



TESIS TE092099

**INFERENSI PADA SISTEM TEMU KEMBALI
INFORMASI DENGAN MENGGUNAKAN JARINGAN
BAYES SEMANTIK**

**JOKO SAMODRA
NRP. 2208205719**

**DOSEN PEMBIMBING
Mochamad Hariadi, S.T., M.Sc., Ph.D.**

**PROGRAM MAGISTER
BIDANG KEAHLIAN JARINGAN CERDAS MULTIMEDIA
KONSENTRASI TEKNOLOGI PERMAINAN
JURUSAN TEKNIK ELEKTRO
FAKULTAS TEKNOLOGI INDUSTRI
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2014**



THESIS TE092099

**INFERENCE IN INFORMATION RETRIEVAL
SYSTEM USING SEMANTIC BAYESIAN NETWORK**

**JOKO SAMODRA
NRP. 2208205719**

**SUPERVISOR
Mochamad Hariadi, S.T., M.Sc., Ph.D.**

**MASTER PROGRAM
INTELLIGENT NETWORKING OF MULTIMEDIA
GAME TECHNOLOGY
DEPARTMENT OF ELECTRICAL ENGINEERING
FACULTY OF INDUSTRIAL TECHNOLOGY
SEPULUH NOPEMBER INSTITUTE OF TECHNOLOGY
SURABAYA
2014**

INFERENSI PADA SISTEM TEMU KEMBALI INFORMASI DENGAN MENGGUNAKAN JARINGAN BAYES SEMANTIK

Nama mahasiswa : Joko Samodra
NRP : 2208205719
Pembimbing : Mochamad Hariadi, S.T., M.Sc., Ph.D.

ABSTRAK

Sistem temu kembali informasi merupakan salah satu model yang dapat membantu pengguna untuk menemukan informasi yang sesuai dengan kebutuhan. Beberapa model yang telah digunakan dalam sistem temu kembali informasi diantaranya yaitu *logical models*, *vector processing models*, *probabilistic models*, dan *cognitive models*. Model-model tersebut sebagian besar bekerja berdasarkan proses statis, sehingga kemungkinan besar akan menemui kesulitan jika digunakan untuk melakukan tugas-tugas yang lebih rumit, misalnya pada saat jumlah dokumen atau informasi yang dikelola semakin banyak dan tingkat kemiripan pola diantara dokumen atau informasi tersebut semakin tinggi. Teknik lain yang juga telah diteliti yaitu sistem temu kembali informasi yang fleksibel dengan menggunakan inferensi berbasis jaringan Bayes dan jaringan Bayes semantik. Pada penelitian ini dibahas penggunaan jaringan Bayes semantik untuk proses inferensi pada sistem temu kembali informasi.

Hasil ujicoba menunjukkan bahwa jika pada kalimat permintaan (*query*) memuat kata kunci nomor seri atau kode model produk elektronik, maka data target dapat ditemukan dengan pencarian berdasarkan pencocokan kata kunci. Sedangkan inferensi menggunakan jaringan Bayes dapat menemukan data target jika pada kalimat permintaan memuat nilai data atribut produk dengan jumlah yang mencukupi. Untuk menemukan data target yang memiliki identitas 3 buah atribut, 39% hasil uji coba membutuhkan input 2 kata kunci, dan 61% hasil ujicoba membutuhkan input 3 kata kunci. Dan untuk menemukan data target yang memiliki identitas 4 buah atribut, 40% hasil uji coba membutuhkan input 3 kata kunci, dan 60% hasil ujicoba membutuhkan input 4 kata kunci. Jika inferensi menggunakan jaringan Bayes belum dapat menemukan data target karena nilai data atribut yang dimasukkan belum mencukupi atau terlalu sedikit, sistem secara otomatis dapat meminta tambahan informasi yang berkaitan erat dengan data target yang sedang dicari dengan menggunakan hubungan semantik objek tersebut pada jaringan semantik.

Kata kunci: temu kembali informasi, inferensi, jaringan bayes semantik

INFERENCE IN INFORMATION RETRIEVAL SYSTEM USING SEMANTIC BAYESIAN NETWORK

Student Name : Joko Samodra
NRP : 2208205719
Instructor : Mochamad Hariadi, S.T., M.Sc., Ph.D.

ABSTRACT

Information retrieval system is one of the model that can help user to find an information. Several models have been used in information retrieval system such as logical models, vector processing models, probabilistic models, and cognitive models . That models are mostly working by a static process, so it most likely would have difficulty if it is used to perform the more complex tasks, such as when the number of document or information are large and the degree of similarity between documents or information pattern is higher . Another technique that has been studied is the flexible information retrieval system that use inference based on Bayesian network and Bayesian semantic network . This research will discuss the using of semantic Bayesian networks for inference process in information retrieval.

The test results show that if the query contains keyword of the serial number or model code of electronic products, the target data can be found by searching based on a keyword matching. While inference using Bayesian networks can find the target data if the query contains the sufficient value of product attribute . To find the target data that has 3 attributes, 39% of the test results require 2 input keywords, and 61% of the test results require 3 input keywords . And to find the target data that has 4 attributes, 40% of the test results require 3 input keywords, and 60% of the test results require 4 input keywords . If the inference using Bayesian networks have not been able to find the target data because the attribute values entered are not enough or too little, the system can automatically ask for additional information which is closely related to the target data using the semantic relations in the semantic network .

Keywords: *information retrieval, inference, semantic bayesian network*

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Teknik (M.T.)
Di Institut Teknologi Sepuluh Nopember

oleh :
Joko Samodra
Nrp. 2208205719

Tanggal Ujian : 13 Januari 2014
Periode Wisuda : Maret 2014

Disetujui oleh :



1. Mochamad Hariadi, S.T., M.Sc., Ph.D. (Pembimbing)
NIP: 196912091997031002



2. Dr. Supeno Mardi SN, S.T., M.T. (Penguji)
NIP: 197003131995121001



3. Christyowidiasmoro, S.T, M.T., M.Sc. (Penguji)
NIP: 198301272009121004

Direktur Program Pasca Sarjana,



Prof. Dr. Ir. Adi Soeprijanto, M.T.
NIP: 196404051990021001

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT serta sholawat kepada Nabi Muhammad SAW atas selesainya penulisan buku tesis dengan judul :

INFERENSI PADA SISTEM TEMU KEMBALI INFORMASI DENGAN MENGGUNAKAN JARINGAN BAYES SEMANTIK.

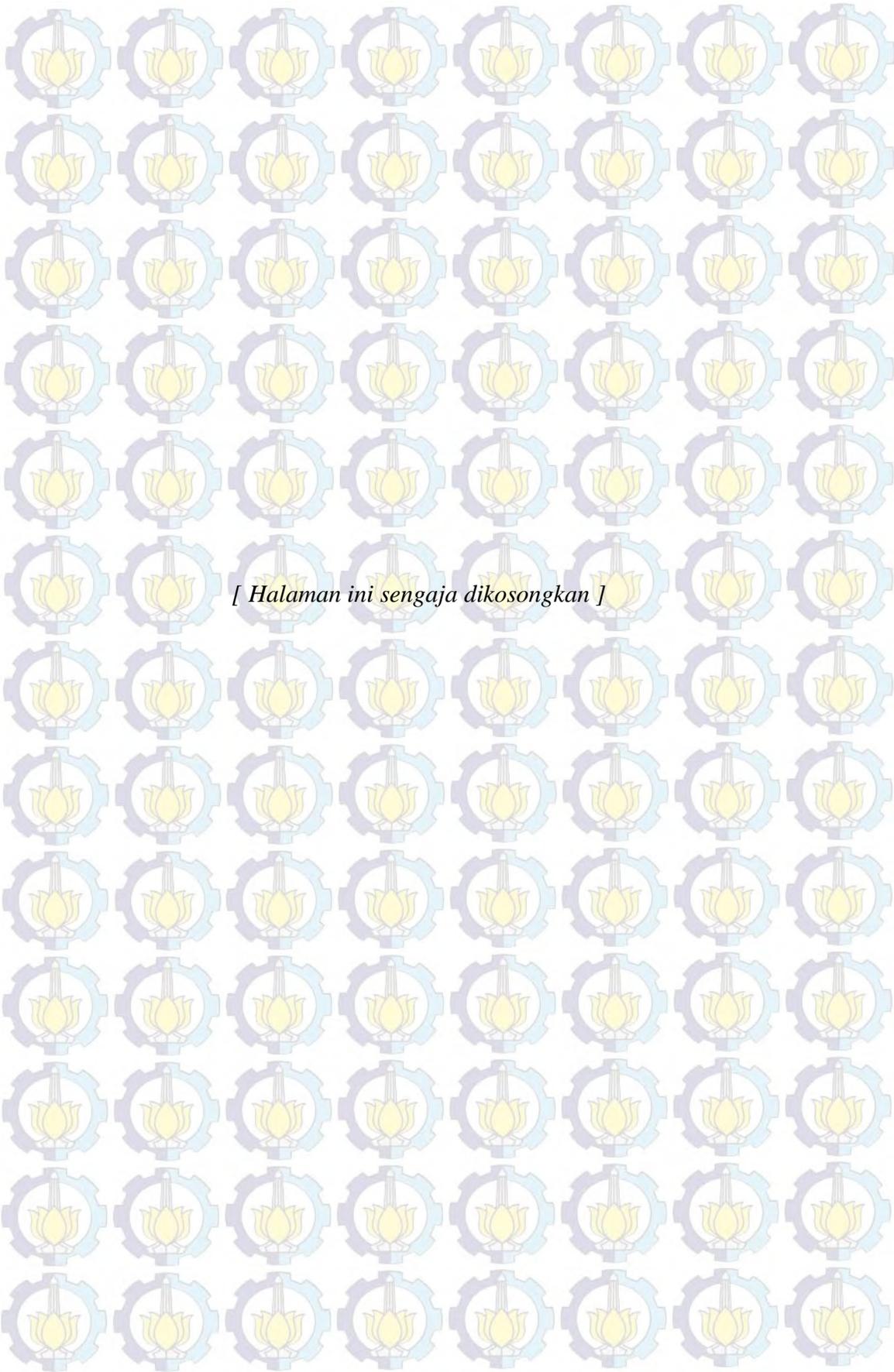
Penulis menyadari bahwa dalam penulisan buku tesis ini masih terdapat berbagai kekurangan dan kelemahan karena keterbatasan ilmu dan kemampuan penulis, oleh karena itu penulis membuka diri seluas-luasnya untuk menerima kritik dan saran yang bersifat membangun dari pembaca.

Ucapan terima kasih penulis sampaikan kepada :

1. Kedua orang tua penulis yang telah almarhum, istri dan kedua buah hati, serta semua keluarga besar penulis, atas semua dukungan, doa dan perhatian yang diberikan kepada penulis hingga saat ini.
2. Bapak Mochamad Hariadi, S.T., M.Sc., Ph.D. selaku dosen pembimbing yang telah memberikan bimbingan dan dukungan kepada penulis dalam menyelesaikan buku tesis ini.
3. Bapak-ibu dosen pengajar bidang keahlian Jaringan Cerdas Multimedia konsentrasi Teknologi Permainan Jurusan Teknik Elektro, atas pengajaran, bimbingan, serta perhatian yang diberikan kepada penulis selama ini.
4. Seluruh pihak yang secara langsung maupun tidak langsung ikut membantu penulis dalam menyelesaikan buku tesis ini.

Terakhir, penulis memohon maaf jika terdapat kesalahan dan kekurangan dalam penyusunan buku tesis ini. Penulis berharap semoga buku ini bermanfaat khususnya bagi penulis sendiri dan umumnya bagi pembaca.

Penulis

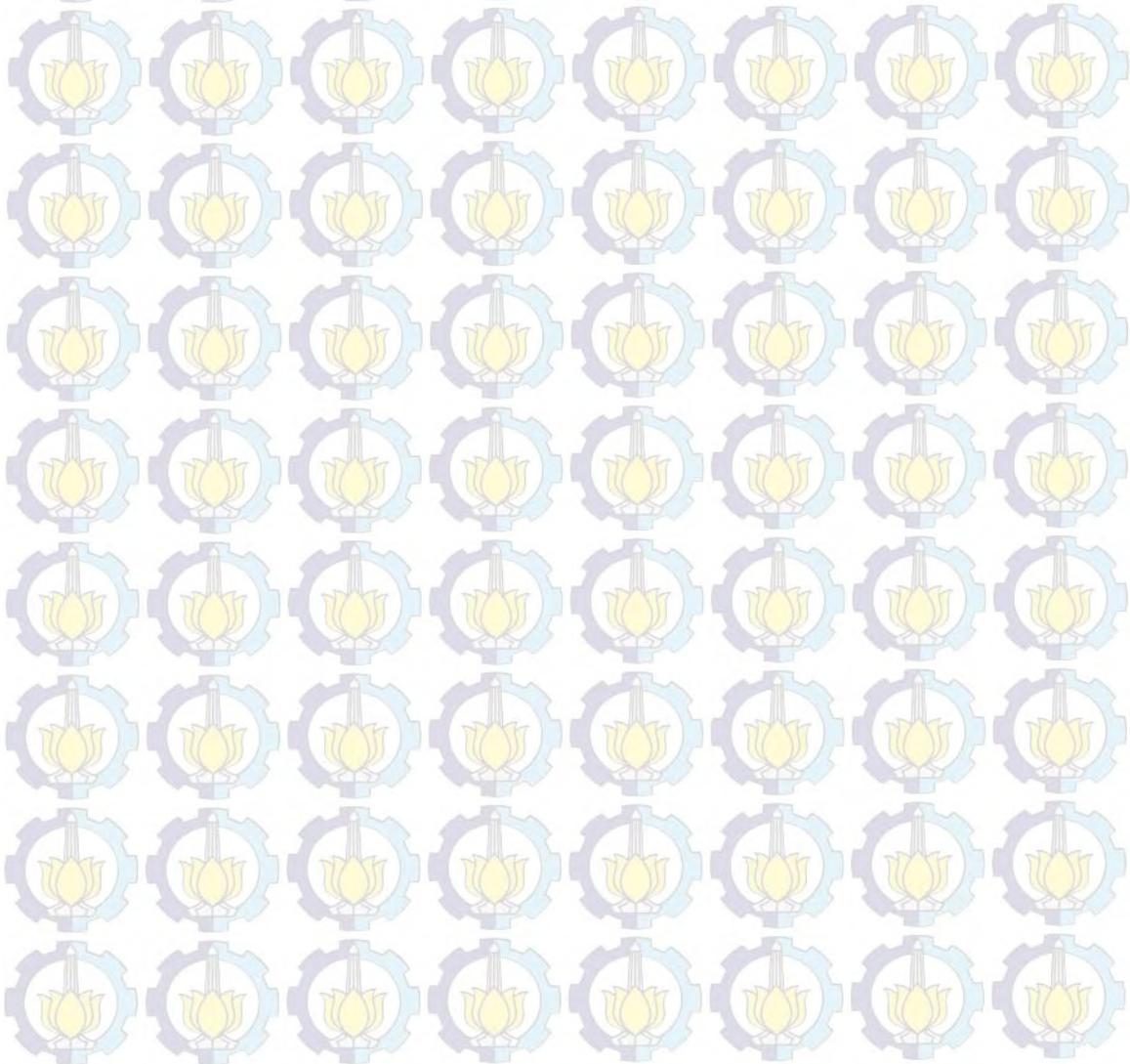


[Halaman ini sengaja dikosongkan]

DAFTAR ISI

Abstrak	i
<i>Abstract</i>	iii
Kata Pengantar	v
Daftar Isi	vii
Daftar Gambar	ix
Daftar Tabel	xi
Bab 1. Pendahuluan	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan dan Manfaat Penelitian	3
1.4. Batasan Masalah	3
1.5. Metodologi	4
1.6. Sistematika Penulisan	5
Bab 2. Kajian Pustaka dan Dasar Teori	7
2.1. Temu Kembali Informasi (<i>Information Retrieval</i>)	7
2.1.1. Komponen Sistem Temu Kembali Informasi	8
2.1.2. Jenis Sistem Temu Kembali Informasi	10
2.2. Jaringan Semantik	10
2.3. Jaringan Bayes	14
2.3.1. Jaringan Bayes untuk Temu Kembali Informasi	15
2.3.2. Jaringan Bayes Semantik untuk Inferensi	16
Bab 3. Metode Penelitian	21
3.1. Bagan Penelitian	21
3.2. Modul Dialog	23
3.3. Modul Inferensi	29
3.3.1. Struktur Jaringan Bayes	33
3.3.2. Struktur Jaringan Semantik	38

Bab 4. Analisa Hasil dan Pembahasan	41
4.1. Percobaan 1	41
4.2. Percobaan 2	42
4.3. Percobaan 3	52
4.4. Percobaan 4	62
4.5. Percobaan 5	69
Bab 5. Kesimpulan dan Saran	85
5.1. Kesimpulan	85
5.2. Saran dan Penelitian Selanjutnya	86
Daftar Pustaka	87



