

## Otomatisasi Konten Statis dan Perilaku Dinamis dari Web *E-Commerce* Menggunakan Metode Content Based Filtering PT Multi Sinar Teknik

*Haryanto* \*, *Abdul Harits*

*Fakultas Sains & Teknologi, Program Studi Sistem Informasi, Universitas Raharja, Tangerang, Indonesia*

### INFORMASI ARTIKEL

Diterima Redaksi: 10 Desember 2023  
Revisi Akhir: 14 Juli 2024  
Diterbitkan *Online*: 30 Juli 2024

### KATA KUNCI

Konten Statis  
Perilaku Dinamis  
Web E-Commerce

### KORESPONDENSI (\*)

Phone: -  
E-mail: [haryanto@raharja.info](mailto:haryanto@raharja.info)

### A B S T R A K



Di dalam web e-commerce sangat bergantung pada analisis perilaku pelanggan, melakukan upaya untuk mempengaruhi Tindakan atau pengguna serta mengoptimalkan metrik keberhasilan. Teknik pembelajaran dan pengembangan data telah diterapkan di bidang ini, sangat mempengaruhi aktivitas pemasaran. Saat ini dengan situs web e-commerce baru, dengan data untuk memulai proses pengumpulan data secara real-time dan historis tentang situs tersebut, menganalisis dan mengubah data ini untuk dengan memahami situs web dan penggunanya. Ilmuwan data biasanya menggunakan pelacakan secara spesifik di domain, sehingga memerlukan modifikasi kode pada halaman web. penelitian ini mengusulkan pendekatan alternatif untuk mengambil informasi dari situs web e-commerce tertentu, mengumpulkan data dari struktur, mengambil informasi semantik di lokasi yang telah ditentukan dan menganalisis log akses pengguna, dengan menggunakan metode Content Based Filtering sehingga memungkinkan pengembangan model yang akurat untuk memprediksi perilaku pengguna di masa depan. Hasil Ada tiga bidang utama dalam pembuatan profil pengguna, terdiri dari latar belakang pengguna, tujuan dan minat pengguna. Hal ini dicapai dengan penerapan proses penambangan web, memahami struktur situs, konten dan penggunaan dalam saluran, menghasilkan grafik web situs web.

### PENDAHULUAN

Salah satu inovasi paling disruptif di dalam perdagangan mengubah cara orang membeli dan menjual barang, diperkirakan memiliki kapitalisasi pasar yang akan tercapai. Ketika pengguna mengunjungi situs web tertentu, beberapa mekanisme pemasaran dan periklanan diaktifkan yang berupaya mempengaruhi tindakan pengguna untuk meningkatkan penjualan serta keuntungan. Mekanisme ini sangat bergantung pada analisis perilaku pelanggan, sehingga memberikan dampak yang signifikan dalam aktivitas pemasaran Internet. Ada beberapa perusahaan saat ini yang mengembangkan dan menyediakan alat serta mekanisme kepada pemilik bisnis e-commerce agar mampu meningkatkan bisnisnya. Alat-alat ini sebagian besar terdiri dari pelacakan dan pembelajaran kebiasaan pelanggan, sering kali dilengkapi dengan sistem. Saat mengakses dengan situs web e-niaga, praktisi pembelajaran biasanya menghabiskan banyak waktu untuk mengumpulkan dan menganalisis data statis serta dinamis situs web. Proses ini penting untuk mengakses informasi berguna tentang struktur situs web, konten, dan perilaku penggunanya. Setelah itu mampu membangun model dan algoritma yang relevan untuk meningkatkan aktivitas pemasaran online. Hal ini merupakan tantangan besar karena sifat data Web yang heterogen, cara penyajiannya yang semi terstruktur atau tidak terstruktur, dan banyaknya jumlah data yang dihasilkan. Untuk menyajikan sebuah proses yang mampu, pada situs web e-commerce, mengekstraksi informasi berguna dari struktur, konten, dan perilaku pengguna pada umumnya menggunakan teknik dan metode terancang dan mengembalikan model yang konsisten yang mewakili konten situs web, hubungan antara halaman dan pola dasar penggunanya. Pada Bagian II kami menyajikan tinjauan literatur dalam e-commerce dan pengembangan web. tentang masalah yang sedang kita hadapi dan memberikan wawasan lebih lanjut mengenai solusi yang diusulkan.

