

Analisis Harga Barang Pada E-Marketplace Dengan Pengambilan Informasi Berupa Url Berbasis Web Crawling Menggunakan BeautifulSoup Dan Django

Analyze The Price Of Goods On The E-Marketplace With Information Retrieval In The Form Of Web-Based Urls Crawling Using BeautifulSoup And Django

1st Novaldi Ibnu Syahban M
Fakultas Ilmu Terapan
Universitas Telkom
Bandung, Indonesia
Novaldiibnusyahbanm@student
.telkomuniversity.ac.id

2nd Indrarini Dyah Irawati
Fakultas Ilmu Terapan
Universitas Telkom
Bandung, Indonesia
indrarini@telkomuniversity.ac.id

3rd Muhammad Iqbal
Fakultas Ilmu Terapan
Universitas Telkom
Bandung, Indonesia
muhammadiqbaaall@elkomuniver
sity.ac.id

Abstrak—*Web Crawling* merupakan sebuah *search engine* yang digunakan dalam menelusuri halaman website yang memungkinkan seseorang untuk mendapatkan data yang dia inginkan pada suatu website tertentu, salah satunya adalah harga barang di E- marketplace. Permasalahan yang seringkali terjadi adalah harga barang yang tidak akurat (Informasi), data barang yang tidak lengkap. Tujuan dalam pembuatan tugas akhir ini adalah dengan membuat sebuah website yang dapat mempermudah melakukan pencari yang di lakukan konsumen dengan mendapatkan informasi pencarian harga barang dengan harga yang akurat. Metode pengumpulan data dilakukan dengan Teknik *Web Crawling* karena tentunya konsumen tidak dapat mengakses ke database masing masing E-Commerce. Hasil dari penelitian ini adalah sebuah Search engine web Crawling yang dapat mengumpulkan data barang barang di Tokopedia, Bukalapak dan Lazada serta sebuah aplikasi search engine dalam mencari informasi harga barang termurah dari sumber data yang telah dikumpulkan dengan menggunakan web crawling ini. Dengan adanya aplikasi ini yang dibuat untuk memudahkan konsumen dalam melakukan pengumpulan informasi dari website e-marketplace yang kemudian dapat dimonitoring harga tiap barangnya. Kemudian informasi ini pun nantinya akan dapat diolah sesuai dengan tujuan awal

dilakukannya web crawling yaitu untuk mempermudah dalam menganalisa harga barang dan kebutuhan konsumen.

Kata kunci—search engine, web crawling, e-marketplace.

I. PENDAHULUAN

Perkembangan di bidang teknologi informasi sekarang ini telah mengalami kemajuan secara pesat, hal ini tidak terlepas dari keberadaan internet. Internet bermanfaat bagi aktivitas kehidupan, salah satunya dalam dunia bisnis. E-marketplace adalah sebuah sistem informasi antar organisasi dimana pembeli dan penjual di pasar mengkomunikasikan informasi tentang harga, produk dan mampu menyelesaikan transaksi melalui saluran komunikasi elektronik. Dengan melalui internet para penjual dimudahkan dalam melakukan promosi dan memasarkan produknya dalam jangkauan yang lebih luas [1]. Salah satu jenis teknologi yang saat ini sudah banyak digunakan untuk mempermudah proses analisa harga barang dan kebutuhan konsumen dengan adanya e-marketplace.

Dengan adanya e-marketplace maka konsumen dapat mempermudah menerima informasi tentang harga barang dan kebutuhan lain dari konsumen dengan di bantu menggunakan web crawling. Salah satu contoh adalah Lazada, Tokopedia, Bukalapak. Dari beberapa E-Commerce ini memiliki beberapa keunggulan yang akan mempermudah konsumen dalam melakukan proses jual beli barang, serta konsumen pun dapat mengetahui informasi harga barang dan kebutuhan konsumen lainnya. Sehingga nantinya E-marketplace akan membantu memberikan informasi antar organisasi dimana pembeli dan penjual di pasar mengkomunikasikan informasi tentang harga, produk dan mampu menyelesaikan transaksi melalui saluran komunikasi elektronik. Suatu e-marketplace merepresentasikan suatu struktur sosial, konsep ekonomi pasar, dan penggunaan teknologi. E-marketplace dapat memberikan peluang untuk melakukan bisnis dan melaksanakan transaksi melalui saluran elektronik, biasanya pada platform yang berbasis internet [1].

