

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Dalam Sistem Informasi, kebenaran dan keakuratan informasi sangatlah penting, agar tidak terjadi kesalahan yang tidak di inginkan, maka di perlukan pendukung yang bisa melakukan hal tersebut dan menghasilkan informasi yang baik, efisien, dan cepat. Tinjauan pustaka ini digunakan sebagai acuan pembimbing antara penelitian yang sudah dilakukan dan yang akan dilakukan peneliti. Penelitian tersebut diantaranya sebagai berikut :

Penelitian yang dilakukan oleh Rakhma dkk (2012) yang dilakukan menghasilkan sebuah sistem pakar untuk merekomendasikan pekerjaan berdasarkan kepribadian menggunakan pendekatan *personality factory*.

Penelitian yang telah dilakukan oleh Darmastuti (2013) yang bertujuan untuk membangun sistem informasi lowongan kerja dan memuat rekomendasi kepada pencari kerja menggunakan metode SAW. Hasil dari penelitian yang dilakukan adalah aplikasi berbasis web yang digunakan oleh para pencari kerja.

Penelitian yang telah dilakukan oleh Limbong (2013) yang dilakukan bertujuan untuk menyajikan dan menguraikan bidang pekerjaan dalam informatika dan komputer dan matakuliah yang mendukung dengan membangun sistem pendukung keputusan pemilihan pekerjaan dibidang informatika.

Penelitian yang dilakukan oleh Sutisna dan Noor (2015) yang dilakukan menggunakan penalaran Logika Fuzzy Mamdani untuk memproses data

masukkan yang selanjutnya digunakan untuk menghasilkan keluaran berupa rekomendasi pekerjaan apa yang tepat bagi seseorang.

Penelitian yang dilakukan oleh Indah dkk (2016) yang bertujuan untuk merekomendasikan pekerjaan berdasarkan karakter dan kepribadian menggunakan pendekatan Personality Factor. Hasil penelitian berupa informasi profesi atau pekerjaan yang direkomendasikan oleh sistem.

Penelitian yang dilakukan oleh Sutisna dan Noor (2015) yang memberikan informasi untuk pengambilan keputusan pemilihan pekerjaan yang disesuaikan dengan kemampuan dan minat yang dimiliki seseorang. Metode yang digunakan adalah Logika Fuzzi untuk memproses data masukan yang selanjutnya digunakan untuk menghasilkan keluaran berupa rekomendasi pekerjaan apa yang tepat bagi seseorang.

Penelitian yang dilakukan oleh Firdaus (2017) menghasilkan sebuah sistem rekomendasi pekerjaan yang memberikan arahan pada lulusan SMKN 1 Pasuruan dalam mencari pekerjaan. Rekomendasi yang dihasilkan bersifat opsional karena tidak mengikat alumni. Sistem rekomendasi tersebut menerapkan metode *Simple Additive Weighting* (SAW) menggunakan data usia, jurusan, jenis kelamin, tinggi badan, nilai sikap dan nilai rata-rata nilai UN.

Penelitian yang dilakukan oleh Puri Dkk (2017) merupakan penelitian yang bertujuan untuk membantu perusahaan dalam melakukan rekrutmen karyawan. Dengan menganalisa kepribadian calon karyawan, perusahaan dapat melihat karakter calon karyawannya untuk dilakukannya tahap seleksi sesuai kebutuhan perusahaan.

Tabel 2.1 : Perbandingan Tinjauan Pustaka

No	Pengarang	Objek	Metode	Hasil
1.	Rakhma dkk (2012)	Pencari Kerja	<i>Personality Factory</i>	Rekomendasi profesi pekerjaan dalam bidang teknologi informasi sebagai sarana mencari referensi pekerjaan.
2.	Darmastuti (2013)	Pencari Kerja	<i>Simple Additive Weighting (SAW)</i>	Sistem pendukung keputusan pencari kerja terbaik yang bertujuan mendapatkan karyawan terbaik.
3.	Limbong (2013)	STMIK Budi Darma Medan	<i>Simple Additive Weighting (SAW)</i>	Sistem pendukung keputusan untuk pemilihan pekerjaan di bidang informatika
4.	Sutisna Dan Noor (2015)	E-Commerce	<i>Decission Tree and Clustering</i>	Sistem rekomendasi e-commerce menggunakan Decission Tree dan Clustering
5.	Indah Dkk (2016)	Pencari Kerja	Personality Factor	Sistem pakar untuk merekomendasikan profesi pekerjaan berdasarkan kepribadian yang dilihat dari faktor kepribadian
6.	Firdaus (2017)	SMKN 1 Pasuruan	<i>Simple Additive Weighting (SAW)</i>	Sistem rekomendasi pekerjaan memiliki presentasi 100% dalam pengujian fungsional dan pengujian algoritma menunjukkan nilai valid.
7.	Puri Dkk (2017)	Perusahaan	Algoritma C4.5 dan PAPI Kostick	Sestem yang menganalisa kepribadian guna melakukan rekrutmen karyawan pada sebuah perusahaan
8.	Restu Julian	Alumni S1 STMIK Akakom Yogyakarta	Algoritma C4.5	Sistem Rekomendasi pekerjaan menggunakan algoritma C4.5 berbasis web