DAFTAR GAMBAR

2.1. Contoh Jaringan Semantik	11
2.2. Jaringan Bayes pada Temu Kembali Informasi	16
2.3. Jaringan Bayes semantik untuk inferensi	17
3.1. Blok diagram penerapan inferensi pada sistem temu kembali informasi	21
3.2. Diagram alir penerapan inferensi pada sistem temu kembali informasi	22
3.3. Diagram alir pemasukan permintaan pengguna	24
3.4. Diagram alir pemberian tanggapan atau jawaban kepada pengguna ..	28
3.5. Diagram alir proses inferensi menggunakan jaringan Bayes semantik untuk menemukan informasi yang relevan dengan permintaan pengguna	31
3.5. Jaringan Bayes yang digunakan oleh modul inferensi	36
3.6. Jaringan semantik yang digunakan oleh modul inferensi	39



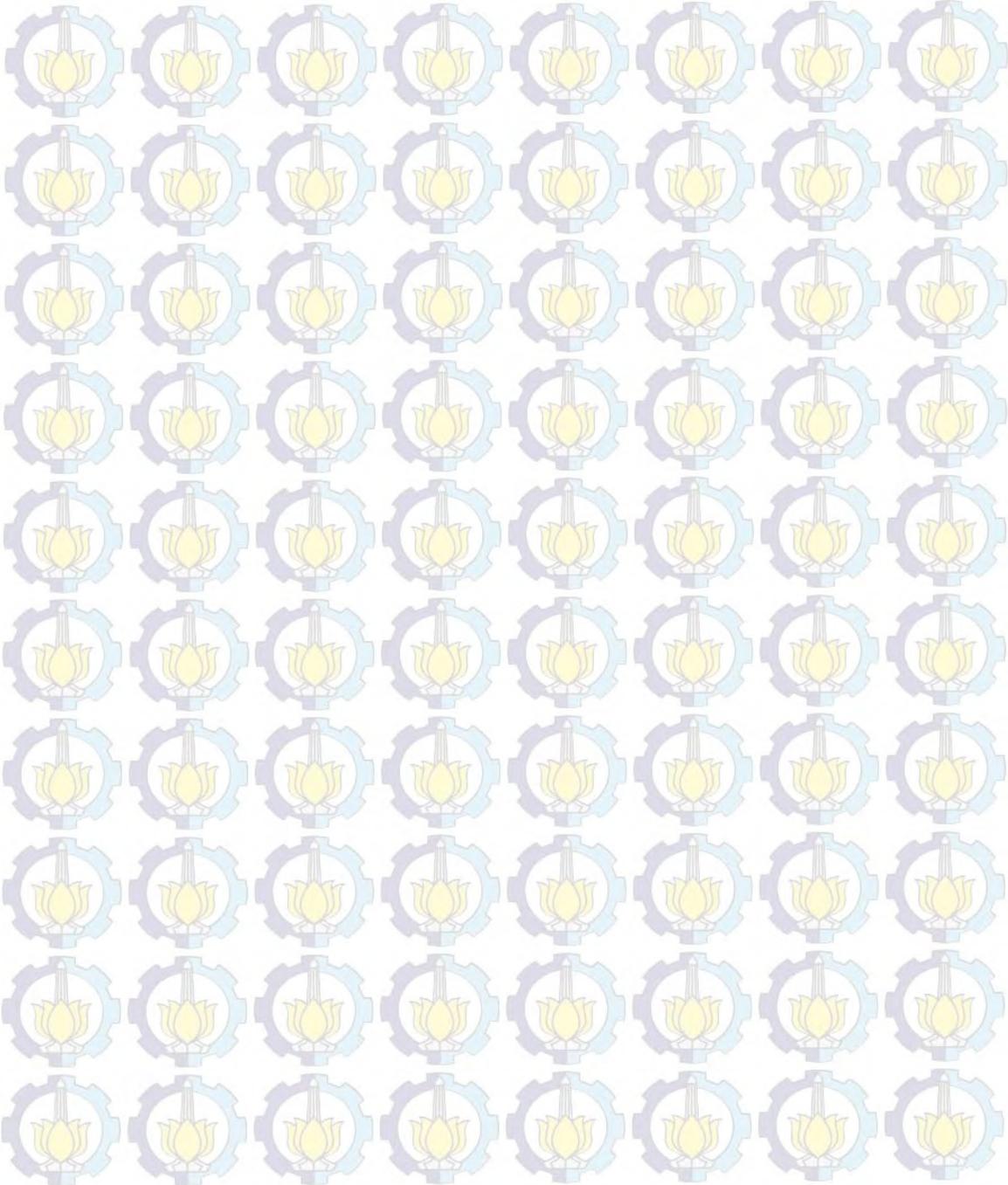
[Halaman ini sengaja dikosongkan]

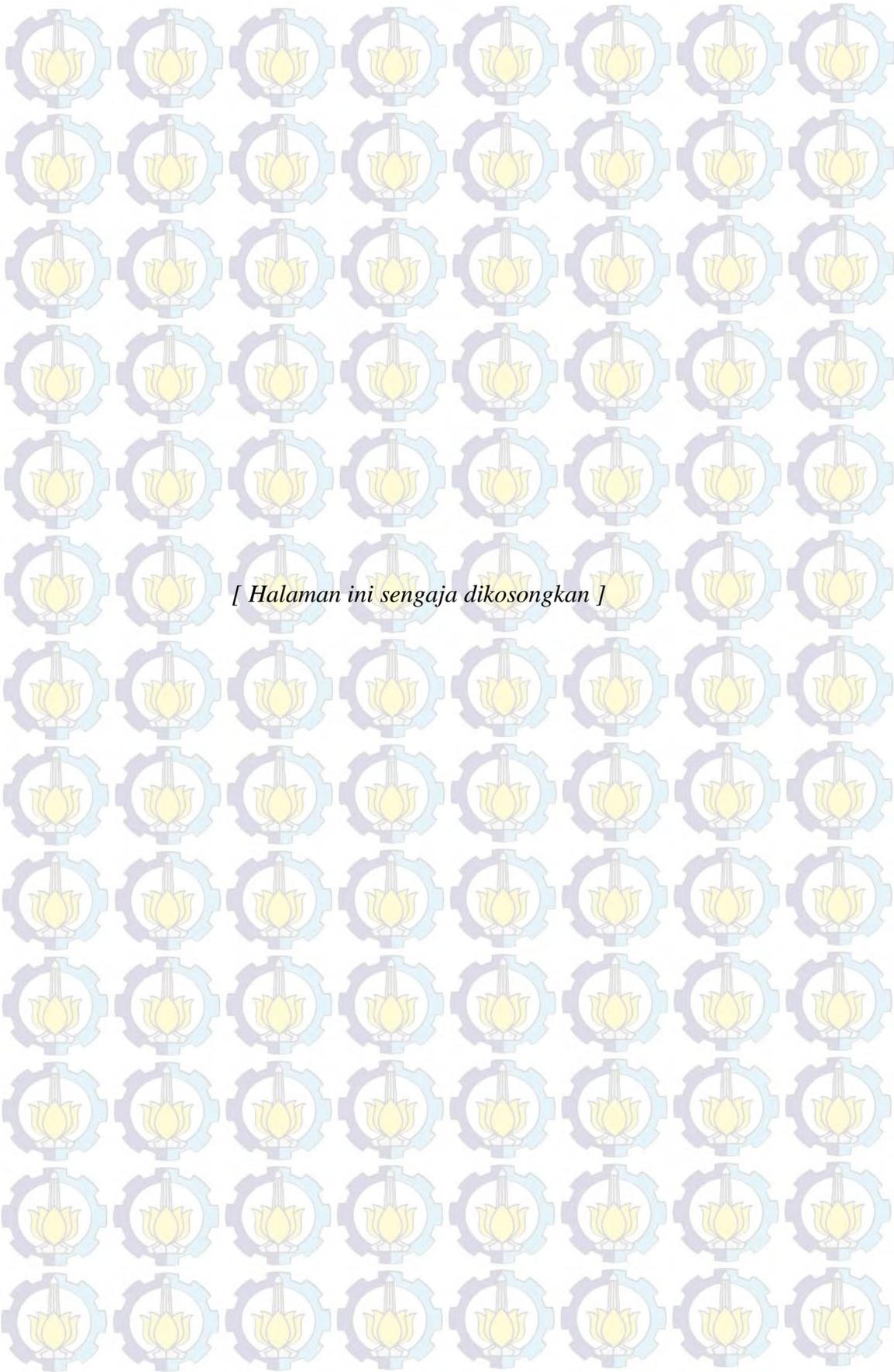
DAFTAR TABEL

3.1. Himpunan Kata Kunci	26
3.2. Struktur Tabel Kata Kunci	26
3.3. Sinonim	27
3.4. Struktur Tabel Sinonim	27
3.5. Struktur Tabel Produk	29
3.6. Probabilitas $P(t_i = 1 k_i = 1)$	35
3.7. Probabilitas $P(t_i = 1 k_i = 0)$	35
3.8. Rata-rata minat terhadap merk tv dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes	36
3.9. Rata-rata minat terhadap ukuran tv dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes	36
3.10. Rata-rata minat terhadap jenis layar tv dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes	36
3.11. Rata-rata minat terhadap merk handycam dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes	36
3.12. Rata-rata minat terhadap media penyimpan handycam dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes	37
3.13. Rata-rata minat terhadap kemampuan zoom handycam dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes	37
3.14. Rata-rata minat terhadap jenis sensor handycam dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes	37
3.15. Nilai probabilitas awal simpul data target pada jaringan Bayes	37
3.16. Struktur Tabel Relasi	40
4.1. Hasil pengujian dengan input awal memuat 2 atribut yaitu “merk_tv” dan “ukuran” layar	43
4.2. Hasil pengujian dengan input awal memuat 2 atribut yaitu “ukuran” dan jenis “layar”	44
4.3. Hasil pencarian berdasarkan pencocokan kata kunci pada percobaan 2	49

4.4. Probabilitas simpul-simpul yang terhubung dengan kata kunci pada percobaan 2	49
4.5. Hasil pengujian dengan input awal memuat satu atribut yaitu “layar”	52
4.6. Hasil pencarian berdasarkan pencocokan kata kunci pada percobaan 3	58
4.7. Probabilitas simpul-simpul yang terhubung dengan kata kunci pada percobaan 3	59
4.8. Simpul jaringan semantik yang memiliki hubungan dengan objek “tv” dan “samsung” pada percobaan 3	60
4.9. Nilai data (<i>value</i>) dari atribut “ukuran” pada percobaan 3.....	60
4.10. Probabilitas simpul-simpul yang terhubung dengan kata kunci awal dan kata kunci tambahan pada percobaan 3	61
4.11. Hasil pengujian dengan input awal memuat dua atribut yaitu “merk” dan “media” penyimpanan	62
4.12. Hasil pencarian berdasarkan pencocokan kata kunci pada percobaan 4	66
4.13. Probabilitas simpul-simpul yang terhubung dengan kata kunci pada percobaan 4	67
4.14. Simpul jaringan semantik yang memiliki hubungan dengan objek “handycam”, “sony”, dan “hdd” pada percobaan 4	67
4.15. Nilai data (<i>value</i>) dari atribut “zoom” pada percobaan 4	68
4.16. Probabilitas simpul-simpul yang terhubung dengan kata kunci awal dan kata kunci tambahan pada percobaan 4	69
4.17. Hasil pengujian dengan input awal memuat satu atribut yaitu optical “zoom’	70
4.18. Hasil pencarian berdasarkan pencocokan kata kunci pada percobaan 5	75
4.19. Probabilitas simpul-simpul yang terhubung dengan kata kunci pada percobaan 5	76
4.20. Simpul yang memiliki hubungan dengan objek “handycam” dan “sony” pada percobaan 5	78
4.21. Nilai data (<i>value</i>) dari atribut “media” pada percobaan 5	79
4.22. Probabilitas simpul-simpul yang terhubung dengan kata kunci awal dan kata kunci tambahan pada percobaan 5	79

4.23. Simpul jaringan semantik yang memiliki hubungan dengan objek “handycam”, “sony” dan “sdcard” pada percobaan 5	80
4.24. Nilai data (<i>value</i>) dari atribut “zoom” pada percobaan 5	81
4.25. Probabilitas simpul-simpul yang terhubung dengan kata kunci awal dan kata kunci tambahan kedua pada percobaan 5	82
4.26. Rekapitulasi hasil ujicoba unjuk kerja sistem dalam menemukan data target berdasarkan jumlah input kata kunci	83





BAB 1

PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi informasi yang semakin pesat memberikan dampak yang sangat luas bagi penyebaran informasi, salah satunya adalah pemanfaatan dokumen digital dan basis data sebagai media untuk menyimpan dan menyebarkan informasi. Jumlah informasi atau dokumen yang tersimpan di media penyimpan digital seperti seperti harddisk, flashdisk, cd, atau dvd, dan dokumen web setiap waktu semakin bertambah hingga mencapai milyaran dokumen. Jika kita dapat memanfaatkan dan mengeksplorasi sedemikian rupa informasi-informasi ini, maka kita akan dapat memiliki gudang informasi yang sangat luas dan senantiasa berkembang secara dinamis.

Menggali informasi dari kumpulan dokumen yang sangat besar seperti basis data, dokumen web, dan dokumen digital yang lainnya tentunya bukan pekerjaan yang mudah. Karena itu diperlukan metode tertentu yang nantinya dapat mengolah dan mengorganisir sedemikian rupa kumpulan dokumen-dokumen yang dimiliki, sehingga dapat membantu kita dalam mendapatkan kembali informasi yang sesuai dengan kebutuhan. Metode ini selanjutnya disebut dengan temu kembali informasi (*information retrieval*).

Dalam penelitian ini, temu kembali informasi akan digunakan sebagai pendekatan untuk mendapatkan informasi yang relevan dari kumpulan informasi yang dimiliki oleh sebuah sistem pencarian atau temu kembali informasi. Sedangkan sistem temu kembali informasi yang dimaksudkan adalah sebuah sistem pencarian informasi yang menyimpan atau memiliki kumpulan informasi dengan domain tertentu sesuai dengan tujuan penggunaan dari sistem temu kembali informasi tersebut, dan memiliki kemampuan untuk menyajikan informasi yang sesuai dengan kebutuhan pengguna.

Beberapa model yang telah digunakan dalam sistem temu kembali informasi diantaranya yaitu *logical models*, *vector processing models*, *probabilistic models*, dan *cognitive models* (Hasugian, 2009). Model-model

tersebut sebagian besar bekerja berdasarkan proses statis, sehingga kemungkinan besar akan menemui kesulitan jika digunakan untuk melakukan tugas-tugas yang lebih rumit, misalnya pada saat jumlah dokumen atau informasi yang dikelola semakin banyak dan tingkat kemiripan pola diantara dokumen atau informasi tersebut semakin tinggi. Teknik lain yang juga telah diteliti oleh para peneliti yaitu sistem temu kembali informasi menggunakan model dialog yang fleksibel dengan menggunakan jaringan Bayes (*Bayesian Networks*) (Hong and Cho, 2003; Horvitz et al., 1998).

Selain itu, para peneliti juga mengembangkan model berbasis perencanaan, yang memiliki kemampuan untuk mempertimbangkan rencana dan memutuskan sebuah tindakan kepada pengguna. Pada pendekatan ini, sistem temu kembali informasi yang bekerja sebagai agen akan mengumpulkan informasi dari setiap *query* yang diberikan oleh pengguna, kemudian sistem dapat memprediksi maksud yang diinginkan oleh pengguna secara berangsur-angsur (Perugini and Ramakrishnan, 2003). Baru-baru ini, Kim, Hong dan Cho telah mengembangkan agen percakapan yang mampu menyimpulkan maksud dari pengguna secara rinci dengan menggunakan metode inferensi berbasis jaringan Bayes semantik dan interaksi inisiatif campuran untuk mengumpulkan informasi tambahan dari pengguna. Dan metode ini memiliki kinerja yang baik dalam memberikan tanggapan yang cocok dengan keinginan pengguna, dibandingkan dengan agen yang berbasis skrip dan berbasis jaringan Bayes (Kim et al., 2007).

Pada penelitian ini mengusulkan penggunaan jaringan Bayes semantik untuk melakukan proses inferensi pada sebuah sistem temu kembali informasi. Metode ini akan diujikan pada sebuah sistem temu kembali informasi yang bertindak sebagai penyedia informasi atau pemandu penjualan produk elektronik, dimana sistem tersebut memiliki basis data berupa kumpulan informasi produk-produk elektronik yang akan dijual, dan akan memberikan informasi yang relevan berupa sebuah produk elektronik yang spesifik berdasarkan *query* yang diberikan oleh pengguna.

1.2. Rumusan Masalah

Saat ini berbagai informasi seperti basis data, dokumen web, dan dokumen digital yang lainnya telah tersedia dalam jumlah yang sangat besar, sehingga diperlukan metode yang dapat digunakan untuk mengakses dan menemukan kembali informasi-informasi tersebut.

