

Penerapan Pohon Keputusan dalam Online Advertising

Ivan Andrianto - 13513039
Program Magister Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
andrianto.ivan@yahoo.co.id

Abstract — Online advertising sudah banyak digunakan oleh perusahaan-perusahaan untuk mempromosikan produk atau layanan mereka agar lebih dikenal oleh pengguna Internet. Perkembangan online advertising cukup pesat dalam beberapa tahun terakhir.

Salah satu kelebihan online advertising dibandingkan dengan iklan melalui media lain adalah dapat menampilkan iklan yang sesuai untuk setiap user. Agar iklan yang ditampilkan sesuai, harus ada suatu sistem yang dapat menentukan iklan apa yang paling sesuai untuk ditampilkan. Oleh karena itu, penulis mencoba menerapkan pohon keputusan dalam sistem online advertising.

Keywords—online advertising, pohon keputusan

I. PENDAHULUAN

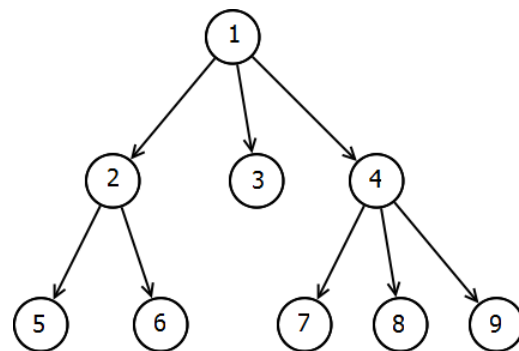
Internet sudah menjadi hal yang umum digunakan oleh masyarakat dunia. Lebih dari satu miliar orang mengakses Internet setiap harinya. Mengakses Internet dapat dilakukan dengan mudah tidak hanya melalui komputer desktop atau laptop. Kini, perangkat mobile seperti tablet dan smartphone juga banyak digunakan untuk mengakses Internet. Dengan tingkat kepopuleran seperti itu, Internet sudah menjadi salah satu media periklanan yang banyak digunakan oleh perusahaan-perusahaan mulai dari skala kecil hingga besar untuk membuat brand serta produk dan/atau jasa mereka lebih dikenal oleh pengguna Internet sehingga bisnis mereka bisa berkembang. Selain perusahaan, pemilik situs atau blog juga dapat meningkatkan kepopuleran situs atau blognya melalui iklan di Internet. Dibandingkan dengan iklan melalui media cetak, radio, atau televisi, biaya yang diperlukan untuk menampilkan iklan di Internet relatif lebih murah.

Namun, menampilkan iklan di Internet dengan memperhatikan isi dari halaman yang menayangkan iklan serta sesuai dengan perilaku dan demografi pengguna akan lebih memberikan hasil yang maksimal bagi pengiklan dan penerbit iklan karena akan menghasilkan CTR (Click Through Rate) lebih tinggi daripada iklan yang ditampilkan secara acak. Penentuan iklan mana yang paling tepat untuk ditampilkan di suatu halaman tertentu untuk setiap pengguna dapat ditentukan dengan menggunakan pohon keputusan.

II. LANDASAN TEORI

1. Pohon

Pohon adalah graf terhubung tidak berarah yang tidak memiliki sirkuit. Karena tidak memiliki sirkuit, pohon tidak dapat mengandung sisi ganda atau loop. Sepasang simpul yang terhubung di dalam suatu pohon selalu terhubung dengan lintasan tunggal. Dengan kata lain, tidak boleh ada lintasan ganda pada suatu pohon.



Gambar 1: Pohon

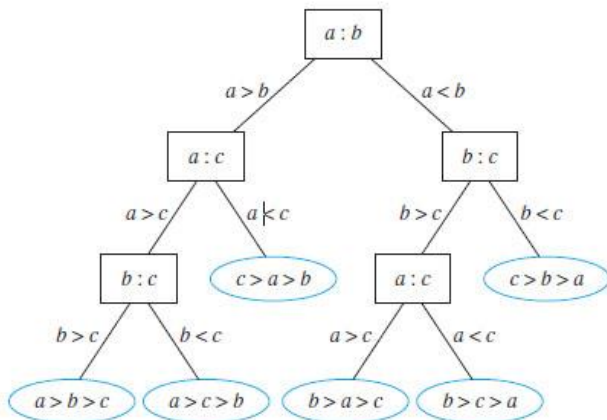
(Sumber gambar: http://1.bp.blogspot.com/-vs0m0CKi4bg/UYeNHPW5DzI/AAAAAAAAArw/Q_Hiphb6sTM/s1600/p1.png. Diakses pada 10 Desember 2014 pukul 13.03)

