

Implementasi Pohon Keputusan Pada Penyaringan Berita Taklayak Baca di Media Sosial

Rahmad Yesa Surya and 13515088¹

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia

¹rahmadyesas@gmail.com

Abstrak— Pohon keputusan merepresentasikan sebuah metode untuk menentukan keluaran dari beberapa kemungkinan. Metode ini digunakan secara luas dalam rekayasa perangkat lunak dan pembelajaran mesin karena sangat mirip dengan cara berpikir komputer. Berkenaan dengan media sosial, pohon keputusan dapat diterapkan untuk menyaring konten yang layak dibaca oleh pengguna, mempertimbangkan jumlah peredaran konten pada saat ini yang tinggi dan kredibilitas konten yang tidak selalu baik.

Kata kunci— Pohon keputusan, media sosial, konten.

I. PENDAHULUAN

Pada akhir bulan November 2017, sebuah fenomena menarik muncul menyusul terpilihnya Donald Trump sebagai presiden Amerika. Fenomena ini bukan merupakan fenomena politik murni, melainkan pengaruh media sosial *Facebook* terhadap penggunaannya, yang merupakan pemilih dalam Pemilu Amerika. Gagasan ini muncul karena tingginya intensitas berita politik pada beranda *Facebook* yang disinyalir memperkuat sentimen terhadap Trump. Alasan lain yang memperkuat adalah banyaknya jumlah pembaca berita tersebut yang merupakan massa mengambang.

Hal serupa juga terjadi di Indonesia dengan subyek masalah Pilkada DKI. Jumlah berita politik yang menyebar di media sosial sangat tinggi, baik berita yang dapat dikategorikan sebagai netral, maupun berita yang mendukung salah satu pihak yang sedang bersaing. Dari permasalahan yang terjadi di Amerika dan Indonesia ini, terdapat satu hal yang menjadi ironi: tidak semua berita tersebut terjamin keasliannya dan memiliki sumber yang kredibel.

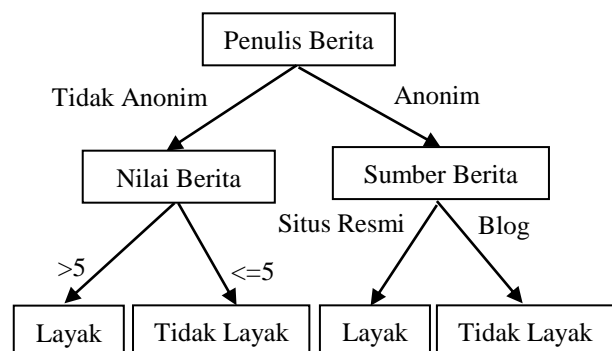
Permasalahan ini akan diulas oleh penulis dalam makalah ini dengan pendekatan matematis, khususnya pohon keputusan yang akan berguna untuk menentukan kelayakan sebuah berita untuk dibaca oleh pengguna. Penulis juga akan mengulas algoritma yang sesuai untuk digunakan dalam konteks ini.

II. LANDASAN TEORI

1. Pohon Keputusan dan Representasinya

Pohon keputusan merupakan salah satu algoritma inferensi induktif yang akan memperkirakan sebuah nilai diskrit dari sebuah fungsi. Obyek-obyek pada pohon keputusan disusun kedalam akar dan daun. Penelusuran terhadap solusi dimulai dari akar pohon dan simpulnya. Dalam setiap simpul terdapat sebuah kondisi yang akan menguji atribut dari beberapa obyek. Untuk setiap hasil benar, penelusuran akan berlanjut pada upapohon dari simpul tersebut hingga mencapai simpul yang tidak lagi memiliki daun. Proses ini merupakan proses berulang. Ketika simpul tak-berdaun telah ditemukan, metode pohon keputusan telah selesai untuk menemukan solusi.