Berbagai metode telah digunakan untuk merancang sistem temu kembali informasi, tetapi sebagian besar bekerja berdasarkan proses statis sehingga kemungkinan besar akan menemui kesulitan jika digunakan untuk melakukan tugas-tugas yang lebih rumit. Karena itu diperlukan metode lain yang dapat berinteraksi dengan pengguna yang secara otomatis dapat meminta tambahan informasi yang diperlukan, sehingga sistem dapat menemukan informasi yang paling sesuai dengan permintaan pengguna.

1.3. Tujuan dan Manfaat Penelitian

Penelitian ini bertujuan untuk menghasilkan sistem temu kembali informasi yang dapat menyimpulkan apa yang diinginkan oleh pengguna dengan menggunakan metode inferensi berbasis jaringan Bayes semantik. Pada pendekatan ini, sistem temu kembali informasi akan mengumpulkan informasi dari *query* yang diberikan oleh pengguna, kemudian akan memprediksi maksud yang diinginkan oleh pengguna. Jika informasi yang diberikan belum mencukupi, maka sistem akan mengajukan pertanyaan untuk meminta tambahan informasi yang berkaitan dengan informasi yang sedang dicari. Proses pengumpulan informasi tambahan ini dilakukan oleh sistem sampai akhirnya sistem dapat menyimpulkan dan memberikan informasi yang paling sesuai dengan permintaan pengguna.

Sedangkan manfaat yang diharapkan pada penelitian ini yaitu didapatkannya sebuah metode yang dapat digunakan untuk melakukan inferensi pada sebuah sistem temu kembali informasi.

1.4. Batasan Masalah

Untuk menyelesaikan permasalahan yang telah dirumuskan di atas, diberikan batasan sebagai berikut :

Sistem temu kembali informasi yang digunakan untuk ujicoba dalam penelitian ini adalah sebuah sistem pencarian atau pemandu penjualan, dimana sistem tersebut memiliki basis data yang berisi kumpulan beberapa jenis produk elektronik.

1.5. Metodologi

Metodologi dalam penyusunan tesis ini adalah sebagai berikut :

1. Studi Literatur

Pengumpulan pustaka untuk dipelajari dalam pengerjaan dan penelitian tesis.

Pustaka yang dikumpulkan berupa *paper*, *text book*, dan tutorial-tutorial tentang teori inferensi, agen percakapan, jaringan Bayes, dan jaringan Bayes semantik.

2. Pengumpulan Data

Pengumpulan data yang diperlukan untuk menguji kinerja sistem yang telah diimplementasikan. Data yang dikumpulkan adalah data teks berbahasa Indonesia yang merupakan deskripsi produk-produk barang elektronik.

3. Perancangan dan Implementasi Sistem

Perancangan dan implementasi sistem akan digunakan dengan mendesain metode-metode yang ada, dan diimplementasikan dengan menggabungkan metode-metode tersebut sehingga didapatkan hasil yang diinginkan.

4. Pengujian Sistem

Pengujian sistem dilakukan dengan memberikan berbagai macam data masukan yang berupa teks berbahasa Indonesia. Pengujian sistem diperlukan untuk mengetahui ketepatan sistem dalam memberikan tanggapan yang sesuai dengan hasil yang diharapkan berdasarkan data masukan yang diberikan ke dalam sistem.

5. Penyusunan Laporan

Melakukan penulisan yang menunjukkan hasil akhir dari tesis. Tema dalam tesis ini adalah menyajikan informasi yang relevan berdasarkan masukan yang berupa teks berbahasa Indonesia.

1.6 Sistematika Penulisan

Sistematika penulisan laporan tesis ini dibagi dalam lima bab, masing-masing bab dapat diuraikan sebagai berikut :

BAB I Bab ini menjelaskan tentang pendahuluan dari tesis ini yang meliputi latar belakang, tujuan dan manfaat, batasan masalah, metodologi penelitian dan sistematika penulisan penyusunan tesis. Bab ini memaparkan berbagai fenomena yang ada disekitar kita sehingga menimbulkan alasan untuk melakukan penelitian dengan tema yang ada dalam tesis ini. Dari bab ini juga dapat dilihat garis besar kegiatan yang dilakukan dan manfaat dari penelitian ini.

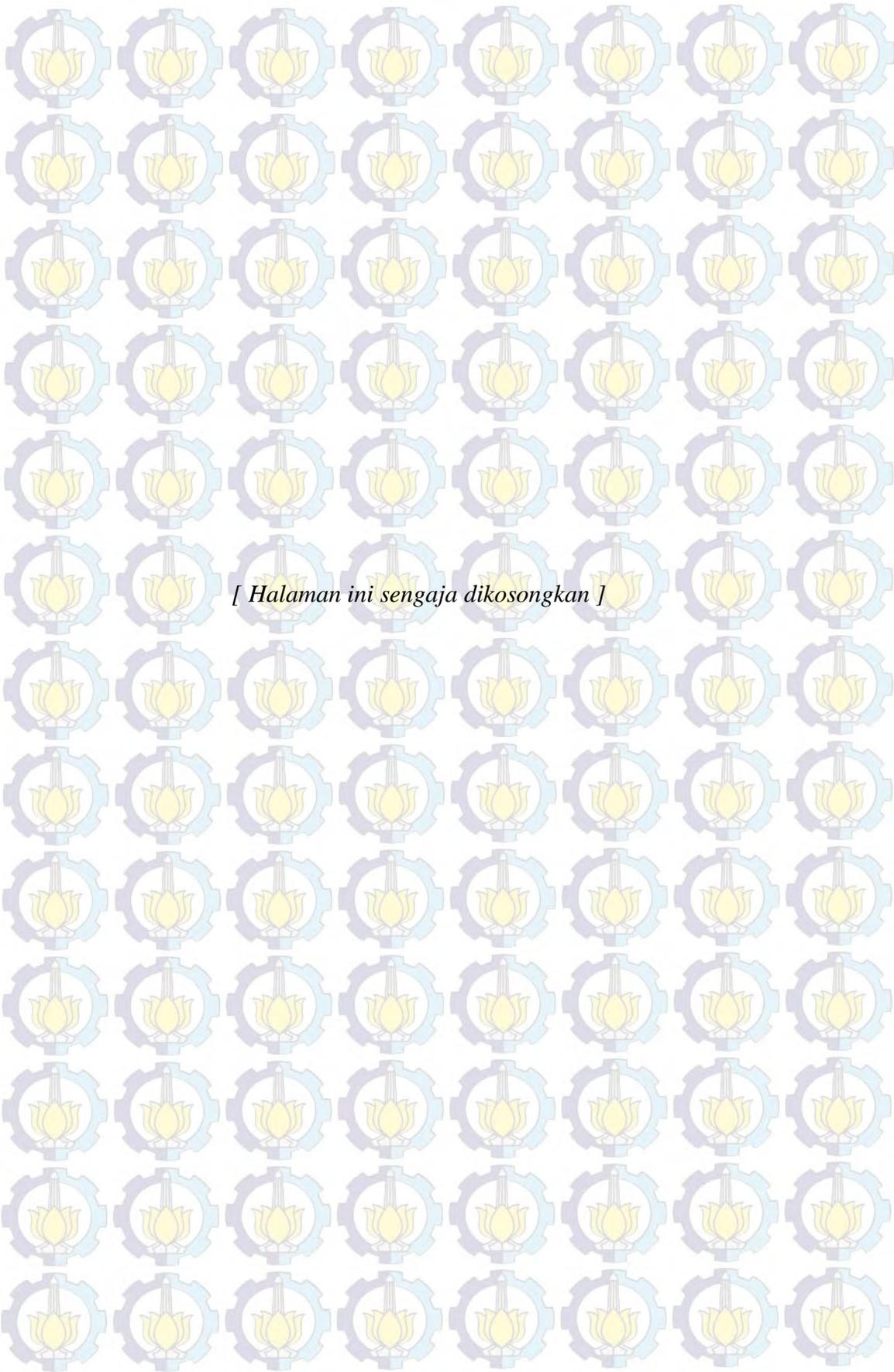
BAB II Pada bab ini, akan dijelaskan teori-teori penunjang yang digunakan sebagai dasar dalam penyusunan tesis ini.

BAB III Pada bab ini, akan membahas perancangan sistem dalam bentuk diagram blok. Diagram ini menggambarkan alur kerja sistem yang digunakan pada penelitian ini, mulai dari data masukan yang diberikan ke dalam sistem, proses inferensi, sampai dengan penyajian informasi yang relevan sebagai hasil akhir dari pengolahan data yang dilakukan oleh sistem.

BAB IV Pada bab ini, akan dijelaskan implementasi penelitian yang dilakukan sesuai dengan rancangan yang sudah dibuat serta pembahasannya.

BAB V Pada bab ini merupakan penutup yang berisi kesimpulan, saran dan rencana ke depan yang diambil dari penelitian ini.

LAMPIRAN Berisi data-data pendukung yang digunakan pada penyusunan buku tesis ini.



BAB 2 KAJIAN PUSTAKA DAN DASAR TEORI

2.1. Temu Kembali Informasi (*Information Retrieval*)

Sistem temu kembali informasi merupakan sebuah sistem informasi yang memiliki kemampuan untuk mencari dan menemukan informasi yang sesuai dengan apa yang dibutuhkan oleh pengguna. Sistem ini dapat berfungsi sebagai media penghubung antara sumber informasi yang tersimpan di sebuah media atau lokasi tertentu dengan pengguna yang membutuhkan berbagai macam informasi sesuai dengan tingkat kebutuhannya masing-masing.

Pengertian temu kembali informasi (*information retrieval*) adalah sebuah metode untuk menemukan materi (biasanya dokumen) dari sesuatu yang tidak terstruktur secara alami (biasanya teks) untuk mendapatkan informasi yang dikehendaki dari kumpulan dokumen yang besar (biasanya tersimpan di komputer) (Manning et al., 2008). Dari definisi di atas dapat dilihat bahwa metode untuk menemukan sebuah informasi menjadi hal yang sangat penting diantaranya karena sumber informasi yang berupa basis data atau kumpulan dokumen digital dapat memiliki ukuran yang sangat besar, dan secara dinamis kumpulan dokumen tersebut akan mengalami penambahan atau peningkatan volume. Sehingga tanpa menggunakan metode yang efektif, maka pengguna akan mengalami kesulitan pada saat ingin menemukan informasi yang dibutuhkan.

Definisi yang lain dari temu kembali informasi yaitu “information retrieval is a communication process. In one sense it is a means by which authors or creators of records communicate with readers, but indirectly and with a possibly long time lag between creation of a message or text and its delivery to the IR system user” (Meadow, 1992). Pada definisi di atas disebutkan bahwa temu kembali informasi merupakan sebuah proses komunikasi, yang berarti ada sebuah pesan yang ingin disampaikan oleh sumber atau pemilik informasi kepada pengguna atau audien yang nantinya akan memanfaatkan pesan atau informasi yang disampaikan tersebut. Proses penyampaian informasi ini dilakukan secara tidak langsung, dimana ada jeda waktu antara pembuatan informasi yang

dilakukan oleh pemilik informasi sampai akhirnya informasi tersebut diterima oleh penggunanya.

Dari kedua definisi di atas dapat disimpulkan bahwa temu kembali informasi merupakan bagian dari sebuah proses komunikasi tidak langsung yang dituangkan dalam bentuk metode yang berfungsi untuk menemukan informasi yang dibutuhkan oleh pengguna dari kumpulan dokumen yang tersedia. Dokumen yang dimaksudkan dapat berupa basis data, dokumen web, atau dokumen digital yang tersimpan di CDROM, harddisk dan media penyimpan lainnya

2.1.1. Komponen Sistem Temu Kembali Informasi

Seperti pada jenis sistem informasi yang lain, sistem temu kembali informasi secara garis besar terdiri dari 3 komponen utama yaitu komponen masukan (*input*), komponen pemroses (*processor*), dan komponen keluaran (*output*), serta komponen eksternal yang tidak dapat dipisahkan dari sebuah sistem informasi yaitu pengguna (*user*). Komponen masukan bertugas untuk memasukkan data-data permintaan pengguna yang nantinya akan diumpungkan dan diolah oleh komponen pemroses. Komponen pemroses bertugas untuk menstrukturkan data-data permintaan pengguna ke dalam bentuk representasi dokumen yang tepat, dan sekaligus melakukan proses temu kembali atau pencarian dengan menggunakan metode tertentu. Dan komponen keluaran adalah produk yang dihasilkan oleh komponen pemroses yang berbentuk kumpulan informasi dari sebuah dokumen, atau acuan ke dokumen yang lain.

Secara lebih spesifik, komponen-komponen dalam sebuah sistem temu kembali informasi juga dapat dijelaskan sebagai berikut :

1. Pengguna (*user*)

Berdasarkan tujuan penggunaannya, jenis pengguna dalam sistem temu kembali informasi dibedakan menjadi 2 macam yaitu pengguna (*user*) dan pengguna akhir (*end user*). Pengguna adalah orang yang mengoperasikan atau menggunakan sistem temu kembali informasi untuk melakukan 2 jenis kegiatan yang berbeda yaitu kegiatan untuk mencari atau menemukan informasi yang sesuai dengan kebutuhan, dan yang kedua yaitu kegiatan yang berhubungan dengan pengelolaan sistem misalnya memasukkan data, memperbarui data,

pemeliharaan sistem, serta kegiatan-kegiatan pengelolaan yang lainnya. Sedangkan pengguna akhir hanya memanfaatkan sistem temu kembali informasi untuk tujuan mencari atau menemukan informasi saja, sesuai dengan tingkat kebutuhannya masing-masing.

2. Kalimat Permintaan (*Query*)

Kalimat permintaan (*query*) merupakan sebuah kata atau kumpulan kata-kata yang disusun oleh pengguna sedemikian rupa dan dimasukkan ke dalam sistem dengan tujuan agar diproses oleh sistem temu kembali informasi untuk menghasilkan informasi yang diinginkan oleh pengguna tersebut. *Query* dapat berbentuk kata tunggal maupun kumpulan kata-kata yang membentuk sebuah frasa. Contoh *query* yang menggunakan kata tunggal misalnya laptop, kamera, buku, musik. Dan contoh *query* yang menggunakan frasa misalnya kamera dslr, tv toshiba 32 in, buku pemrograman java, musik keroncong.

3. Dokumen

Yang dimaksud dokumen dalam sistem temu kembali informasi adalah seluruh sumber informasi yang dimiliki oleh sistem, dimana dokumen tersebut dapat berupa kumpulan dokumen web, buku elektronik, atau dokumen digital lainnya yang telah dimasukkan ke dalam sistem dan diorganisasikan secara sistematis ke dalam sebuah basis data (*database*). Selanjutnya kumpulan dokumen-dokumen tersebut disimpan ke dalam sebuah media penyimpan elektronik misalnya cdrom atau harddisk.

4. Indeks Dokumen

Indeks adalah kumpulan kata atau istilah yang digunakan sebagai sarana untuk merepresentasikan atau menggambarkan karakteristik dari dokumen yang dimiliki oleh sistem. Misalnya kata kunci yang merujuk ke sebuah nama produk, judul berita, ukuran layar, atau merk barang. Selain berfungsi sebagai representasi dari dokumen, indeks juga berguna untuk mempermudah dan mempercepat proses pencarian informasi.

5. Mesin Pemroses

Mesin pemroses bertugas untuk melakukan proses pencarian informasi dan mencocokkan kata-kata kunci yang terdapat pada *query* pengguna dengan indeks dokumen yang dimiliki oleh sistem. Setelah menemukan informasi atau dokumen yang sesuai dengan permintaan pengguna, mesin pemroses akan menampilkan hasil keluaran sebagai tanggapan atas permintaan dari pengguna.