## TINJAUAN PUSTAKA

Perdagangan elektronik juga sebagai cara untuk melibatkan mereka. Seseorang tidak dapat berbicara tentang e-commerce, yang merupakan salah satu perusahaan Internet besar pertama yang menghadirkan e-commerce secara luas kepada penggunanya. Namun, saat ini, transaksi semacam ini tersebar luas. Saat ini semakin umum untuk menemukan kampanye pemasaran dan periklanan di Internet, dengan perusahaan e-commerce yang menggunakan teknik personalisasi Web sebagai cara melakukan pemasaran target terhadap pengunjungnya dengan tujuan meningkatkan kemungkinan pengunjung menghasilkan keuntungan. Untuk mengukur keterlibatan dan keberhasilan situs web e-niaga, pemilik layanan menggunakan metode ini e-metrik. Metrik ini berupaya memberikan gambaran kinerja bisnis online yang lebih baik dan langsung, yaitu: Rasio adalah tingkat pengguna yang mengunjungi satu halaman dan keluar dari situs web; Tingkat konversi adalah jumlah pembeli dibandingkan jumlah total pengguna situs web; Rasio klik-tayang singkatan dari rasio antara pengguna yang mengklik link tertentu dengan jumlah total pengunjung; Retensi pelanggan menunjukkan kemampuan dalam mempertahankan pelanggan dalam jangka panjang.

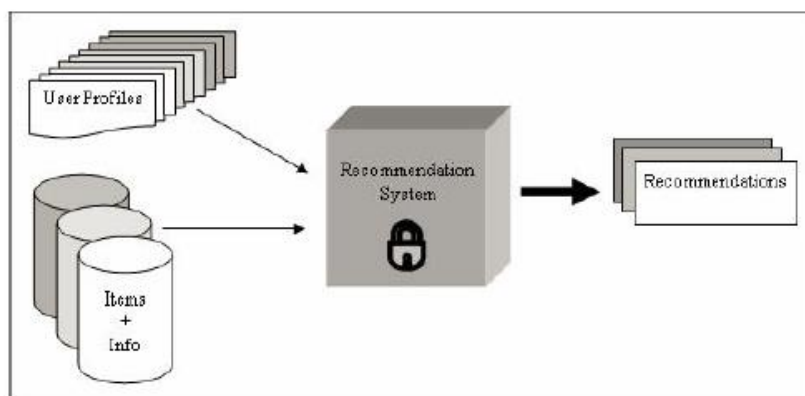
Di dalam keranjang belanja memberikan berapa kali item ditambahkan ke keranjang belanja tetapi esanan tidak diselesaikan. Salah satu alat inti untuk meningkatkan e-metrik adalah sistem rekomendasi. Sistem ini mengandalkan proses tiga langkah, dimulai dengan (i) pengambilan data, biasanya menggunakan teknik pengembangan Web, di mana kita mengetahui preferensi pengguna dengan menganalisis konten statis dan data perilaku pengguna, diikuti dengan (ii) komputasi dan validasi rekomendasi dengan menggunakan teknik yang tepat, dan finalisasinya dengan (iii) menyajikan hasil rekomendasi kepada pelanggan. Sistem ini umumnya dapat menggunakan tiga pendekatan berbeda: Filteran kolaboratif berbasis konten, dan hybrid.

Heterogenitas Web, kita dapat menyimpulkan struktur umum situs web e-commerce karena tujuannya yang sama, yaitu (i) menampilkan produk kepada pelanggan, (ii) memberi mereka kemungkinan untuk berpindah antar produk yang berbeda, dan (iii) memungkinkan pembelian. Struktur dan desain umum ini secara de facto pedoman yang diikuti oleh pemilik situs web membantu pengguna beradaptasi dengan mudah ke situs web e-niaga mana pun, sehingga meningkatkan pengalamannya secara keseluruhan.

## METODOLOGI

### *Penambangan Web*

Penambangan data telah digunakan oleh perusahaan untuk fokus pada informasi paling penting yang ada dalam data mereka. Komunitas penambangan data mengidentifikasi tiga jenis pengembangan yang berbeda: pengembangan data, pengembangan web. Pengembangan web menggabungkan teknik pengembangan data dan teks yang menerapkan konsep ini. Penambangan web ini menggunakan sebuah metode yang dinamakan *Content Based Filtering*.



