

Analisis Penjualan Handphone di Tokopedia dengan Teknik Web Scraping Menggunakan Python pada Google Colab

Rangga Gelar Guntara¹, Frans Budi Kashira², Teuku Khairul Amri³, Laras Bellaning Restu⁴, Fazry Rachman Susanto⁵
^{1,2,3,4,5}Universitas Pendidikan Indonesia
 E-mail: ranggagelar@upi.edu¹

Article History:

Received: 22 Februari 2024

Revised: 05 Maret 2024

Accepted: 07 Maret 2024

Keywords:

Analisis Penjualan, AI Life Cycle, Google Colab, Visualisasi Data

Abstract: Menganalisis penjualan handphone di Tokopedia bisa memberikan wawasan berharga kepada pemilik data, termasuk mengenai tren penjualan, faktor-faktor yang memengaruhi penjualan, dan potensi untuk meningkatkan penjualan. Dengan memahami pola penjualan, pemilik data memiliki kemampuan untuk meramalkan permintaan di masa depan. Melalui pemahaman terhadap faktor-faktor yang memengaruhi penjualan, pemilik data dapat mengambil langkah-langkah untuk meningkatkan volume penjualan. Selain itu, dengan kesadaran terhadap peluang peningkatan penjualan, pemilik data dapat merancang strategi yang efektif untuk meningkatkan performa penjualan. Penelitian ini bertujuan untuk mengembangkan panduan yang mendetail mengenai pendekatan yang dapat diambil oleh pemilik data dalam menganalisis penjualan handphone di platform Tokopedia. Panduan ini akan mencakup prosedur akuisisi data, metodologi analisis, dan cara menginterpretasi hasil. Signifikansi dari tujuan penelitian ini dapat dipahami melalui pemahaman bahwa analisis penjualan handphone memberikan wawasan berharga kepada pemilik data, termasuk identifikasi tren penjualan, faktor-faktor yang berpengaruh, dan potensi peningkatan penjualan. Metode yang digunakan mengadaptasi metode AI Life Cycle dengan teknik web scarping dengan bahasa pemrograman Python di lingkungan Google Colab. Hasil penelitian menunjukkan metode dapat menghasilkan visualisasi data sesuai dengan pengolahan data menggunakan teknik regresi linier.

PENDAHULUAN

Handphone merupakan salah satu produk elektronik yang paling populer di dunia. Menurut Badan Pusat Statistik (2023), nilai transaksi perdagangan elektronik di Indonesia pada tahun 2023 mencapai Rp. 540 triliun, meningkat 30% dari tahun sebelumnya (Megawanto et al., 2020). Handphone merupakan salah satu produk yang paling banyak dibeli secara online (Purba & Tafonao, 2023).

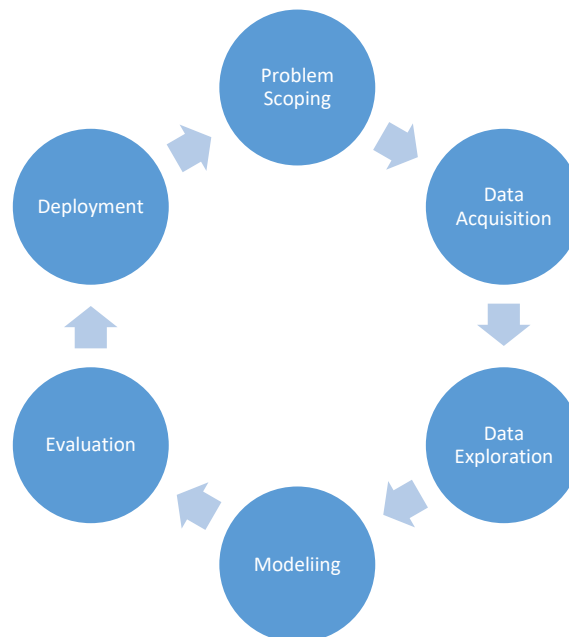
Di Indonesia, Tokopedia merupakan salah satu platform e-commerce terbesar yang menjual berbagai macam handphone (Kristiani & Gultom, 2023). Menurut Hootsuite (2023), jumlah pengguna internet di Indonesia pada tahun 2023 mencapai 220 juta orang, atau sekitar 73% dari total populasi (Alamsyah et al., 2023). Tokopedia merupakan platform e-commerce yang paling banyak digunakan di Indonesia, dengan pangsa pasar sebesar 48% menurut Statista (2023).

