

ANALISIS PENDUDUK MISKIN BERDASARKAN PENJUALAN PULSA TELEPON GENGAM DENGAN MENGUNAKAN WEB MINING

¹Muhamad Derajat
²Essa Jiwa Permana
³Asep Suandi
⁴Intan Fitri Wulandari
⁵I Wayan Simri Wicaksana

^{1,2,3,4}Perangkat Lunak Sistem Informasi, Fakultas Ilmu Komputer, Universitas Gunadarma
⁵Pusat Studi Teknologi Sistem Informasi, Universitas Gunadarma
¹muhamad.derajat@gmail.com, ²essa.jp@gmail.com, ³asep.suandi@fremantlemedia.com,
⁴intan_fitri@yahoo.co.id, ⁵iwayan@staff.gunadarma.ac.id

ABSTRAK

Pada setiap penyajian data dari Pemerintah kerap kali menimbulkan ketidakpuasan dari berbagai pihak, salah satunya adalah data mengenai penduduk miskin. Banyak pihak yang mengatakan bahwa sebenarnya penduduk miskin lebih besar dari pada data yang diinformasikan dari pemerintah, seperti dari BPS.

Mengingat hal ini, maka pendekatan web mining bisa memberikan sebuah gambaran berbagai sumber data yang menyatakan besaran nilai dari penduduk miskin dan dapat dilakukan perbandingan nilai data. Dan untuk lebih menarik, perlu dilihat juga data pendamping seperti besarnya nilai belanja pulsa telepon genggam.

Web mining akan memberikan gambaran kelengkapan data dari internet lebih banyak walaupun data yang dimiliki bisa berbeda secara ekstrim.

At each presentation of data from the government often lead to dissatisfaction from various parties, one of which is data about the poor people. Many parties who say that poor people actually bigger than the data that informed by the government, such as from the BPS.

Refer to this issue, the web mining approach can provide a variety of data sources that states the amount of the value of the population poor and can do a comparison of data values. And for more interesting, also need to see data assistants such as the size of the value of free mobile phones.

Web mining will give you a completeness of data from the internet although more data that can be owned by a different extreme.

Kata Kunci : bps, internet, penduduk miskin, pulsa, web mining, web content mining,

PENDAHULUAN

Pada saat ini internet sudah menjadi tempat penyimpanan informasi dan data yang luar biasa besarnya. Walaupun dengan internet memiliki kendala seperti sifat-sifat penyajian yang berbeda, heterogen dan tidak terstruktur. Melimpahnya data dan informasi yang disediakan oleh internet telah membuat banyak orang mengandalkan internet dalam pencarian data. Bukan tidak

mungkin di internet kita bisa mendapatkan data penduduk miskin dengan berbagai cara, semisal dengan mengetikkan kata kunci 'penduduk miskin' maka akan keluar banyak sekali hasil pencarian yang mengandung kata tersebut. Walaupun hasilnya tidak akurat dan dipertanyakan keabsahannya.

Mengapa penulis mencari data penduduk miskin? Karena hal ini sangat menarik. Badan Pusat Statistik Rabu

(30/1/2009) mengumumkan, jumlah penduduk miskin mengalami penurunan. Data tentang jumlah penduduk miskinpun menjadi isu yang sangat menarik yang digunakan pada kampanye pemilihan presiden tahun 2009 ini. Dimana masing-masing calon presiden, mempunyai data masing-masing dan terjadi perbedaan dari segi jumlah penduduk miskin dan kategori rakyat miskin di Indonesia.

Menurunnya jumlah penduduk miskin yang diumumkan oleh BPS ternyata berbanding lurus dengan menjamurnya bisnis penjualan pulsa telpon genggam, hal ini dapat dibuktikan jika kita menggunakan mesin pencari di internet dan memasukan kata 'penjualan pulsa' maka kita akan mendapatkan informasi bisnis voucher yang luar biasa banyak. Fakta yang menarik lainnya adalah pembelian voucher itu dilakukan setiap kali pulsa di telpon genggamnya habis, bahkan satu bulan bisa lebih dari satu kali melakukan pembelian voucher telpon genggam tersebut.