Gambar 1. Suatu Proses sebuah Rekomendasi Content Based Filtering

### ***Pengembangan Website***

Dalam pengembangan data web menghadapi beberapa tantangan karena penyimpanan data yang masih ada, beragam, dinamis, terus berkembang, dan sebagian besar tidak terstruktur, yang memberikan sejumlah besar informasi dan meningkatkan kompleksitas penanganannya dari sudut pandang ilmuwan data. Seperti yang disebutkan oleh Patel dkk (2013) . Penambangan web dibagi menjadi beberapa kategori yang melambangkan penekanan yang berbeda dan cara yang berbeda untuk memperoleh informasi, meskipun perbedaan di antara keduanya semakin menyempit karena kategori kategori tersebut saling berhubungan. Pertama, kami memiliki penambangan konten Web yang berkaitan dengan pengambilan informasi dari situs web dan halaman Web-nya. Ini terdiri dari mengubah data asli menjadi bentuk yang lebih terstruktur, mengindeks informasi untuk mengambilnya dengan lebih mudah dan cepat. Pada penggunaan Web tujuan utamanya adalah menemukan pola akses pengguna dari penggunaan Web dan log akses. Terakhir, Penambangan Struktur Web adalah Teknik menganalisis struktur hyperlink di dalam Web itu sendiri, dengan mempertimbangkan halaman web sebagai node dan hyperlink di antara keduanya sebagai untuk menghasilkan Grafik Web.

### ***Pengumpulan dan Pemrosesan Data***

Dengan tujuan utama dalam tahap pengumpulan data adalah untuk mengumpulkan sumber daya, mengambil data dalam halaman situs web dan catatan penggunaan Web.

### ***Log Penggunaan Web***

Di dalam log server web menyimpan catatan peristiwa interaksi pengguna di situs web. Data ini umumnya menampilkan data yang tidak lengkap, berlebihan, Untuk proses yang lebih efisien, penting untuk memfilter data yang berisik ini menggunakan teknik pra-pemrosesan, sehingga menghasilkan data yang lebih akurat dan ringkas. Pemrosesan awal data terdiri dari pembersihan data, identifikasi pengguna unik, identifikasi sesi pengguna, penambahan jalur akses, dan identifikasi transaksi. Selain itu, terdapat informasi yang diambil di tingkat aplikasi, dengan tujuan untuk mencatat informasi tentang interaksi pengguna dengan situs web. Dengan bantuan alat yaitu, durasi sesi Dan waktu rata-rata per halaman.

### ***Perayap Web***

Saat mengumpulkan data dari suatu halaman serta struktur situs web tertentu, dengan pendekatan yang paling simple serta umum dengan penggunaan perayap Web. Dengan Perayap Web ini program yang, dengan satu atau lebih URL dasar, serta dalam melihat halaman Web serta terkait dengan URL, serta mengekstrak dalam suatu hyperlink yang terlihat dana ada di dalamnya serta belum pernah ditemukan sebelumnya.

### ***Ekstraksi Data Web***

Data yang ada pada konten halaman website sebagian besar merupakan data semi terstruktur atau tidak terstruktur. Untuk memahami dan mengambil informasi berguna dari halaman ini dan mengubahnya menjadi format data terstruktur, diperlukan penggunaan teknik otomatis tertentu untuk mengekstrak, menafsirkan, dan menyajikan data, proses ini disebut ekstraksi data web. Proses ekstraksi data web dilakukan dengan menggunakan pembungkus. mencantumkan tiga pendekatan berbeda untuk menghasilkan pembungkus: pendekatan manual, induksi pembungkus, dan ekstraksi otomatis.

## **HASIL DAN PEMBAHASAN**

### ***Penemuan dan Analisis Pola***

Dengan penemuan pola terdiri dari pengambilan informasi serta pengetahuan yang efektif serta berguna dan dapat dipahami menggunakan algoritma penambangan. Metodenya antara lain mencakup analisis data eksplorasi, serta analisis klasifikasi, penemuan aturan asosiasi, penemuan pola sekuensial, dan analisis klaster.

### ***Profil Pengguna***

Dalam pembuatan profil pengguna adalah proses yang mengacu pada konstruksi profil melalui ekstraksi informasi dari data penggunaan situs web. Ini adalah bagian penting dari sistem rekomendasi yang dipersonalisasi. Kualitas suatu profil pengguna mempengaruhi kualitas dan rekomendasi secara langsung. Hanya sistem yang memahami kebutuhan dan kepentingan pengguna yang mampu merekomendasikan informasi yang memuaskan pengguna. Ada tiga bidang utama

dalam pembuatan profil pengguna, seperti yang, terdiri dari latar belakang pengguna (pengetahuan yang diperoleh dalam berbagai mata pelajaran), tujuan minat pengguna.

