amazon-scraper Documentation

Release 0.1

picorana

Nov 04, 2017

Contents:

1 Indices and tables 3

Python Module Index 5

class app.**AmazonScraper**(***kwargs*)

Hello, this is AmazonScraper!

static **get_logger**(*level=10*, *verbose=False*)

Returns a logger

get_proxies()

Retrieves a list of proxies

static **parse_asins_from_file**(*path*)

Reads a list of asins from a file

retrieve_page(*asin*)

Requests the main product page, saves it, and returns an url for the reviews

retrieve_questions(*asin*)

Scrapes the questions pages and returns a list of dicts containing questions and respecitive answers

retrieve_reviews(*main_reviews_url*, *review_pages_number*)

Collects the review given the url of the reviews page

scrape()

Manages the whole scraping process

scrape_page_questions(*asin*, *page_num*, *fails*, *results*)

threads to request and scrape a single questions page

scrape_page_reviews(*url*, *fails*, *results*)

threads to request and scrape a single review page

CHAPTER 1

Indices and tables

- genindex
- modindex
- search

Python Module Index

a

app, 1

Index

A

AmazonScraper (class in app), 1
app (module), 1

G

get_logger() (app.AmazonScraper static method), 1
get_proxies() (app.AmazonScraper method), 1

P

parse_asins_from_file() (app.AmazonScraper static method), 1

R

retrieve_page() (app.AmazonScraper method), 1
retrieve_questions() (app.AmazonScraper method), 1
retrieve_reviews() (app.AmazonScraper method), 1

S

scrape() (app.AmazonScraper method), 1
scrape_page_questions() (app.AmazonScraper method), 1
scrape_page_reviews() (app.AmazonScraper method), 1