Permasalahan yang muncul dalam pencarian data dan informasi didalam internet adalah masalah ketepatan data dan informasi selain itu masalah keabsahan data yang didapatkan. Untuk memecahkan masalah tersebut salah satu pendekatannya adalah menggunakan metode web mining. Proses pencarian data menggunakan web mining adalah proses pencarian pola dalam mencari informasi atau data dengan menggunakan metode tertentu.

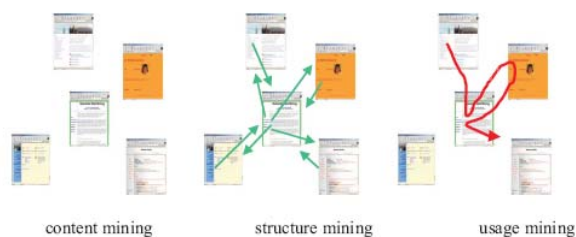
Dengan menggunakan metode web mining penulis membandingkan jumlah penduduk miskin yang di keluarkan oleh Badan Pusat Statistik dengan penjualan voucher telpon genggam berdasarkan data-data yang didapat dari hasil pencarian di internet.

Web mining itu sendiri merupakan salah satu teknologi yang dewasa ini berkembang dengan pesat.

Dan telah diaplikasikan di berbagai domain penelitian seperti web mining dipakai untuk menemukan informasi spatio temporal penyebaran penyakit dari data yang tersedia di internet[1], untuk mengelola *customer relationship* dalam *e-commerce*[2] dan masih banyak lagi.

Bettina Berendt, Andreas Hotho dan Gerd Stumme [8] menyatakan secara garis besar web mining dibagi dalam tiga macam yaitu:

1. Web Content Mining (WCM) yaitu pencarian informasi dan data dari isi web.
2. Web Usage Mining (WUM) yaitu pencarian informasi dan data dari penggunaan web.
3. Web Structure Mining (WSM) adalah pencarian dan informasi yang berkaitan dengan struktur hubungan web.



Gambar 1. Tiga macam dari web mining

Sumber: Bettina Berendt, Andreas Hotho dan gerd Stumme [8].

Jiye Ai dan James Laffey [3] menggunakan Web usage mining dengan cara mengklasifikasikan data untuk dapat memprediksi kinerja mahasiswa yang mengikuti WebCT e-learning. Data yang digunakan adalah dengan mengambil log data dari webserver. Christopher C. Yang dan Y. C. ong [9] menggunakan web content mining untuk mengetahui opini tentang fitur dari produk yang dipasarkan dari para pengunjung web-ecommerce, pendekatan yang digunakan adalah dengan menggunakan class association rules mining dan natural

language processing untuk dapat memecahkan problem yang timbul akibat penulisan opini yang menggunakan bahasa manusia yang tidak terstruktur dan banyak gramatikalnya yang salah. Peter Toth [6] menggunakan metode webusage mining ini untuk menganalisa kebiasaan para siswa dalam mengakses materi kursus yang disimpan dalam sebuah web elearning, dengan pendekatan yang dilakukan adalah menganalisa kebiasaan siswa dalam mengakses web elearning. Liwen Vaughan dan Justin You [7] menggunakan metode Web Structure Mining dalam journalnya yaitu untuk mengetahui link antar website dan menggunakan tool mesin pencari yahoo dalam penelitiannya.

METODE WEB CONTENT MINING

Penelitian ini termasuk dalam kategori studi banding dimana data penduduk miskin yang dikeluarkan BPS dibandingkan dengan subjek penelitian dalam paper ini yaitu data dari penjualan pulsa telepon genggam di internet. Teknik pengumpulan data di internet akan dilakukan dengan menggunakan web mining.

Dan di dalam paper ini akan menggunakan metode *Web Content Mining* untuk pengumpulan data penjualan pulsa telepon genggam di internet. Metode yang digunakan adalah menggunakan pendekatan pencarian data dengan menggunakan mesin pencari seperti google. Pencarian terus dilakukan sampai beberap kali sehingga mendapatkan data atau informasi dari website tertentu. Biasanya walaupun data dan informasinya sama akan terjadi perbedaan pada saat penyampaian data dan juga karakteristik datanya. Oleh karena itu data atau informasi yang didapatkan akan dicoba dibandingkan dengan data yang berasal dari website yang mengeluarkannya pertama kali.