### ***Representasi Profil Pengguna***

Dalam pemodelan profil pengguna terdiri dari perancangan struktur untuk menyimpan semua informasi yang menjadi ciri pengguna dan menggambarkan minatnya serta tujuannya. Profil pengguna umumnya direpresentasikan sebagai kumpulan kata kunci yang berbobot, serta jaringan semantik, atau konsep berbobot, atau aturan asosiasi.

### ***Konstruksi Profil Pengguna***

Dalam profil pengguna dibuat dari sumber informasi menggunakan berbagai teknik berdasarkan pembelajaran atau pengambilan informasi, mencoba membangun profil yang mencerminkan minat dan karakteristik pengguna. Bergantung pada representasi profil pengguna yang dipilih, diperlukan pendekatan berbeda untuk membangunnya. Terkadang hal ini menjadi lebih rumit, karena, di satu sisi, pengguna mungkin tidak yakin dengan kepentingannya sendiri, dan, di sisi lain, sering kali pengguna tidak mau atau bahkan tidak bisa melakukan upaya untuk menciptakan kepentingannya sendiri. Profil.

### ***Ikhtisar Tingkat Tinggi***

Saat pembelajar dihadapkan dengan situs web e-niaga baru yang sebelumnya tidak diketahui, para ilmuwan data harus menelusuri situs web tersebut, memahami tidak hanya bisnisnya tetapi juga bagaimana situs itu dibuat, halaman - halamannya saling berhubungan, dan bagaimana kontennya. disajikan kepada pengguna. Mereka juga perlu menganalisis log dari sistem pencatatan peristiwa apa pun yang mungkin diaktifkan oleh situs web, untuk memahami cara pengguna menavigasi situs web, menemukan pola dan kebiasaan penggunaan. Dalam beberapa kasus, namun kurang umum, data scientist bahkan perlu mengaktifkan sistem pelacakan di situs web, untuk menemukan pola akses pengguna dan jalur navigasi umum. Artikel ini mencoba membuat pendekatan baru terhadap pengumpulan dan representasi data situs web e-commerce, dengan tujuan sebagai berikut:

1. Kumpulkan struktur situs web, yang mewakili halaman dan interkoneksinya, sebagai grafik;
2. Mengekstraksi informasi yang relevan dari halaman situs web dan mengelompokkannya ke dalam jenis yang telah ditentukan sebelumnya;
3. Kumpulkan informasi tentang pengguna situs web:
  - a. Identifikasi pengguna dan sesi yang berbeda, termasuk informasi tentang alur navigasi pengguna di situs web;
  - b. Tentukan preferensi pengguna serta jenis halaman yang paling banyak dikunjungi;
  - c. Karakterisasikan sesi pengguna berdasarkan metrik yang ditetapkan;
4. Berbagai sumber informasi guna menemukan pola dan memperkaya data yang ada;
5. Temukan pola dasar pengguna situs web;
6. Model konsisten yang mewakili semua informasi situs web yang dirancang agar mudah beradaptasi.

Kami menganggap bahwa ilmuwan data setelah memiliki informasi ini mampu memberikan model dan algoritma yang lebih halus kepada pemilik situs e-commerce. Untuk mencapai tujuan kami, mengusulkan penggunaan teknik Web Mining yang terkenal secara tepat dan konsisten semua dalam satu proses. Dengan mempertimbangkan kebutuhan mengambil konten dari halaman situs web, hubungan antara pengetahuan tentang pola dasar pengguna, pendekatan yang berbeda dipilih untuk setiap sumber informasi, dan model data informasi yang tepat dikembangkan. Dengan mempertimbangkan keberagaman website e-commerce, maka rangkaian langkah tersebut dirancang agar tidak hanya diterapkan pada satu jenis atau website tertentu saja, namun mampu.