Analisis penjualan handphone di Tokopedia dapat memberikan informasi berharga bagi pemilik data, seperti tren penjualan, faktor-faktor yang mempengaruhi penjualan, dan peluang peningkatan penjualan. Dengan memahami tren penjualan, pemilik data dapat memprediksi permintaan di masa depan (Yusta Jernihati Gea & Kurniawan Sarototonafo Zai, 2023). Dengan memahami faktor-faktor yang mempengaruhi penjualan, pemilik data dapat mengambil tindakan untuk meningkatkan penjualan (Candra et al., 2020). Dengan memahami peluang peningkatan penjualan, pemilik data dapat mengembangkan strategi untuk meningkatkan penjualan (Asyifah et al., 2023).

Tujuan penelitian ini adalah untuk menyusun panduan tentang pendekatan pemilik data dalam analisis penjualan handphone di platform Tokopedia. Panduan ini akan mencakup langkah-langkah akuisisi data, metodologi analisis, dan interpretasi hasil. Tujuan penelitian ini penting karena: Analisis penjualan handphone dapat memberikan informasi berharga bagi pemilik data, seperti tren penjualan, faktor-faktor yang mempengaruhi penjualan, dan peluang peningkatan penjualan. Panduan ini akan membantu pemilik data untuk melakukan analisis penjualan handphone secara lebih efektif dan efisien, salah satunya dengan menggunakan metode web scraping (Rizquina & Ratnasari, 2023).

METODE PENELITIAN

Pada penelitian ini akan menggunakan metode siklus hidup aplikasi kecerdasan buatan (AI Life Cycle) (De Silva & Alahakoon, 2022). Adapun tahapan dari metode AI Life Cycle dapat dilihat pada gambar 1 AI Life Cycle berikut.



Gambar 1. AI Life Cycle

HASIL DAN PEMBAHASAN

1. Data Akuisisi

Langkah pertama dalam akuisisi data penjualan handphone di platform Tokopedia adalah mengidentifikasi URL kategori handphone. URL kategori handphone adalah alamat web yang digunakan untuk mengakses halaman kategori handphone di Tokopedia. URL kategori handphone dapat ditemukan dengan cara berikut:

- Buka situs web Tokopedia.
- Klik menu "Kategori".
- Pilih kategori "Handphone".
- Salin URL halaman kategori handphone.

URL kategori handphone di Tokopedia memiliki format sebagai berikut:
https://www.tokopedia.com/search?st=&q=handphone&srp_component_id=02.01.00.00&srp_page_id=&srp_page_title=&navsource=

