

Aplikasi Pohon Keputusan sebagai Alat Penggalian Data pada Toko Belanja Online

Vanya Deasy Safrina - 13513035
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13513035@itb.ac.id

Abstrak—Makalah ini akan mendemonstrasikan sebuah aplikasi dari pohon keputusan sebagai alat *data mining* atau dikenal dengan sebutan penggalian data pada toko belanja online. Penggalian data memiliki kemampuan untuk mengklasifikasi, memprediksi, mengestimasi, dan mengenali pola dengan menggunakan basis data yang ada. Basis data yang ada tersebut terdiri dari informasi penting yang akan digunakan dalam pengambilan keputusan. Pohon keputusan akan digunakan untuk mengidentifikasi nilai dari kumpulan informasi pada basis data. Pada makalah ini, pohon keputusan akan digunakan untuk memprediksi kebutuhan pelanggan sebuah toko belanja online.

Kata Kunci—Pohon, pohon keputusan, data mining.

I. PENDAHULUAN

Pada zaman berteknologi canggih ini, toko belanja *online* bukanlah sesuatu yang aneh. Ada ratusan bahkan ribuan toko belanja online yang dapat dinikmati di dalam negeri. Mulai dari pakaian hingga perabotan rumah tangga sekarang terjual dengan bebas di internet.

Hal yang perlu dilakukan oleh seorang pelanggan adalah mengetik alamat laman toko dan selanjutnya yang perlu dilakukan hanya mencari barang yang diinginkan lalu mememesannya. Pelanggan yang malas dengan kemacetan kota atau pelanggan yang sibuk di kantor dapat dengan mudah dan cepat memesan barang yang diinginkannya tapi harus pergi ke toko untuk membelinya.

Tapi dengan banyaknya toko online yang beredar, tentu saja akan ada kesulitan yang dihadapi. Pemilik toko harus siap dengan persaingan yang ketat. Penyebab dari persaingan yang ketat ini adalah tidak adanya batas waktu dan jarak antara pelanggan dan toko. Pelanggan dapat memesan kapan saja dan dari mana saja. Dengan begitu, sebuah toko online bukan saja hanya memiliki pesaing dari satu kota maupun satu provinsi. Toko online memiliki pesaing dalam jangkauan sebuah negara.

Untuk mengatasi persaingan yang ketat, maka pemilik toko online harus memiliki strategi agar toko mereka dapat terus bertahan. Strategi yang dilakukan pemilik toko sangat beragam. Mulai dari memasang iklan tentang toko di berbagai macam media hingga memberikan diskon besar-besaran dengan beberapa syarat dan

ketentuan.

Salah satu strategi yang dapat dilakukan untuk mengatasi persaingan dalam industri ini juga adalah dengan penggunaan data mining.

Data mining dapat didefinisikan sebagai proses untuk menemukan sebuah pola dan tren dalam basis data dan menggunakan informasi yang didapatkan dari pola tersebut untuk membuat model prediksi. Data mining juga dapat didefinisikan sebagai proses pemilihan dan pembuatan model menggunakan data yang ada. Data mining bukanlah sebuah metode baru. Metode ini telah digunakan di berbagai bidang, contohnya pada institusi keuangan.

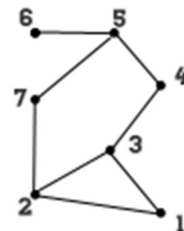
Toko belanja online juga dapat mengolah data yang didapat dari data registrasi dan pemesanan pelanggan. Dari data tersebut, pohon keputusan dapat dibuat. Pohon keputusan berguna untuk pengambilan keputusan yang tepat bagi toko tersebut. Keputusan dapat berupa pengurangan salah satu item yang dianggap tidak terlalu diminati dan sebagainya.

II. LANDASAN TEORI

A. Graf

Graf dapat didefinisikan sebagai pasangan himpunan (V,E) , dimana V adalah himpunan tidak kosong dari simpul-simpul yang ada di dalam graf dan E adalah himpunan dari sisi di dalam graf yang menghubungkan sepasang simpul^[X].

Contoh dari graf dapat dilihat sebagai berikut.