II. DASAR TEORI

A. Web Crawling

Web Crawling adalah proses di mana *search engine* menemukan konten yang di- update di sebuah situs atau halaman baru, perubahan situs, atau link yang mati. *Web Crawling* juga merupakan suatu proses di mana mesin pencari mengirimkan tim robot (crawler atau spider) untuk menemukan konten-konten baru dan konten yang telah di- update. Ketika ada pengguna yang mencari sebuah konten di search engine dengan keyword tertentu, search engine akan mencarinya di indeks dan menentukan konten mana yang paling sesuai untuk pengguna tersebut. Proses web crawling tidak dapat dilakukan secara manual. Ada beragam pilihan tools yang harus digunakan. Tools untuk web crawling tersebut adalah web crawler yang sering juga disebut sebagai web robot atau web spider.

1. Etika Web Crawler

Proses crawling secara terus-menerus dapat menyebabkan beban berlebih pada server suatu web, bukan hanyadari proses penelusuran link tapi juga dari proses pengun- duhan halaman web yang bisa berjumlah ratusan. Selainitu pada beberapa web terdapat bagian web yang diharaptidak dimasuki proses crawling. Untuk mengatasi masalah-masalah tersebut, dibuatlah suatu protokol yang disebut Robot Exclusion Protocol [4]. Protokol pertama kali dia- jukan oleh anggota mailing list robot di tahun

1994 dan ke-mudian menjadi Internet Draft di tahun 1997. Protokol ini sampai saat ini tidak memiliki standar resmi untuk struk-turnya, sehingga protokol ini menjadi semacam aturan ti-dak tertulis untuk proses crawling.

2. Algoritma Web Crawler

Konsep dari *algoritma Breadth-First Crawler* yang di-pakai dalam proses crawling di makalah ini mengecek se-tiap link dalam suatu halaman web sebelum berpindah kehalaman lain [5]. Breadth-First Crawler akan menelusuri setiap link pada halaman pertama, lalu menelusuri setiap link pada halaman dari link pertama di halaman pertama, dan seterusnya. Penelusuran dilakukan sampai tidak adalagi link baru yang dapat ditelusuri atau jika jumlah hala-man yang ditelusuri sudah mencapai batas maksimal yang ditentukan. Algoritma Depth-First Crawler memiliki ke-balikan proses dari Breadth-First Crawler. Depth-First Crawler menelusuri semua kemungkinan jalur dari suatu link sampai mencapai suatu dasar sebelum melanjutkan penelusuran ke link berikutnya.

B. Search Engine

Search engine atau apabila diartikan dalam bahasa Indonesia adalah mesin pencari pada dasarnya merupakan program berbasis web yang diperuntukkan untuk mencari informasi di dalam *World Wide Web (www)*. Pencarian informasi melalui search engine bisa didapatkan dengan menyesuaikan dengan kata kunci yang pengguna masukkan.

Teknologi dari search engine dapat memberikan informasi yang diinginkan pengguna dengan daftar pencarian terbaik yang tersedia. Proses menghasilkan informasi tersebut biasa disebut sebagai SERP atau *search engine result page*. Sejak pertama kali diciptakan, banyak sekali search engine yang ada di dunia saat ini, salah satu yang sangat terkenal tentu saja Google. Pengguna yang ingin menggunakan Google bisa melalui berbagai perangkat yang dimiliki, dari browser yang ada smartphone, tablet, komputer, dan berbagai perangkat browser lainnya.

1. Fungsi yang Dimiliki Search Engine

Search engine pada dasarnya memiliki fungsi utama sebagai alat untuk menyediakan informasi bagi semua orang. Ketika hampir semua orang menggunakan mesin pencarian, lantas apa sebenarnya fungsi yang ditawarkan oleh search engine itu sendiri. Ketika menggunakan search engine, pengguna yang ingin atau membutuhkan suatu informasi hanya perlu masukkan kata kunci dalam sistem mesin pencarian. Selanjutnya, berbagai daftar web yang berkaitan dengan kata yang dimasukkan akan ditampilkan kepada pengguna. Langkah seperti biasa disebut dalam dunia komputer sebagai crawling atau proses mengumpulkan data atau mengindeks.