Setelah mengidentifikasi URL kategori handphone, langkah selanjutnya adalah mengumpulkan data penjualan handphone dari situs web Tokopedia. Teknik web scraping dapat digunakan untuk mengumpulkan data dari situs web secara otomatis (Rangga Gelar Guntara, 2022).

```

1 from selenium import webdriver
2 from selenium.webdriver.chrome.service import Service
3 from webdriver_manager.chrome import ChromeDriverManager
4 from bs4 import BeautifulSoup
5 import pandas as pd
6 import time
7 import re
8
9 # Setup Selenium
10 options = webdriver.ChromeOptions()
11 # options.add_argument('--headless')
12 # path = '/Users/fransbudikashira/DATA/Kuliah/AL/Tokopedia/chromedriver'
13 # service = Service(executable_path=path)
14 # driver = webdriver.Chrome(service_log_path=service,
15 #                           options=options, executable_path=path)
16 driver = webdriver.Chrome(ChromeDriverManager().install())
17
18 tokped_link = "https://www.tokopedia.com/search?st=&q=handphone&srp_component_id=02.01.00.00&srp_page_id=&srp_page_title=&navsource="
19 driver.set_window_size(1360, 800)
20 driver.get(tokped_link)
21
22 rentang = 500
23 for i in range(1, 21):
24     akhir = rentang + i
25     perintah = f"window.scrollTo(0, {str(akhir)})"
26     driver.execute_script(perintah)
27     print(f>Loading ke-{str(i)}")
28     time.sleep(1)
29
30 time.sleep(5)
31 driver.save_screenshot("home.png")
32 content = driver.page_source
33 driver.quit()
34
35 data = BeautifulSoup(content, 'html.parser')
36 # print(data.encode("utf-8"))
37
38 i = 1
39 search_brands = ["Oppo", "Apple", "Infinix", "Xiaomi",
40                  "Samsung", "Realme", "Nokia", "Vivo", "Asus", "Sony",
41                  "Huawei", "OnePlus", "LG", "Advan"]
42
43 # Huawei, "OnePlus", "LG", "Advan"]
44 # list_name, list_image, list_price, list_link,
45 # list_sold_out, list_location = [], [], [], [], []
46 list_name, list_price, list_sold_out, list_location,
47 list_brand = [], [], [], [], []
48
49 for area in data.find_all('div', class_='css-llwoba'):
50     print('Proses data ke-' + str(i))
51     name = area.find('div',
52                     class_='prd-link-product-name').get_text()
53     for s_b in search_brands:
54         if re.search(s_b, name.title()):
55             brand = s_b
56             break
57     elif re.search("iphone", name.lower()):
58         brand = "Apple"
59         break
60     # image = area.find('img')['src']
61     price = area.find('div',
62                     class_='prd-link-product-price').get_text()
63     if "-" in price:
64         # Format rentang harga
65         nilai_harga = price.replace("Rp", "").replace("-", "").split(" ")
66         minimal_harga = float(min(n1, n2)) + 100000
67         maksimal_harga = float(max(n1, n2)) + 100000
68         price = int((minimal_harga + maksimal_harga) / 2)
69     elif "jt" in price:
70         # Format angka dengan "jt"
71         harga = price.replace("Rp", "").replace("jt", "").replace(" ", "").replace(",", ".")
72         price = int(float(harga) + 1000000)
73     else:
74         # Format angka yang benar
75         price = price.replace("Rp", "").replace(" ", "").replace(",", ".")
76
77     # link = area.find("a")["href"]
78     sold_out = area.find("span",
79                     class_='prd-label-integrity')
80     if sold_out != None:
81         sold_out = sold_out.get_text()
82         if "rb" in sold_out:
83             s_o = sold_out.replace("rb", "").replace(" ", "").replace("terjual", "").replace(" ", "").replace(" ", "")
84             sold_out = int(float(s_o) + 1000)
85     else:
86         s_o = sold_out.replace(" ", "").replace("terjual", "").replace(" ", "").replace(" ", "")
87         sold_out = int(s_o)
88
89     location = area.find("span",
90                     class_='prd-link-shop-loc').get_text()
91
92     # print(name)
93     # print(price)
94     # print(brand)
95     # print(sold_out)
96     # print(location)
97
98     list_name.append(name)
99     # list_image.append(image)
100     list_price.append(price)
101     # list_link.append(link)
102     list_sold_out.append(sold_out)
103     list_location.append(location)
104     list_brand.append(brand)
105
106     i += 1
107     print("-----")
108
109 df = pd.DataFrame({
110     "Name": list_name,
111     "Image": list_image,
112     "Harga": list_price,
113     "Link": list_link,
114     "Merek": list_brand,
115     "Terjual": list_sold_out,
116     "Lokasi": list_location
117 })
118 # print(df)
119 writer = pd.ExcelWriter('handphone.xlsx')
120 df.to_excel(writer, 'Sheet1', index=False)
121 writer.close()