2.1.2. Jenis Sistem Temu Kembali Informasi

Beberapa teknik atau model yang telah digunakan dalam sistem temu kembali informasi diantaranya yaitu *logical models*, *vector processing models*, *probabilistic models*, dan *cognitive models* (Hasugian, 2009). Penjelasan dari tiap-tiap model tersebut adalah sebagai berikut :

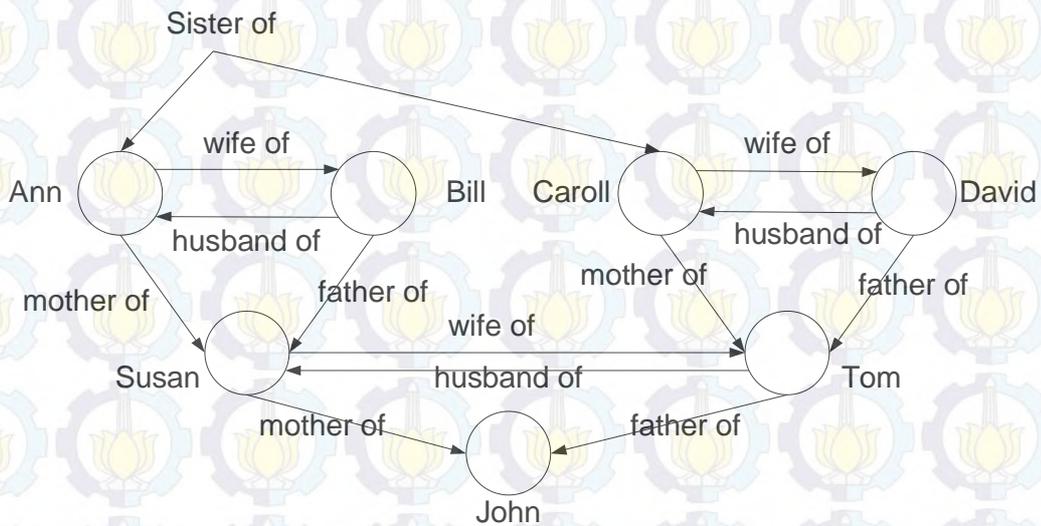
1. Pada *Logical models*, sistem temu kembali informasi menterjemahkan apa yang diinginkan oleh pengguna ke dalam sebuah kalimat permintaan (*query*) yang dioperasikan menggunakan operator-operator logika (*boolean*) yaitu operator *and*, *or*, dan *not*.
2. Pada *vector processing models*, sistem temu kembali informasi menggunakan nilai-nilai vektor pada indeks dokumen untuk menunjukkan hubungan kedekatan antara dokumen dan kalimat permintaan (*query*) pengguna.
3. Pada *probabilistic models*, sistem temu kembali informasi menggunakan perhitungan probabilitas untuk mengetahui tingkat relevansi dokumen dan membuat ranking atau urutan dokumen sesuai dengan relevansinya dengan permintaan pengguna.
4. Dan pada *cognitive models*, sistem temu kembali informasi menitikberatkan pada bagaimana agar sistem dapat berinteraksi dengan pengguna melalui antarmuka yang komunikatif.

2.2. Jaringan Semantik

Jaringan semantik merupakan usaha memodelkan bagaimana pengetahuan disimpan dan dipanggil kembali dalam ingatan manusia (Danenberg, 1988). Dan dalam aplikasinya, jaringan semantik umumnya disajikan dalam sebuah graf berarah, dimana simpul-simpul dalam jaringan menggambarkan sebuah konsep,

objek atau situasi tertentu yang dapat berupa objek tunggal maupun sebuah kelas yang melingkupi sebuah area tema tertentu, dan diantara simpul-simpul tersebut dihubungkan dengan garis-garis penghubung yang menunjukkan hubungan diantara konsep, objek atau situasi dari simpul-simpul tersebut.

Berikut ini adalah contoh sebuah jaringan semantik :

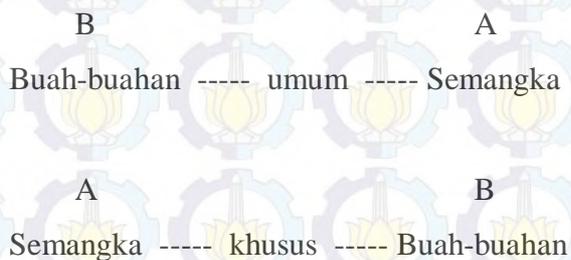


Gambar 2.1. Contoh jaringan semantik

Menurut Danenberg (1988), jenis jaringan semantik dapat dibedakan menjadi tujuh jenis hubungan, yaitu hubungan umum-khusus, teknik-aplikasi, prasyarat-kelanjutan, sistem-komponen, sinonim, perbedaan, dan berhubungan. Adapun penjelasan dari tiap-tiap jenis hubungan dalam jaringan semantik adalah sebagai berikut :

1. Hubungan Umum-Khusus, yaitu jenis hubungan yang bersifat inversi, dimana objek umum menyatakan kategori yang abstrak dan objek khusus menyatakan kategori yang lebih konkrit.

Contoh :



2. Hubungan Teknik-Aplikasi, yaitu hubungan yang merupakan pasangan inversi, dimana objek teknik menyatakan pengetahuan pokok atau kumpulan ketrampilan teknis dan metode yang digunakan untuk tujuan khusus atau aplikasi. Jika topik A adalah teknik yang digunakan dalam aplikasi topik B, maka topik B adalah aplikasi dari topik A.

Contoh :

A

B

Program Linier ----- teknik ----- Metode Grafik

A

B

Sistem Persamaan ----- aplikasi ----- Program Linier

3. Hubungan Prasyarat-Kelanjutan, yaitu pasangan relasi inversi, dimana prasyarat menunjukkan sebuah pengetahuan atau ketrampilan tertentu yang harus dikuasai sebelum boleh mempelajari kelanjutannya. Jika topik A adalah prasyarat untuk topik B, maka topik B adalah kelanjutan dari topik A.

Contoh :

A

B

Berenang ----- prasyarat ----- Menyelam

A

B

Menyelam ----- kelanjutan ----- Berenang

4. Hubungan Sistem-Komponen, yaitu pasangan relasi inversi, dimana sistem adalah kumpulan komponen yang interdependen sedemikian sehingga tiap komponen memiliki pengaruh terhadap setiap komponen lainnya dan secara keseluruhan komponen-komponen itu membentuk sebuah sistem. Jika topik A adalah sistem yang memuat komponen B, maka topik B adalah komponen dari topik A.

Contoh :

A

B

Persamaan ----- sistem ----- Model Matematika

A

B

Model Matematika ----- komponen ----- Persamaan

5. Sinonim adalah hubungan simetrik yang menyatakan kesamaan atau ekivalensi antara dua topik. Jika topik A sinonim topik B, maka topik B adalah sinonim dengan topik A.

Contoh :

A

B

Fungsi ----- sinonim ----- Pemetaan

6. Perbedaan adalah hubungan simetrik yang menyediakan informasi dengan contoh berlawanan dan mengakibatkan dua topik cukup berbeda untuk dipertimbangkan sebagai antonim satu dengan lainnya. Jika topik A berbeda dengan topik B, maka topik B berbeda dengan topik A.

Contoh :

A

B

Persamaan ----- berbeda ----- Pertidaksamaan

7. Berhubungan adalah hubungan simetrik yang digunakan jika terdapat hubungan antara dua topik tetapi belum bisa dinyatakan. Selanjutnya, hubungan bisa diidentifikasi sebagai relasi berhubungan. Jika topik A berhubungan dengan topik B, maka topik B berhubungan dengan topik A.

Contoh :

A

B

Pemecahan Masalah ----- berhubungan ----- Soal cerita matematika

Secara lebih sederhana, jenis hubungan dari simpul-simpul pada jaringan semantik dapat dibedakan menjadi 3 jenis sebagai berikut :

1. IS-A berarti “contoh/bagian dari” dan merupakan anggota tertentu dari sebuah kelas yang lebih luas.
2. A KIND OF (AKO) berarti “jenis dari” dan menghubungkan antara suatu kelas dengan kelas lainnya. AKO menghubungkan kelas individu ke kelas induk dari kelas-kelas dimana individu tersebut merupakan kelas anak.
3. HAS-A berarti “mempunyai” yang menghubungkan suatu kelas menjadi subkelas. HAS-A berlawanan dengan AKO dan sering digunakan untuk merelasikan suatu objek ke bagian-bagian yang lebih kecil dari objek tersebut.

Sedangkan berdasarkan karakteristik informasi yang ada di dalamnya, simpul-simpul pada jaringan semantik dapat berisi objek, atribut, atau value. Dimana objek dapat berupa sebuah bentuk fisik atau konsepsi, atribut merupakan karakteristik yang dimiliki oleh sebuah objek, dan value merupakan ukuran spesifik dari atribut dalam situasi tertentu.

Contoh :

Objek	Atribut	Nilai
Mobil	Warna	Biru
Mobil	Jumlah	25
Penggaris	Panjang	30

2.3. Jaringan Bayes

Jaringan Bayes (*Bayesian Networks*) merupakan sebuah metode yang efisien digunakan untuk merepresentasikan distribusi probabilitas bersama atas sekumpulan variabel acak. Jaringan Bayes memberikan representasi grafis yang secara eksplisit merepresentasikan independensi diantara variabel-variabel pada sebuah domain tertentu, serta spesifikasi singkat dari distribusi probabilitas bersama (Pearl, 1998).

Sebuah jaringan Bayes terdiri dari dua komponen yaitu sebuah *directed acyclic graph* (DAG) dan satu set tabel probabilitas bersyarat (CPT) yang menspesifikasikan distribusi probabilitas bersyarat dari setiap variabel yang diberi nilai dari induknya dalam DAG. Node dalam jaringan Bayes mewakili variabel-

variabel acak, dan ujung-ujungnya menunjukkan ketergantungan diantara variabel-variabel tersebut. Garis tepi menghasilkan sebuah distribusi probabilitas bersama, sehingga induk memiliki probabilitas *prior* $P(p)$ dan anak memiliki probabilitas bersyarat $P(c|p)$.

Dengan menggunakan independensi bersyarat, maka distribusi probabilitas bersama dari $P(x_1, x_2, \dots, x_n)$ dapat diperhitungkan sebagai berikut :

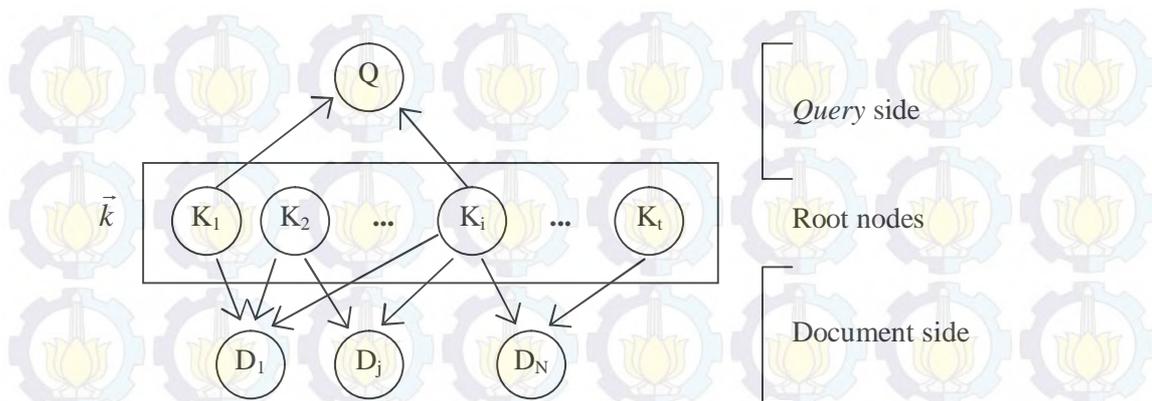
$$P(x_1, x_2, \dots, x_n) = \prod_i P(x_i | Parents(x_i)) \quad (2.1)$$

2.3.1. Jaringan Bayes untuk Temu Kembali Informasi

Model jaringan Bayes pertama kali digunakan dalam temu kembali informasi oleh Turtle dan Croft, dimana mereka menunjukkan bahwa model temu kembali informasi yang diusulkan bekerja lebih baik dari beberapa model probabilistik tradisional untuk meranking dokumen (Turtle and Croft, 1990).

Model yang berbeda diusulkan oleh Ribeiro-Neto dan Muntz (1996), mereka menunjukkan bahwa kombinasi fakta dari *query* di masa lalu dengan perangkingan ruang vektor menghasilkan hasil yang lebih baik daripada penggunaan perangkingan ruang vektor sendirian. Selanjutnya Acid, de Campos, Fernandez-Luna, and Huete (2003) menyajikan model dengan topologi jaringan yang didefinisikan sedemikian rupa sehingga algoritma propagasi yang tepat dapat digunakan untuk menghitung probabilitas relevansi dokumen secara efisien.

Jaringan Bayes juga telah diterapkan pada masalah-masalah temu kembali informasi yang lainnya, diantaranya untuk penanganan umpan balik yang relevan untuk melakukan konstruksi atau membangun hypertext secara otomatis sistem untuk melakukan ekspansi *query*, sistem penyaringan informasi, dan klasifikasi dokumen (Samodra et al., 2009).



Gambar 2.2. Jaringan Bayes pada Temu Kembali Informasi

Menurut model dasar jaringan Bayes dalam sistem temu kembali informasi, kalimat permintaan (*query*), kumpulan dokumen, dan kumpulan kata kunci dianggap sebagai peristiwa (Calado et al., 2004). Pada jaringan Bayes seperti yang ditunjukkan dalam gambar 2.1 di atas, simpul D_j menunjukkan kumpulan dokumen yang dikelola dan disimpan oleh sistem. Simpul Q menunjukkan kalimat permintaan (*query*) yang dimasukkan oleh pengguna, dan simpul K_i menyatakan kumpulan kata kunci atau istilah-istilah yang digunakan untuk merepresentasikan objek-objek yang terdapat di dalam domain. Vektor \vec{k} mengacu pada setiap keadaan yang mungkin pada simpul akar K_i . Hubungan antara dokumen D_j dan *query* Q ditafsirkan sebagai probabilitas dokumen D_j yang terjadi pada *query* Q .

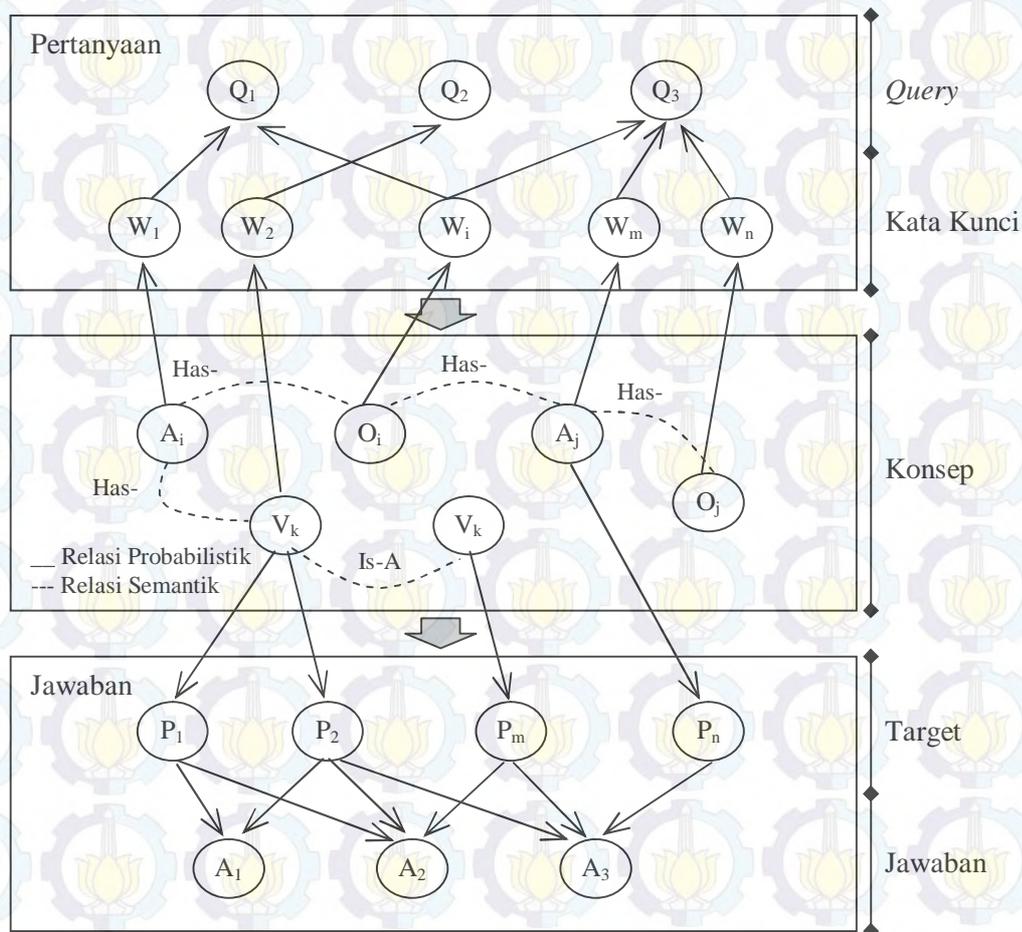
Dengan menggunakan hukum dan aturan probabilitas Bayes, maka probabilitas $P(D_j|Q)$ dapat dihitung sebagai berikut :

$$P(D_j | Q) = \eta \sum_k P(D_j | \vec{k}) P(Q | \vec{k}) P(\vec{k}) \quad (2.2)$$

2.3.2. Jaringan Bayes Semantik Untuk Inferensi

Jaringan Bayes semantik yang merupakan kombinasi dari jaringan Bayes dan jaringan semantik adalah sebuah pengembangan penggunaan jaringan Bayes pada sistem temu kembali informasi. Metode ini telah digunakan oleh Kim, Hong, dan Cho untuk mendapatkan kesimpulan secara efisien melalui dua tahapan pemrosesan yaitu dengan menyusun probabilitas kesimpulan dan kesimpulan

semantik. Model ini terdiri dari tiga lapisan (*layer*) yang memiliki tujuan penggunaan dan fungsi yang berbeda yaitu lapisan kata kunci, lapisan konsep, dan dan lapisan target (Kim et al., 2007). Secara garis besar, arsitektur dari jaringan Bayes semantik dapat dilihat pada gambar 2. 3 :



Gambar 2.3. Jaringan Bayes semantik untuk inferensi (Kim et al., 2007)

Lapisan kata kunci terdiri dari kumpulan kata-kata atau istilah-istilah yang sering digunakan oleh pengguna untuk membentuk kalimat permintaan (*query*), dimana sebuah *query* merupakan media yang digunakan oleh pengguna untuk berkomunikasi dengan sistem untuk menyatakan informasi apa yang dibutuhkan dari sistem. Lapisan konsep terdiri dari entitas domain yang dimiliki oleh sistem dan hubungan semantik diantara mereka. Domain informasi yang terdapat pada lapisan konsep dibentuk dari tiga komponen utama yaitu objek, atribut yang

dimiliki oleh objek, dan nilai data (*value*) dari atribut objek. Sedangkan lapisan target merepresentasikan kumpulan informasi atau dokumen target yang atribut-atributnya telah didefinisikan pada lapisan konsep, dan nantinya akan disajikan kepada pengguna sebagai tanggapan atas permintaan pengguna yang diberikan melalui kalimat permintaan (*query*).

