

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Decision Support System (DSS) atau Sistem Pengambilan Keputusan merupakan suatu sistem yang digunakan dalam pengambilan keputusan dengan memecahkan masalah yang tidak terstruktur dengan memanfaatkan data dan model (Qammaddin, 2016). Salah satu metode yang sudah banyak digunakan dalam permasalahan pengambilan keputusan yaitu AHP.

Metode AHP telah banyak yang diimplementasikan dalam berbagai penelitian, salah satu penelitian yang mengimplementasikan AHP pada bidang manajemen sumber daya manusia yaitu digunakan dalam deskripsi dan analisis kriteria yang digunakan dalam penilaian pegawai yang digunakan untuk promosi jabatan dengan cara mengumpulkan data pada Lembaga Kebijakan Pengadaan Barang/Jasa Jakarta yang kemudian dilakukan proses dengan metode AHP sehingga didapatkan hasil pembobotan kriteria dan subkriteria yang digunakan dalam penilaian promosi jabatan (Ayuningtyas, dkk, 2017). Pada bidang pariwisata yaitu penelitian sistem rekomendasi kuliner berbasis aplikasi *mobile*, pada penelitian ini membuat suatu sistem yang dapat merekomendasikan wisata kuliner yang berada di wilayah Malang dengan menggunakan beberapa kriteria yang kemudian diproses menggunakan metode AHP (Dewi, dkk, 2018). Hasil dari penelitian tersebut menunjukkan bahwa metode AHP dapat melakukan pemeringkatan pada alternatif yang dinilai dan pembobotan pada kriteria atau subkriteria dengan baik, tetapi mempunyai kelemahan yaitu tidak dapat melakukan klasifikasi pada setiap alternatif yang dinilai.

Klasifikasi merupakan sebuah metode dari penambangan data yang digunakan dalam memprediksi kategori atau kelas dari suatu data berdasarkan sekumpulan atribut-atribut dari data tersebut, atau merupakan metode pembelajaran terpadu (*supervised learning*) yang memproses dari sekumpulan data dan menghasilkan model klasifikasi yang digunakan untuk mengklasifikasikan *instance* baru (Malik, dkk, 2021).

Decision Tree merupakan salah satu metode klasifikasi yang banyak digunakan dalam membuat klasifikasi pada data, dengan membuat suatu pohon keputusan yang terdiri dari atribut yang digunakan sebagai akar pohon serta cabang yang dihasilkan dari pembagian kasus pada setiap atribut, sehingga menghasilkan klasifikasi data yang ditentukan.

Metode klasifikasi menggunakan *decision tree* C 4.5 telah banyak digunakan dalam berbagai penelitian. Salah satu penerapan *decision tree* C 4.5 pada bidang pendidikan yaitu proses evaluasi beasiswa pada mahasiswa perguruan tinggi, sistem ini akan secara otomatis menghasilkan penghargaan untuk setiap mahasiswa sesuai dengan capaian prestasi pada perguruan tinggi (Wanga, dkk, 2019). Pada bidang energi, *decision tree* C 4.5 digunakan dalam melakukan klasifikasi pada penilaian stabilitas tegangan *online*, proses dalam penelitian ini mencakup tiga langkah, yaitu akuisisi sampel, pemilihan atribut, dan konstruksi *decision tree* (Meng, dkk, 2020). Hasil dari penelitian tersebut menunjukkan bahwa metode *decision tree* C 4.5 dapat melakukan klasifikasi dengan baik, tetapi membutuhkan kombinasi dari metode lain agar dapat membuat bobot pada setiap kriteria dan alternatif.

Kombinasi metode AHP dengan *decision tree* C 4.5 dapat saling melengkapi kekurangan dari masing-masing metode yaitu dapat melakukan pemeringkatan dan pembobotan pada kriteria dan alternatif pilihan yang dilakukan oleh metode AHP, dan melakukan klasifikasi yang dilakukan oleh metode *decision tree* C 4.5, seperti pada penelitian kombinasi metode AHP dan *decision tree* C 4.5 dalam klasifikasi siswa *drop out* (Gustian, dkk, 2017). Pada penerapan klasifikasi *decision tree* C 4.5 dapat ditambahkan metode optimasi untuk meningkatkan nilai akurasi pada klasifikasi *decision tree* C 4.5.

Optimasi yang sudah digunakan dalam meningkatkan akurasi klasifikasi *decision tree* C 4.5 salah satunya yaitu PSO. Pada penelitian klasifikasi menggunakan *decision tree* C 4.5 dalam menentukan klasifikasi penggunaan pembayaran *digital*, sudah menggunakan optimasi PSO guna mendapatkan hasil akurasi baik yang lebih baik (Ariyati, dkk, 2020). Pada penelitian tersebut, PSO digunakan dalam seleksi fitur

pada kriteria yang akan digunakan dalam penentuan model C 4.5 yang akan dibuat dengan menghitung nilai entropi dan *gain*, sehingga didapatkan dapat meningkatkan nilai akurasi dalam pengklasifikasian penggunaan pembayaran digital dibandingkan dengan algoritma C 4.5.