Secara umum, pohon keputusan merepresentasikan sebuah disjungsi dari konjungsi pada beberapa atribut. Setiap jalan yang ditempuh dari akar hingga daun mengindikasikan konjungsi dari ujicoba atribut terhadap suatu kondisi, sedangkan pohon atau upapohon yang dibentuk merupakan konjungsi. Ilustrasi berikut akan menggambarkan bagaimana disjungsi dan konjungsi yang dimaksud dalam konteks penyaringan berita layak baca.



Gambar 2.1 Ilustrasi Pohon Keputusan Berita Layak Baca
Sumber: cs.princeton.edu (diakses pada 8 Desember 2016) dengan pengubahan.

Berdasarkan gambar 2.1, jika sebuah berita digolongkan kepada sebuah berita yang layak baca, maka pernyataan logika yang menggunakan konjungsi dan disjungsi adalah

(Penulis Berita = Tidak Anonim AND Nilai Berita > 5)
OR (Penulis Berita = Anonim AND Sumber Berita = Situs Resmi)

2. Kesesuaian Masalah dengan Penggunaan Pohon Keputusan

Terdapat beberapa domain permasalahan yang sesuai untuk diselesaikan dengan pohon keputusan. Kriteria-kriteria yang membuat sebuah permasalahan dapat diselesaikan dengan pohon keputusan adalah:

- Obyek-obyek yang didefinisikan memiliki atribut yang dapat digunakan sebagai pengukuran. Berdasarkan gambar 2.1, contoh obyek adalah penulis berita dengan atributnya yakni anonim dan tidak anonim.
- Fungsi target memiliki sebuah keluaran yang merupakan nilai diskrit. Berdasarkan gambar 2.1, keluaran yang diharapkan adalah apakah sebuah berita layak atau tidak untuk dibaca. Dengan demikian, keluaran pada masalah ini adalah nilai yang diskrit. Beberapa permasalahan yang mengharuskan keluaran berupa nilai non-diskrit tidak sesuai untuk diselesaikan dengan pendekatan pohon keputusan.
- Deskripsi logika dalam bentuk konjungsi dan disjungsi mungkin dibutuhkan. Hal ini bertujuan supaya permasalahan dapat dipahami dengan mudah oleh pengguna informasi.
- Permasalahan yang ingin diselesaikan memiliki beberapa sampel yang akan digunakan untuk melatih kerja dari pohon keputusan. Sampel yang dimaksud adalah rekaman mengenai masukan dan keluaran yang telah ada sebelumnya.

3. Keandalan Pohon Keputusan untuk Penyaringan Berita Layak Baca

Pohon keputusan sejatinya merupakan salah satu bagian dari analisis banyak peubah. Segala jenis analisis banyak peubah membuat pengguna dapat memprediksi, mendeskripsikan, atau mengklasifikasikan sebuah keluaran atau target. Dalam pengaplikasiannya dalam penyaringan berita layak baca, pohon keputusan dapat dengan mudah menentukan apakah sebuah berita layak atau tidak untuk dibaca berdasarkan obyek-obyek yang telah disusun dalam akar dan daun. Analisis banyak peubah seperti ini membuat pengguna dapat percaya terhadap keluaran karena analisis yang dilakukan, dalam konteks ini menggunakan pohon keputusan, telah melewati banyak uji kondisi. Jumlah peubah yang banyak merupakan hal yang penting untuk penyelesaian masalah karena hal ini menggambarkan kondisi nyata yang mana

banyak faktor mempengaruhi sebuah keluaran untuk terjadi. Walaupun sebenarnya analisis yang tidak melibatkan pohon keputusan mudah untuk dilakukan, yakni dengan cara *one-cause* atau *one-effect relationship*, pengguna akan menyadari bahwa cara seperti ini dapat berdampak pada keluaran yang tidak kredibel.