Salah satu jenis pohon adalah pohon berakar, yang merupakan pohon dengan salah satu simpulnya ditetapkan sebagai akar dan setiap sisi menjauhi akar. Gambar di atas adalah contoh pohon berakar. Pada pohon berakar di atas, simpul 5 dan simpul 6 merupakan anak dari simpul 2, sedangkan simpul 2 merupakan orang tua dari simpul 5 dan simpul 6. Simpul 5 merupakan saudara kandung dari simpul 6, begitu pula sebaliknya. Derajat suatu simpul merupakan jumlah anak pada simpul tersebut. Simpul 2 berderajat 2, sedangkan simpul 5 berderajat 0. Simpul yang berderajat 0 atau simpul yang tidak mempunyai anak disebut daun, sedangkan simpul yang berderajat lebih besar dari 0 atau simpul yang mempunyai anak disebut simpul internal. Pada pohon berakar di atas, yang merupakan daun adalah simpul 3, 5, 6, 7, 8, dan 9. Setiap simpul pada suatu pohon berakar berada pada suatu level. Simpul 1 yang merupakan akar berada pada level 0; simpul 2, 3, dan 4 berada pada level 1; sedangkan simpul 5, 6, 7, 8, dan 9 berada pada level 2.

Kedalaman suatu pohon adalah level maksimum dari pohon tersebut. Kedalaman pohon di atas adalah 2.

2. Pohon Keputusan

Pohon keputusan merupakan pohon berakar yang seluruh simpul internalnya merupakan suatu pertimbangan dan setiap sisi yang keluar dari simpul internal tersebut merepresentasikan kemungkinan hasil keputusan. Pohon keputusan dapat diterapkan untuk memecahkan berbagai masalah yang memerlukan pengambilan-pengambilan keputusan sehingga diperoleh suatu solusi.



Gambar 2: Pohon Keputusan

(Sumber gambar: Discrete Mathematics and Its Applications 7th Edition halaman 761)

Gambar di atas adalah contoh pohon keputusan yang membandingkan mengurutkan 3 buah variabel dari yang paling besar menuju yang paling kecil. Setiap simpul pada pohon keputusan dapat merepresentasikan suatu Boolean test, perbandingan, atau pengukuran lainnya. Pada contoh di atas, setiap simpul membandingkan 2 buah variabel. Setiap sisi yang keluar dari suatu simpul merepresentasikan berbagai kemungkinan yang ada dari hasil pengujian, perbandingan, atau pengukuran pada simpul tersebut. Setiap sisi akan menuju ke simpul anak yang dapat merupakan simpul internal atau daun. Jika simpul anak tersebut merupakan simpul internal, akan dilakukan pengujian, perbandingan, atau pengukuran seperti pada simpul orang tuanya untuk selanjutnya menuju ke simpul pada tingkat yang lebih tinggi. Demikian seterusnya sampai mencapai daun.

III. Tipe-tipe Online Advertising

1. Display Advertising

Pada tipe ini, iklan ditampilkan dalam bentuk banner yang ditempatkan pada situs-situs pihak ketiga. Iklan yang ditampilkan dapat berupa teks, gambar, kombinasi teks dan gambar, atau media-media lainnya seperti animasi, audio, atau video.

2. Affiliate Marketing

Affiliate marketing merupakan jenis online advertising yang menampilkan produk/jasa yang dijual oleh pengiklan pada

situs-situs di Internet. Pengiklan hanya perlu membayar jika produk/jasa yang dijual dibeli oleh orang melalui iklan.

3. Social Network Advertising

Iklan ditampilkan pada situs-situs jejaring sosial seperti Facebook.

4. Search Engine Marketing

Bentuk iklan ini meningkatkan kemunculan suatu situs pada hasil-hasil pencarian search engine. Cara yang digunakan ada berbagai macam seperti paid placement, contextual advertising, atau menggunakan teknik search engine optimization untuk meningkatkan posisi di hasil-hasil pencarian.

5. Mobile Advertising

Iklan ditampilkan pada aplikasi-aplikasi untuk ponsel. Salah satu alasan menggunakan mobile advertising adalah lebih banyak orang yang menggunakan perangkat mobile seperti smartphone sepanjang hari dibandingkan menggunakan komputer desktop atau laptop sepanjang hari. Teknologi location-based pada perangkat mobile dapat menentukan dimana lokasi dari seorang pengguna sehingga iklan yang ditampilkan dapat sesuai dengan lokasi pengguna.