2.2 Penambangan Data

Penambangan data adalah proses yang digunakan untuk menemukan informasi yang signifikan atau hilang yang tersimpan dalam jumlah data yang mempunyai skala besar (Saha, dkk, 2018), penambangan data merupakan aktivitas dalam eksplorasi data untuk mendapatkan pengetahuan yang bermanfaat (Lee, dkk, 2009) dengan menganalisis pola data pada basis data. Penambangan data merupakan langkah penting dalam proses *Knowledge Discovery From Data* (KDD) (Han, dkk, 2012) yang mempunyai beberapa tahapan.

Tahap pertama adalah pembersihan data yaitu tahapan yang digunakan untuk menghilangkan data yang tidak konsisten dan *noise*. Kedua, integrasi data yaitu menggabungkan banyak sumber data. Ketiga yaitu pemilihan data merupakan penggunaan data yang sesuai dengan tugas analisis yang diperoleh dari basis data. Keempat, transformasi data yaitu data diubah dan dikonsolidasikan ke dalam bentuk yang sesuai untuk penambangan dengan melakukan operasi ringkasan atau agregasi. Kelima yaitu penambangan data merupakan proses dari metode yang diterapkan untuk mengekstrak pola data. Keenam yaitu evaluasi pola digunakan untuk mengidentifikasi pola yang mewakili pengetahuan berdasarkan pada langkah-langkah tertentu. Yang terakhir yaitu presentasi pengetahuan merupakan teknik visualisasi dan representasi pengetahuan yang digunakan untuk menyajikan pengetahuan yang diolah kepada pengguna.

Fungsi dasar dari proses penambangan data yaitu terdiri dari fitur pemilihan, peringkasan, asosiasi, pengelompokan, prediksi, dan klasifikasi.

2.3 Metode Klasifikasi

Metode klasifikasi adalah sebuah metode dari penambangan data yang digunakan untuk memprediksi kategori atau kelas dari suatu data *instance* berdasarkan sekumpulan atribut-atribut dari data tersebut. Klasifikasi merupakan metode pembelajaran terpadu yang memproses dari sekumpulan data yang sudah mempunyai label untuk pembelajaran dan menghasilkan suatu model klasifikasi yang digunakan untuk mengklasifikasikan *instance* baru (Malik, dkk, 2021). Klasifikasi merupakan metode yang paling sering digunakan dalam penambangan data yang menggunakan teknik *machine-learning*, klasifikasi mempelajari pola-pola dari data historis atau sekumpulan informasi, seperti ciri-ciri, variabel-variabel, fitur-fitur pada berbagai karakteristik item-item yang sudah diberi label dengan tujuan untuk menempatkan objek-objek baru dengan label yang tak diketahui sebelumnya ke dalam kelompok atau kelas masing-masing.

Metode klasifikasi pada dasarnya mempunyai 2 tahap dalam prosesnya. Tahap pertama adalah pembelajaran atau pelatihan. Pada langkah ini, *classifier* dibangun berdasarkan sekumpulan kelas atau kategori yang sudah ditentukan dari data. Langkah ini disebut tahap pembelajaran atau tahap pelatihan, dimana sebuah algoritma klasifikasi membangun *classifier* dengan menganalisis data pelatihan. Langkah kedua adalah pengujian. Pada langkah ini, *classifier* yang sudah dibangun akan digunakan untuk mengklasifikasi data. Kemudian melakukan uji coba kemampuan atau akurasi dari model *classifier* tersebut dengan menggunakan data pengujian.

2.4 Decision Tree C 4.5

Algoritma *Decision Tree C4.5* seperti dengan namanya merupakan salah satu algoritma yang termasuk dalam metode klasifikasi pohon keputusan. *Decision Tree C4.5* merupakan arus proses yang digunakan dalam pohon keputusan dengan memilih atribut dari dataset sebagai akar dan membuat cabang pada setiap nilai atribut yang digunakan dalam membagi kasus pada setiap cabang, kemudian melakukan proses yang sama pada setiap cabang hingga semua kasus pada cabang mempunyai kelas yang

sama (Ariyati, dkk. 2020). Algoritma ini mempunyai masukan berupa data sampel latih dan sampel. Data sampel latih merupakan data yang akan digunakan untuk membangun sebuah pohon yang telah diuji kebenarannya. Sedangkan sampel merupakan data yang akan dijadikan parameter dalam klasifikasi data (Tarigan, dkk. 2019).

Algoritma C4.5 secara umum mempunyai langkah-langkah dalam membangun pohon keputusan yaitu sebagai berikut (Tarigan, dkk. 2019). Langkah pertama yaitu pilih atribut sebagai akar. Kedua, buat cabang untuk setiap nilai. Ketiga, menentukan kasus di cabang. Terakhir, ulangi proses untuk setiap cabang sampai semua kasus di cabang memiliki kelas yang sama.