Berdasarkan sebuah riset dalam bidang psikologi kognitif oleh Miller (1956), kemampuan untuk mendapatkan sebuah informasi secara konseptual dan memanipulasinya dibatasi oleh kemampuan fisik dan kognitif pada ingatan jangka pendek di otak. Hal ini menjadi dasar mengapa penggunaan pohon keputusan dirasa sangat tepat karena pohon keputusan menyajikan sebuah cara dimana keseluruhan informasi dipertimbangkan sehingga keterbatasan manusia untuk menyeleksi secara keseluruhan dapat dibantu oleh pohon keputusan. Aspek lain yang menjadikan pohon keputusan sebagai metode yang tepat dalam menyeleksi adalah keterbacaannya yang tinggi oleh pengguna sehingga pengguna, walaupun tidak dibantu oleh komputer dapat memahami bagaimana sebuah pohon keputusan akan bekerja. Pengguna juga dapat memanfaatkan pernyataan logika konjungsi dan disjungsi yang dapat diturunkan dari sebuah pohon keputusan.

Dalam pengaplikasiannya untuk menyaring berita layak baca, pohon keputusan dapat mengumpulkan seluruh informasi mengenai berita tersebut. Istilah *metadata* dapat digunakan untuk merepresentasikan tentang hal ini. Seluruh *metadata* atau informasi mengenai berita, yang jumlahnya mungkin puluhan atau ratusan akan dikumpulkan dan digolongkan oleh pohon keputusan dan diseleksi menggunakan uji kondisi yang terdapat pada akarnya.

4. Evaluasi Pada Pohon Keputusan

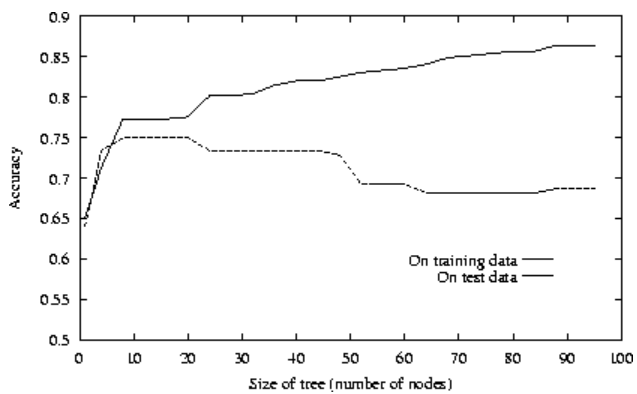
Sebagaimana jenisnya yang merupakan metode inferensi induktif, pohon keputusan dapat dikarakterisasi dengan pencarian ruang hipotesis yang sesuai dengan contoh latihan (*training example*). Ruang hipotesis merupakan kumpulan dari pohon keputusan yang mungkin. Ruang hipotesis akan digunakan untuk mengevaluasi pohon keputusan yang dibuat. Hal ini dilakukan dengan secara perlahan memberikan masukan yang bervariasi kepada pohon keputusan. Seiring semakin banyak masukan yang diterima oleh pohon keputusan, pengguna yang awalnya sudah tahu mengenai masukan dan jenis keluarannya, akan memperbaiki pohon keputusan sehingga pada akhirnya didapatkan pohon keputusan yang baik. Fase ini disebut juga evaluasi pohon keputusan. Setelah sebuah pohon keputusan dievaluasi, diharapkan pohon keputusan dapat bekerja dengan baik pada data nyata yang bukan merupakan contoh latihan lagi.

5. Kebagusan Pohon Keputusan

Sebuah pohon keputusan tidak bisa serta merta dikatakan bagus hanya dengan melihat banyaknya indikator yang digunakan sebagai penyeleksi keluaran. Seringkali pohon keputusan yang terlalu banyak melibatkan indikator dinilai terlalu berlebihan sehingga tidak bisa bekerja dengan baik ketika dihadapkan dengan masukan yang bervariasi. Hal inilah yang kemudian menentukan kebagusan dari sebuah pohon keputusan, selain dari kemungkinannya menghasilkan bias.

Dalam pembahasan yang lebih dalam pada topik pembelajaran mesin, bias ini kemudian akan dibagi lagi menjadi bias preferen (*preference bias*) dan bias restriksi (*restrictive bias*). Namun demikian, dalam makalah ini tidak akan dibahas lebih lanjut mengenai perbedaan kedua bias ini. Yang perlu menjadi perhatian adalah pohon keputusan tidak perlu melibatkan seluruh indikator berita kedalam penyeleksian keluaran karena hal ini dapat berdampak pada pohon keputusan yang tepat-berlebihan atau dalam istilah asingnya *overfitting*. Permasalahan ini ingin dihindari karena pohon keputusan yang *overfitting* menyebabkan ia tidak bisa bekerja dengan baik pada masukan yang bervariasi.