PEMBAHASAN

Data BPS

Badan Pusat Statistik Rabu (30/1/2009) mengumumkan, jumlah penduduk miskin pada Maret 2009 sebesar 32,53 juta atau 14,15 persen. Jumlah itu terendah sejak 1996 yang tercatat 34,01 juta penduduk miskin. Selama periode Maret 2008-Maret 2009 penduduk miskin di daerah pedesaan berkurang 1,57 juta sementara daerah perkotaan berkurang 0,86 juta orang (tempointeraktif).

Pengumpulan Data Dengan Web Content Mining

Pengumpulan data dimulai dengan memberikan daftar URL yang selanjutnya akan di jelajahi dan dikumpulkan datanya.

Salah satu teknik yang dilakukan dalam *web content mining* adalah *template detection and page segmentation* [4]. Dan teknik itulah yang akan dipakai untuk mendapatkan data dari artikel-artikel yang ada di internet baik itu dari blog ataupun online news dan majalah online.

Karena hampir semua blog ataupun website-website yang ada di internet ini pasti mempunyai template untuk pembangunan halaman-perhalaman-nya.

Dan biasanya template ini mempunyai blok-blok yang sudah umum yaitu[4]:

- a. Blok isi utama
- b. Blok navigasi
- c. Blok service
- d. Blok iklan

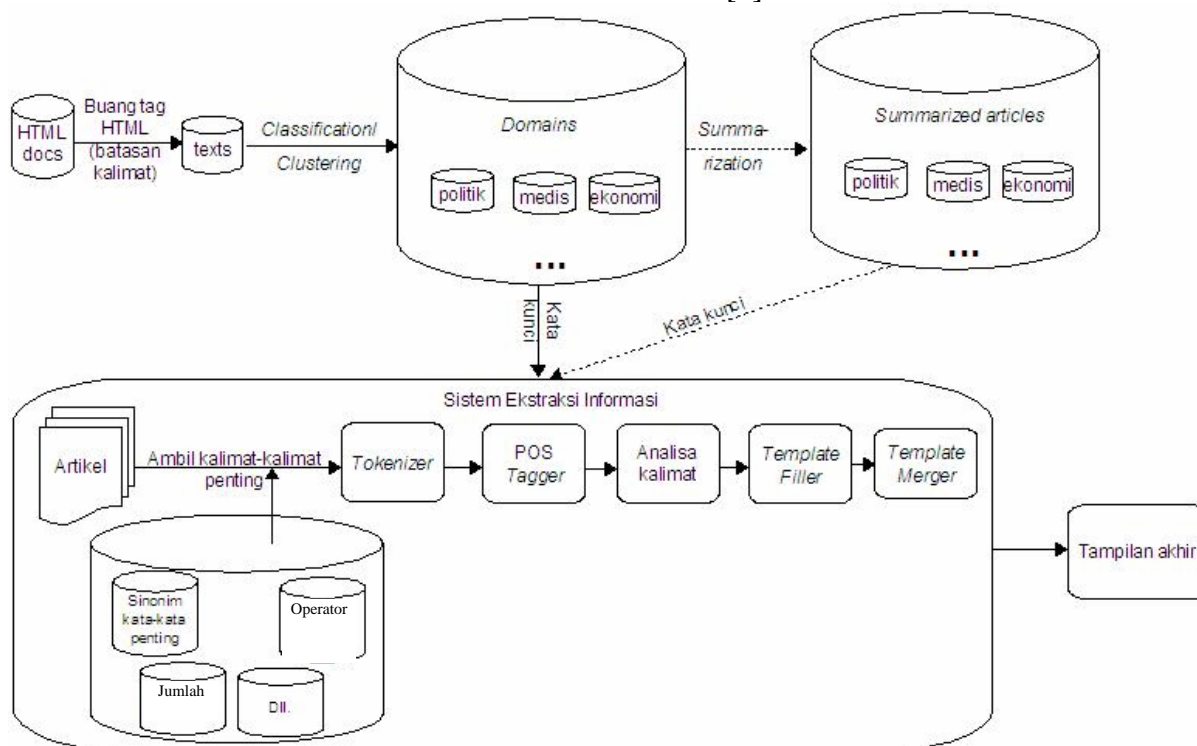
Dengan teknik ini, maka pengumpulan data akan lebih terfokus kepada blok isi utama berita atau tulisan di badan-nya saja dalam sebuah website atau blog, sehingga data yang dikumpulkan untuk dianalisa tidak akan terlalu banyak yang tidak sesuai, dalam hal ini mengabaikan

isi dari blok navigasi, blok servis dan blok iklan.