Hubungan probabilistik pada jaringan Bayes semantik secara garis besar hampir sama dengan jaringan Bayes untuk temu kembali informasi tradisional. Langkah pertama yang dilakukan yaitu menghitung probabilitas antara lapisan kata kunci dan lapisan konsep. *Query* pengguna $U = \{k_1, k_2, \dots, k_i\}$ merupakan himpunan istilah atau kata kunci k_i yang diterjemahkan sebagai kumpulan kata dasar dalam lapisan kata kunci, dimana sebuah simpul kata kunci diberi nilai 1 bila kata yang diberikan dalam lapisan kata kunci terdapat dalam *query* Q , dan jika sebaliknya maka diberi nilai 0.

$$P(w_i) = \begin{cases} 1, w_i \in U \\ 0, w_i \notin U \end{cases} \quad (2.3)$$

Setelah semua variabel fakta yang berupa kumpulan kata kunci di lapisan kata kunci selesai ditetapkan nilainya, maka tahap berikutnya yaitu menyimpulkan probabilitas dari setiap simpul pada lapisan konsep dengan menggunakan kata kunci W di lapisan kata kunci sebagai fakta. Probabilitas simpul-simpul di lapisan konsep yaitu $P(c|W)$ didefinisikan sebagai berikut :

$$\begin{aligned} P(c | W) &= P(c | w_1, w_2, \dots, w_n) = \frac{P(c) \times P(w_1, w_2, \dots, w_n | c)}{P(w_1, w_2, \dots, w_n)} \approx P(c) \times P(w_1, w_2, \dots, w_n | c) \\ &= P(c) \times P(w_1 | c) \times P(w_2 | c) \times \dots \times P(w_n | c) = P(c) \prod_{i=1}^n P(w_i | c) \end{aligned} \quad (2.4)$$

Dimana n adalah jumlah simpul pada himpunan kata kunci, dan $c \in O \cup A \cup V$ dimana W adalah himpunan kata kunci, O adalah himpunan objek, A adalah himpunan atribut, dan V adalah himpunan nilai data atribut.

Setelah semua simpul pada lapisan konsep dihitung probabilitasnya, tahap berikutnya yaitu menyimpulkan probabilitas setiap objek target p di lapisan target

dengan menggunakan konsep C di lapisan konsep sebagai fakta. Probabilitas $P(p|C)$ didefinisikan sebagai berikut :

$$P(p|C) = P(p|c_1, c_2, \dots, c_L) = \frac{P(p) \times P(c_1, c_2, \dots, c_L | p)}{P(c_1, c_2, \dots, c_L)} \approx P(p) \times P(c_1, c_2, \dots, c_L | p)$$

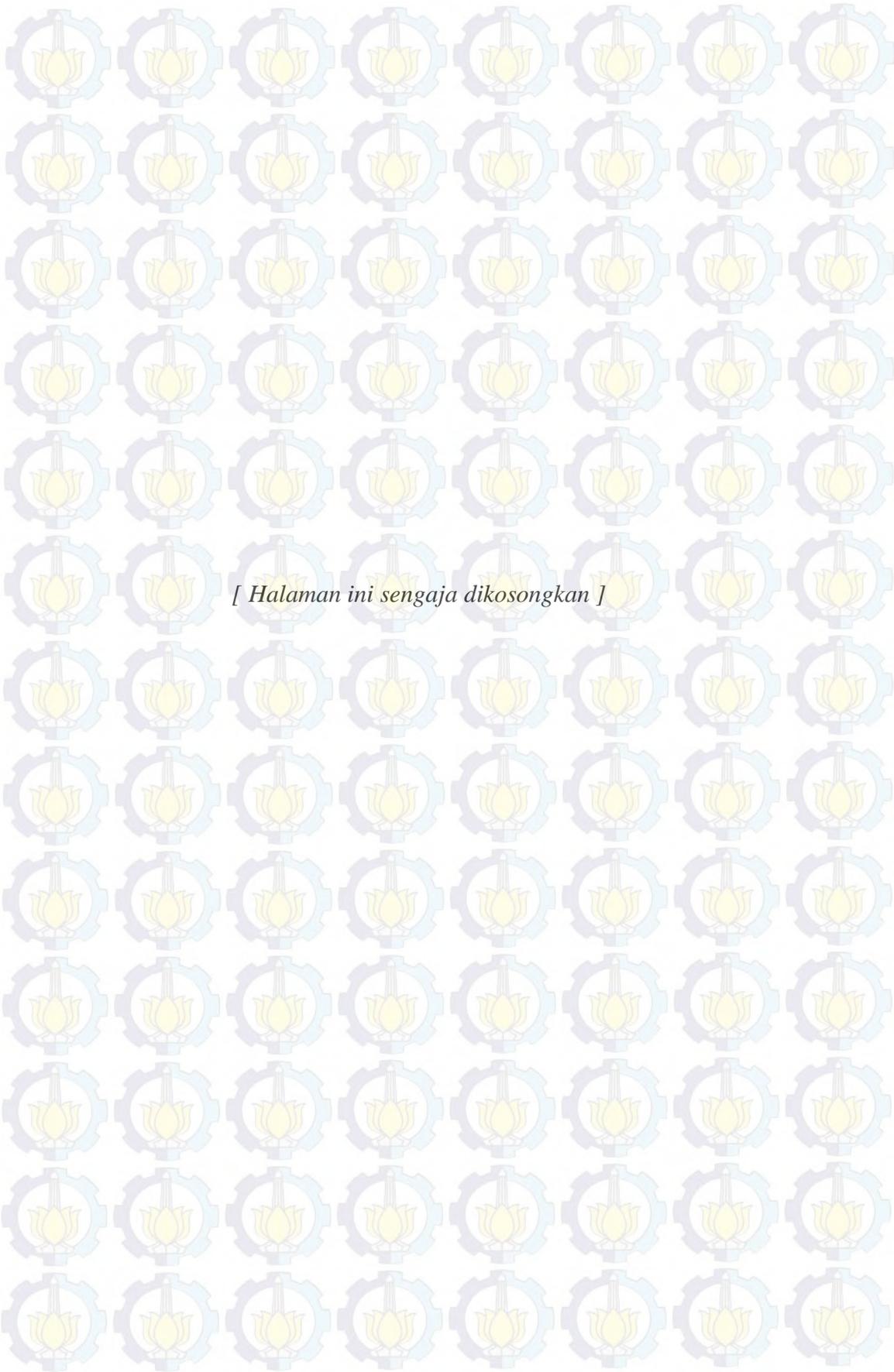
$$= P(p) \times P(c_1 | p) \times P(c_2 | p) \times \dots \times P(c_L | p) = P(p) \prod_{i=1}^L P(c_i | p) \quad (2.5)$$

Dimana L adalah jumlah simpul dalam himpunan konsep, dan $C = O \cup A \cup V$ dimana C adalah himpunan konsep, O adalah himpunan objek, A adalah himpunan atribut, dan V adalah himpunan nilai data atribut.

Dan sebagai kesimpulan akhir, probabilitas $P(p|W)$ antara lapisan konsep dan lapisan target didefinisikan sebagai berikut :

$$P(p|W) = \eta \sum_{i=1}^L (p|c_i) P(W|c_i) P(c_i) \quad (2.6)$$

Sebagai tahapan yang terakhir, setelah semua simpul pada lapisan target dihitung probabilitasnya, maka sistem akan memilih sebuah simpul yang nilai probabilitasnya lebih tinggi dari ambang batas yang telah ditetapkan. Dan selanjutnya sistem akan memberikan informasi yang relevan dengan simpul target yang telah dipilih tersebut sebagai tanggapan atas permintaan dari pengguna.



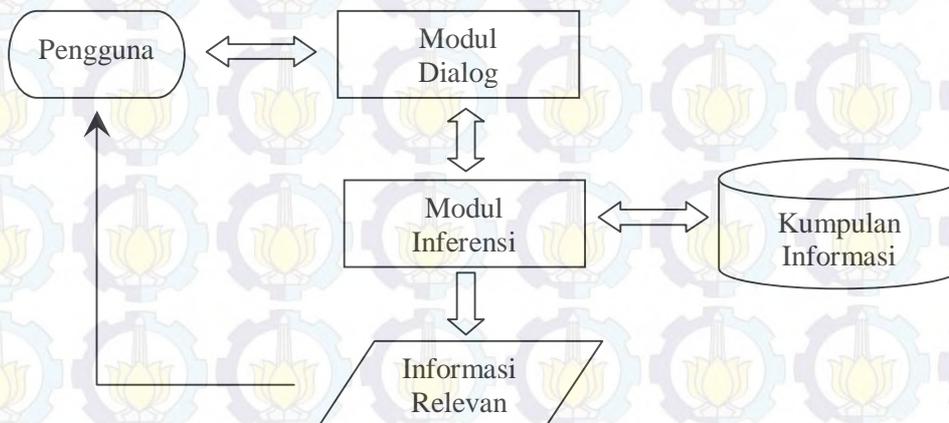
[Halaman ini sengaja dikosongkan]

BAB 3 METODE PENELITIAN

Pada bab ini akan dijelaskan metode dan cara kerja yang dilakukan pada kegiatan penelitian ini, dimana metode yang digunakan adalah metode eksploratif dan eksperimentatif. Metode eksploratif digunakan pada proses pengumpulan landasan teori dan kajian pustaka yang dilakukan melalui studi pustaka, baik dari jurnal, *paper*, buku teks, tutorial, dan artikel-artikel lain dari internet. Dari kegiatan diatas maka akan didapatkan ide-ide yang selanjutnya akan dikaji lebih lanjut dan rencana penelitian yang dapat dikerjakan dengan mempertimbangkan waktu dan biaya implementasi yang wajar. Untuk memperkirakan hasil yang akan dicapai, maka sebelumnya telah dilakukan penelitian pendahuluan berupa eksperimen-eksperimen yang berhubungan dengan penelitian yang sebenarnya. Hasil yang diharapkan dari penelitian ini adalah berupa tesis yang berasal dari hipotesis yang telah teruji dan memiliki nilai kebaruan, serta dapat memberikan manfaat bagi pengembangan bidang ilmu sejenis dimasa yang akan datang.

3.1. Bagan Penelitian

Model yang diajukan untuk menerapkan inferensi pada sistem temu kembali informasi dalam penelitian ini ditunjukkan pada gambar 3.1 :

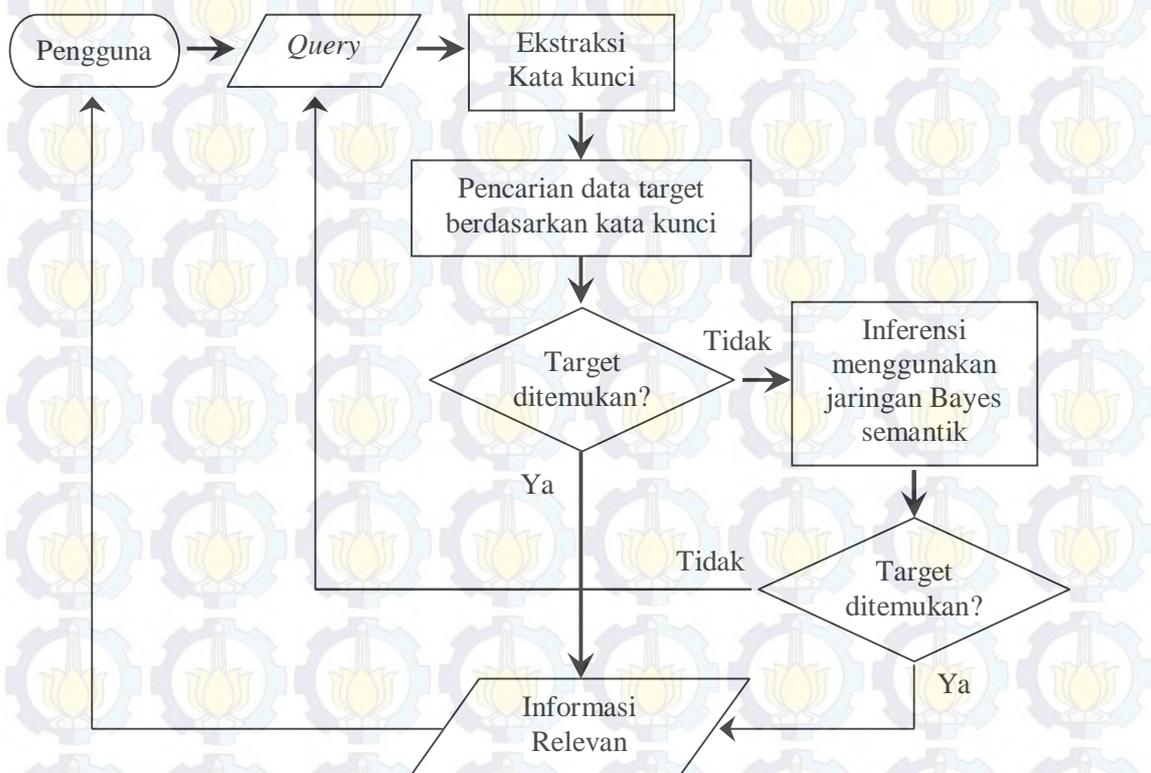


Gambar 3.1. Blok diagram penerapan inferensi pada sistem temu kembali informasi

Seperti ditunjukkan dalam gambar 3.1, sistem temu kembali informasi yang diajukan dalam penelitian ini terdiri dari dua bagian utama yaitu modul dialog dan modul inferensi.

Modul dialog digunakan sebagai antarmuka untuk menerima masukan dan melakukan interaksi dengan pengguna, sedangkan modul inferensi digunakan untuk menganalisa dan menyimpulkan apa yang diinginkan oleh pengguna.

Adapun diagram alir penerapan inferensi pada sistem temu kembali informasi dapat dilihat pada gambar 3.2 :



Gambar 3.2. Diagram alir penerapan inferensi pada sistem temu kembali informasi

Pengguna melakukan interaksi dengan sistem dengan cara memberikan masukan atau mengetikkan kalimat permintaan (*query*). Selanjutnya akan dilakukan proses ekstraksi kata kunci dengan cara mengekstraksi kalimat permintaan (*query*) menjadi kumpulan kata dasar, dan dilakukan proses seleksi sehingga menghasilkan kumpulan atau himpunan kata kunci.

Berdasarkan himpunan kata kunci yang telah didapatkan pada proses ekstraksi kata kunci tersebut, pertama-tama sistem akan melakukan pencarian data target menggunakan metode pencocokan kata kunci. Jika target yang diinginkan oleh pengguna dapat ditemukan, maka sistem akan memberikan tanggapan atau jawaban kepada pengguna berupa informasi yang relevan atau mendekati dengan permintaan pengguna, tanpa melakukan proses inferensi lagi. Tetapi jika pada proses pencarian menggunakan pencocokan kata kunci gagal menemukan data target, maka sistem akan melakukan inferensi menggunakan jaringan Bayes dan jaringan semantik untuk mencari atau menemukan data target yang sesuai dengan yang diinginkan oleh pengguna. Setelah proses inferensi dijalankan, akan dilakukan pengecekan apakah data target yang diinginkan oleh pengguna sudah dapat ditemukan oleh sistem atau belum.

Jika data target yang diinginkan oleh pengguna belum dapat ditemukan, maka sistem akan meminta kepada pengguna untuk memberikan tambahan informasi yang diperlukan melalui kalimat permintaan (*query*) yang baru. Sistem akan mengulangi proses ekstraksi kata kunci dan proses inferensi seperti yang dilakukan pada kalimat permintaan (*query*) sebelumnya, sampai ditemukan data target yang sesuai dengan yang diinginkan oleh pengguna.