Ketika diaplikasikan untuk menyeleksi berita yang layak baca, jumlah berita yang beredar sangat banyak dan tidak semua berita memiliki indikator yang nilainya tetap, misalkan nilai berita (*rating*). Hal seperti inilah yang penulis maksudkan sebagai masukan yang bervariasi. Jika sebuah pohon keputusan *overfitting*, maka masukan yang semacam ini dapat mempengaruhi keluaran yang akan dihasilkan. Berita yang seharusnya layak, hanya karena baru muncul dan masih memiliki nilai (*rating*) dibawah 5, tidak bisa serta merta dikategorikan sebagai berita yang tidak layak. Tentu saja ini tidak akan bekerja dengan pohon keputusan yang *overfitting* karena ia akan menentukan bahwa berita tersebut sebagai berita yang tidak layak. Dengan demikian, pohon keputusan yang baik harusnya hanya memilih indikator-indikator penting dan krusial dalam menentukan kelayakan sebuah berita.



Gambar 2.2 Ilustrasi Jumlah Simpul Pada Pohon Keputusan dan Akurasi Keluaran

Sumber: eecs.wsu.edu (diakses pada 8 Desember 2016)

Terdapat permasalahan lain yang harus diperhatikan, yakni ketika masukan pada pohon keputusan tidak

memiliki atribut yang disyaratkan oleh pohon keputusan. Dalam kasus nyata, misalkan berita tersebut tidak memuat identitas penulis, padahal pohon keputusan menyaratkan adanya sebuah penulis berita, maka apa yang akan dilakukan oleh pohon keputusan? Alih-alih membiarkan masukan ini dikategorikan sebagai masukan yang tidak valid, terdapat cara yang lebih baik dalam menangani kasus ini. Cara ini dikemukakan oleh Mingers.

Menurut Mingers (1989), ketika sebuah masukan tidak memiliki atribut yang diharuskan, maka dilakukan penambahan atribut terhadap masukan tersebut dengan mengacu kepada nilai atribut yang paling sering muncul. Sebagai contoh, ketika sebuah berita tidak memuat identitas penulis, maka pohon keputusan akan menambahkan atribut identitas penulis kepadanya, dengan mempertimbangkan jumlah nilai atribut identitas penulis sebelumnya. Jika sebelumnya, nilai anonim mendominasi, maka pohon keputusan akan menambahkan atribut identitas penulis dengan nilai anonim kepada masukan tersebut. Begitupun untuk yang sebaliknya. Dengan demikian, pohon keputusan dapat meneruskan kembali penelusurannya hingga ditemukan keluaran yang tepat tanpa harus diterminasi karena masukan yang tidak valid.

III. ALGORITMA POHON KEPUTUSAN

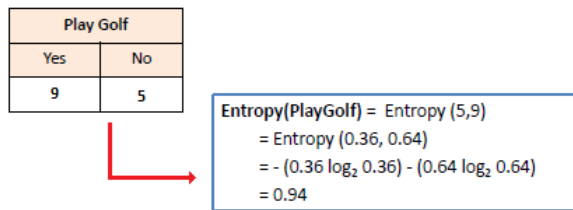
1. Pengenalan Terhadap Algoritma ID3

Algoritma ID3 (*Iterative Dichotomiser 3*) adalah algoritma yang digunakan secara luas dalam bidang pembelajaran mesin dan menjadi inti dari pohon keputusan. Algoritma ID3 dikembangkan oleh J. R. Quinlan dan bekerja dengan mengandalkan mekanisme *top-down* dan *greedy search* pada cabang-cabang yang mungkin tanpa melakukan *backtracking*. Algoritma ini menggunakan entropi dan *information gain* untuk menentukan sebuah keluaran pada pohon keputusan. Secara garis besar, algoritma ini akan bermula dari akar sebuah pohon. Setiap melakukan iterasi, ID3 akan menghitung entropi dan *information gain* dari sebuah atribut. ID3 kemudian akan memilih atribut yang menghasilkan entropi paling kecil atau nilai *information gain* paling besar untuk membentuk sebuah himpunan, kemudian melanjutkan penelusurannya.