2. BeautifulSoup

BeautifulSoup adalah sebuah Python database yang berdasarkan penemuan mesin analisa HTML dan XML, digunakan untuk melakukan ekstraksi, analisa, dan mengubah informasi di dalam pohon DOM dari sebuah situs (Chunmei, Guomei, &

Zuojie, 2015). BeautifulSoup akan menganalisa semua dokumen yang diberikan, termasuk semua jenis dokumen HTML dan XML. Gambar 2.1 merupakan salah satu contoh penggunaan fungsi yang disediakan oleh BeautifulSoup.

```
from bs4 import BeautifulSoup
import csv

soup = BeautifulSoup(open("43rd-congress.html"), features="lxml")

final_link = soup.p.a
final_link.decompose()

f = csv.writer(open("43rd_Congress.csv", "w"))
f.writerow(["Name", "Link"]) # Write column headers as the first line

links = soup.find_all('a')
for link in links:
    names = link.contents[0]
    fullLink = link.get('href')
    f.writerow([names, fullLink])
```

C. Phyton

Python adalah Bahasa pemrograman interpretatif multiguna dengan filosofi perancangan yang berfokus pada tingkat keterbacaan kode. Python diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif. Python juga didukung oleh komunitas yang besar.

D. Django

Django adalah web framework Python yang didesain untuk membuat aplikasi web yang dinamis, kaya fitur dan aman. Django yang dikembangkan oleh Django Software Foundation terus mendapatkan perbaikan sehingga membuat web framework yang satu ini menjadi pilihan utama bagi banyak pengembang aplikasi web.

Django bertujuan untuk memudahkan pengembangan situs web dan basis data yang kompleks (Django Software Foundation, 2017). Salah satu keunggulan django adalah, framework ini telah merepresentasikan ORM (Object Relational Mapper) sehingga tidak perlu menyesuaikan query jika terjadi perubahan database yang digunakan. Berikut ini adalah beberapa keuntungan menggunakan Django:

1. Object-Relational Mapping (ORM) Support.
2. Automatic Admin Interface.
3. Elegant URL Design.
4. Template System.
5. Cache system.
6. Internationalization.
7. A light weight web server for development and test.

E. E-Marketplace

E-Marketplace merupakan media online berbasis internet (web based) tempat melakukan kegiatan bisnis dan transaksi antara pembeli dan penjual. Pembeli dapat mencari supplier sebanyak mungkin

dengan kriteria yang diinginkan, sehingga memperoleh sesuai harga pasar. Sedangkan bagi supplier/penjual dapat mengetahui perusahaan-perusahaan yang membutuhkan produk/jasa mereka.

E-MarketPlace merupakan model E-Business yang berhubungan dengan penjual dan pembeli (seller & buyer). E-MarketPlace di Indonesia merupakan salah satu media penggerak ekonomi nasional dalam rangka menghadapi era globalisasi.

F. Framework

Framework adalah suatu alat yang digunakan untuk keperluan pengembangan suatu software atau aplikasi. Jadi dengan menggunakan framework, maka para developer ataupun programmer sudah tidak perlu lagi membuat kodingan dari awal lagi. Di dalam framework juga sudah tersedia berbagai dasar kode yang diperlukan untuk membuat suatu aplikasi. Singkatnya, Anda bisa menganggap framework sebagai suatu template atau format dalam membuat suatu aplikasi. Selain itu, banyak yang bilang framework sama seperti API atau application programming interface.

G. CRUD

Crud adalah aspek penting dalam SQL. Crud merupakan sebuah singkatan dari create, read, update serta delete. Dalam konsep interface atau tampilan antar muka, Crud adalah fasilitator terkait dengan tampilan dari pencarian, perubahan informasi dalam bentuk tabel, laporan, atau formulir.