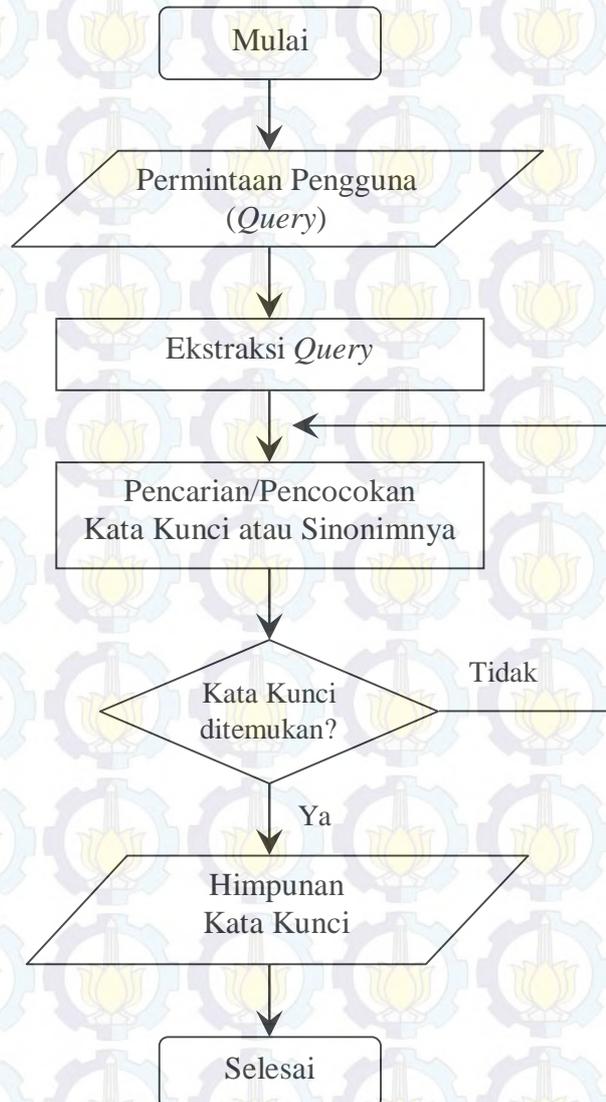
Jika data target yang diinginkan oleh pengguna sudah dapat ditemukan, maka sistem akan memberikan tanggapan atau jawaban kepada pengguna berupa informasi yang relevan dengan permintaan pengguna. Dalam konteks penelitian ini, informasi yang dimaksud dan akan diberikan kepada pengguna yaitu berupa deskripsi atau informasi lengkap tentang sebuah produk elektronik yang paling sesuai dengan apa yang diinginkan oleh pengguna, sesuai dengan kalimat permintaan (*query*) yang diberikan oleh pengguna kepada sistem.

3.2. Modul Dialog

Modul dialog merupakan sebuah subsistem yang berfungsi sebagai antarmuka untuk melakukan interaksi dengan pengguna. Modul ini secara garis besar bertugas untuk menerima masukan atau permintaan dari pengguna. Selain itu, modul dialog juga bertugas untuk melakukan ekstraksi atau mengolah data

masukan yang diberikan oleh pengguna, dan sekaligus sebagai sarana untuk memberikan tanggapan atau jawaban kepada pengguna.

Adapun tahapan yang dilakukan oleh modul dialog untuk menerima masukan dan memproses data permintaan (*query*) dari pengguna dapat dilihat pada diagram alir yang ditunjukkan pada gambar 3.3 :



Gambar 3.3 Diagram alir pemasukan permintaan pengguna

Penjelasan dari tiap-tiap tahapan proses dalam diagram alir di atas dijelaskan pada uraian berikut ini :

1. Proses awal yang dilakukan oleh sistem yaitu melakukan pembacaan data masukan atau permintaan dari pengguna. Data masukan atau permintaan akan

diberikan oleh pengguna dengan cara mengetikkan kalimat permintaan (*query*) yang berbentuk sebuah kalimat atau kumpulan kata-kata pada kotak teks yang disediakan di halaman utama.

2. Sistem akan mengekstraksi kalimat permintaan (*query*) yang telah diketikkan oleh pengguna dengan memisahkan seluruh kata-kata yang terdapat pada kalimat permintaan menjadi kumpulan kata dasar.
3. Selanjutnya akan dilakukan proses pencarian atau pencocokan kata kunci untuk menyeleksi kumpulan kata dasar yang dihasilkan dari proses ekstraksi kalimat permintaan di atas. Proses ini bertujuan untuk membuang kata-kata yang tidak berguna dan menghasilkan himpunan kata kunci yang akan digunakan sebagai dasar untuk melakukan proses inferensi.

Proses seleksi kata kunci dilakukan dengan cara mencocokkan setiap kata dasar yang didapatkan dari proses ekstraksi kalimat permintaan (*query*) dengan himpunan kata kunci yang tersimpan di tabel “kata kunci”. Jika sebuah kata dasar terdapat pada tabel “kata kunci”, maka kata dasar tersebut akan dimasukkan ke himpunan kata kunci. Jika sebuah kata dasar tidak terdapat pada tabel “kata kunci”, maka sistem akan mencari dan mencocokkan kata dasar tersebut dengan himpunan sinonim dari kata kunci yang tersimpan di tabel “sinonim”. Jika kata dasar tersebut terdapat pada tabel “sinonim”, maka kata dasar tersebut akan dimasukkan ke himpunan kata kunci. Dan sebaliknya, jika kata dasar tersebut tidak ditemukan pada tabel “sinonim”, maka kata tersebut akan dibuang karena tidak memiliki hubungan dengan informasi yang tersimpan di basis pengetahuan sistem temu kembali informasi. Proses ini diulang sampai seluruh kata yang ada di himpunan kata dasar diseleksi seluruhnya.

Tabel “kata kunci” merupakan sebuah tabel yang di dalamnya diisi dengan kumpulan kata kunci yang memiliki hubungan erat dengan informasi produk elektronik yang tersimpan sebagai basis pengetahuan yang dimiliki oleh sistem temu kembali informasi. Kata-kata yang dimasukkan ke tabel ini adalah kata-kata yang umum dan sering digunakan oleh pengguna pada saat ingin mencari atau menemukan sebuah produk yang diinginkan. Berikut ini adalah kumpulan kata kunci yang tersimpan di tabel “kata kunci” :

Tabel 3.1. Himpunan Kata Kunci

tv	50
samsung	50in
lg	50 in
sharp	lcd
canon	led
sony	plasma
toshiba	handycam
jvc	sdcard
19	minidv
19in	hdd
19 in	dvd
32	10x
32in	20x
32 in	40x
40	ccd
40in	cmos
40 in	3ccd
46	
46in	
46 in	

Pengisian data pada tabel ini dilakukan melalui sebuah modul pengisian data, yang selanjutnya akan disimpan ke database. Jika diperlukan, data-data yang telah tersimpan pada tabel ini dapat diperbarui sesuai dengan kebutuhan misalnya pada saat terdapat kata-kata baru yang digunakan oleh pengguna untuk melakukan pencarian, atau terdapat penambahan atau pengurangan data produk yang dimiliki oleh sistem temu kembali informasi. Adapun struktur dari tabel “kata kunci” dapat dilihat pada tabel 3.2 :

Tabel 3.2. Struktur Tabel Kata Kunci

Atribut	Tipe Data	Panjang	Keterangan
Kunci	VarChar	30	Kata Kunci

Selain tabel “kata kunci”, pada sistem ini juga dilengkapi dengan tabel “sinonim” yang berisi himpunan kata kunci yang merupakan sinonim dari kata kunci yang terdapat di tabel “kata kunci”. Tabel ini diperlukan karena pada beberapa kata kunci tertentu, terdapat beberapa istilah atau kata dasar yang

penulisannya berbeda tetapi memiliki arti dan maksud yang sama. Berikut ini adalah beberapa sinonim yang tersimpan di tabel “sinonim” :

Tabel 3.3. Sinonim

Kata Kunci	Sinonim
Handycam	Camcorder
Hdd	Harddisk
Hdd	Hard disk
Hdd	Disk
In	Inch
In	Inchi
In	Inci
In	“
Memory	Memori
Tv	televisi

Dan sama seperti pada tabel “kata kunci”, pengisian data pada tabel “sinonim” dilakukan melalui sebuah modul pengisian data, yang selanjutnya akan disimpan ke database. Jika diperlukan, data-data yang telah tersimpan pada tabel ini dapat diperbarui sesuai dengan kebutuhan misalnya pada saat terdapat penambahan atau pengurangan kata kunci dan sinonimnya. Adapun struktur dari tabel “sinonim” dapat dilihat pada tabel 3.4 :

Tabel 3.4. Struktur Tabel Sinonim

Atribut	Tipe Data	Panjang	Keterangan
Kunci	VarChar	30	Kata Kunci
Sinonim	VarChar	30	Sinonim dari Kata Kunci

- Setelah proses pencarian atau pencocokan kata kunci selesai dijalankan, maka tahapan pemrosesan data permintaan dari pengguna telah selesai. Hasil dari tahap ini yaitu akan didapatkan himpunan kata kunci yang nantinya akan digunakan sebagai dasar untuk melakukan proses inferensi oleh modul inferensi. Jadi, himpunan kata kunci ini selanjutnya akan diumpungkan sebagai data masukan ke modul inferensi.

Selain berfungsi sebagai sarana untuk menerima dan memproses data permintaan dari pengguna, modul dialog juga bertugas untuk memberikan tanggapan atau jawaban kepada pengguna. Adapun tahapan yang dilakukan untuk memberikan tanggapan atau jawaban kepada pengguna dapat dilihat pada diagram alir yang ditunjukkan pada gambar 3.4 :



Gambar 3.4. Diagram alir pemberian tanggapan atau jawaban kepada pengguna

Penjelasan dari tiap-tiap tahapan proses dalam diagram alir di atas dapat dijelaskan pada uraian berikut ini :

1. Modul dialog menerima masukan yang berupa kata kunci produk target. Kata kunci ini dihasilkan dari proses inferensi yang dilakukan oleh modul inferensi, yang merupakan kata kunci produk target yang paling sesuai dengan permintaan pengguna.
2. Selanjutnya akan dilakukan proses pencarian dan pembacaan data produk target yang sesuai dengan kata kunci yang diterima. Dalam sistem ini, seluruh

data produk yang disediakan oleh sistem temu kembali informasi disimpan pada tabel “produk” dengan struktur yang ditunjukkan pada tabel 3.5 :

Tabel 3.5. Struktur Tabel Produk

Atribut	Tipe Data	Panjang	Keterangan
Kode	VarChar	20	Kode Produk
NodeBN	VarChar	20	Kode simpul di jaringan Bayes
Nama	VarChar	50	Nama, merk dan tipe produk
Kunci	VarChar	40	Himpunan kata kunci produk
Spesifikasi	Blob		Deskripsi / spesifikasi produk
Foto	Blob		Foto / gambar produk

Seluruh data produk yang tersimpan di tabel ini diisikan melalui sebuah modul pengisian data, yang selanjutnya akan disimpan ke database. Jika diperlukan, data-data yang telah tersimpan pada tabel ini dapat diperbarui sesuai dengan kebutuhan misalnya pada saat terdapat penambahan atau pengurangan data produk.

3. Setelah data produk target dibaca dari tabel, selanjutnya deskripsi atau spesifikasi dari produk tersebut akan ditampilkan di layar sebagai tanggapan atau jawaban kepada pengguna.

3.3. Modul Inferensi

Modul inferensi merupakan sebuah subsistem yang berfungsi melakukan inferensi menggunakan jaringan Bayes dan jaringan semantik untuk mencari atau menemukan data target yang sesuai dengan permintaan pengguna, berdasarkan himpunan kata kunci yang telah diekstraksi dari kalimat permintaan (*query*) pengguna. Setelah proses inferensi dijalankan, akan dilakukan pengecekan apakah data target yang sesuai dengan permintaan pengguna sudah dapat ditemukan oleh sistem atau belum.

Secara garis besar, modul inferensi akan melakukan pencarian data target melalui dua tahapan sebagai berikut :

1. Tahapan pertama.

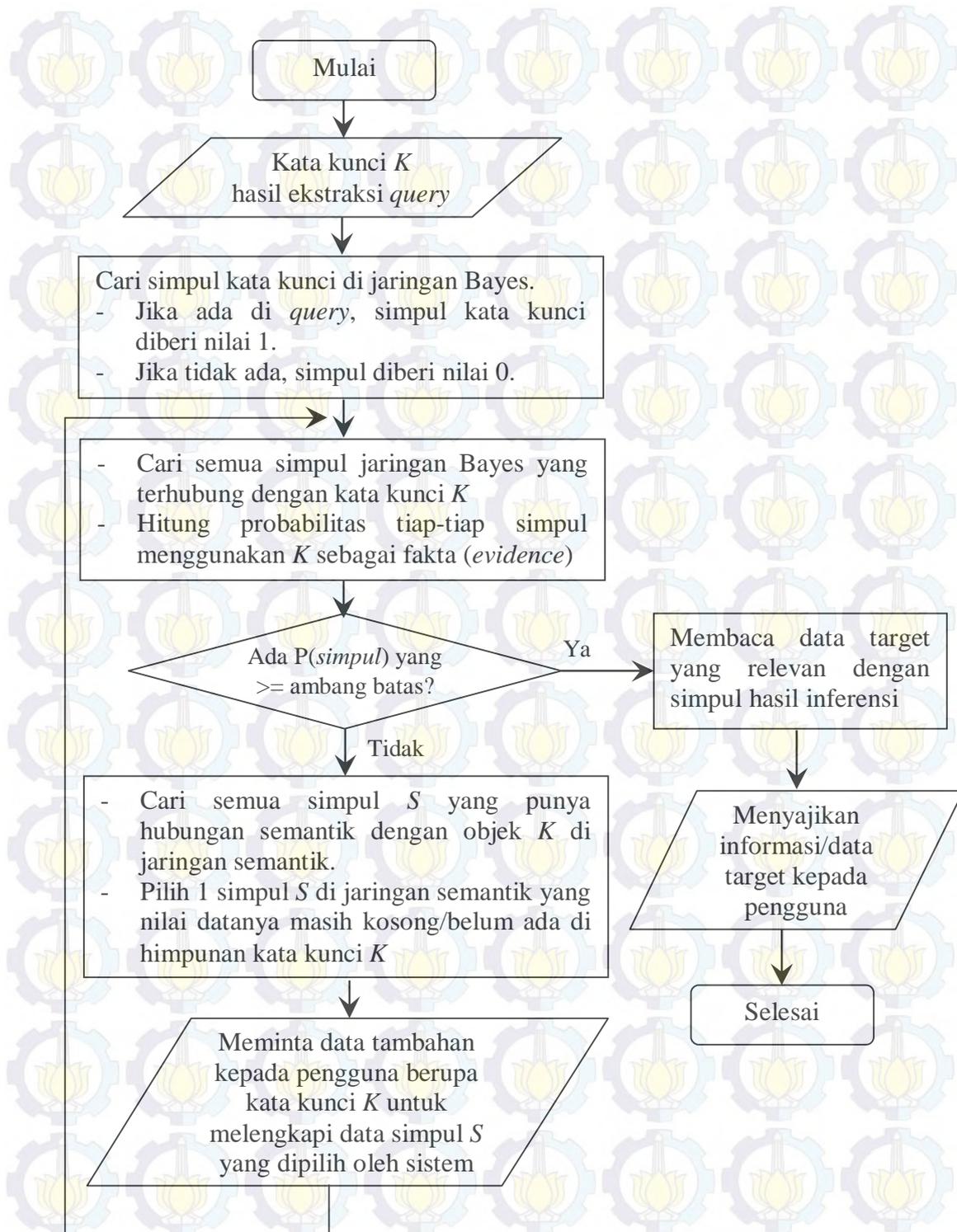
Sistem akan melakukan pencarian data target menggunakan metode pencocokan kata kunci, yaitu dengan cara mencari dan mencocokkan himpunan kata kunci yang diekstraksi dari kalimat permintaan pengguna dengan pola kata kunci yang dimiliki oleh tiap-tiap data target, yang dalam hal ini berupa data produk elektronik.

Jika dari proses pencocokan kata kunci ini menemukan satu data target yang memiliki pola kata kunci yang sesuai, maka proses pencarian dikatakan berhasil dan sistem akan memberikan tanggapan atau jawaban kepada pengguna berupa informasi yang relevan dengan permintaan pengguna, tanpa melakukan proses inferensi menggunakan jaringan Bayes semantik.

Tetapi jika dari proses pencocokan kata kunci ini tidak menemukan data target, atau data target yang memiliki pola kata kunci yang sama lebih dari satu data, maka proses pencarian dikatakan gagal, dan sistem akan melakukan tahapan kedua yaitu melakukan inferensi menggunakan jaringan Bayes semantik untuk mencari atau menemukan data target yang sesuai dengan permintaan oleh pengguna.