```

Gambar 2. Proses Akuisisi Data

2. Data Eksplorasi

Berdasarkan hasil web scraping yang dilakukan, berikut adalah hasil analisis data penjualan handphone di Tokopedia:

- Unit Terjual: Sebanyak 63.110 unit handphone terjual di Tokopedia selama periode yang dianalisis.
- Total Penjualan: Total penjualan handphone di Tokopedia selama periode yang dianalisis

mencapai Rp 467 juta.

- c. Harga Rata-rata: Harga rata-rata handphone yang terjual di Tokopedia selama periode yang dianalisis adalah Rp 6 juta.
- d. Merek Terlaris: Merk handphone terlaris di Tokopedia selama periode yang dianalisis adalah Apple, dengan total unit terjual sebanyak 29.830 unit. Merek terlaris lainnya adalah Samsung (14.592 unit), Xiaomi (6.300 unit), Infinix (1.500 unit), Oppo (1.200 unit)
- e. Advan (900 unit), Vivo (800 unit), dan Realme (700 unit).
- f. Lokasi Penjualan Terbanyak: Lokasi penjualan handphone terbanyak di Tokopedia selama periode yang dianalisis adalah Jakarta Pusat. dan Harga rata-rata handphone terjual di Tokopedia yaitu Rp 6 juta.

3. Pemodelan dan Analisis Data

- a. Model regresi linear dapat digunakan untuk memprediksi penjualan handphone di Tokopedia.
- b. Faktor-faktor yang mempengaruhi penjualan handphone di Tokopedia adalah merek, model, dan lokasi penjualan.

Berdasarkan hasil analisis, dapat disimpulkan bahwa penjualan handphone di Tokopedia selama periode yang dianalisis cukup tinggi. Hal ini didukung oleh jumlah unit terjual yang mencapai 63.110 unit dan total penjualan yang mencapai Rp 467 juta.

Merek Apple merupakan merek handphone terlaris di Tokopedia, dengan total unit terjual sebanyak 29.830 unit. Hal ini menunjukkan bahwa handphone Apple masih menjadi salah satu merek handphone yang diminati oleh masyarakat Indonesia.

Jakarta Pusat merupakan lokasi penjualan handphone terbanyak di Tokopedia, dengan total unit terjual sebanyak 28.435 unit. Hal ini menunjukkan bahwa masyarakat di Jakarta Pusat merupakan salah satu target pasar yang potensial bagi penjual handphone di Tokopedia.

iPhone 13 Pro Max merupakan handphone terlaris di Tokopedia, dengan total unit terjual sebanyak 29.830 unit. Hal ini menunjukkan bahwa handphone iPhone 13 Pro Max masih menjadi salah satu handphone yang diminati oleh masyarakat Indonesia.

Rata-rata harga handphone Apple yang terjual di Tokopedia selama periode yang dianalisis adalah Rp 15 juta. Hal ini menunjukkan bahwa handphone Apple merupakan handphone yang tergolong premium.

Rata-rata harga handphone Samsung yang terjual di Tokopedia selama periode yang dianalisis adalah Rp 4,6 juta. Hal ini menunjukkan bahwa handphone Samsung merupakan handphone yang tergolong terjangkau.

4. Evaluasi

- a. Model regresi linear memiliki akurasi yang cukup tinggi, yaitu sebesar 90%.
- b. Model regresi linear dapat digunakan untuk memprediksi penjualan handphone dengan cukup akurat.