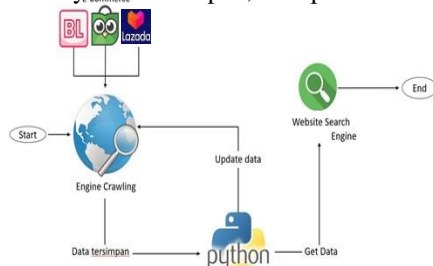
H. E-Commerce

Menurut Kotler & Amstrong (2012) E-commerce adalah saluran online yang dapat dijangkau seseorang melalui komputer, yang digunakan oleh pebisnis dalam melakukan aktifitas bisnisnya dan digunakan konsumen untuk mendapatkan informasi dengan menggunakan bantuan komputer yang dalam prosesnya diawali dengan memberi jasa informasi pada konsumen dalam penentuan pilihan. Menurut Wong (2010) e-commerce adalah proses jual beli dan memasarkan barang serta jasa melalui sistem elektronik, seperti radio, televisi dan jaringan komputer atau internet. Maka dapat disimpulkan bahwa e-commerce merupakan kumpulan dinamis antara teknologi, aplikasi dan proses bisnis yang menghubungkan perusahaan dan konsumen serta komunitas tertentu dimana pertukaran barang antara pengecer dan konsumen dari berbagai komoditi dalam skala luas dan suatu transaksi elektronik, dan dalam proses pengiriman barang dari pengecer menggunakan transportasi dari suatu wilayah ke wilayah lain hingga sampai ke tangan konsumen dan hubungan yang terjadi adalah hubungan yang saling menguntungkan kedua belah pihak.

III. PERANCANGAN SISTEM

A. Blok Diagram Sistem

Berikut merupakan penjelasan dari perancangan sistem dengan Pengambilan informasi berupa URL berbasis web crawling menggunakan BeautifulSoup dan Django pada E-marketplace. Sistem ini menganalisa harga barang/produk dari beberapa E-commerce. Dari E-commerce yang di telah ambil untuk percobaan pada proyek akhir ini yaitu Bukalapak, Tokopedia dan Lazada. Dapat kita lihat pada gambar 3.1 diatas bahwa pertama kita memulai untuk melakukan proses di mana *engine crawling* akan menemukan konten yang di-update di sebuah situs atau halaman baru, perubahan situs, atau link yang mati. Kemudian data akan tersimpan berupa data harga yang akan di ambil dari beberapa E-commerce yaitu Bukalapak, Tokopedia dan Lazada.



Gambar 3. 1 Sistem Analisa harga barang

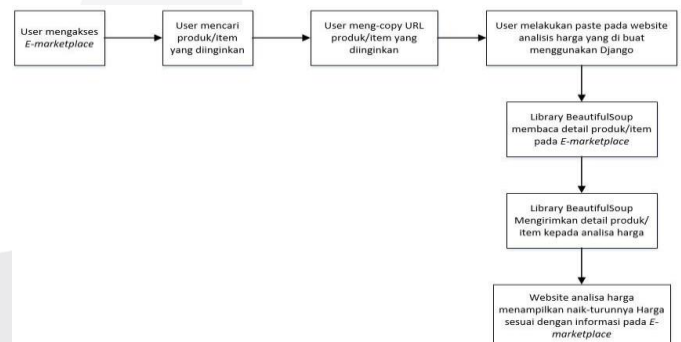
Dari data yang tersimpan dari beberapa E-commerce ini maka data akan diolah oleh Bahasa pemrograman Python. Setelah data di olah oleh Bahasa pemrograman python maka data yang diinginkan berupa data harga barang/produk dapat di akses oleh website. Namun jika data akan di update maka data akan di olah Kembali oleh search crawling atau web crawling kemudian akan di teruskan Kembali menggunakan Bahasa python. Selanjutnya data dapat diakses oleh konsumen untuk mencari harga barang yang diinginkan dari E-commerce yang dituju. Setelah itu proses akan selesai.

Setelah itu dari data harga barang yang telah tersimpan berupa website. Maka dapat kita lihat dari Diagram Penelitian Web Crawling pada gambar 3.2 di bawah ini yang menjelaskan bagaimana Bahasa python mengolah data menggunakan Library BeautifulSoup dan Django.

Gambar 3. 2 Diagram Penelitian Web Crawling

Pada bab ini akan dijelaskan mengenai perancangan sistem analisis harga barang pada e-marketplace dengan pengambilan informasi berupa URL berbasis web crawling menggunakan BeautifulSoup dan Django.