2. Tahapan kedua.

Sistem akan melakukan inferensi dengan menghitung probabilitas dari simpul-simpul di jaringan Bayes dengan menggunakan himpunan kata kunci yang dihasilkan dari proses ekstraksi kalimat permintaan pengguna sebagai fakta (*evidence*). Langkah-langkah yang dilakukan pada tahapan ini dapat dilihat pada diagram alir yang ditunjukkan pada gambar 3.5 :



Gambar 3.5. Diagram alir proses inferensi menggunakan jaringan Bayes semantik untuk menemukan informasi yang relevan dengan permintaan pengguna

Penjelasan dari tiap-tiap tahapan proses pada diagram alir di atas dijelaskan pada uraian berikut ini :

- a. Modul inferensi menerima masukan berupa himpunan kata kunci yang dihasilkan dari proses ekstraksi *query* pengguna.
- b. Langkah pertama yaitu memberikan nilai awal simpul-simpul kata kunci yang terdapat pada kalimat permintaan (*query*) pengguna, dimana kalimat permintaan $Q = \{k_1, k_2, \dots, k_t\}$ merupakan himpunan kata kunci k_i yang berbentuk kata dasar. Sebuah simpul kata kunci diberi nilai 1 bila kata kunci tersebut terdapat dalam *query* Q , dan jika sebaliknya maka diberi nilai 0.
- c. Setelah semua variabel fakta yang terkait dengan kata kunci telah ditetapkan nilainya, maka tahap berikutnya yaitu menghitung probabilitas dari setiap simpul yang terhubung dengan kata kunci, dengan menggunakan kata kunci K sebagai fakta.
- d. Setelah semua simpul yang terhubung dengan kata kunci dihitung probabilitasnya, jika ditemukan satu simpul yang nilai probabilitasnya sama atau lebih besar dari ambang batas yang ditetapkan, maka proses inferensi dikatakan berhasil. Selanjutnya sistem akan membaca data target tersebut dan memberikan tanggapan atau jawaban kepada pengguna berupa informasi yang relevan dengan permintaan pengguna.

Tetapi jika dari proses inferensi ini tidak dapat menemukan simpul yang nilai probabilitasnya sama atau lebih besar dari ambang batas yang ditetapkan, maka sistem akan meminta informasi tambahan kepada pengguna melalui kalimat permintaan yang baru.

- e. Tambahan informasi yang dibutuhkan oleh sistem berupa satu atau beberapa nilai atribut yang berkaitan erat dengan objek yang sedang dicari, yang didapatkan dari simpul-simpul di jaringan semantik yang memiliki hubungan dengan objek yang sedang dicari.
- f. Setelah pengguna memberikan tambahan informasi yang dibutuhkan, sistem akan mengulang proses inferensi dengan menghitung ulang probabilitas simpul-simpul di jaringan Bayes, sampai akhirnya ditemukan satu simpul yang nilai probabilitasnya sama atau lebih besar dari ambang batas yang ditetapkan.

Dalam penelitian ini, ada 2 simpul jenis produk yang digunakan untuk ujicoba yaitu simpul produk tv dan handycam. Jenis produk tv memiliki 3 buah atribut yaitu merk, ukuran, dan jenis layar, dimana tiap-tiap atribut dari produk tersebut memiliki nilai data (*value*) yang ditetapkan sebagai simpul dalam jaringan Bayes yaitu :

1. Atribut merk tv memiliki 5 buah nilai data (*value*) yaitu samsung, lg, sharp, sony, dan toshiba.
2. Atribut ukuran layar memiliki 5 buah nilai data (*value*) yaitu 19 in, 32 in, 40 in, 46 in, dan 50 in.
3. Atribut jenis layar memiliki 3 buah nilai data (*value*) yaitu lcd, led, dan plasma.

Sedangkan untuk jenis produk handycam memiliki 4 buah atribut yaitu merk, media penyimpanan, zoom, dan jenis sensor, dimana tiap-tiap atribut dari produk tersebut memiliki nilai data (*value*) yang ditetapkan sebagai simpul dalam jaringan Bayes yaitu :

1. Atribut merk handycam memiliki 3 buah nilai data (*value*) yaitu sony, canon, dan jvc.
2. Atribut media penyimpanan memiliki 4 buah nilai data (*value*) yaitu sdcard, minidv, hdd, dan dvd.
3. Atribut optical zoom memiliki 3 buah nilai data (*value*) yaitu 10x, 20x, dan 40x.
4. Atribut jenis sensor memiliki 3 buah nilai data (*value*) yaitu ccd, cmos, dan 3ccd.

Struktur dan banyaknya simpul jaringan Bayes yang digunakan dalam sistem temu kembali informasi ini bersifat statis, artinya jika diperlukan adanya penambahan, pengurangan atau perubahan jumlah simpul, maka struktur jaringan harus diperbarui secara manual oleh pengelola sistem. Misalnya pada saat ingin menambahkan data-data jenis produk baru, ingin menambah atau mengurangi nilai data atribut sebuah produk, atau pada saat ada kata-kata kunci baru yang banyak digunakan oleh pengguna pada saat melakukan pencarian informasi.

Setelah struktur dan seluruh simpul jaringan Bayes selesai ditetapkan, maka selanjutnya dilakukan penetapan nilai probabilitas awal dari tiap-tiap simpul

pada jaringan Bayes, dimana nilai probabilitas awal ini ditetapkan secara manual berdasarkan karakteristik yang dimiliki oleh simpul-simpul tersebut. Rentang nilai probabilitas awal yang digunakan dalam sistem temu kembali informasi ini dapat dilihat pada tabel 3.6 dan 3.7 :

Tabel 3.6. Probabilitas $P(t_i = 1 | k_i = 1)$

Kondisi	Probabilitas
t_i sering ada jika diberikan fakta k_i	0,7 – 0,95
t_i kadang-kadang ada jika diberikan fakta k_i	0,5 – 0,69
t_i jarang/tidak pernah ada jika diberikan fakta k_i	0,01 – 0,49

Tabel 3.7. Probabilitas $P(t_i = 1 | k_i = 0)$

Kondisi	Probabilitas
t_i sering ada jika diberikan fakta selain k_i	0,5 – 0,9
t_i kadang-kadang ada jika diberikan fakta selain k_i	0,1 – 0,4
t_i jarang/tidak pernah ada jika diberikan fakta selain k_i	0,01 – 0,09

Secara lebih spesifik, setiap simpul dalam jaringan Bayes yang merupakan kumpulan kata kunci yang digunakan untuk melakukan pencarian oleh pengguna, akan diberikan nilai probabilitas awal berdasarkan tingkat popularitas atau sering tidaknya kata kunci tersebut digunakan dalam proses pencarian. Untuk uji coba dalam penelitian ini, tingkat minat atau popularitas kata kunci menggunakan data yang disediakan oleh Google Trend yang dapat diakses melalui url <http://www.google.co.id/trends/>. Sumber data yang diambil dibatasi untuk area penelusuran Indonesia, dengan rentang waktu penelusuran selama bulan Januari hingga Desember tahun 2013. Selanjutnya data rata-rata minat atau popularitas kata kunci yang memiliki rentang nilai 0 hingga 100 akan dikonversi menjadi nilai probabilitas awal pada jaringan Bayes dengan rentang nilai 0,70 hingga 0,95.

Tabel 3.8. Rata-rata minat terhadap merk tv dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes

Merk tv	Rata-rata Minat	P (merk tv)
samsung	73	0,89
lg	55	0,84
sharp	36	0,79
sony	20	0,75
toshiba	21	0,75

Tabel 3.9. Rata-rata minat terhadap ukuran tv dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes

Ukuran tv	Rata-rata Minat	P (ukuran tv)
19 " / in	27	0,77
32 " / in	79	0,90
40 " / in	13	0,73
46 " / in	0	0,70
50" / in	5	0,71

Tabel 3.10. Rata-rata minat terhadap jenis layar tv dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes

Jenis Layar	Rata-rata Minat	P (layar tv)
lcd	50	0,83
led	73	0,89
plasma	6	0,71

Tabel 3.11. Rata-rata minat terhadap merk handycam dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes

Merk handycam	Rata-rata Minat	P (merk handycam)
sony	82	0,91
canon	48	0,82
jvc	1	0,70

Tabel 3.12. Rata-rata minat terhadap media penyimpanan handycam dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes

Media Penyimpanan	Rata-rata Minat	P (media handycam)
sdcard	63	0,86
minidv	29	0,77
hdd	0	0,70
dvd	0	0,70

Tabel 3.13. Rata-rata minat terhadap kemampuan zoom handycam dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes

Zoom	Rata-rata Minat	P (zoom handycam)
10 x	2	0,70
20 x	4	0,71
40 x	19	0,75

Tabel 3.14. Rata-rata minat terhadap jenis sensor handycam dan hasil konversi menjadi nilai probabilitas awal pada jaringan Bayes

Jenis Sensor	Rata-rata Minat	P (sensor handycam)
ccd	25	0,76
cmos	62	0,86
3ccd	0	0,70

Tabel 3.15. Nilai probabilitas awal simpul data target pada jaringan Bayes

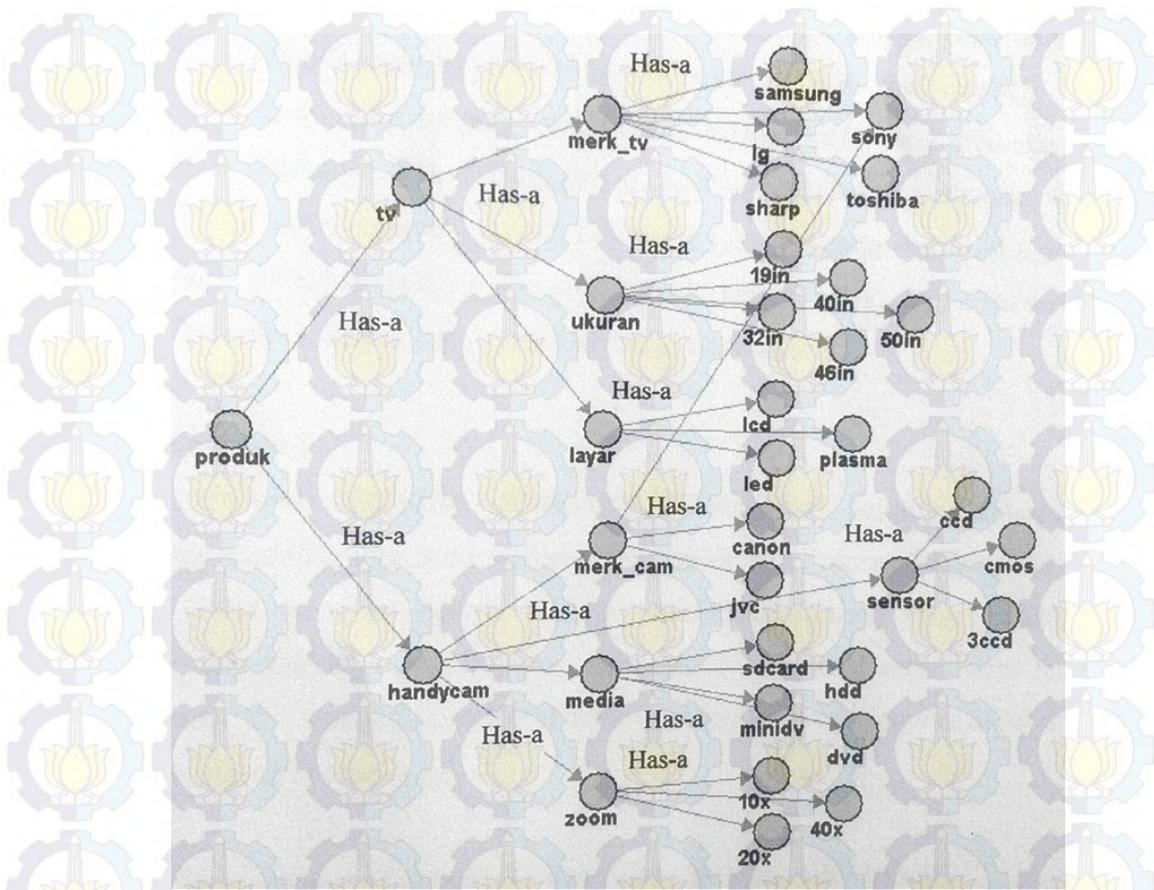
Simpul T_i	Probabilitas T_i
$P(T merk-tv1, ukuran1, layar1)$	0,95
$P(T merk-tv1, ukuran1, layar0)$	0,60
$P(T_i merk-tv1, ukuran0, layar1)$	0,60
$P(T_i merk-tv1, ukuran0, layar0)$	0,30
$P(T_i merk-tv0, ukuran1, layar1)$	0,40
$P(T_i merk-tv0, ukuran1, layar0)$	0,20
$P(T_i merk-tv0, ukuran0, layar1)$	0,20
$P(T_i merk-tv0, ukuran0, layar0)$	0,05
$P(H_i merk-h1, media1, zoom1, sensor1)$	0,95
$P(H_i merk-h1, media1, zoom1, sensor0)$	0,70
$P(H_i merk-h1, media1, zoom0, sensor1)$	0,70
$P(H_i merk-h1, media1, zoom0, sensor0)$	0,30
$P(H_i merk-h1, media0, zoom1, sensor1)$	0,70
$P(H_i merk-h1, media0, zoom1, sensor0)$	0,30

$P(H_i merk-h1, media0, zoom0, sensor1)$	0,30
$P(H_i merk-h1, media0, zoom0, sensor0)$	0,10
$P(H_i merk-h0, media1, zoom1, sensor1)$	0,60
$P(H_i merk-h0, media1, zoom1, sensor0)$	0,30
$P(H_i merk-h0, media1, zoom0, sensor1)$	0,30
$P(H_i merk-h0, media1, zoom0, sensor0)$	0,10
$P(H_i merk-h0, media0, zoom1, sensor1)$	0,30
$P(H_i merk-h0, media0, zoom1, sensor0)$	0,10
$P(H_i merk-h0, media0, zoom0, sensor1)$	0,10
$P(H_i merk-h0, media0, zoom0, sensor0)$	0,05

Sedangkan nilai batas toleransi yang digunakan untuk pengambilan kesimpulan ditetapkan sebesar 0,05 dari probabilitas maksimal yang digunakan dalam jaringan Bayes yaitu sebesar 0,95. Artinya sebuah simpul dalam jaringan Bayes dikatakan memenuhi syarat atau dapat diterima sebagai hasil inferensi jika nilai probabilitas dari simpul tersebut berkisar antara 0,90 – 0,95. Penetapan nilai ambang batas yang cukup tinggi ini bertujuan agar kesimpulan yang dihasilkan dari proses inferensi memiliki relevansi atau kesesuaian yang tinggi terhadap permintaan pengguna. Pertimbangan yang lainnya yaitu dikarenakan metode pembaruan nilai probabilitas simpul kata kunci yang terdapat pada *query* menggunakan nilai probabilitas maksimum yaitu 1, sehingga probabilitas simpul-simpul lain yang terhubung dengan kata kunci tersebut akan meningkat cukup besar dan semakin mendekati nilai 1.

3.3.2. Struktur Jaringan Semantik

Struktur jaringan semantik yang digunakan untuk ujicoba dalam penelitian ini dapat dilihat pada gambar 3.7 :



Gambar 3.7. Jaringan semantik yang digunakan oleh modul inferensi

Dari gambar 3.7 di atas dapat dilihat bahwa komponen utama jaringan semantik yang digunakan terdiri dari simpul-simpul jenis produk, atribut produk, dan nilai data (*value*) dari atribut produk. Hubungan yang terbentuk dari ketiga komponen tersebut dapat dilihat pada penjelasan di bawah ini :

1. Jenis produk mempunyai (*has-a*) 2 buah data produk yaitu tv dan handycam. Jenis hubungan yang terbentuk adalah objek-atribut (OA).
2. Produk tv mempunyai (*has-a*) 3 buah data atribut yaitu merk, ukuran, dan jenis layar. Jenis hubungan yang terbentuk adalah objek-atribut (OA).
3. Atribut merk tv mempunyai (*has-a*) 5 buah nilai data (*value*) yaitu samsung, lg, sharp, sony, dan toshiba. Jenis hubungan yang terbentuk adalah atribut-value (AV).
4. Atribut ukuran layar mempunyai (*has-a*) 5 buah nilai data (*value*) yaitu 19 in, 32 in, 40 in, 46 in, dan 50 in. Jenis hubungan yang terbentuk adalah atribut-value (AV).

5. Atribut jenis layar mempunyai (*has-a*) 3 buah nilai data (*value*) yaitu lcd, led, dan plasma. Jenis hubungan yang terbentuk adalah atribut-value (AV).
6. Produk handycam mempunyai (*has-a*) 4 buah atribut yaitu merk, media penyimpanan, zoom, dan jenis sensor. Jenis hubungan yang terbentuk adalah objek-atribut (OA).
5. Atribut merk handycam mempunyai (*has-a*) 3 buah nilai data (*value*) yaitu sony, canon, dan jvc. Jenis hubungan yang terbentuk adalah atribut-value (AV).
6. Atribut media penyimpan mempunyai (*has-a*) 4 buah nilai data (*value*) yaitu sdcard, minidv, hdd, dan dvd. Jenis hubungan yang terbentuk adalah atribut-value (AV).
7. Atribut optical zoom mempunyai (*has-a*) 3 buah nilai data (*value*) yaitu 10x, 20x, dan 40x. Jenis hubungan yang terbentuk adalah atribut-value (AV).
8. Atribut jenis sensor mempunyai (*has-a*) 3 buah nilai data (*value*) yaitu ccd, cmos, dan 3ccd. . Jenis hubungan yang terbentuk adalah atribut-value (AV).