IV. Jenis Online Advertising Berdasarkan Model Pendapatan

1. Pay Per View

Model pendapatan ini menghitung berdasarkan berapa kali suatu halaman atau file yang diiklankan diakses melalui iklan yang ditayangkan.

2. Cost Per Mil (CPM)

Pengiklan wajib membayar sejumlah tertentu uang untuk setiap 1.000 halaman atau file yang diakses meskipun tidak ada yang mengklik iklannya. Model ini tidak mempedulikan jumlah klik yang dilakukan oleh pengunjung.

3. Cost Per Click (CPC)

CPC merupakan salah satu jenis model pendapatan yang paling sering digunakan pada online advertising saat ini. Pengiklan wajib membayar sejumlah uang setiap kali iklannya diklik oleh pengunjung situs.

4. Cost Per Action

Model pembayaran ini mewajibkan pengiklan untuk membayar hanya jika pengunjung yang mengklik suatu iklan melakukan suatu aksi seperti membeli suatu produk/jasa, melakukan registrasi pada website pengiklan, atau melakukan aksi-aksi lain sesuai kesepakatan. Jika iklan hanya ditampilkan saja tetapi tidak diklik oleh pengunjung, pengiklan tidak perlu membayar.

V. Ad Targeting

Ad targeting adalah menentukan kepada siapa iklan tersebut ditampilkan. Setiap pengiklan dapat menentukan di

mana dan kepada siapa iklannya boleh ditayangkan. Dengan demmikian, sistem periklanan tidak sembarangan dalam menampilkan iklan. Ada beberapa jenis ad targeting yang umum digunakan, diantaranya:

1. No targeting

Menampilkan iklan yang sama atau menampilkan iklan secara acak tanpa mempedulikan isi dari halaman tempat iklan tersebut ditampilkan atau informasi yang berkaitan dengan pengunjung. Sistem ini kurang baik karena iklan yang ditampilkan belum tentu sesuai dengan isi dari halaman yang memuat iklan tersebut maupun topik-topik yang disukai oleh pengunjung sehingga kurang efektif.

2. Contextual Advertising

Menggunakan analisis frekuensi kemunculan suatu keyword, dan strkuktur link pada suatu website untuk menentukan isi dari suatu halaman dan iklan mana yang cocok untuk halaman tersebut. Setiap pengiklan akan menentukan keyword-keyword yang ditarget untuk suatu iklan.. Dengan sistem ini, iklan yang ditampilkan pada suatu halaman akan terlihat sesuai dengan isi dari halaman tersebut.

3. Placement Targeting

Pengiklan memilih menempatkan iklan dengan menetapkan di mana saja iklannya boleh ditempatkan. Dengan cara ini, iklan yang ditampilkan belum sesuai dengan isi dari suatu halaman. Namun, mungkin pengiklan telah melakukan analisis mengenai tempat yang paling sesuai untuk iklannya.

4. Interest-Based Targeting

Sistem ini menggunakan informasi dari setiap pengunjung yang meliputi interests dan demografik. Untuk menerapkan interest-based targeting, sistem periklanan harus mempunyai data apa saja yang pernah dicari oleh seorang pengguna sebelumnya dan menggunakan data tersebut untuk menampilkan iklan yang berbeda-beda kepada setiap pengguna.

5. Language Targeting

Dengan menggunakan teknologi untuk menentukan bahasa dari suatu halaman web. Pengiklan dapat menentukan pada halaman yang berbahasa apa iklannya boleh ditampilkan.

Kebanyakan advertising network saat ini menggunakan kombinasi dari berbagai macam tipe targeting seperti contextual, placement, interest-based, dan language dalam menentukan iklan yang paling tepat untuk ditampilkan. Walaupun demikian, masih banyak sistem periklanan yang tidak menggunakan targeting.