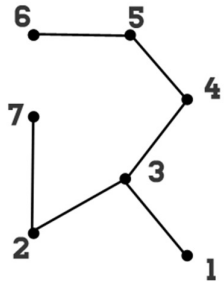


Gambar 2.1 Graf
(sumber : dokumen penulis)

B. Pengertian Pohon

Pohon adalah sebuah graf tak berarah, terhubung, yang tidak mengandung sirkuit. Terhubung artinya setiap verteks dalam graf terhubung dengan yang lainnya. Pohon juga tidak mengandung sirkuit. Graf dikatakan memiliki sirkuit jika ada lintasan yang berawal dan berakhir pada simpul yang sama.

Contoh dari pohon dapat dilihat sebagai berikut.



Gambar 2.2 Pohon
(sumber : dokumen penulis)

C. Pohon Keputusan

Di dalam kehidupan manusia sehari-hari, manusia akan selalu dihadapkan dengan berbagai macam masalah. Mulai dari masalah yang sangat sederhana hingga yang kompleks. Masalah sederhana adalah sebuah masalah yang dapat dengan mudah dan cepat diselesaikan karena faktor-faktor yang berkaitan dengan masalah itu tidak begitu banyak. Sedangkan masalah kompleks adalah sebuah masalah yang penyelesaiannya harus memperhatikan setiap faktor-faktor permasalahan yang saling berkaitan. Untuk menghadapi masalah-masalah tersebut, manusia mulai mengembangkan sebuah sistem yang dapat membantu manusia agar dapat dengan mudah mampu untuk menyelesaikan masalah-masalah tersebut.

Pohon Keputusan adalah salah satu contoh pohon yang sangat populer karena mudah untuk diinterpretasi oleh manusia. Konsep dari pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan.

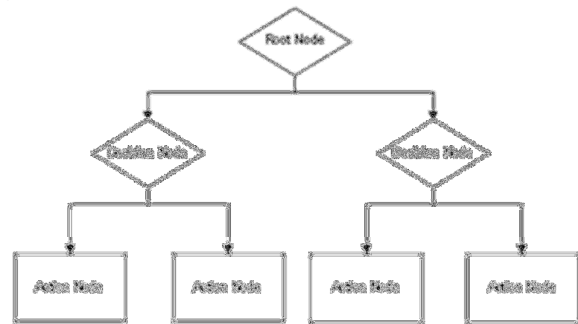
Dengan pohon keputusan, manusia dapat dengan mudah melihat mengidentifikasi dan melihat hubungan antara faktor-faktor yang mempengaruhi suatu masalah dan dapat mencari penyelesaian terbaik dengan memperhitungkan faktor-faktor tersebut. Pohon keputusan ini juga dapat menganalisa nilai resiko dan nilai suatu informasi yang terdapat dalam suatu alternatif pemecahan masalah. Peranan pohon keputusan ini sebagai alat Bantu dalam mengambil keputusan (decision support tool) telah dikembangkan oleh manusia sejak perkembangan teori pohon yang dilandaskan pada teori graf. Kegunaan pohon keputusan yang sangat banyak ini membuatnya telah dimanfaatkan oleh manusia dalam berbagai macam sistem pengambilan keputusan.

Pohon keputusan memiliki banyak kelebihan dalam memodelkan persoalan, diantaranya:

1. mudah untuk diinterpretasikan,
2. mudah untuk dimengerti,
3. memungkinkan untuk menjelajah semua kemungkinan,
4. lebih akurat dalam hal pemodelan, dan
5. pengambilan keputusan yang rumit dapat diubah menjadi lebih sederhana.

Walaupun pohon keputusan memiliki banyak kelebihan, setiap metode tentu ada kelemahannya. Diantara kelemahan dari pemodelan persoalan menggunakan pohon keputusan adalah

1. hasil atau solusi dari pohon keputusan bergantung kepada desain dari pohon itu dibuat,
2. adanya kesulitan dalam pembuatan pohon keputusan yang optimal, dan
3. adanya kesulitan membuat pohon keputusan yang lebih umum.



