Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

Team project ©2017
Dony Pratidana S. Hum | Bima Agus Setyawan S. IIP

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

PT Gramedia adalah salah satu perseroan terbatas yang ada di Kompas Gramedia, dimana unit Corporate IT & IS, Corporate Human Resource, Corporate Secretary, dan Group of Manufacture berada di dalam PT tersebut. Corporate Human Resource, selaku unit yang bertanggung jawab atas data seluruh karyawan yang ada di Kompas Gramedia termasuk PT Gramedia, membutuhkan sebuah aplikasi yang dapat mengelompokan profil karyawan ke dalam kategori Penilaian Kinerja yang berlaku. Terdapat lima kategori Penilaian Kinerja yang saat ini berlaku di Kompas Gramedia, yaitu Unacceptable, Need Improvement, Meet Expectation, Exceed Expectation, dan Outstanding. Pengelompokan ini diperlukan untuk mencari kecenderungan kriteria karyawan di masing-masing kategori. Pada saat ini, data-data karyawan yang ada di *database* Kompas Gramedia belum dikelompokan ke dalam kategori apapun dan jika dilakukan pengelompokan secara manual, pengelompokkan ini akan memakan waktu yang lama (Cahyadi, 2017).

Berdasarkan kebutuhan Corporate Human Resource, metode yang dapat digunakan adalah metode klasifikasi. Metode klasifikasi adalah salah satu metode dalam data *mining* yang digunakan untuk mengklasifikasikan *item* (data) ke dalam salah satu kelas dari kelas-kelas yang telah ditentukan sebelumnya (Weiss & Kulikowski, 1991). Data *mining* sendiri adalah sebuah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* (Kusrini & Luthfi, 2009).

Penggunaan metode klasifikasi bertujuan untuk mengklasifikasikan suatu target ke dalam kategori yang dipilih, berbeda dengan klasterisasi yang tujuannya adalah hanya mengelompokan atribut berdasarkan similaritas (MacLennan, Tang, & Crivat, 2009). MacLennan, Tang & Krivat (2009) dalam bukunya menyatakan metode regresi mirip dengan metode klasifikasi akan tetapi metode regresi memberikan hasil yang tidak terbatas sedangkan klasifikasi memberikan hasil yang diskrit.

Terdapat beberapa algoritma yang dapat digunakan untuk metode klasifikasi dan berdasarkan hasil survei "10 Well-Known Algorithm in Data Mining" oleh IEEE pada tahun 2006, terdapat 4 algoritma dari metode klasifikasi yang mendapatkan nominasi tersebut, yaitu C4.5 (digunakan sebanyak 6.907 kali), CART (digunakan sebanyak 6.078 kali), Naïve Bayes (digunakan sebanyak 498 kali) dan K Nearest Neighbor (digunakan sebanyak 183 kali). Algoritma Naïve Bayes digunakan untuk mencari probabilitas sebuah *class*, hal ini dinyatakan dalam penelitian yang dilakukan oleh Oktafia & Pardede (2014) dan juga Wicaksana (2015). Algoritma CART (Classification and Regression Tree) digunakan apabila kategori yang dihasilkan juga bersifat kontinu dan akan menghasilkan *regression tree* bukan *classification tree*, hal ini dinyatakan oleh Zain & Pratiwi pada penelitiannya pada tahun 2014. Jika ingin menentukan nilai jarak pada pengujian *data testing* dengan *data training* berdasarkan nilai terkecil dari nilai ketetanggaan terdekat, maka digunakanlah algoritma k-NN (Goujon, Chaoqun, & Jianhong, 2013). Algoritma C4.5 merupakan algoritma klasifikasi pohon keputusan yang banyak digunakan karena memiliki kelebihan utama dari algoritma yang lainnya. Kelebihan algoritma C4.5 adalah dapat menghasilkan pohon keputusan yang

mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisien dalam menangani atribut bertipe diskret dan dapat menangani atribut bertipe diskret dan numerik (Han & Kamber, 2006). Algoritma C4.5 sendiri adalah pengembangan dari algoritma ID3 (Rahmanita & Kustiyahningsih, 2016). Rahmanita & Kustiyaningsih (2016) dalam penelitiannya juga menyatakan bahwa algoritma C4.5 merupakan suksesor dari ID3 dengan menggunakan *gain ratio* untuk memperbaiki *information gain*, karena jika menggunakan ID3 akan ada partisi yang tidak beguna karena atribut yang mememiliki banyak nilailah yang dipilih sebagai *splitting* atribut.