2.2 Dasar Teori

Dasar teori merupakan acuan dalam mengerjakan penelitian dengan sumber terpercaya seperti buku, jurnal, atau situs web yang terverifikasi. Penelitian ini menggunakan referensi dari jurnan dan buku serta website sebagai referensi dan pembanding dengan penelitian lainnya.

2.2.1 Sistem Rekomendasi

Sistem rekomendasi adalah sebuah sistem yang dirancang dengan tujuan membantu penggunaannya dalam mencari sesuatu yang mungkin mereka sukai dengan cara memberikan rekomendasi kepada pengguna. Rekomendasi itu berkaitan dengan berbagai proses pengambilan keputusan, seperti barang apa yang harus dibeli, musik apa yang didengar, atau berita *online* apa yang harus dibaca (Ricci et.al., 2011). Cara pencarian *item* yang akan direkomendasikan dapat dilakukan berdasarkan kemiripan baik berupa kemiripan suatu *item* dengan *item* lainnya berdasarkan konten atau kemiripan selera suatu pengguna dengan pengguna lainnya berdasarkan *rating* yang diberikan pada item (A. Djamal et.al., 2010).

2.2.2 Pekerjaan

Pekerjaan merupakan suatu kegiatan yang wajib dilakukan oleh setiap orang demi kelangsungan hidupnya atau untuk memenuhi berbagai macam kebutuhan hidupnya. Setiap orang melakukan pekerjaan salah satunya untuk memenuhi kebutuhan pokoknya, karena kebutuhan pokok merupakan kebutuhan yang harus dipenuhi dan tidak bisa di tunda-tunda.

Kebutuhan tersebut misalnya pokok seperti makan, minum, pakaian, pendidikan dan lain-lain. Untuk dapat memenuhi berbagai kebutuhannya, maka seseorang membutuhkan uang dan umumnya uang di dapatkan dari bekerja, saat ini banyak sekali pekerjaan yang dilakukan seseorang untuk menghasilkan uang.

2.2.3 Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) adalah proses menentukan informasi yang berguna serta pola-pola yang ada dalam data. Informasi ini terkandung dalam basis data yang berukuran besar yang sebelumnya tidak diketahui dan potensial bermanfaat (Han & Kamber, 2006 dalam Baskoro, 2010).

Berikut merupakan tahapan dari KDD :

1. *Data selection* (di mana data yang relevan dengan tugas analisis dikembalikan ke dalam *database*).
2. *Data transformation* (di mana data berubah atau bersatu menjadi bentuk yang tepat untuk menambang dengan ringkasan performa atau operasi agresif).
3. *Knowledge Discovery* (proses esensial di mana metode yang intelektual digunakan untuk mengekstrak pola data).
4. *Pattern evaluation* (untuk mengidentifikasi pola yang benar-benar menarik yang mewakili pengetahuan berdasarkan atas beberapa tindakan yang menarik).

5. *Knowledge presentation* (di mana gambaran teknik visualisasi dan pengetahuan digunakan untuk memberikan pengetahuan yang telah ditambah kepada user).

Dalam salah satu tahap KDD terdapat *data mining*, *data mining* merupakan suatu kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menentukan keteraturan, pola atau hubungan dalam set data berukuran besar (Santosa, 2007). Beberapa teknik yang sering disebut-sebut dalam literatur data mining antara lain yaitu *associationrule mining*, *clustering*, *classification*, *neural network*, *genetic algorithm* dan lain-lain.

2.2.4 Data Mining

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. *Data Mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *mechine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar (Turban, dkk. 2015).

Menurut *Garter Group data mining* adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Larose, 2005).

Data mining memiliki fungsi seperti menemukan pola dari data yang sedang dipakai lalu menggunakan variabel untuk memprediksi variabel lainnya

serta menemukan karakteristik penting dalam data yang sedang dipakai. Selain itu, *data mining* berfungsi untuk menemukan model dalam menggambarkan konsep kelas atau data dan menemukan hubungan nilai atribut dari beberapa data.