Gambar 2.3 Pohon Keputusan

(sumber :

<http://pzuh.blogspot.com/2011/10/actionsript-30-decision-tree.html> diakses pada 10 Desember 2014 pukul 14:27 WIB)

D. Penggalan Data

Penggalan data (bahasa Inggris: data mining) adalah ekstraksi pola yang menarik dari data dalam jumlah besar [6]. Suatu pola dikatakan menarik apabila pola tersebut tidak sepele, implisit, tidak diketahui sebelumnya, dan berguna. Pola yang disajikan haruslah mudah dipahami, berlaku untuk data yang akan diprediksi dengan derajat kepastian tertentu, berguna, dan baru.

Kehadirannya dilatar belakangi dengan problema data explosion yang dialami akhir-akhir ini dimana banyak organisasi telah mengumpulkan data sekian tahun lamanya (data pembelian, data penjualan, data nasabah, data transaksi dsb.). Alasan utama mengapa penggalan data atau dalam bahasa Inggris dikenal sebagai *data mining* sangat menarik perhatian industri informasi dalam beberapa tahun belakangan ini adalah karena tersedianya data dalam jumlah yang besar dan semakin besarnya kebutuhan untuk mengubah data tersebut menjadi informasi dan pengetahuan yang berguna. Banyak persoalan yang dapat diselesaikan dengan penggalan data tersebut.

III. Aplikasi Pohon Keputusan pada Toko Belanja Online

Pohon keputusan pada bidang industri pada umumnya digunakan untuk pengambilan keputusan strategi untuk bertahan di pasar. Sama halnya dengan toko belanja online. Pada satu sisi, toko belanja online memiliki keuntungan yang tidak dimiliki oleh toko belanja pada umumnya. Toko belanja online memiliki basis data pelanggan yang cukup lengkap dibandingkan dengan toko belanja offline.

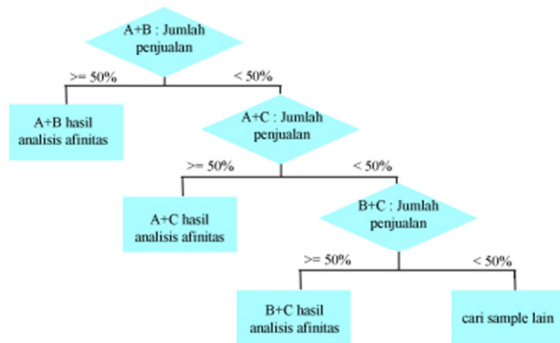
Setiap kali pelanggan ingin membeli suatu barang pada suatu toko belanja online, pelanggan tersebut harus terlebih dahulu mengisi data diri yang diharuskan oleh pihak toko. Data diri tersebut dapat berupa nama, alamat, umur, dan lain sebagainya. Meskipun tidak semua toko belanja online melakukan itu karena manusia sekarang ini lebih menyukai hal praktis.

Walaupun data diri yang dimiliki oleh pihak toko hanya berupa email dan nama pelanggan, ada satu basis data lagi yang dimiliki oleh pihak toko. Data tersebut adalah data pembelian pelanggan. Pihak toko memiliki data rincian lengkap mengenai data pembelian pelanggan, mulai dari tanggal pembelian, hingga harga barang tersebut. Dengan data tersebut, pihak toko dapat melakukan penggalan data dengan memanfaatkan pohon keputusan untuk menaikkan daya saing tokonya. Beberapa keuntungan yang didapat jika melakukan penggalan data adalah peningkatan kesetiaan pelanggan, penemuan keuntungan yang tersembunyi, dan mengurangi keluhan pelanggan.

Data yang akan dibahas kali ini adalah sebagai berikut.

A. Analisis Afinitas

Analisis Afinitas atau yang biasa disebut analisis keranjang merupakan analisis atau teknik penggalan data yang berguna untuk mendapatkan hubungan yang terdapat dari benda-benda yang dibeli oleh pelanggan. Contoh, analisis afinitas dapat memberitahu pihak toko bahwa pelanggan sering membeli produk A dan produk B secara bersamaan atau dalam jangka waktu yang singkat. Dengan contoh seorang pelanggan yang telah berbelanja sebanyak 3 item di toko online tersebut. 3 item tersebut adalah A, B, dan C. Pohon keputusan yang dapat dibuat untuk analisis afinitas adalah sebagai berikut.