Kemudian data-data textual ini akan disimpan di dalam database yang selanjutnya akan dianalisa untuk mendapatkan informasi penjualan pulsa.

Ekstraksi Informasi Penjualan Pulsa

Seteleh data yang berasal dari web didapatkan tahapan selanjutnya adalah dengan mengekstraksi data yang berupa text tersebut dengan text mining. Dan pada paper ini akan menggunakan rancangan system ekstraksi dalam penelitian untuk menemukan informasi spatio temporal penyebaran penyakit menular [1].



Gambar 2 Rancangan Sistem Ekstraksi
Sumber: Prasetyo et al (2008)

Sebagaimana tertera pada Gambar 2, berikut adalah keterangan ringkas mengenai sistem ekstraksi informasi yang akan dikembangkan. Sistem ekstraksi informasi memiliki input berupa kumpulan artikel-artikel teks yang diterima dari hasil pencarian data dengan *Web Content Mining* dan peringkasan dokumen (*summarization module*). Kalimat-kalimat dalam artikel-artikel tersebut kemudian akan dipilih lebih lanjut dengan mencari kalimat-kalimat yang mempunyai kata-kata penting yang dibutuhkan untuk mengisi tabel *template* dengan merujuk

sekumpulan *data base* sinonim kata (yang dirasa penting), nama operator, jumlah pembelian, tanggal penjualan dan lokasi penjualan.

Kalimat-kalimat yang terpilih untuk diproses kemudian diproses oleh modul pemarka jenis kata untuk menentukan jenis kata. Tugas modul *tokenizer* akan dijalankan oleh modul pemarka jenis kata karena kalimat-kalimat dalam Bahasa Indonesia menggunakan spasi sebagai pemisah kata, Daftar kata berikut informasi mengenai jenis kata akan dimasukkan

dalam kamus elektronik modul analisa kalimat.

Hasil keluaran modul analisa kalimat akan diterima oleh modul *Template Filler* yang akan menentukan bagian-bagian kalimat yang dirasakan perlu dimasukkan dalam tabel *template*. Selanjutnya modul *Template Merger* akan menggabungkan tabel-tabel *template* yang data-datanya mempunyai kemiripan. Hasil akhir dari sistem ekstraksi informasi adalah tabel-tabel berisikan data-data mengenai penjualan pulsa menurut operator, tanggal dan lokasi penjualan yang kemudian akan dianalisis untuk menghasilkan jumlah penduduk yang mampu membeli pulsa di suatu daerah untuk kurun waktu tertentu.

Mining Data Penduduk Miskin.

1. [http://www.bps.go.id/releases/files/kemiskinan-01jul09.pdf?](http://www.bps.go.id/releases/files/kemiskinan-01jul09.pdf) sebagai website yang dimiliki oleh pemerintah yang bertugas untuk mengumpulkan data statistik untuk kepentingan pemerintah dan umum. BPS dalam websitenya telah mengeluarkan data tentang jumlah penduduk miskin pada bulan maret 2009. Data yang disajikan sangat lengkap yaitu berisi jumlah penduduk miskin terkahir dan perbandingannya dengan pada bulan yang sama pada tahun 2008. Dalam laporan tersebut BPS menyebutkan bahwa jumlah penduduk miskin pada maret 2009 adalah sebesar 32.53 juta (14.15 persen dari total penduduk Indonesia) angka ini mengalami penurunan jika dibandingkan dengan bulan yang sama pada tahun 2008 yaitu berjumlah 34.96 juta (15.42 persen dari total penduduk Indonesia). Serta mengeluarkan juga data penduduk miskin untuk perprovinsi juga jumlah penduduk miskin dipedesaan dan diperkotaan. Data

dalam website ini lebih bersifat statistik.