Penelitian terkait tentang penggunaan algoritma C4.5 telah dilakukan oleh Mabrur (2011) dan juga Maris (2015). Pada penelitian Maris (2015), dengan menggunakan 135 data *training* dan 15 data *testing*, nilai *accuracy* dari algoritma C4.5 yang digunakan mencapai 93%, dan pada penelitian Mabrur (2011), hasil penggunaan algoritma C4.5 dikatakan dapat memenuhi kebutuhan pengklasifikasian.

Berdasarkan penjelasan di atas dan juga kebutuhan Corporate Human Resource Kompas Gramedia, maka dibutuhkanlah penerapan algoritma C4.5 pada aplikasi klasifikasi profil karyawan.

1.2 Rumusan Masalah

Rumusan masalah dari penelitian ini adalah sebagai berikut.

1. Bagaimana cara mengimplementasikan algoritma C4.5 pada aplikasi klasifikasi profil karyawan berbasis web?
2. Bagaimana hasil akurasi yang didapatkan dari aplikasi klasifikasi profil karyawan dengan mengimplementasikan algoritma C4.5?

1.3 Batasan Masalah

Batasan masalah dari penelitian ini adalah sebagai berikut.

1. Profil karyawan yang akan digunakan hanya profil karyawan dari unit-unit yang ada di PT Gramedia, yaitu Corporate IT & IS, Corporate Human Resource, Corporate Secretary dan Group of Manufacture.
2. Data-data yang digunakan hanya data yang tersedia di Corporate Human Resource Kompas Gramedia.
3. Profil karyawan PT Gramedia akan diklasifikasikan ke masing-masing kategori PK yang berlaku di Kompas Gramedia.
4. Data-data yang akan digunakan adalah data personal seluruh karyawan PT Gramedia dan kategori nilai Penilaian Kinerja (PK) yang dikeluarkan pada Maret 2017.
5. Data personal yang dimaksud dalam penelitian ini adalah jenis kelamin, marital status, jumlah anak, status usia, status generasi, tingkat pendidikan terakhir, provinsi tempat tinggal karyawan, jumlah pelatihan yang telah diikuti, dan masa kerja karyawan.
6. Pengguna yang dimaksud dalam penelitian ini adalah Corporate Human Resource Kompas Gramedia.
7. Penggunaan dari aplikasi untuk ke depannya adalah keputusan dari Corporate Human Resource Kompas Gramedia.

UNIVERSITAS
MULTIMEDIA
NUSANTARA

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan dari penelitian ini adalah sebagai berikut.

1. Mengimplementasikan algoritma C4.5 pada aplikasi klasifikasi profil karyawan berbasis web.
2. Mengetahui hasil akurasi yang didapatkan dari aplikasi klasifikasi profil karyawan dengan mengimplementasikan algoritma C4.5.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut.

1. Mendapatkan kriteria karyawan dari masing-masing kategori Penilaian Kinerja.
2. Mendapatkan kriteria *recruitment* dengan lebih spesifik.

1.6 Sistematika Penulisan

Sistematika penulisan laporan penelitian disusun dan dibagi atas 5 (lima) bab sebagai berikut.

1. BAB I PENDAHULUAN

   Bab pertama ini menjabarkan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

2. BAB II LANDASAN TEORI

   Bab kedua membahas uraian teori dari penelitian yang dilakukan. Teori-teori yang dibahas adalah profil karyawan, nilai Penilaian Kinerja, data *mining*,

klasifikasi, algoritma C4.5, pengukuran hasil akurasi, Technology Acceptance Model (TAM), dan Skala Likert.

3. BAB III METODOLOGI DAN PERANCANGAN APLIKASI

   Bab ketiga menjelaskan metode penelitian yang dilakukan dan perancangan aplikasi. Perancangan aplikasi yang dimaksud terdiri dari perancangan *flowchart*, dan desain antarmuka.

4. BAB IV IMPLEMENTASI DAN UJI COBA

   Bab keempat memuat implementasi dan hasil dari uji coba aplikasi.

5. BAB V SIMPULAN DAN SARAN

   Bab kelima merupakan bab terakhir yang berisi simpulan dari hasil pengujian aplikasi dan juga saran untuk pengembangan aplikasi di masa mendatang.

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA