до бази даних та заповнити її даними

Прийняття рішень на основі історичних даних

Оскільки веб-скрапінг передбачає завантаження даних з веб сторінок, необхідна база даних (БД), яка зберігатиме історичні цінові значення різних стоків.

В основу прийняття рішень закладені скрипти, які визначають ту чи іншу модель поведінки ціни на графіку: проста чаша, перевернута чаша, потрійна вершина і т.д. Після перебору всіх історичних даних всіх стоків в базі даних за допомогою скриптів ідентифікуються точки, що сигналізують про відкриття позиції на buy чи sell. Дані точки, що означають певну модель поведінки ціни, допоможуть трейдеру приймати рішення самому – шляхом проведення аналізу графіків та перевірки даних, або ж використовувати автоматичний режим, який, хоч і матиме меншу прибутковість, проте надаватиме більше часу для оцінки інших потенційних «виграшних» стоків.

При правильній розробці програмного забезпечення, можна забезпечити подвійну швидкість завантаження даних в БД та більш точне і вчасне виявлення шуканих моделей на графіках.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Фундаментальний аналіз [Електронний ресурс]. Режим доступу: https://uk.wikipedia.org/wiki/%D0%A4%D1%83%D0%BD%D0%B4%D0%B0%D0%BC%D0%B5%D0%BD%D1%82%D0%B0%D0%BB%D1%8C%D0%BD%D0%B8%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D1%96%D0%B7 – назва з екрану.
2. Фігури технічного аналізу [Електронний ресурс]. Режим доступу: <http://forextrading.org.ua/figuri-tehnichngo-analizu.html> - назва з екрану.
3. Інтернет трейдинг в Україні [Електронний ресурс]. Режим доступу: <https://uk.wikipedia.org/wiki/%D0%86%D0%BD%D1%82%D0%B5%D1%80%D0%BD%D0%B5%D1%82-%D1%82%D1%80%D0%B5%D0%B9%D0%B4%D0%B8%D0%BD%D0%B3>

Денис Анатолійович Ткачик — студент групи 2СІ-146, факультет комп'ютерних систем та автоматики, Вінницький національний технічний університет, Вінниця, e-mail: true.tkachyk@gmail.com;

Науковий керівник: **Кветний Роман Наумович** — д-р. техн. наук, професор, завідувач кафедри АІВТ, Вінницький національний технічний університет, м. Вінниця.

Denys A. Tkachyk — Department Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia, email: true.tkachyk@gmail.com;

Supervisor: **Kvyetnyu Roman N.** — Dr. Sc. (Eng.), Professor, Head of the Chair of Automation and Information Measuring Devices, Vinnytsia National Technical University, Vinnytsia, email: rkvetny@mail.ru.