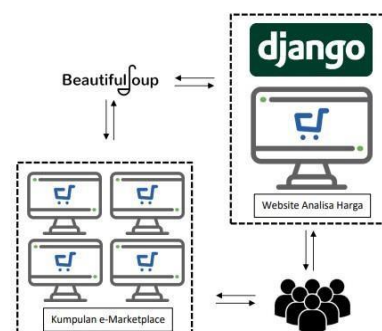
BeautifulSoup merupakan library dari bahasa pemrograman Python yang dapat digunakan untuk melakukan web crawling dengan cara mengidentifikasi struktur dari website yang menjadi target crawling dan kemudian membaca isi dari website tersebut sesuai dengan struktur yang telah diidentifikasi sebelumnya. Django merupakan framework berbasis web dari bahasa pemrograman Python yang nantinya akan digunakan sebagai pengaplikasian untuk menjalankan serta menampilkan data crawling dengan sistem CRUD. Setelah crawling selesai dilakukan, User dapat mengamati perbandingan dan naik turunnya harga pada tiap e-marketplace hanya pada satu tempat. Jika digambarkan dalam bentuk blok diagram, maka model sistem analisis harga barang pada e-marketplace dengan pengambilan informasi berupa URL berbasis web crawling menggunakan BeautifulSoup dan Django dapat dilihat pada Gambar 3.3 di bawah ini.

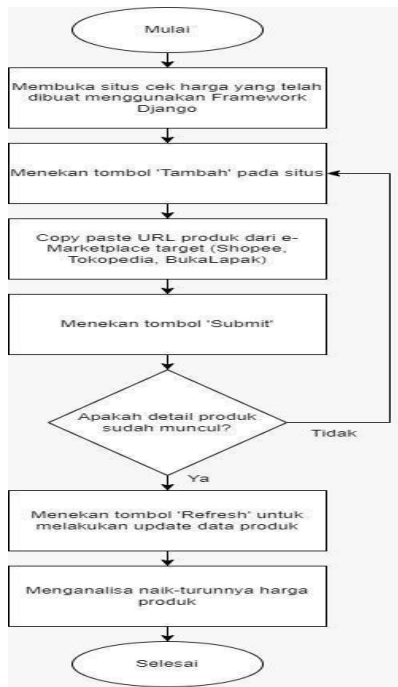


Gambar 3. 3 Blok Diagram Penelitian Web Crawling

B. Flowchart Perancangan Sistem

Berikut merupakan diagram alur sistem pengambilan informasi berupa URL berbasis web crawling menggunakan BeautifulSoup dan Django.





Gambar 3. 2 Flowchart Sistem

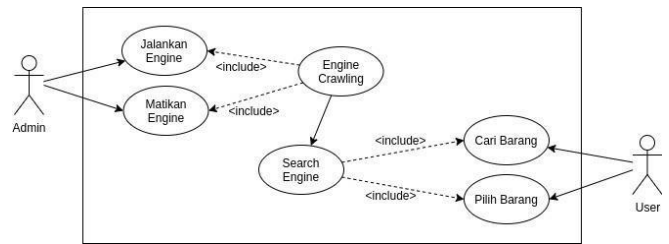
Pada Gambar 3.2 dijelaskan bahwa Sistem menganalisa harga barang/produk pada beberapa E-commerce yaitu Bukalapak, Tokopedia dan Lazada ini dimulai dengan membuka situs cek harga yang telah di buat menggunakan framework Django. Selanjutnya, sistem akan menekan tombol 'Tambah' pada situs. Setelah itu konsumen dapat meng-copy paste URL produk dari E-marketplace target (Bukalapak, Tokopedia dan Lazada). Kemudian menekan tombol 'Submit' yang ada pada website yang telah tersedia. Setelah itu, jika detail produk sudah muncul maka proses akan dilanjutkan dengan menekan tombol refresh untuk melakukan update data produk dan website akan menganalisa naik-turunnya harga produk.

Dan Jika proses detail produk tidak muncul maka proses akan diulang Kembali ke menekan tombol 'tambah' pada situs, kemudian meng-copy paste URL produk dari E-marketplace target (Bukalapak, Tokopedia dan Lazada dan selanjutnya menekan tombol 'submit'. Jika proses detail produk sudah muncul maka proses sudah berhasil dan bisa menekan tombol refresh untuk melakukan update data produk dan website akan menganalisa naik-turunnya harga produk. Dan proses selesai, selanjutnya website dapat di akses oleh konsumen untuk mencari informasi harga barang/produk.

C. Perancangan Sistem

Sistem ini membuat dua aplikasi, Satu buah aplikasi web crawling untuk mengumpulkan data barang dan satu buah aplikasi search engine untuk digunakan user dalam melakukan pencarian barang seperti yang telah dijelaskan pada sistem usulan.