Gambar 3.1 Analisis Afinitas (sumber : dokumen penulis)

Gambar 3.1 menjelaskan bahwa pohon keputusan pertama-tama akan menggabungkan produk A dan B. Jika produk A dan B dibeli oleh lebih dari atau sama dengan 50% jumlah total pelanggan, maka A dan B merupakan hasil dari analisis afinitas tersebut. Jika produk A dan B dibeli oleh kurang dari 50% jumlah total pelanggan, maka akan dilakukan pengecekan pada produk A dan C. Jika produk A dan C dibeli oleh kurang dari 50% jumlah total pelanggan, maka dilakukan pengecekan pada produk B dan C. Jika produk B dan C dibeli oleh lebih dari atau sama dengan 50% jumlah total pelanggan, maka B dan C adalah hasil dari analisis afinitas tersebut. Jika produk B dan C dibeli oleh kurang dari 50% jumlah total pelanggan, maka akan dilakukan sample pelanggan lain.

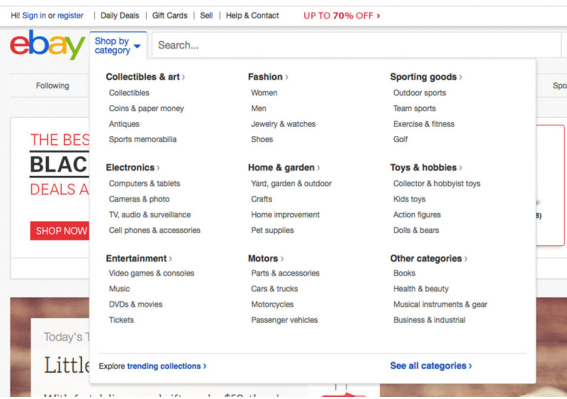
Setelah pihak toko mengetahui hasil dari analisis afinitas dengan menggunakan pohon keputusan yang telah tertera pada gambar 3., maka pihak toko dapat mengambil beberapa keputusan dari data yang didapatkan itu. Berikut merupakan pohon keputusan yang didapat dari analisis afinitas.



Gambar 3.2 Pohon Keputusan Hasil Analisis Afinitas (sumber : dokumen penulis)

Penjelasan dari gambar 3.1 adalah sebagai berikut. Setelah menggunakan analisis afinitas, didapatkan bahwa produk A dan B sering dibeli secara bersamaan atau dalam waktu berdekatan oleh pelanggan. Produk A dan B masuk ke dalam salah satu kategori seperti contoh yang tertera dalam gambar 3.2. Jika produk A dan B berada pada kategori yang sama, maka keputusan terbaik yang diambil oleh pihak toko adalah menjual produk A dan B dalam satu paket dan memberikan produk tersebut diskon. Hal ini tepat untuk dilakukan karena pemotongan harga yang dilakukan dalam satu paket lebih kecil dibandingkan pemotongan harga per produk. Selain itu, pelanggan merasa bahwa mereka membutuhkan keduanya sehingga tidak akan masalah untuk membeli kedua produk tersebut dalam satu paket.

Jika produk A dan B tidak berada dalam satu kategori yang sama, maka akan mengherankan jika produk A dan B dijual dalam satu paket sehingga keputusan tersebut tidak dapat diambil. Keputusan yang dapat diambil disini adalah menyarankan pelanggan yang membeli produk A agar membeli produk B atau sebaliknya. Jika pelanggan membeli produk B maka produk A akan disarankan untuk dibeli.



Gambar 3.3 Contoh Kategori Produk (sumber : <http://www.ebay.com/> diakses pada 10 Desember 2014 pukul 13:40 WIB)

B. Prediksi Penjualan

Prediksi penjualan adalah sebuah metode dimana pihak toko mengira atau memprediksi produk apa yang akan laku pada jangka waktu tertentu. Contoh pada saat musim hujan, produk seperti payung, jas hujan, dan jaket akan memiliki daya jual lebih tinggi dibandingkan yang lain. Tentu saja prediksi tidak selalu tepat tapi dengan penggalan data, dapat diputuskan teknik apa yang harus digunakan agar pihak toko dapat menerima keuntungan yang besar.