2. Perhitungan Entropi dan *Information Gain*

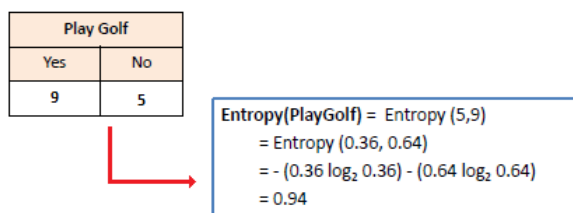
Entropi digunakan untuk menghitung homogenitas dari sampel. Jika sampel yang ada benar benar homogen, maka entropi akan bernilai nol, sedangkan jika sampel terbagi secara rata, maka nilai entropi adalah satu. Dalam pengaplikasian algoritma ID3, entropi akan dihitung menggunakan tabel frekuensi seperti yang diilustrasikan pada gambar dibawah.

$$E(S) = \sum_{i=1}^c -p_i \log_2 p_i$$



Gambar 2.3 Perhitungan Entropi Menggunakan Tabel Frekuensi Dari Satu Atribut
Sumber: saedsayad.com (diakses pada 8 Desember 2016)

$$E(S) = \sum_{i=1}^c -p_i \log_2 p_i$$



Gambar 2.4 Perhitungan Entropi Menggunakan Tabel Frekuensi Dari Dua Atribut
Sumber: saedsayad.com (diakses pada 8 Desember 2016)

Selain menggunakan nilai dari entropi untuk membangun sebuah pohon keputusan, algoritma ID3 juga menggunakan nilai *information gain*. Nilai ini didapatkan dengan beberapa langkah sebagai berikut:

- Langkah 1. Menghitung entropi dari target
- Langkah 2. Membagi himpunan data menjadi beberapa kelompok berdasarkan atributnya. Setiap entropi dari cabangnya kemudian dihitung. Selanjutnya, nilai dari perhitungan ditambahkan untuk mendapatkan nilai entropi total. Nilai entropi hasil diperoleh kemudian dikurangkan dari nilai entropi sebelum dipisah. Selisih dari nilai inilah yang dikatakan sebagai nilai *information gain*.
- Langkah 3. Memilih atribut dengan nilai *information gain* yang paling besar dan menjadikannya sebuah simpul keputusan (*decision node*).
- Langkah 4. Cabang yang memiliki nilai entropi nol adalah daun.
- Langkah 5. Cabang yang memiliki nilai entropi bukan nol perlu dibagi lagi sesuai dengan Langkah 2.
- Langkah 6. Langkah-langkah ini dilakukan secara rekursif dan akan berhenti ketika seluruh data telah diklasifikasikan.

Walaupun algoritma ID3 dinilai andal dan bagus dalam penyelesaian permasalahan ini, namun algoritma ini

juga masih memiliki beberapa kelemahan. ID3 tidak menjamin solusi yang optimal karena ia dapat terjebak dalam optima local. ID3 menggunakan pendekatan *greedy* dengan memilih atribut terbaik untuk membagi himpunan data pada setiap iterasi. Pun, ID3 tidak menjamin pohon keputusan yang dibentuk adalah pohon keputusan yang paling kecil. Dengan demikian, tidak dapat dikatakan sebuah pohon keputusan yang dihasilkan selalu pohon keputusan yang tidak *overfitting*.

IV. ANALISIS BERITA LAYAK BACA

Menentukan kelayakan sebuah berita membutuhkan parameter-parameter yang tepat. Selain memudahkan algoritma bekerja, pemilihan parameter yang tepat juga akan mengurangi kemungkinan pohon keputusan yang dibentuk untuk *overfitting*, seperti yang telah dijelaskan pada bab sebelumnya.