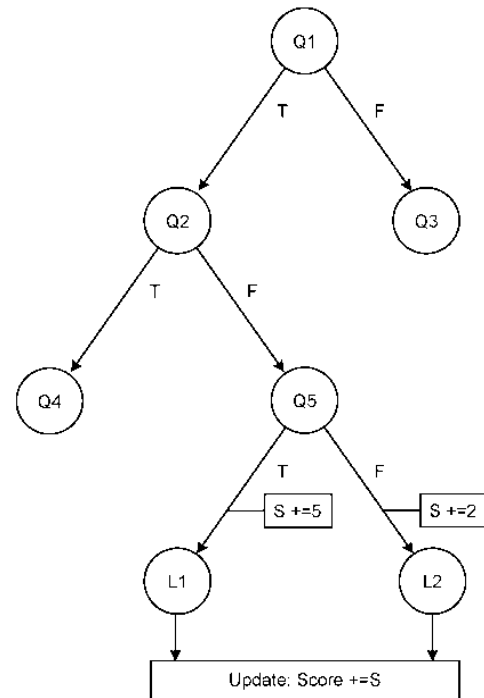
VI. Penerapan Pohon Keputusan Dalam Online Advertising

Berikut ini adalah penerapan pohon dalam menentukan iklan mana yang paling sesuai untuk ditampilkan.

Pengiklan yang ingin memasang iklan terlebih dahulu memberi informasi mengenai keyword yang ditarget serta

kepada siapa iklan tersebut ditujukan yang meliputi umur, jenis kelamin, lokasi, interest, dan kriteria-kriteria lainnya. Ada beberapa pohon keputusan yang digunakan untuk menentukan relevansi suatu iklan. Setiap iklan mempunyai decision tree yang berbeda-beda berdasarkan berbagai parameter yang ditarget oleh pengiklan.

a. Pohon Keputusan Untuk Menentukan Relevansi Iklan dengan Lokasi dan Demografi Pengguna



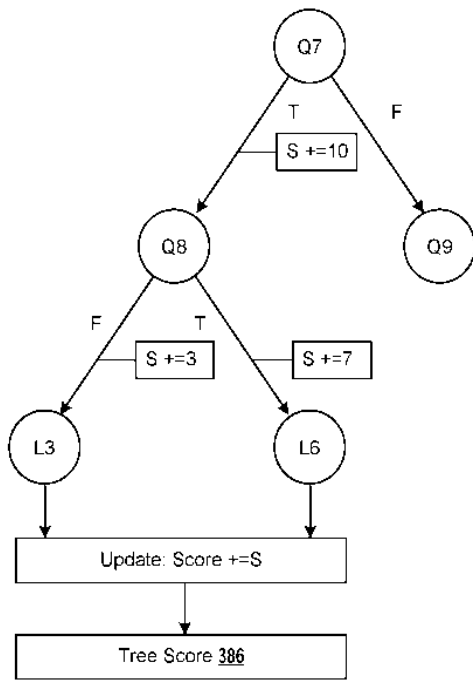
Gambar 3: Pohon keputusan untuk menentukan relevansi suatu iklan dengan lokasi dan demografi pengguna
(Sumber: <http://patentimages.storage.googleapis.com/US8571924B2/US08571924-20131029-D00003.png>. Diakses pada 8 Desember 2014 pukul 17.30)

Setiap simpul yang berlabel Q1, Q2, ... Q5 merepresentasikan suatu pengujian yang dapat berupa sebuah Boolean test, pengukuran, atau jenis pengujian lain yang berkaitan dengan lokasi dan demografi pengguna. Setiap sisi merepresentasikan traversal yang mungkin dari suatu simpul ke simpul lain. Contoh data yang termasuk ke dalam demografi pengguna misalnya jenis kelamin, lokasi, kelompok umur, level pendidikan, dan lain-lain. Data yang berkaitan dengan lokasi dan demografi pengguna sangat penting dalam meningkatkan relevansi iklan. Misalkan, perusahaan yang menawarkan jasa perbaikan komputer di Surabaya hanya akan ditampilkan kepada pengunjung yang berada di Surabaya karena tidak ada gunanya menampilkan iklan itu kepada pengunjung di luar Surabaya. Untuk mengetahui lokasi seorang pengguna, dapat diketahui dari IP address pengguna tersebut atau menggunakan GPS pada perangkat mobile. Sedangkan data demografi pengguna seperti jenis kelamin, umur, dan lain-lain biasanya hanya diketahui jika pengguna telah

melakukan sign-in pada suatu akun yang menyimpan informasi mengenai jenis kelamin, umur, dan data lainnya dari pengguna tersebut, serta dapat diakses oleh sistem periklanan.

