

РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ЗБОРУ ФІНАНСОВИХ ІСТОРИЧНИХ ДАНИХ З ВИКОРИСТАННЯМ ВЕБ-СКРАПІНГУ

Вінницький національний технічний університет

Анотація

Запропоновано метод збору інформації для аналізу фінансових історичних даних за допомогою web-scraping для подальшого їх прогнозування і використання у торгових системах

Ключові слова: програмування, веб-скрапінг, технічний аналіз, фундаментальний аналіз, історичні дані, фінанси, торгові системи.

Abstract

The method for collecting and analysis of historical data through web-scraping and their subsequent use in forecasting and trading systems was proposed

Keywords: programming, web scraping, technical analysis, fundamental analysis, historical data, finance, trading system.

З розвитком інтернет-трейдингу виникає потреба в створенні та оптимізації різних торгових систем для більш ефективної роботи з цінними паперами та фінансовими даними. Це зумовлено тим, що трейдери намагаються зафіксувати якомога більший прибуток від своїх інвестицій та прогнозувати подальший рух ринку для відкриття нових позицій в майбутньому.

Використовуючи фундаментальний та технічний аналіз трейдери створюють моделі які будуть прогнозувати поведінку ринку та надавати актуальну інформацію про події та зміни в ньому.

Технічний аналіз — прогнозування зміни ціни в майбутньому на основі аналізу зміни ціни у минулому. У його основі лежить аналіз часових рядів ціни — «чартів». Окрім цінових рядів, в технічному аналізі використовується інформація про об'єми торгів та інші статистичні дані.

Фундаментальний аналіз — підхід до аналізу фінансових ринків на основі вивчення фінансово-економічної інформації, яка, імовірно, впливає на динаміку активу або фінансового інструменту. Особливістю фундаментального аналізу є те, що його украй складно формалізувати.

Оскільки об'єм даних постійно збільшується, доцільним стає використання програмного забезпечення, яке буде автоматизувати 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 [1].

Основні мови програмування, що допомагають в розробці програмного забезпечення для веб-скрапінгу:

- 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 [3]:

- 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 [2].

Requests – бібліотека Python, яка дозволяє відправляти http-запити на сервер.

Такі можливості Python дозволяють швидко проаналізувати відповідну веб-сторінку та зібрати з неї потрібні дані, занести їх в базу даних, та використовувати для подальшої обробки.

Завдяки таким програмним засобам, трейдери матимуть можливість без ризику працювати зі своїми активами та збільшувати прибуток. За рахунок аналізу отриманих даних можна створити нові моделі та прогнозувати ринок більш успішно. Можливість оптимізувати критерії для фільтрації даних покращить роботу із таким програмним забезпеченням.

В результаті роботи було представлено метод збору фінансових історичних даних, використовуючи веб-скрапінг. Описано переваги даного методу та практичне застосування серед потенційних споживачів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ/REFERENCES

1. Web scraping [Електронний ресурс]/ «Wikipedia» – Електрон. текстові данні -,2017. – Режим доступу: https://uk.wikipedia.org/wiki/Web_scraping, вільний.
2. McKinney W. Python for Data Analysis/ Wes McKinney. – «O'Reilly Media», 2012 – 466 p.
3. Lutz M. Programming Python, 4th Edition/ Mark Lutz. – «O'Reilly Media», 2010 – 1632 p.

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

Науковий керівник: **Володимир Юрійович Коцюбинський** — к.т.н., доцент, Вінницький національний технічний університет, м. Вінниця.

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

Supervisor: **Volodymyr Y. Kotsiubynskyi**— Ph.D., Vinnytsia National Technical University, Vinnytsia.