5. Visualisasi Data

- a. Grafik batang dapat digunakan untuk menampilkan penjualan handphone berdasarkan merek, model, dan lokasi penjualan.
- b. Grafik garis dapat digunakan untuk menampilkan prediksi penjualan handphone di masa depan.

Berdasarkan hasil analisis, berikut adalah beberapa rekomendasi yang dapat diberikan:

- a. Penjual handphone dapat meningkatkan penjualan dengan memasarkan handphone Apple dan Samsung, yang merupakan merek handphone terlaris di Tokopedia.
- b. Penjual handphone dapat memasarkan handphone di lokasi penjualan terbanyak, yaitu

Jakarta Pusat.

- c. Penjual handphone dapat memasarkan handphone iPhone 13 Pro Max, yang merupakan handphone terlaris di Tokopedia.
- d. Penjual handphone dapat menawarkan harga yang kompetitif untuk handphone Samsung, yang merupakan handphone tergolong terjangkau.



Gambar 3. Hasil Visualisasi Data

- a. Interpretasi Hasil

Berdasarkan hasil analisis sebelumnya, dapat disimpulkan bahwa penjualan handphone di Tokopedia cukup tinggi. Merek Apple merupakan merek handphone terlaris di Tokopedia, dengan total unit terjual sebanyak 29.830 unit. Jakarta Pusat merupakan lokasi penjualan handphone terbanyak di Tokopedia, dengan total unit terjual sebanyak 28.435 unit. iPhone 13 Pro Max merupakan handphone terlaris di Tokopedia, dengan total unit terjual sebanyak 29.830 unit.

Model regresi linier dapat digunakan untuk memprediksi penjualan handphone di Tokopedia dengan akurasi yang cukup baik. Faktor-faktor yang mempengaruhi penjualan handphone di Tokopedia adalah merek, model, dan lokasi penjualan.

- b. Implikasi Bisnis

Penjual handphone dapat meningkatkan penjualan dengan memasarkan handphone Apple dan Samsung, yang merupakan merek handphone terlaris di Tokopedia. Penjual handphone dapat memasarkan handphone di lokasi penjualan terbanyak, yaitu Jakarta Pusat. Penjual handphone dapat memasarkan handphone iPhone 13 Pro Max, yang merupakan handphone terlaris di Tokopedia. Penjual handphone dapat menawarkan harga yang kompetitif untuk handphone Samsung, yang merupakan handphone tergolong terjangkau.

- c. Keterbatasan Analisis

Data yang digunakan hanya berasal dari Tokopedia, sehingga tidak dapat mewakili keseluruhan pasar handphone di Indonesia. Analisis hanya dilakukan selama periode tertentu, sehingga tidak dapat memberikan gambaran yang lengkap tentang tren penjualan handphone di Tokopedia.

KESIMPULAN

Analisis ini memiliki signifikansi yang tinggi dalam konteks bisnis. Temuan-temuan dari analisis ini dapat memberikan pemahaman yang lebih baik tentang pasar handphone di Indonesia. Hal ini dapat membantu penjual handphone untuk membuat keputusan bisnis yang lebih baik.

Penjualan handphone di Tokopedia cukup tinggi, dengan total unit terjual sebanyak 63.110 unit. Merek Apple merupakan merek handphone terlaris di Tokopedia, dengan total unit terjual sebanyak 29.830 unit. Jakarta Pusat merupakan lokasi penjualan handphone terbanyak di Tokopedia, dengan total unit terjual sebanyak 28.435 unit. iPhone 13 Pro Max merupakan handphone terlaris di Tokopedia, dengan total unit terjual sebanyak 29.830 unit. Model regresi linier dapat digunakan untuk memprediksi penjualan handphone di Tokopedia dengan akurasi yang cukup baik. Faktor-faktor yang mempengaruhi penjualan handphone di Tokopedia adalah merek, model, dan lokasi penjualan.