Misalkan, Q1 merepresentasikan query “apakah pengguna tinggal di Indonesia”. Jika ya, maka akan dilanjutkan ke sisi kiri menuju Q2. Jika tidak, akan dilanjutkan sisi kanan menuju Q3. Misalkan Q2 merepresentasikan query “apakah pengguna adalah seorang laki-laki”. Jika ya, maka akan dilanjutkan ke sisi kiri menuju Q4. Jika tidak akan dilanjutkan ke Q5. Misalkan Q5 merepresentasikan query “apakah pengguna berusia dibawah 21 tahun”. Jika ya, maka akan dilanjutkan ke sisi kiri menuju L1 serta menambahkan 2 poin. Jika tidak, akan dilanjutkan ke sisi kanan menuju L2 serta menambahkan 5 poin.

b. Pohon Keputusan Untuk Menentukan Relevansi Iklan dengan Perilaku Pengguna



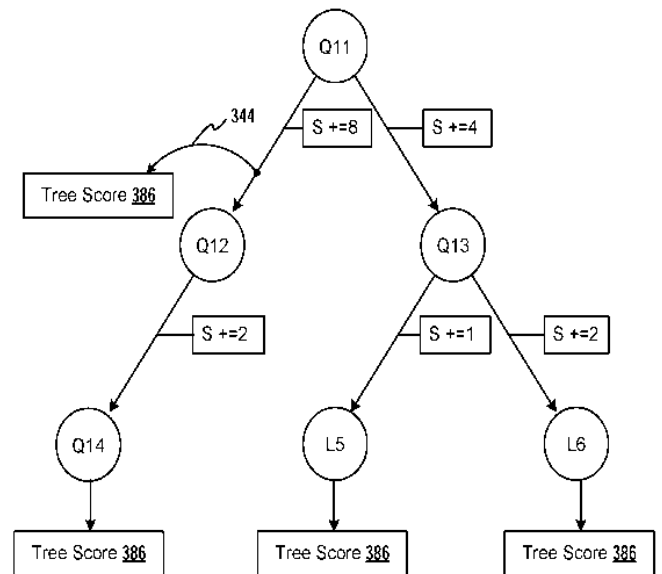
Gambar 4: Pohon keputusan untuk menentukan relevansi suatu iklan dengan perilaku pengguna
 (Sumber: <http://patentimages.storage.googleapis.com/US8571924B2/US08571924-20131029-D00004.png>. Diakses pada 8 Desember 2014 pukul 17.30)

Pohon keputusan di atas digunakan untuk menentukan relevansi suatu iklan dengan perilaku pengguna. Setiap simpul merepresentasikan suatu pengujian yang dapat berupa sebuah Boolean test, pengukuran, atau jenis pengujian lain yang berkaitan dengan perilaku pengguna. Setiap sisi merepresentasikan traversal yang mungkin dari suatu simpul ke simpul lain. Sistem periklanan menggunakan teknologi untuk menyimpan topik-topik yang pernah dicari, iklan yang pernah diklik oleh seorang

pengguna serta mendeteksi perangkat yang sedang digunakan oleh seorang pengguna. Dengan cara ini, iklan yang ditampilkan pada suatu halaman bisa berkaitan dengan apa yang pernah dibuka sebelumnya oleh pengguna, bukan berkaitan dengan isi dari halaman yang sedang dibuka.

Misalkan Q7 merepresentasikan “apakah pengguna menggunakan perangkat mobile”. Jika ya, maka akan dilanjutkan ke sisi kiri menuju Q8 dan menambahkan 10 poin. Jika tidak, maka akan dilanjutkan ke sisi kanan menuju Q9. Misalkan Q8 merepresentasikan “apakah pengguna memiliki interest terhadap smartphone”. Jika ya, maka akan dilanjutkan ke sisi kiri menuju L3 dan menambahkan 3 poin. Jika tidak, maka akan dilanjutkan ke sisi kanan menuju L4 dan menambahkan 7 poin.

c. Pohon Keputusan Untuk Menentukan Relevansi Iklan dengan Konteks Pengguna



Gambar 5: Pohon keputusan untuk menentukan relevansi suatu iklan dengan konteks pengguna
 (Sumber: <http://patentimages.storage.googleapis.com/US8571924B2/US08571924-20131029-D00005.png>. Diakses pada 8 Desember 2014 pukul 17.30)

Pohon keputusan yang ketiga digunakan untuk menentukan relevansi suatu iklan dengan konteks pengguna. Setiap simpul merepresentasikan suatu pengujian yang dapat berupa sebuah Boolean test, pengukuran, atau jenis pengujian lain. Setiap sisi merepresentasikan traversal yang mungkin dari suatu simpul ke simpul lain. Sistem periklanan menggunakan teknologi untuk membaca isi dari halaman dan menentukan topik dari halaman tersebut.