Dengan tahapan utama yang diidentifikasi hampir beroperasi di semua situs web e-niaga, menangani berbagai struktur, konten, atau format yang mungkin dimilikinya. Selain pengumpulan data dari berbagai sumber dan menangani format yang berbeda, proses ini bertujuan untuk mampu membangun hubungan antara semua sumber informasi yang mencoba melengkapi data yang terkadang tidak lengkap, dan menyoroti hubungan antara data yang mungkin ada. hilang. Prosesnya dibagi menjadi tiga tahap utama. Yang pertama terdiri dari proses Web Mining yang khas, mengumpulkan data dari situs web e-commerce dengan masing-masing pra dan pasca pemrosesan. Tahap kedua terdiri dari membangun hubungan antar data, mengambil informasi berguna seperti kumpulan pengguna dasar di situs web e-commerce.

Ada berbagai gambaran umum suatu proses, kita dapat mempertimbangkan tiga tahapan berbeda, yaitu tahap pengumpulan dan pemrosesan data, tahap penyeberangan data dan pasca pemrosesan, serta penemuan dan analisis pola.

Dengan mempertimbangkan aliran data serta transformasi, mengetahui bahwa tahapan terdiri dari sekelompok langkah yang dapat dibagi lagi menjadi lebih dari satu tugas, atau tahap, kami menetapkan alur sebagai berikut:

### ***Tahap pengumpulan dan pengolahan data***

1. Pemilihan sumber data: Memilih dan memilih sumber data mana yang digunakan, dalam hal ini terdiri dari situs web itu sendiri dan catatan situs web yang terkait dengan interaksi pengguna di halaman.
2. Struktur situs web serta pengambilan pra-pemrosesan dalam konten halaman: Untuk mengumpulkan informasi tentang struktur situs web dan halaman-halamannya, digunakan yang menavigasi melalui domain situs web tertentu. Data dikumpulkan dan diproses terlebih dahulu untuk menjaga konsistensi grafik yang dihasilkan.
3. Pengumpulan log penggunaan situs web dan pra-pemrosesan: Selama langkah ini, catatan penggunaan, yang berasal dari server atau lapisan aplikasi diurai menjadi struktur yang telah ditentukan sebelumnya, dan diproses terlebih dahulu untuk menghilangkan gangguan seperti entri yang tidak konsisten dalam catatan atau lalu lintas tidak valid seperti yang dihasilkan oleh mesin pencari.
4. Pasca-pemrosesan halaman situs web: Untuk setiap halaman situs web yang ditemukan, kode HTML halaman tersebut diproses untuk mengklasifikasikan halaman berdasarkan jenisnya dan kemudian, tergantung pada jenisnya, berbagai bidang .
5. Pasca pemrosesan data penggunaan situs web: Catatan penggunaan ditafsirkan dan pengguna unik yang berbeda diidentifikasi. Sesi pengguna diidentifikasi.
6. Representasi dan penyimpanan data: Pada akhir tahap ini, data disimpan dalam model adaptif, yang mewakili grafik struktur situs web, informasi di suatu halaman, dan data pengguna.

### ***Tahap pasca pemrosesan***

1. Persilangan data penggunaan dan konten halaman: Untuk meningkatkan informasi yang ada di profil pengguna, informasi yang berasal dari halaman situs web pasca-pemrosesan disilangkan dengan alur penggunaan, sehingga memungkinkan ekstraksi data relatif terhadap jenis dan kategori halaman yang dikunjungi, sehingga memungkinkan pembuatan kata kunci profil pengguna.
2. Pasca-pemrosesan sesi pengguna: Karena beragamnya jenis sesi, dengan karakteristik yang berbeda, setiap sesi dapat diklasifikasikan sebagai jenis tertentu dan kemudian menggeneralisasi informasi ini kepada pengguna itu sendiri, sehingga memungkinkan untuk membedakan, misalnya, pengguna dengan jumlah halaman yang dikunjungi berdasarkan sesi lebih banyak dari pengguna yang memiliki jumlah halaman yang dikunjungi lebih sedikit.
3. Representasi dan penyimpanan data: Menyelesaikan tahap ini, kami menyimpan data yang sesuai dengan pohon kategori situs web dan profil pengguna berbasis kata kunci ke dalam model adaptif.

### ***Penemuan dan analisis pola***

1. Pembuatan profil pengguna tipikal: Karena profil pengguna sekarang lebih kaya, tidak hanya berisi informasi dari log penggunaan karena persilangan data, profil pengguna tipikal dapat dibuat, dalam hal ini menggunakan algoritma pengelompokan.
2. Representasi dan penyimpanan data: Menyelesaikan tahap ini, data yang dihasilkan dari proses clustering disimpan dalam model data yang adaptif.