Pertama yang perlu dilakukan adalah menentukan tema dari jangka waktu tertentu. Hal tersebut dapat dibagi berdasarkan musim, hari raya, mode, dan lain sebagainya. Pada pembahasan kali ini, akan dibahas secara umum.

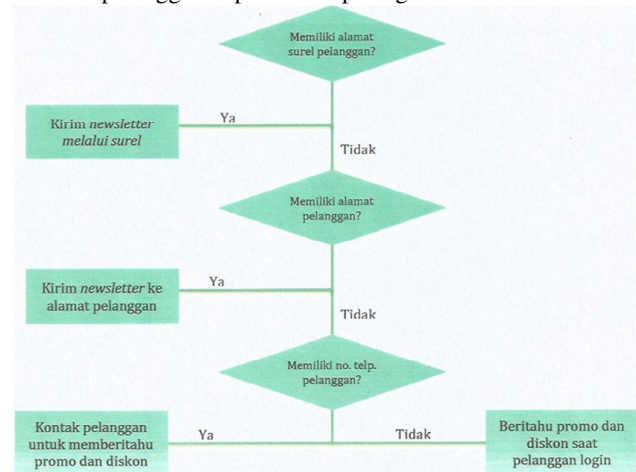
Seperti yang terlihat pada gambar 3.3, jika produk A berada dalam kategori barang yang cocok untuk jangka waktu tertentu, maka akan dilakukan perbandingan penjualan produk A pada waktu tersebut dengan waktu sebelumnya. Jika daya jual produk A pada waktu tersebut meningkat, maka hal yang perlu dilakukan pihak toko adalah menyarankan pembelian produk tersebut kepada pelanggan. Jika daya jual produk A menurun, maka perlu dilakukan promo atau diskon harga pada produk A untuk menarik pelanggan.

Jika produk A tidak berada dalam kategori barang yang cocok untuk jangka waktu tersebut, akan dilakukan keputusan sebagai berikut. Jika produk A berada dalam kategori produk yang cocok untuk jangka waktu dekat, jangan lakukan diskon pada produk A. Produk A dapat diiklankan sehingga pelanggan dapat membelinya lebih awal. Iklan dapat dilakukan dengan memberikan *newsletter* ke surel pelanggan ataupun dengan memasang iklan di berbagai sosial media yang ada. Jika produk A tidak berada dalam kategori produk yang cocok untuk jangka waktu dekat, pihak toko tidak perlu memberikan diskon atau semacamnya.

C. Pemanfaatan Data Diri Pelanggan

Prediksi penjualan adalah sebuah metode dimana pihak toko mengira atau memprediksi produk apa yang akan laku pada jangka waktu tertentu. Contoh pada saat musim hujan, produk seperti payung, jas hujan, dan jaket akan memiliki daya jual lebih tinggi dibandingkan yang lain. Tentu saja prediksi tidak selalu tepat tapi dengan penggalan data, dapat diputuskan teknik apa yang harus digunakan agar pihak toko dapat menerima keuntungan yang besar.

Keputusan dapat didapatkan dengan melihat data diri apa saja yang dimiliki oleh pihak toko. Beberapa toko belanja online mengharuskan pendaftaran yang cukup lengkap meliputi nama, alamat, nomor telepon, dan sebagainya. Beberapa toko belanja online lainnya hanya mengharuskan pendaftaran berupa email. Sehingga pohon keputusan perlu dibuat. Pohon keputusan pemanfaatan data diri pelanggan dapat dilihat pada gambar berikut.



