

Abstract

Amazon Web Scraper

Synopsis: This project develops an Amazon web scraper to automate the extraction of key product details such as titles, prices, ratings, reviews, and availability. Web scraping is an essential technique in e-commerce analytics, providing insights into market trends and competitive intelligence.

Objective: To design a reliable and efficient web scraper for Amazon that automates data extraction while adhering to ethical and legal guidelines. The collected data will be used for price comparison, trend analysis, and business intelligence.

Methodology: The scraper is built using Python, employing libraries such as BeautifulSoup and Scrapy for parsing HTML content. Selenium is incorporated to handle dynamic content and JavaScript-based elements. Proxy rotation and user-agent switching techniques enhance anonymity and prevent detection. Extracted data is structured and stored in a database for easy querying and analysis.

Expected Outcome: A robust web scraping solution capable of efficiently extracting and structuring Amazon product data. The scraper will provide accurate insights into pricing trends, customer preferences, and market competition, facilitating informed decision-making.

Impact: This project provides a foundation for data-driven decision-making in e-commerce, benefiting businesses and consumers alike. By offering real-time insights, it aids businesses in pricing strategies and inventory management while enabling consumers to make informed purchasing decisions. Future enhancements include integrating machine learning models for price prediction and sentiment analysis, further extending the scraper's analytical capabilities.

Name: - Ronak Kumar Bhambu

B tech Ai and Data Science.