### ***Detail Implementasi***

Dengan tujuan untuk menunjukkan kelayakan dan penerapan proses yang dikembangkan, disusun untuk bukti dari konsep. Implementasinya didasarkan pada prinsip pola perangkat lunak dan filter. Dengan cara ini, implementasi terdiri dari rangkaian proses yang disusun sedemikian rupa sehingga keluaran dari setiap elemen mata rantai merupakan masukan dari elemen berikutnya. Proses model rantai ini. Menganalisis proses yang disajikan secara mendalam, pertama-tama kami memiliki sumber data, yang terdiri dari situs web e-commerce itu sendiri dan log penggunaan yang terkait dengannya. Di satu sisi, kami memiliki situs web yang umumnya dibuat menggunakan teknologi web umum dan sumber daya lainnya. Di sisi lain, kami mendapatkan log yang terkait dengan situs web. Log ini berisi informasi tentang permintaan atau peristiwa yang terkait dengan interaksi pengguna dengan situs web. Jika data diambil di lapisan server, log ini berisi informasi tentang permintaan HTTP pengguna saat ia navigasi antar halaman situs web. Jika data diambil pada lapisan aplikasi, log ini berisi data yang lebih kaya tentang interaksi pengguna, tidak hanya berisi informasi tentang

halaman yang dikunjungi tetapi juga informasi tentang interaksi pengguna. Keluaran dari pendekatan yang diusulkan, dimulai dengan sumber data yang digunakan di sisi kiri gambar, hingga keluaran proses akhir di sisi kanan.

### ***Evaluasi***

Proses evaluasi terdiri dari dua situs e-commerce berbeda yang diambil sebagai contoh masukan, yang mencerminkan dua pasar dan pengguna yang berbeda. Penting untuk mengevaluasi bagaimana proses berperilaku dalam keadaan yang berbeda, sehingga situs web ini dipilih karena struktur, konten, dan sistem pencatatan penggunaannya berbeda-beda. Salah satu eksperimen didasarkan pada situs web e-niaga dengan tujuan umum, yang menggabungkan produk dari banyak kategori berbeda dan tanpa target pasar. Di sini menjadi penting untuk memahami bagaimana konten ditempatkan di situs web, dan mengetahui kategori atau produk yang paling banyak dikunjungi untuk meningkatkan aktivitas pemasaran di area ini. Eksperimen lainnya didasarkan pada situs web e-niaga yang didedikasikan untuk ceruk pasar, dalam kasus ini para penggemar industri game. Di sini juga berguna bagi data scientist untuk memahami bagaimana data didistribusikan di situs web, dan menemukan preferensi paling khusus dari penggunanya. Menerapkan crawler pada Situs Web E-niaga Khusus Niche, kami mendapatkan grafik web yang terdiri dari:

1. Total 2687 halaman (grafik webnode);
2. Total 361.344 tautan valid ditemukan (grafik web tepi).

Dengan menyilangkan informasi dari halaman dan struktur situs web, kami mendapatkan pohon kategori dengan 24 kategori dasar dan total 102 subkategori. Untuk situs web ini kami tidak memiliki akses ke catatan penggunaan, dan, sebagai alternatif, kami menggunakan data sintesis untuk memvalidasi komponen pengguna dalam proses tersebut. Untuk mencapai evaluasi ini, kami mendefinisikan alur navigasi dan menganalisis catatan, membandingkan alur ini dengan pola dasar yang ditemukan pengguna, menemukan kecocokan langsung sebagai cara validasi.

Menerapkan crawler pada Situs Web E-niaga Tujuan Umum, kami mendapatkan grafik web yang terdiri dari:

1. Total halaman yang di rayapi (node grafik): 621.303 halaman;
2. Total tepi valid yang ditemukan: 10.044.225 tautan.

Dengan menyilangkan informasi dari jenis halaman dan struktur situs web, kami mendapatkan pohon kategori dengan 14 kategori dasar dan total 1618 subkategori. Dari catatan penggunaan 4,5 GB yang dianalisis, kami dapat mengidentifikasi:

1. 111.141 pengguna unik dan 130.056 sesi;
2. Rata-rata halaman yang dikunjungi oleh pengguna: 4,6;
3. Durasi sesi rata-rata: 125,07 detik;
4. Total tampilan halaman: 511.354.