Gambar 3.4 Pohon Keputusan Pemanfaatan Data Diri Pelanggan (sumber : dokumen penulis)

Seperti yang terlihat pada gambar 3.4, jika pihak toko memiliki identitas berupa alamat surel pelanggan dan pihak toko memiliki otoritas untuk mengirimkan iklan ke surel pelanggan maka pihak toko dapat mengirimkan *newsletter* secara berkala ke surel pelanggan. Dengan begitu, pelanggan akan dapat mengetahui promo dan diskon apa saja yang ada di toko tanpa membuka website toko belanja online. Selain itu, dilihat dari segi biaya, pengiriman surel tidak memakan biaya sama sekali. Jika pihak toko tidak memiliki otoritas untuk mengirimkan iklan ke surel pelanggan, maka pihak toko harus mencari jalan lain. Jika pihak toko memiliki nomor telepon pelanggan dan memiliki otoritas untuk menelepon pelanggan, maka gunakan *customer service* untuk memberitahu pelanggan mengenai promo dan diskon yang ada melalui jaringan telepon. Jika pihak toko tidak memiliki nomor telepon pelanggan atau tidak memiliki otoritas untuk menelepon pelanggan, maka pihak toko harus mengambil bentuk keputusan lain. Jika pihak toko memiliki alamat tempat tinggal pelanggan, maka pihak toko dapat mengirimkan *newsletter* melalui surat.



Gambar 3.3 Pohon Keputusan berdasarkan Prediksi Penjualan
(sumber : dokumen penulis)

IV. KELEMAHAN ANALISIS

Dalam analisis yang digunakan pada makalah ini, tidak diperhitungkan masalah biaya secara detil. Seperti pada bagian prediksi penjualan, pihak toko dianjurkan untuk memberikan promo ataupun diskon tanpa memperhatikan biaya pemotongan dan sebagainya.

V. KESIMPULAN

Pohon keputusan dapat diterapkan sebagai alat penggalian data pada toko belanja online. Penggalian data dapat dilakukan dengan tiga cara yaitu analisis afinitas, prediksi penjualan, dan pemanfaatan data diri pelanggan. Analisis afinitas memanfaatkan data pembelian pelanggan beserta waktu pembelian. Prediksi penjualan memanfaatkan data pembelian pelanggan secara keseluruhan. Dan pemanfaatan data diri pelanggan memanfaatkan data diri pelanggan yang dimiliki.

VII. UCAPAN TERIMA KASIH

Pertama-tama, penulis mengucapkan terima kasih kepada Tuhan Yang Maha Esa. Selanjutnya ucapan terima kasih saya berikan kepada pak Rinaldi Munir dan bu Harlili yang telah memberikan tugas membuat makalah ini sehingga dapat bermanfaat untuk penulis pribadi maupun orang lain. Penulis juga berterima kasih kepada seluruh orang yang telah mendukung serta memberikan informasi sehingga makalah ini dapat diselesaikan.

REFERENSI

- Dr. Zaki S. Towfik, "Journal of Kufa for Mathematics and Compute" Vol.1, No.5, May, 2012, pp. 102-109
<https://fairuzelsaid.wordpress.com/2009/11/24/data-mining-konsep-pohon-keputusan/> diakses pada 8 Desember 2014 pukul 20:31 WIB
http://www.unej.ac.id/files/pdf2/Ari_Fadli_Konsep_Data_Mining.pdf diakses 8 Desember 2014 pukul 20:31 WIB
<http://dua7an.blogspot.com/2013/12/tentang-pohon-keputusan-decision-tree.html> diakses pada 9 Desember 2014 pukul 20:42 WIB
 Pramudiono, Iko. Pengantar Data Mining: Menambang Permata Pengetahuan di Gunung Data. <http://www.ilmukomputer.com>
 Kurt Thearling, An Introduction to Data Mining

- <http://www.amazon.com/Data-Mining-Concepts-Techniques-Management/dp/1558609016/qid=1278582726> diakses pada 9 Desember 2014 pukul 22:06 WIB
<https://blog.kissmetrics.com/data-mining/> diakses pada 9 Desember 2014 pukul 22:50 WIB
 J. Han et al., 2006, Data Mining: Concepts and Techniques
<http://www.hasbihtc.com/kekurangan-dan-kelebihan-toko-online.html> diakses pada 10 Desember 2014 pukul 14:30 WIB

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 7 Desember 2014

Vanya Deasy Safrina
13513035