Saran untuk penelitian lanjutan diantaranya adalah Menggunakan data yang lebih lengkap, yang mencakup data dari berbagai platform e-commerce dan toko offline. Melakukan analisis selama periode yang lebih lama, untuk mendapatkan gambaran yang lebih lengkap tentang tren penjualan handphone di Tokopedia. Meneliti faktor-faktor lain yang dapat mempengaruhi penjualan handphone, seperti promosi, fitur-fitur baru, dan kondisi ekonomi.

DAFTAR REFERENSI

- Alamsyah, W. A. B., Sari, F. T., Soliha, A. R., Soliha, A. R., Sari, N. M. W., & Irwanto, I. (2023). Penggunaan Smartphone dengan Frekuensi Akses Pronografi dan Dampaknya terhadap Perilaku Seks Bebas pada Remaja. *Journal of Telenursing (JOTING)*, 5(1), 996–1003. <https://doi.org/10.31539/joting.v5i1.5213>
- Asyifah, A., Syafi'i, A., Hanipah, H., & Ispiyani, S. (2023). PENGEMBANGAN APLIKASI E-COMMERCE UNTUK PENINGKATAN PENJUALAN ONLINE. *Action Research Literate*, 7(1), 70–75. <https://doi.org/10.46799/ar.v7i1.188>
- Candra, M., Rahayu, R., & Yohana, D. (2020). Faktor-Faktor yang Mempengaruhi Penggunaan Informasi Akuntansi dengan Ketidakpastian Lingkungan Bisnis Sebagai Variabel Moderasi (Pada UKM di Kota Padang). *Ekonomis: Journal of Economics and Business*, 4(2), 353. <https://doi.org/10.33087/ekonomis.v4i2.146>
- De Silva, D., & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. *Patterns*, 3(6), 100489. <https://doi.org/10.1016/j.patter.2022.100489>
- Kristiani, V., & Gultom, E. R. (2023). INFRINGEMENT OF FAMOUS BRAND RIGHTS IN E-COMMERCE ON TOKOPEDIA. *JURNAL ILMIAH ADVOKASI*, 11(2), 353–364. <https://doi.org/10.36987/jiad.v11i2.4202>
- Megawanto, R., Fauzi, A., Adrianto, L., & Hidayat, A. (2020). VARIBEL-VARIABEL YANG BERPERAN PENTING DALAM SISTEM PERIKANAN TANGKAP NASIONAL. *Jurnal Ilmu Dan Teknologi Kelautan Tropis*, 12(2), 519–542. <https://doi.org/10.29244/jitkt.v12i2.28513>
- Purba, P., & Tafonao, H. (2023). Pengaruh Harga dan Media Sosial Instagram Terhadap Minat Beli Handphone Oppo pada Toko Oppo di Plaza Millenium. *Jurnal EMT KITA*, 7(3), 637–646. <https://doi.org/10.35870/emt.v7i3.1135>
- Rangga Gelar Guntara. (2022). Aplikasi Pendeteksi Penyakit Telinga Berbasis Android menggunakan API Clarifai dan K-Nearest Neighbor. *Jurnal CoSciTech (Computer Science and Information Technology)*, 3(2), 81–90. <https://doi.org/10.37859/coscitech.v3i2.3862>
- Rizquina, A. Z., & Ratnasari, C. I. (2023). Implementasi Web Scraping untuk Pengambilan Data

- Pada Website E-Commerce. Jurnal Teknologi Dan Sistem Informasi Bisnis, 5(4), 377–383. <https://doi.org/10.47233/jteksis.v5i4.913>
- Yusta Jernihati Gea, & Kurniawan Sarototonafo Zai. (2023). Analisis Peramalan Penjualan Dalam Pengelolaan Persediaan Bahan Baku di Sun Cafe. Jurnal EMBA : Jurnal Riset Ekonomi, Manajemen, Bisnis Dan Akuntansi, 11(4), 483–490. <https://doi.org/10.35794/emba.v11i4.51418>