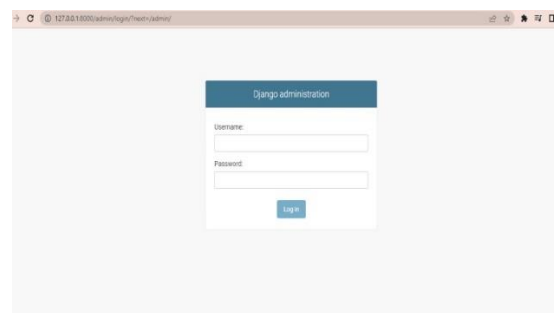
Adapun fase perancangan sistem penulis menggunakan tools Unified Modeling Language (UML) yang didalamnya terdapat usecase diagram, activity diagram, sequence diagram dan class diagram.



Gambar 3. 3 Usecase Diagram Sistem

D. Desain Antarmuka

Desain yang digunakan pada web crawling adalah dengan menggunakan bahasa pemrograman python tanpa menggunakan database. Kemudian dibantu dengan library yaitu BeautifulSoup yang berfungsi untuk ekstraksi file dengan format XML atau HTML agar bisa membaca stuktur data pada e-commerce dan Django sebagai framework aplikasi website pada pemrograman python, Berikut tampilan desain pada web crawling yang digunakan :

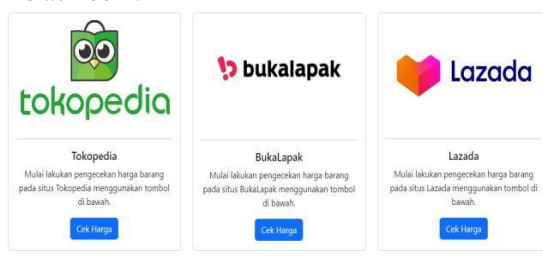


Gambar 3. 4 Halaman login website Admin Django

IV. PENGUJIAN DAN ANALISIS

A. Pengujian Sistem

Pengujian sistem dilakukan untuk mengetahui apakah sistem sudah berjalan sesuai dengan sasaran pembuatan aplikasi. Pengujian sistem menggunakan pengujian black box testing. Pengujian dilakukan oleh Novaldi Ibnu Syahban M selaku pengguna ingin membeli barang online. Berikut adalah hasil pengujian sistem dengan menggunakan metode black box :

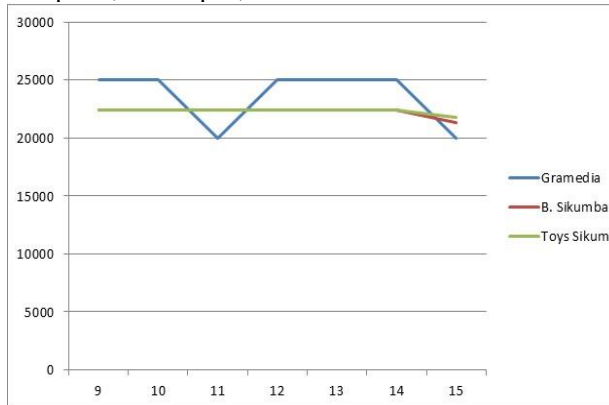


D3 Teknologi Telekomunikasi | Universitas Telkom

Tabel 4. 1 Tampilan awal web crawling

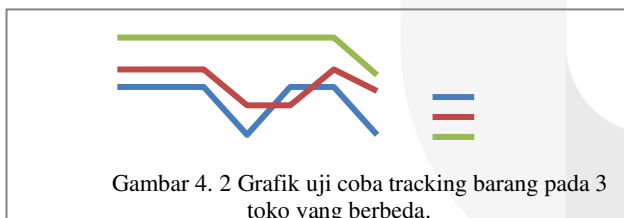
B. Pengujian Hasil

Setelah dilakukan pengujian pada web crawling, diperlukan hasil dan analisis dari barang yang ditracking harganya dari setiap E-marketplace. Untuk mengetahui apakah hasil data produk yang ditampilkan di web crawling sesuai atau tidak, kemudian kita bisa mengetahui harga termurah dari setiap E-marketplace menggunakan web crawling. Berikut adalah grafik uji coba tracking harga di tokopedia, bukalapak, lazada



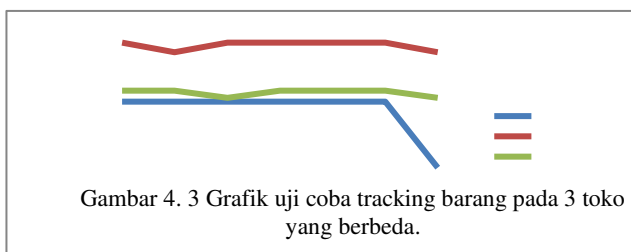
Gambar 4. 1 Grafik uji coba tracking barang pada 3 toko yang berbeda

Gambar diatas merupakan hasil grafik dari uji coba perbandingan 1 barang yaitu comic one piece vol 90 pada 3 toko yang berbeda dalam 1 e-commerce didapatkan harga termurah yaitu pada toko Gramedia dengan harga Rp. 20000.