2.2.5 Pohon Keputusan

Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Model pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang mempresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti *Structured Query Language* untuk mencari *record* pada kategori tertentu.

Sebuah pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Dengan masing-masing rangkaian pembagian, anggota himpunan hasil menjadi mirip dengan yang lain (Berry & Linoff, 2004).

Sebuah model pohon keputusan terdiri dari sekumpulan aturan untuk membagi sejumlah populasi yang *heterogen* menjadi lebih kecil, lebih *homogen* dengan memperhatikan pada variabel tujuan.

Sebuah pohon keputusan mungkin dibangun dengan saksama secara manual atau dapat tumbuh secara otomatis dengan menerapkan salah satu atau beberapa algoritma pohon keputusan untuk memodelkan himpunan data yang belum terklarifikasi.

Variabel tujuan biasanya dikelompokkan dengan pasti dan model keputusan lebih mengarah pada perhitungan *probabilitas* pada tiap-tiap *record* terhadap kategori-kategori tersebut atau untuk mengklasifikasi *record* dengan mengelompokkannya dalam satu kelas.

Pohon keputusan juga dapat digunakan untuk mengestimasi nilai dari variabel *continue* meskipun ada beberapa teknik yang lebih sesuai untuk kasus ini.

2.2.6 Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan (Kusrini, Emha. 2009).

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut.

1. Pilih atribut sebagai akar
2. Membuat cabang untuk tiap-tiap nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus berikut.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2.1)$$

Keterangan :

S : himpunan kasus

A : atribut

- n : jumlah partisi atribut A
 |Si| : jumlah kasus pada partisi ke-i
 |S| : jumlah kasus dalam S

Sementara itu untuk mendapatkan nilai *entropy* digunakan rumus sebagai berikut.

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad (2.2)$$

Keterangan :

- S : himpunan kasus
 A : fitur
 n : jumlah partisi S
 pi : properti dari Si terhadap S

2.2.7 PAPI Kostick

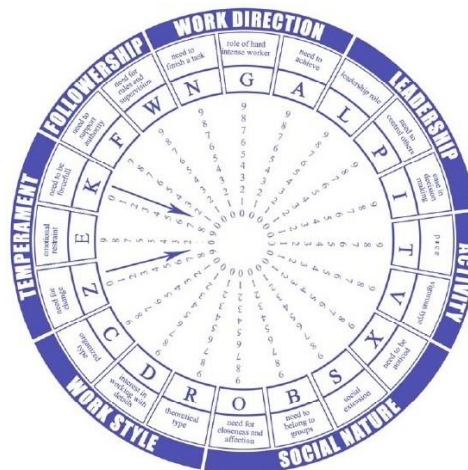
Inventory) adalah “personality assessment” (alat tes penilaian kepribadian) terkemuka yang digunakan oleh para profesional HR (Human Resource) dan manajer terkait untuk mengevaluasi perilaku dan gaya kerja individu pada semua tingkatan. Personality and Preference Inventory (PAPI) dibuat oleh Guru Besar Psikologi Industri dari Massachusetts, Amerika, yang bernama Dr. Max Martin Kostick pada awal tahun 1960-an. Versi Swedia lebih dulu diperkenalkan di awal 1980-an dan versi ini diperkenalkan pada tahun 1997 dengan versi ipsative (PAPI-I) dan normative (PAPI-N). Versi ipsative, PAPI-I, dirancang untuk digunakan untuk pengembangan pribadi, sedangkan normative versi, PAPI-N, yang dimaksudkan untuk digunakan untuk perbandingan dan seleksi. Dasar pemikiran

untuk desain dan formulasi PAPI didasarkan pada penelitian dan teori kepribadian “needs-press” oleh Murray (1938). Adapun aspek yang diungkap Tes PAPI Kostick sebagai berikut :

1. *Work Direction* (Arah Kerja)
 - a. Need to finish task (N) – Kebutuhan menyelesaikan tugas secara mandiri
 - b. Hard intense worked (G) – Peran pekerja keras
 - c. Need to achieve (A) – Kebutuhan berprestasi
2. *Leadership* (Kepemimpinan)
 - a. Leadership role (L) – Peran kepemimpinan
 - b. Need to control others (P) – Kebutuhan mengatur orang lain
 - c. Ease in decision making (I) – Peran membuat keputusan
3. *Activity* (Aktivitas Kerja)
 - a. Pace (T) – Peran sibuk
 - b. Vigorous type (V) – Peran penuh semangat
4. *Social Nature* (Relasi Sosial)
 - a. Need for closeness and affection (O) – Kebutuhan kedekatan dan kasih sayang
 - b. Need to belong to groups (B) – Kebutuhan diterima dalam kelompok
 - c. Social extension (S) – Peran hubungan sosial
 - d. Need to be noticed (X) – Kebutuhan untuk diperhatikan
5. *Work Style* (Gaya Kerja)
 - a. Organized type (C) – Peran mengatur
 - b. Interest in working with details (D) – Peran bekerja dengan hal – hal rinci