Misalkan Q11 merepersentasikan “apakah halaman yang dibuka pengguna berkaitan dengan teknologi”. Jika ya, maka akan dilanjutkan ke sisi kiri menuju Q12 dan menambahkan 8 poin. Jika tidak (halaman tidak berkaitan

dengan teknologi), maka akan dilanjutkan ke sisi kanan menuju Q13 dan hanya menambahkan 4 poin. Demikian seterusnya sampai mencapai daun dari pohon keputusan.

Proses pada masing-masing pohon keputusan berhenti ketika mencapai daun. Setiap proses dapat dijalankan pada core prosesor yang berbeda sehingga lebih cepat. Teknik tersebut telah dipatenkan oleh Google. Iklan yang dianggap paling relevan adalah iklan yang memiliki jumlah poin total tertinggi berdasarkan pohon keputusan-pohon keputusan yang digunakan.

Tidak semua sistem periklanan mengurutkan iklan berdasarkan relevansi iklan. Pada sistem periklanan Google Adwords, terdapat istilah Quality Score yang merupakan rating Google terhadap relevansi dan kualitas dari keyword dan iklan yang berbasis CPC (Cost Per Click). CPC akan ditentukan berdasarkan Quality Score. Untuk mengurutkan ranking iklan, hasil Quality Score dikalikan dengan maximum bid. Faktor-faktor yang mempengaruhi Quality Score adalah:

- Relevansi keyword-keyword terhadap grup iklan
- Relevansi dari teks pada iklan
- Click-Through Rate (CTR)
- Kualitas dan relevansi landing page, yaitu halaman yang terbuka jika iklan diklik
- Performa dari akun AdWords seorang pengguna

VII. KESIMPULAN

Pohon keputusan dapat digunakan dalam menentukan iklan mana yang paling tepat untuk ditampilkan pada suatu halaman agar sesuai dengan isi dari halaman yang menampilkan, perilaku pengguna, dan demografi pengguna. Dengan demikian iklan menjadi lebih relevan sehingga CTR (Click Through Rate) akan meningkat. Hal itu akan menguntungkan para pengiklan dan para pemilik situs yang menampilkan iklan. Selain itu, user experience terhadap situs yang menampilkan iklan relevan bisa menjadi lebih baik.

Jika diterapkan, pohon keputusan yang dibuat akan lebih kompleks daripada contoh di atas, bisa mencapai ratusan atau ribuan tingkat. Namun, karena harus menelusuri pohon keputusan, iklan yang ditayangkan menggunakan pohon keputusan akan memerlukan waktu yang lebih lama karena harus menelusuri pohon keputusan-pohon keputusan dari banyak iklan.

Walaupun lebih kompleks dan memerlukan waktu yang lebih lama untuk menampilkan suatu iklan daripada menampilkan iklan secara acak, jika penggunaan pohon keputusan berjalan dengan baik, hasilnya akan jauh lebih maksimal. Sistem periklanan yang memperhatikan relevansi iklan akan lebih disukai dan dipercaya oleh para pengiklan karena iklan mereka hanya akan ditampilkan kepada pengguna yang tepat sehingga mereka berani membayar dengan harga lebih tinggi kepada sistem periklanan atau penerbit iklan.

REFERENSI

- [1]. Rosen, Kenneth H., 2012, Discrete Mathematics and Its Applications Seventh Edition, New York: McGraw-Hill
- [2]. <http://www.smartinsights.com/digital-marketing-strategy/online-business-revenue-models/online-revenue-model-options-internet-business/>. Diakses 6 Desember 2014 pukul 20.40
- [3]. http://www.marketingterms.com/dictionary/cost_per_action/. Diakses 6 Desember 2014 pukul 21.15
- [4]. <https://www.boundless.com/marketing/textbooks/boundless-marketing-textbook/social-media-marketing-15/introduction-to-social-media-digital-marketing-98/types-of-internet-advertising-483-10593/>. Diakses 6 Desember 2014 pukul 22.21
- [5]. <http://www.google.com/patents/US8571924>. Diakses pada 7 Desember 2014 pukul 17.42.
- [6]. <http://www.wordstream.com/quality-score>. Diakses pada 10 Desember 2014 pukul 12.32.

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, 11 Desember 2014



Ivan Andrianto, 13513039