Dalam proses kerjanya untuk menentukan apakah sebuah berita layak untuk dibaca atau tidak, pengguna perlu memiliki himpunan sampel terlebih dahulu yang dapat dikatakan sebagai *training sample*. Himpunan sampel ini kemudian diproses menggunakan algoritma ID3 untuk menentukan sebuah pohon keputusan yang akan digunakan selanjutnya untuk himpunan data yang nyata. Himpunan sampel ini seharusnya adalah jenis-jenis berita yang bervariasi sehingga akan memperkaya indikator yang akan diproses oleh algoritma ID3. Dengan demikian, pohon keputusan yang dihasilkan diharapkan dapat mengakomodasi indikator-indikator penting yang digunakan untuk penentuan kelayakan sebuah berita.

Setelah pohon keputusan dihasilkan, alangkah baiknya pengguna tidak langsung menggunakannya untuk himpunan data yang asli, namun mengetesnya terlebih dahulu pada himpunan sampel. Hal ini dilakukan untuk menentukan seberapa akurat hasil yang dikeluarkan oleh pohon keputusan tersebut. Jika pada tahap pengetesan terhadap himpunan sampel, rata-rata hasil yang didapatkan menunjukkan akurasi yang tinggi, maka pohon keputusan dapat dikatakan berhasil dan siap untuk digunakan pada himpunan data yang asli. Jika pohon keputusan belum menunjukkan akurasi yang tinggi, maka pengguna harus melakukan evaluasi terhadap pohon keputusan tersebut, seperti yang telah dijelaskan pada bab II.

Dalam tabel berikut ini merupakan indikator-indikator berita yang mungkin menjadi atribut untuk diletakkan pada pohon keputusan. Atribut-atribut ini haruslah berbentuk nilai yang diskrit sehingga nilai-nilai seperti terkenal atau tidak harus dikonversi terlebih dahulu menjadi nilai diskrit, misalkan 0 atau 1, ketika diimplementasikan. Pun, terdapat beberapa metode lain yang digunakan dalam penentuan atribut, misalkan penghitungan frekuensi kemunculan sebuah kata dalam

sebuah berita.

No	Parameter	Alasan Digunakan
1	Penulis Berita - Ada - Anonim	Berita yang tidak memiliki penulis cenderung merupakan berita yang tidak layak baca karena mengandung konten yang tidak dapat dipercaya keasliannya.
2	Sumber Berita - Blog pribadi - Portal berita resmi	Berita yang berasal dari blog pribadi biasanya merupakan pandangan pribadi penulis berkaitan mengenai sebuah topik, berbeda dengan berita yang berasal dari portal resmi yang cenderung lebih obyektif. Maka dari itu, perlu dilakukan cek terhadap penulis berita jika berita berasal dari blog pribadi.
3	Nilai (<i>Rating</i>) Berita dalam skala 10 - >5 - <= 5	Berita yang nilai rendah cenderung tidak layak untuk dibaca. Namun demikian, perlu dicek lagi apakah berita merupakan berita baru. Jika ia merupakan berita baru, kemungkinan nilainya memang masih rendah karena belum banyak yang memberikan penilaian.
4	Alamat Situs - Terkenal - Tidak Terkenal	Berita yang berasal dari situs yang tidak terkenal dan situs yang baru muncul biasanya berita yang tidak layak baca karena cenderung menyajikan konten yang sedang panas di mata publik saja.
5	Frekuensi Sebuah Kata di Berita	Metode ini mengecek frekuensi kata-kata yang dicurigai mengandung unsur provokatif atau kata-kata yang menjelekkan seseorang atau sebuah pihak. Jika ditemukan frekuensi yang melebihi nilai <i>threshold</i> , berita kemungkinan besar tidak layak baca.

Tabel 4.1 Indikator-Indikator Yang Mungkin Menjadi Atribut Pada Pohon Keputusan

Dalam implementasi pohon keputusan, memang dimungkinkan terdapat cara-cara lain yang diperlukan

seperti perhitungan frekuensi kata. Dengan demikian, untuk digunakan dalam analisis berita layak baca, pohon keputusan alangkah baiknya dikombinasikan dengan metode-metode lain yang bersesuaian.