- c. Theoretical type (R) – Peran orang yang teoritis
6. *Temprament (Sifat tempramen)*
- a. Need for change (Z) – Kebutuhan untuk berubah
 - b. Emotional resistant (E) – Peran pengendalian emosi
 - c. Need to be forceful (K) – Kebutuhan untuk agresif
7. *Followership (Posisi atasan – bawahan)*
- a. Need to support authority (F) – Kebutuhan membantu atasan
 - b. Need for rules and supervision (W) – Kebutuhan mengikuti aturan dan pengawasan

PAPI Kostick tersusun atas 90 pasang pernyataan pendek berhubungan pada situasi kerja, menyangkut 20 aspek kepribadian yang dikelompokkan dalam 7 bidang :



Gambar 2.1 : Aspek Kepribadian PAPI Kostick

Menghitung skor peran, yaitu dengan menjumlahkan anak panah yang dilingkari, baik yang horizontal maupun vertical sesuai dengan arah tanda panah. Menuliskan jumlah skor pada masing – masing kotak skor dibawah huruf G, L, I,

T, V, S, R, D, C, E yang telah tersedia pada lembar jawab. Menghitung jumlah skor pada seluruh kotak skor peran secara horizontal, dan jumlah skor harus 45. Menghitung skor “kebutuhan” yaitu dengan menjumlahkan anak panah yang dilingkari baik yang horizontal maupun yang vertical sesuai dengan arah tanda panah. Menjumlahkan jumlah skor pada masing – masing kotak dibawah huruf N, A, P, X, B, O, Z, K, F, W yang telah tersedia pada lembar jawaban. Mengitung jumlah skor pada seluruh kotak skor kebutuhan secara vertical, dan jumlah skor harus 45. Memindahkan setiap skor pada lembar jawaban ke lembar scoring sesuai dengan setiap huruf pada aspek “peran” dan “kebutuhan” dengan cara melingkari angka di dalam lingkaran. Membuat garis penghubung antara angka yang satu dengan angka lainnya sehingga terbentuklah sebuah diagram pada lembar psikogram yang telah tersedia. Gambar alat tes yang digunakan untuk melakukan skoring ditunjukkan pada gambar

PAPI

NOMOR : _____
 NAMA : _____
 TANGGAL TES _____
 TANGGAL PEMERIKSAAN _____

Total	G	L	I	T	V	S	R	D	C	E
41	71	42	51	42	31	21	11	1	1	1
42	72	43	52	43	32	22	12	1	1	1
43	73	44	53	44	33	23	13	1	1	1
44	74	45	54	45	34	24	14	1	1	1
45	75	46	55	46	35	25	15	1	1	1
46	76	47	56	47	36	26	16	1	1	1
47	77	48	57	48	37	27	17	1	1	1
48	78	49	58	49	38	28	18	1	1	1
49	79	50	59	50	39	29	19	1	1	1
50	80	51	60	51	40	30	20	1	1	1
										Total
	N	A	P	X	B	O	Z	K	F	W

Gambar 2.2 : Alat tes PAPI Kostick

PAPI disusun sebagai dua aspek yang terpisah, yaitu ; Pengukuran kebutuhan (needs) dan pengukuran persepsi (roles), yaitu persepsi keadaan

individu di tempat kerja. PAPI Kostick untuk menjabarkan kepribadian dalam 20 aspek yang masing – masing mewakili need dan role tertentu.

2.2.8 Confusion Matrix

Confusion Matrix merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi. Pada dasarnya *confusion matrix* mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} * 100\% \quad (2.3)$$

$$Presisi = \frac{TP}{FP+TP} * 100\% \quad (2.4)$$

$$Recall = \frac{TP}{FN+TP} * 100\% \quad (2.5)$$

Pada pengukuran kinerja menggunakan *confusion matrix*, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN). Nilai *True Negative* (TN) merupakan jumlah data negative yang terdeteksi dengan benar, sedangkan *False Positive* (FP) merupakan data negatif namun terdeteksi sebagai data positif. Sementara itu, *True Positive* (TP) merupakan data positif yang terdeteksi benar. *False Negative* (FN) merupakan kebalikan dari *True Positive* (TP), sehingga data positif, namun terdeteksi sebagai data negatif. Klasifikasi dengan jumlah 2 (dua) keluaran kelas maka digunakan rumus berikut (Solichin, 2017).