Gambar 4. 2 Grafik uji coba tracking barang pada 3 toko yang berbeda.

Gambar diatas merupakan hasil grafik dari uji coba perbandingan 1 barang yaitu oxihom tempat beras kacang pada 3 toko yang berbeda dalam 1 e-commerce didapatkan harga termurah yaitu pada toko Oxihom dengan harga Rp. 40900.



Gambar 4. 3 Grafik uji coba tracking barang pada 3 toko yang berbeda.

Gambar diatas merupakan hasil grafik dari uji coba perbandingan 1 barang yaitu oxihom tempat beras kacang pada 3 toko yang berbeda dalam 1 e-commerce didapatkan harga termurah yaitu pada toko Oxihom dengan harga Rp. 78500.

V. KESIMPULAN

Berdasarkan hasil perancangan, pengujian dan analisa yang telah dilakukan maka dapat diambil beberapa kesimpulan sebagai berikut :

- A. Berdasarkan hasil pengujian fungsionalitas terhadap fitur-fitur pada sistem web crawling, seperti menampilkan produk dari E-marketplace, menampilkan harga barang, harga sebelumnya dan selisih harga, tingkat keberhasilan sistem web crawling telah berhasil digunakan dengan akurat. Dapat disimpulkan bahwa semua fungsi 100% berjalan dengan baik sebagaimana semestinya.
- B. Dari hasil pengujian dengan perbandingan satu produk pada 3 toko yang berbeda dalam satu E-marketplace dan perbandingan 1 produk pada 3 E-marketplace yang berbeda cukup efektif untuk bisa mendapatkan perbandingan harga termurah.
- C. Hasil analisis harga barang di 3 E-marketplace cukup signifikan perubahan harganya.

REFERENSI

- [1] Castrounis, A. (2017). Artificial Intelligence, Deep Learning, and Neural Networks, Explained. [online] Kdnuggets.com. Available, <http://www.kdnuggets.com/2016/10/artificial-intelligence-deep-learning-neural-networks-explained.html> [Accessed 1 Agustus.2022].
- [2] L. A. A. Mahmuddah, S. A. Wibowo and G. Budiman, "Analisis Performansi pada Pengambilan URL Berbasis Web Crawling dengan Menggunakan Teknologi Pengenalan Wajah YOLOv3," 2021.
- [3] F. Ridho, Rancang Bangun Aplikasi Web Crawling Untuk Mencari Harga Barang Termurah dari Berbagai e-Marketplace (Studi Kasus Tokopedia, Bukalapak, Shopee), Jakarta: Universitas Islam Negeri, 2020.
- [4] F. Ridho, E. Hanafi Fernando, H. Sagala and A. Elbert Budiman, "Ekstrasi dan Analisis Produk di Marketplace Secara Otomatis dengan Memanfaatkan Teknologi Web Crawling," 2019.
- [5] L. B. Ilmawan, "Membangun Web Crawler Berbasis Web Service untuk Data Crawling pada Website Google Play Store," 2018.
- [6] D. and E. Mailoa, "Implementasi Web Crawling Untuk Pencarian Harga Sparepart Pada PT Asuransi Sinar Mas," 2018.

- [7] I.Ramadhan and H. Sastramihardja,
"Pemanfaatan Web Crawler dalam
Mengumpulkan Informasi Melalui Internet,"
2018.
- [8] F. Suharno and L. Listiyoko, "Aplikasi Berbasis
Web dengan Metode Crawling sebagai Cara
Pengumpulan Data untuk Mengambil
Keputusan," 2018
- [9] R. Hanifah and I. S. Nurhasanah,
"Implementasi Web Crawling untuk
Mengumpulkan Informasi Wisata Kuliner di
Bandar Lampung.