Dari total pengguna yang ditemukan, sampel acak sebanyak 50.000 pengguna diekstraksi, dan berarti algoritma pengelompokan diterapkan. Dengan menggunakan informasi sesi sebagai variabel, kita mendapatkan 5 cluster dan 26 pelanggan diabaikan sebagai outlier. Dari data ini kita dapat melihat banyak hal rasio pentalan dalam pelanggan situs web. Menerapkan berarti algoritma atas preferensi pengguna menghasilkan 5 cluster, dengan 26 pelanggan diabaikan sebagai outlier. Dari sini, kita dapat melihat bahwa, meskipun dalam sebagian besar kasus pengguna mengunjungi produk di semua kategori, preferensi pengguna tertentu juga cukup terwakili.

### **KESIMPULAN DAN SARAN**

Dalam penerapan teknik Web Mining ke web, dan e-commerce, untuk meningkatkan keuntungan, bukanlah hal baru, dan banyak penelitian telah dilakukan di bidang ini, khususnya ketika berhubungan dengan penggunaan data. Dalam artikel ini, kami mengusulkan sebuah atau semua dalam satu proses untuk mengumpulkan dan menyusun data dari konten, struktur, dan pengguna situs web e-niaga. Proses ini diperkaya dengan persilangan data yang dikumpulkan dari berbagai sumber untuk menemukan hubungan yang tidak eksplisit, sehingga memperkaya proses keluaran. Akhirnya disajikan proposal model informasi untuk situs web e-niaga, yang berisi informasi dan data yang dikumpulkan dan terstruktur yang dihasilkan dari persilangan berbagai sumber dan tugas penemuan pola. Pengembangan lebih lanjut dapat diterapkan pada setiap fase proses yang disajikan. Dimulai dengan langkah semacam indeks relevansi dapat dihitung dan dikaitkan dengan setiap tautan, yang dapat berguna saat menggabungkan jalur navigasi pengguna dengan relevansi yang terkait dengan setiap tautan. Pendekatan secara manual digunakan, namun masih ada ruang untuk bereksperimen dengan pendekatan

lain. Selain itu, terdapat dua persilangan data, namun terdapat ruang untuk bereksperimen dengan persilangan data dari sumber lain. Akhirnya, untuk menemukan pengguna dasar situs web, kami menggunakan algoritma untuk pengelompokan teknik penemuan pola lainnya dapat dilakukan.

## DAFTAR PUSTAKA

- [1] Schafer, J.B., Konstan, J., Riedl, J. Recommender systems in e-commerce. In: Proceedings of the 1st ACM Conference on Electronic Commerce; EC '99. New York, NY, USA: ACM. H.163
- [2] Mobasher, B. Data mining for web personalization. In: The adaptive web. Springer; p. 90–135. 2007,
- [3] Zhang, Q., Segall, R. Web Mining: a Survey of Current Research, Techniques, and Software. International Journal of Information Technology & Decision Making .2008
- [4] Eisingerich, A., Kretschmer, T. In e-commerce, more is more. Havard Business Review 2008;86(3).
- [5] Bresnahan, T., Davis, J.P., Yin, P.L. Economic value creation in mobile applications 2014;.
- [6] Mohapatra, S. *E-commerce strategy: text and cases*. Springer Science & Business Media;89, 2012.
- [7] Hydzik, J. *The revolution is just beginning*. Total Telecom 32. 2005
- [8] Sterne, J., Cutler, M. E-metrics: business metrics for the new economy. URL (consulted March 2005): <http://www.emetrics.org/articles/whitepaper.html> 2000
- [9] Peska, L., Vojtas, P. Evaluating various implicit factors in e-commerce. CEUR Workshop Proceedings;910(Rue):51–55. 2012
- [10] Gefen, D. Customer loyalty in e-commerce. Journal of the association for information systems 2002;3(1):2.
- [11] Wei, K., Huang, J., Fu, S. A survey of e-commerce recommender systems. In: Service Systems and Service Management, International Conference on. IEEE; p. 1–5. 2007
- [12] Purwati, Y. Standard features of e-commerce user interface for the web. Researchers World 2011;2(3):77.
- [13] Kosala, R., Blockeel, H. Web mining research: A survey. ACM Sigkdd Explorations Newsletter;2(1):1–15. 2000
- [14] Haryanto, A Rufai, Model Manajemen Hubungan Pelanggan dan Sistem Bisnis Intelijen untuk Katalog Secara Online, dalam Jurnal Snartisi , 2018