2. <http://www.kesimpulan.co.cc/2009/07/badan-pusat-statistik-atau-bps-laporkan.html>

Website didapatkan informasi tentang penduduk miskin pada tahun maret 2009. Data dan informasi yang dikeluarkan BPS. Namun demikian data website ini mencoba menjabarkan penyebab penurunan angka kemiskinan tersebut dan menjelaskan perhitungan menjadi dasar perhitungan angka kemiskinan. Data dalam website ini lebih bersifat deskripsi informasi.

Diatas adalah hasil dari pencarian penduduk miskin menggunakan webmining. Hampir dari banyak website yang menyediakan data tentang penduduk miskin menggunakan data yang dirilis oleh Badan Pusat Statistik (BPS). Jadi penulis hanya mengambil dua saja yaitu data yang dikeluarkan oleh BPS dan Sebuah website yang mencoba menganalisa data tersebut untuk berbagai keperluan dan dari berbagai sudut pandang.

Mining Data Penjualan Pulsa Telpon Genggam

<http://www.koran-jakarta.com/ver02/detail-news.php?id=1221&&idkat=60>.

Website ini memberikan informasi bahwa perusahaan-perusahaan operator sellular telah mengalami penurunan penjualan pulsa. Dari tiga operator besar yaitu XL, Indosat dan telkomsel, hanya telkomsel saja yang tidak mengalami penurunan pendapatannya dari pulsa isi ulang. Sedangkan XL dan indosat mengalami penurunan pendapatan.

KESIMPULAN

Pada penelitian ini kami menggunakan teknik webmining untuk mencari data-data yang kami perlukan dari internet. Pendekatan webmining yang kami gunakan adalah webcontent mining dengan memanfaatkan mesin pencari yaitu google.com dan melakukan ekstraksi atau pemilihan data terhadap website-website yang menyediakan informasi atau data yang terkait dengan penelitian ini.

Hasil dari web content mining menunjukkan bahwa data yang dikeluarkan oleh BPS mengundang banyak tanya. Hal ini dapat dilihat dari banyaknya website yang mencoba membahas tentang penurunan jumlah penduduk miskin dari berbagai sudut pandang. Membandingkan jumlah penduduk miskin dengan penjualan pulsa telepon genggam menjadi fokus pada penelitian ini. Hasil dari penelitian ini menunjukkan bahwa menurunnya jumlah penduduk miskin yang seharusnya dapat mendongkrak daya beli masyarakat ternyata tidak mempengaruhi penjualan pulsa telepon genggam yang malah mengalami penurunan. Hal ini dapat dibuktikan dari hasil webmining penjualan pulsa, dimana hasil webmining tersebut menunjukkan bahwa dua dari tiga operator besar mengalami penurunan penjualan pulsa telepon genggam.

Web mining merupakan bagian dari data mining yang fungsinya menggali data atau mencari informasi dari internet.

DAFTAR PUSTAKA

- [1] Abdurrahman et al.2006. "Pemodelan Web Usage Mining

Untuk Mengelola E-Commerce". Prosiding Konferensi Nasional Teknologi Informasi & Komunikasi untuk Indonesia, Bandung.

- [2] Achmad Solichin. 2008. "Web Usage Mining dengan Google Analytics: Studi Kasus Situs Achmatim.Net".
- [3] Ai, Jiye. Laffey, James. 2007. "Web Mining as a Tool for Understanding Online Learning". Vol. 3, No. 2. MERLOT Journal of Online Learning and Teaching.
- [4] Bettina Berendt, Andreas Hotho, and Gerd Stumme. 2002. "Towards Semantic Web Mining".
- [5] Christopher C. Yang, Y. C. Wong, 2006." MINING CONSUMER OPINIONS FROM THE WEB".
- [6] D. Hand, H. Mannila, and P. Smyth, 2001. "*Principles of Data Mining*". Cambridge, MA: MIT Press.
- [7] Liu, Bing. 2005. "Web Content Mining". WWW-2005, Chiba, Japan.
- [8] Liwen Vaughan, Justin You. 2007. "Keyword Enhanced Web Structure Mining for Business Intelligence".
- [9] Mining".pp.44-49. Proc. Of Konferensi Nasional Sistem & Informatika 2008, Bali.
- [10] Peter Toth. 2007. "Applying Web-mining Methods for Analysis in Virtual Learning Environment".
- [11] Prasetyo, Bowo et al.2008. "Desain Sistem Analisa Spatio-Temporal Penyebaran Penyakit Tropis Memakai Web