Kategori-kategori berita juga memungkinkan untuk dijadikan penentuan atribut dalam pohon keputusan. Berdasarkan observasi yang dilaksanakan oleh Prof. Melissa Zimdars dari Merrimack College, berita-berita yang tidak layak baca memiliki beberapa karakteristik umum dan dapat digolongkan kedalam beberapa kategori. Observasi ini dilaksanakan terhadap berita-berita yang beredar di berbagai sosial media pada saat masa Pemilu Amerika dihelat. Kategori-kategori tersebut ialah:

Kategori 1. Berita yang bertujuan untuk mendapatkan *like* atau komentar dari pengguna media sosial, yang kemudian akan dikomersialisasikan untuk tujuan mendapatkan keuntungan.

Kategori 2. Berita yang berisi konten tak bersumber sehingga informasi yang dikandungnya berpotensi tidak dapat dipercaya.

Kategori 3. Berita yang menggunakan judul kontroversi atau penggunaan tatabahasa yang tidak baik.

Kategori 4. Berita yang sengaja dibuat untuk menyindir seseorang atau sebuah pihak, khususnya yang bermuatan politis dan mempunyai potensi untuk disebarluaskan.

Hal yang perlu diperhatikan adalah bahwa setiap parameter yang digunakan untuk penyeleksian masukan pada pohon keputusan harus dalam bentuk diskrit. Dengan demikian, jika parameter tersebut memiliki nilai non-diskrit, harus diubah terlebih dahulu menjadi nilai diskrit untuk kemudian diproses menggunakan algoritma ID3 dan menghasilkan sebuah pohon keputusan.

V. HASIL YANG DIHARAPKAN

Pengguna media sosial berjumlah sangat banyak dan mereka menggunakan media sosial mereka secara aktif dengan frekuensi waktu yang tinggi. Hal inilah yang membuat para pembuat berita palsu menjadikan mereka sebagai target. Dengan jumlah berita palsu yang banyak, pengguna seringkali tidak dapat menyaringnya dengan benar karena keterbatasan kemampuan otak manusia. Maka dari itu, penyaringan berita taklayak secara otomatis perlu dilakukan oleh penyedia layanan media sosial. Salah satu cara yang dapat diterapkan adalah pengimplementasian pohon keputusan ini.

Penulis berharap pengotomatisan proses penyaringan berita taklayak baca dapat mengurangi intensitas penyebarannya. Dengan demikian, pengguna media sosial benar-benar mendapatkan esensi dari keberadaan media sosial: menjalin pertemanan dan mendapatkan berita-berita tepercaya, baik yang disebarkan oleh temannya atau dari portal berita yang diikutinya.

VII. UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih kepada Bapak Rinaldi Munir dan Ibu Harlili selaku dosen matakuliah Matematika Diskrit yang telah memberikan pengetahuan dasar mengenai topik-topik terkait selama satu semester. Penulis juga mengucapkan terimakasih kepada orangtua dan rekan-rekan yang memberikan motivasi dan semangat dalam penyelesaian penulisan makalah ini.

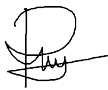
DAFTAR PUSTAKA

- [1] Hunt, E. B. (1975), *Artificial Intelligence*. New York: Academic Press.
- [2] Lopez de Mantaras, R. (1991). A distance-based attribute selection measure for decision tree induction. *Machine Learning*.
- [3] Mingers, J. (1989). An empirical comparison of selection measures for decision-tree induction. *Machine Learning*.
- [4] Pagallo, G. & Haussler, D. (1990). Boolean feature discovery in empirical learning. *Machine Learning*.
- [5] Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*.
- [6] Utgoff, P. E. & Brodley, C. E (1991). *Linear machine decision trees*, (COINS Technical Report), University of Massachusetts.

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, 9 Desember 2016



Rahmad Yesa Surya, 13515088