Nilai *gain* tertinggi dari atribut merupakan dasar yang digunakan dalam memilih atribut sebagai akar. Sebelum melakukan penghitungan nilai *gain* pada atribut, terlebih dahulu dilakukan penghitungan nilai entropi, yaitu dengan rumus sebagai berikut :

$$H(S) = -\sum_{i=1}^n p_i \log_2 p_i \quad (2.1)$$

Dengan :

- H : Entropi
- S : kumpulan (set data)
- n : jumlah *record*
- p_i : probabilitas diperoleh dari ya atau tidak total kasus

Untuk menghitung nilai *gain*, digunakan rumus sebagai berikut :

$$IG(A) = H(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} H(S_i) \quad (2.2)$$

Dengan :

- IG : *Gain*
- H : Entropi
- S : kumpulan (set data)

- A : atribut yang akan digunakan
n : jumlah atribut partisi A
 $|S_i|$: jumlah kasus pada partisi i
 $|S|$: jumlah kasus di S

2.5 Particle Swarm Optimization (PSO)

PSO merupakan suatu algoritma pencarian berbasis populasi dengan menginisialisasi populasi solusi acak dan digunakan untuk menyelesaikan masalah optimasi (Abraham, dkk. 2006). PSO merupakan suatu metode optimasi heuristik global yang diperkenalkan oleh Dr. Kennedy dan Eberhart berdasarkan penelitian kawanan burung dan ikan pada tahun 1995 (Haupt, dkk. 2004), untuk mendapatkan posisi yang tepat untuk mencapai tujuan tertentu. Pada penerapan PSO, setiap partikel memiliki posisi dan bergerak dengan kecepatan yang diperbarui pada setiap perulangan. Setiap partikel dalam suatu populasi memiliki nilai *fitness* yang dihitung dari fungsi *fitness*. Ciri-ciri utama partikel dalam PSO adalah posisi, kecepatan, dan kemampuan bertukar informasi dengan partikel tetangganya, kemampuan mengingat posisi sebelumnya, dan kemampuan menggunakan informasi untuk mengambil keputusan (Chen, dkk, 2014).

Setiap partikel dalam PSO juga diinterpretasikan oleh kecepatan partikel yang terbang melalui ruang pencarian dengan kecepatan yang secara dinamis disesuaikan dengan perilaku historisnya. Oleh karena itu, partikel memiliki kecenderungan untuk terbang menuju area pencarian yang lebih baik dan lebih baik selama proses pencarian (Abraham, dkk, 2006). Berikut merupakan rumus yang digunakan untuk menghitung perpindahan posisi dan kecepatan partikel:

$$V_i^t = \omega V_i^{t-1} + c_1 r_1 (P_{best\ i} - X_i^t) + c_2 r_2 (G_{best} - X_i^t) \quad (2.3)$$

$$X_i^t = X_i^{t-1} + V_i^t \quad (2.4)$$

Dengan :

V_i^t : kecepatan partikel i selama iterasi t

X_i^t : posisi partikel i selama iterasi t

ω : bobot inersia

c_1 dan c_2 : koefisien akselerasi

r_1 dan r_2 : angka acak antara 0 sampai 1

$P_{best\ i}$: posisi terbaik untuk partisan i

G_{best} : posisi terbaik global

2.6 Analytical Hierarchy Process (AHP)

AHP adalah metode pendukung keputusan yang dikembangkan oleh Thomas L. Saaty yang merupakan seorang professor matematika University of Pittsburgh kelahiran Irak. Metode AHP merupakan salah satu metode yang dapat digunakan untuk menyelesaikan suatu masalah *Multiple Attribute Decision Making* (MADM). AHP merupakan salah satu metode dalam sistem pendukung keputusan yang menggunakan beberapa kriteria dan alternatif dengan proses analisis bertingkat (Xiulin, dkk, 2014), atau suatu metode pengambilan keputusan dengan melakukan penilaian terhadap kriteria-kriteria terkait dalam suatu urutan hierarki (Saaty, 2008).

Prinsip dasar dalam penerapan AHP yaitu terdiri dari dekomposisi, perbandingan penilaian, dan sintesis prioritas. Pada tahap dekomposisi ini, setiap struktur masalah dibagi menjadi bagian-bagian, yaitu tujuan, kriteria, dan alternatif. Pada himpunan alternatif memungkinkan untuk dibagi menjadi tingkatan yang lebih detail, mencakup lebih banyak kriteria yang lain.

Pada tahap perbandingan penilaian dilakukan perbandingan berpasangan dari elemen kriteria maupun alternatif. Penilaian dari perbandingan berpasangan ini menggunakan skala perbandingan Saaty, dengan menggunakan nilai dari 1 s.d 9. Hasil dari perbandingan berpasangan tersebut akan menghasilkan bobot prioritas.

Pada tahap sintesis prioritas, dari hasil matriks perbandingan berpasangan dilakukan pengolahan untuk menentukan suatu alternatif terbaik. Proses penghitungan dengan beberapa tahap, yaitu pertama menghitung bobot prioritas, kedua menghitung rasio konsistensi, terakhir pengambilan keputusan berdasarkan perbandingan hasil penghitungan bobot alternatif.



SEKOLAH PASCASARJANA