Data-data jaringan semantik yang digunakan dalam sistem temu kembali informasi ini disimpan pada tabel “relasi” dengan struktur yang ditunjukkan pada tabel 3.16 :

Tabel 3.16. Struktur Tabel Relasi

Atribut	Tipe Data	Panjang	Keterangan
Objek1	VarChar	20	Objek pertama
Relasi	VarChar	20	Hubungan antara objek pertama dan kedua
Objek2	VarChar	20	Objek kedua
Jenis	VarChar	10	Jenis hubungan: objek-atribut (OA) atau atribut-value (AV)

Seluruh data jaringan semantik yang tersimpan di tabel ini diisikan melalui sebuah modul pengisian data, yang selanjutnya akan disimpan ke database. Jika diperlukan, data-data yang telah tersimpan pada tabel ini dapat diperbarui sesuai dengan kebutuhan misalnya pada saat terdapat penambahan atau pengurangan data produk, atribut produk, maupun nilai-nilai data dari sebuah atribut produk.

BAB 4

ANALISA HASIL DAN PEMBAHASAN

Setelah menetapkan metode dan rancangan sistem seperti yang diuraikan pada bab 3, selanjutnya akan dibahas hasil penelitian yang telah dilakukan untuk menerapkan inferensi pada sistem temu kembali informasi. Uji coba dilakukan dengan memberikan beberapa kombinasi kalimat permintaan (*query*) yang diinputkan melalui modul dialog, dan selanjutnya akan dianalisa hasil inferensi yang dilakukan oleh modul inferensi.

Sedangkan untuk data target yaitu berupa data-data produk elektronik yang digunakan sebagai sampel dalam uji coba ini, yang dibedakan menjadi dua macam berdasarkan karakteristik identitas yang dimilikinya. Kelompok pertama adalah data produk elektronik yang memiliki identitas berupa tiga buah atribut penting dari produk tersebut. Dan kelompok kedua adalah data produk elektronik yang memiliki identitas berupa empat buah atribut penting dari produk tersebut.

Berikut ini akan disajikan dan dibahas hasil uji coba yang dilakukan pada kedua jenis data target di atas. Pada percobaan 1 sampai dengan percobaan 3, akan dilakukan uji coba terhadap produk kelompok pertama yang memiliki tiga buah atribut. Dan pada percobaan 4 dan percobaan 5, akan dilakukan uji coba terhadap produk kelompok kedua yang memiliki empat buah atribut.

4.1. Percobaan 1

Percobaan dilakukan terhadap produk kelompok pertama yaitu produk “tv” yang memiliki 3 buah atribut yaitu “merk_tv”, “ukuran” layar, dan jenis “layar”. Data masukan berupa kalimat permintaan (*query*) yang didalamnya memuat kata kunci yang merupakan nomor seri atau kode model dari sebuah produk. Langkah-langkah dan hasil pengujian selengkapnya dapat dilihat pada penjelasan di bawah ini.

Kalimat permintaan : “saya ingin membeli tv samsung LN32B360”

Hasil ekstraksi kata kunci : tv, samsung, LN32B360

Hasil pencarian berdasarkan pencocokan kata kunci : 1 data

Produk target :

“Samsung LN32B360 32 in LCD HDTV”

Dari data hasil percobaan 1 di atas dapat dilihat bahwa pencarian data target menggunakan kata kunci “tv, samsung, LN32B360” menghasilkan 1 data tunggal yaitu “Samsung LN32B360 32 in LCD HDTV”, sehingga dapat disimpulkan bahwa data target telah ditemukan. Pencarian menggunakan kata kunci “LN32B360” dapat menemukan data target karena kata kunci tersebut merupakan nomor seri atau kode model dari sebuah produk elektronik yang bersifat unik, artinya setiap produk akan memiliki kode model yang berbeda.

Dari percobaan 1 ini dapat disimpulkan bahwa jika pada himpunan kata kunci terdapat nomor seri atau kode model dari sebuah produk, maka data target akan dapat ditemukan dengan melakukan pencarian berdasarkan pencocokan kata kunci nomor seri atau kode model tersebut, dan tidak perlu melakukan proses inferensi menggunakan jaringan Bayes semantik.

4.2. Percobaan 2

Percobaan dilakukan terhadap produk kelompok pertama yaitu produk “tv” yang memiliki 3 buah atribut yaitu “merk_tv”, “ukuran” layar, dan jenis “layar”. Data masukan berupa kalimat permintaan (*query*) dengan berbagai kombinasi yang didalamnya memuat kata kunci yang merupakan nilai data (*value*) dari 2 buah atribut produk. Hasil pengujian selengkapnya disajikan pada tabel 4.1 dan tabel 4.2 di bawah ini :

Tabel 4.1. Hasil pengujian dengan input awal memuat 2 atribut yaitu “merk_tv” dan “ukuran” layar

No.	Input dari Pengguna	Probabilitas Tertinggi		Tanggapan dari Sistem
		Simpul	P	
1	[tanya] tv toshiba 19 in	t19e	0,9115	[jawab] Toshiba t19e 19 in LED HDTV
2	[tanya] tv toshiba 32 in	t32e	0,9115	[jawab] Toshiba t32e 32 in LED HDTV
3	[tanya] tv toshiba 40 in	t40e	0,9115	[jawab] Toshiba t40e 40 in LED HDTV
4	[tanya] tv toshiba 46 in	t46e	0,9115	[jawab] Toshiba t46e 46 in LED HDTV
5	[tanya] tv toshiba 50 in	t50e	0,9115	[jawab] Toshiba t50e 50 in LED HDTV
6	[tanya] tv samsung 19 in	s19e	0,9115	[jawab] Samsung s19e 19 in LED HDTV
7	[tanya] tv samsung 32 in	s32e	0,9115	[jawab] Samsung s32e 32 in LED HDTV
8	[tanya] tv samsung 40 in	s40e	0,9115	[jawab] Samsung s40e 40 in LED HDTV
9	[tanya] tv samsung 46 in	s46e	0,9115	[jawab] Samsung s46e 46 in LED HDTV
10	[tanya] tv samsung 50 in	s50e	0,9115	[jawab] Samsung s50e 50 in LED HDTV
11	[tanya] tv lg 19 in	l19e	0,9115	[jawab] LG l19e 19 in LED HDTV
12	[tanya] tv lg 32 in	l32e	0,9115	[jawab] LG l32e 32 in LED HDTV
13	[tanya] tv lg 40 in	l40e	0,9115	[jawab] LG l40e 40 in LED HDTV
14	[tanya] tv lg 46 in	l46e	0,9115	[jawab] LG l46e 46 in LED HDTV
15	[tanya] tv lg 50 in	l50e	0,9115	[jawab] LG l50e 50 in LED HDTV
16	[tanya] tv sharp 19 in	sh19e	0,9115	[jawab] Sharp sh19e 19 in LED HDTV
17	[tanya] tv sharp 32 in	sh32e	0,9115	[jawab] Sharp sh32e 32 in LED HDTV
18	[tanya] tv sharp 40 in	sh40e	0,9115	[jawab] Sharp sh40e 40 in LED HDTV
19	[tanya] tv sharp 46 in	sh46e	0,9115	[jawab] Sharp sh46e 46 in LED HDTV
20	[tanya] tv sharp 50 in	sh50e	0,9115	[jawab] Sharp sh50e 50 in LED HDTV

21	[tanya] tv sony 19 in	so19e	0,9115	[jawab] Sony so19e 19 in LED HDTV
22	[tanya] tv sony 32 in	so32e	0,9115	[jawab] Sony so32e 32 in LED HDTV
23	[tanya] tv sony 40 in	so40e	0,9115	[jawab] Sony so40e 40 in LED HDTV
24	[tanya] tv sony 46 in	so46e	0,9115	[jawab] Sony so46e 46 in LED HDTV
25	[tanya] tv sony 50 in	so50e	0,9115	[jawab] Sony so50e 50 in LED HDTV

Tabel 4.2. Hasil pengujian dengan input awal memuat 2 atribut yaitu “ukuran” dan jenis “layar”

No	Input dari Pengguna	Probabilitas Tertinggi		Tanggapan dari Sistem
		Simpul	P	
1	[tanya] tv 19 in lcd	s19c	0,8895	[tanya] “merk_tv” yang anda inginkan?
	[jawab] lg	l19c	0,9500	[jawab] LG l19c 19 in LCD HDTV
2	[tanya] tv 19 in lcd	s19c	0,8895	[tanya] “merk_tv”?
	[jawab] sharp	sh19c	0,9500	[jawab] Sharp sh19c 19 in LCD HDTV
3	[tanya] tv 19 in lcd	s19c	0,8895	[tanya] “merk_tv”?
	[jawab] toshiba	t19c	0,9500	[jawab] Toshiba t19c 19 in LCD HDTV
4	[tanya] tv 19 in led	s19e	0,8895	[tanya] “merk_tv”?
	[jawab] samsung	s19e	0,9500	[jawab] Samsung s19e 19 in LED HDTV
5	[tanya] tv 19 in led	s19e	0,8895	[tanya] “merk_tv”?
	[jawab] sharp	sh19e	0,9500	[jawab] Sharp sh19e 19 in LED HDTV
6	[tanya] tv 19 in led	s19e	0,8895	[tanya] “merk_tv”?
	[jawab] sony	so19e	0,9500	[jawab] Sony so19e 19 in LED HDTV
7	[tanya] tv 19 in plasma	s19p	0,8895	[tanya] “merk_tv”?
	[jawab] lg	l19p	0,9500	[jawab] LG l19p 19 in Plasma HDTV
8	[tanya] tv 19 in	s19p	0,8895	[tanya] “merk_tv”?

	plasma				
	[jawab] sony	so19p	0,9500	[jawab] Sony so19p 19 in Plasma HDTV	
9	[tanya] tv 19 in plasma	s19p	0,8895	[tanya] “merk_tv”?	
	[jawab] toshiba	t19p	0,9500	[jawab] Toshiba t19p 19 in Plasma HDTV	
10	[tanya] tv 32 in lcd	s32c	0,8895	[tanya] “merk_tv”?	
	[jawab] samsung	s32c	0,9500	[jawab] Samsung s32c 19 in LCD HDTV	
11	[tanya] tv 32 in lcd	s32c	0,8895	[tanya] “merk_tv”?	
	[jawab] lg	l32c	0,9500	[jawab] LG l32c 32 in LCD HDTV	
12	[tanya] tv 32 in lcd	s32c	0,8895	[tanya]	
	[jawab] sharp	sh32c	0,9500	[jawab] Sharp sh32c 32 in LCD HDTV	
13	[tanya] tv 32 in led	s32e	0,8895	[tanya] “merk_tv”?	
	[jawab] lg	l32e	0,9500	[jawab] LG l32e 32 in LED HDTV	
14	[tanya] tv 32 in led	s32e	0,8895	[tanya] “merk_tv”?	
	[jawab] sony	so32e	0,9500	[jawab] Sony so32e 32 in LED HDTV	
15	[tanya] tv 32 in led	s32e	0,8895	[tanya] “merk_tv”?	
	[jawab] toshiba	t32e	0,9500	[jawab] Toshiba t32e 32 in LED HDTV	
16	[tanya] tv 32 in plasma	s32p	0,8895	[tanya] “merk_tv”?	
	[jawab] sharp	sh32p	0,9500	[jawab] Sharp sh32p 32 in Plasma HDTV	
17	[tanya] tv 32 in plasma	s32p	0,8895	[tanya] “merk_tv”?	
	[jawab] Sony	so32p	0,9500	[jawab] Sony so32p 32 in Plasma HDTV	
18	[tanya] tv 32 in plasma	s32p	0,8895	[tanya] “merk_tv”?	
	[jawab] toshiba	t32p	0,9500	[jawab] Toshiba t32p 32 in Plasma HDTV	
19	[tanya] tv 40 in lcd	s40c	0,8895	[tanya] “merk_tv”?	
	[jawab] sharp	sh40c	0,9500	[jawab] Sharp sh40c 40 in LCD	

				HDTV
20	[tanya] tv 40 in lcd	s40c	0,8895	[tanya] “merk_tv”?
	[jawab] sony	so40c	0,9500	[jawab] Sony so40c 40 in LCD HDTV
21	[tanya] tv 40 in lcd	s40c	0,8895	[tanya] “merk_tv”?
	[jawab] toshiba	t40c	0,9500	[jawab] Toshiba t40c 40 in LCD HDTV
22	[tanya] tv 40 in led	s40e	0,8895	[tanya] “merk_tv”?
	[jawab] lg	l40e	0,9500	[jawab] LG l40e 40 in LED HDTV
23	[tanya] tv 40 in led	s40e	0,8895	[tanya] “merk_tv”?
	[jawab] sharp	sh40e	0,9500	[jawab] Sharp sh40e 40 in LED HDTV
24	[tanya] tv 40 in led	s40e	0,8895	[tanya] “merk_tv”?
	[jawab] sony	so40e	0,9500	[jawab] Sony s40e 40 in LED HDTV
25	[tanya] tv 40 in plasma	s40c	0,8895	[tanya] “merk_tv”?
	[jawab] samsung	s40p	0,9500	[jawab] Samsung s40p 40 in Plasma HDTV
26	[tanya] tv 40 in plasma	s40p	0,8895	[tanya] “merk_tv”?
	[jawab] lg	l40p	0,9500	[jawab] LG l40p 40 in Plasma HDTV
27	[tanya] tv 40 in plasma	s40p	0,8895	[tanya] “merk_tv”?
	[jawab] toshiba	t40p	0,9500	[jawab] Toshiba 40 in Plasma HDTV
28	[tanya] tv 46 in lcd	s46c	0,8895	[tanya] “merk_tv”?
	[jawab] lg	l46c	0,9500	[jawab] LG l46c 46 in LCD HDTV
29	[tanya] tv 46 in lcd	s46c	0,8895	[tanya] “merk_tv”?
	[jawab] sharp	sh46c	0,9500	[jawab] Sharp sh46c 46 in LCD HDTV
30	[tanya] tv 46 in lcd	s46c	0,8895	[tanya] “merk_tv”?
	[jawab] toshiba	t46c	0,9500	[jawab] Toshiba t46c 46 in LCD HDTV
31	[tanya] tv 46 in	s46e	0,8895	[tanya] “merk_tv”?

	led			
	[jawab] samsung	s46e	0,9500	[jawab] Samsung s46e 46 in LED HDTV
32	[tanya] tv 46 in led	s46e	0,8895	[tanya] “merk_tv”?
	[jawab] sharp	sh46e	0,9500	[jawab] Sharp sh46e 46 in LED HDTV
33	[tanya] tv 46 in led	s46e	0,8895	[tanya] “merk_tv”?
	[jawab] sony	so46e	0,9500	[jawab] Sony so46e 46 in LED HDTV
34	[tanya] tv 46 in plasma	s46p	0,8895	[tanya] “merk_tv”?
	[jawab] sharp	sh46p	0,9500	[jawab] Sharp sh46p 46 in Plasma HDTV
35	[tanya] tv 46 in plasma	s46p	0,8895	[tanya] “merk_tv”?
	[jawab] sony	so46p	0,9500	[jawab] Sony so46p 46 in Plasma HDTV
36	[tanya] tv 46 in plasma	s46p	0,8895	[tanya] “merk_tv”?
	[jawab] toshiba	t46p	0,9500	[jawab] Toshiba t46p 46 in Plasma HDTV
37	[tanya] tv 50 in lcd	s50c	0,8895	[tanya] “merk_tv”?
	[jawab] sony	so50c	0,9500	[jawab] Sony so50c 50 in LCD HDTV
38	[tanya] tv 50 in lcd	s50c	0,8895	[tanya] “merk_tv”?
	[jawab] lg	l50c	0,9500	[jawab] LG l50c 50 in LCD HDTV
39	[tanya] 50 in lcd	s50c	0,8895	[tanya] “merk_tv”?
	[jawab] samsung	s50c	0,9500	[jawab] Samsung s50c 50 in LCD HDTV
40	tanya] 50 in led	s50e	0,8895	[tanya] “merk_tv”?
	[jawab] sharp	sh50e	0,9500	[jawab] Sharp sh50e 50 in LED HDTV
41	[tanya] 50 in led	s50e	0,8895	[tanya] “merk_tv”?
	[jawab] toshiba	t50e	0,9500	[jawab] Toshiba t50e 50 in LED HDTV
42	[tanya] 50 in led	s50e	0,8895	[tanya] “merk_tv”?
	[jawab] lg	l50e	0,9500	[jawab] LG l50e 50 in LED HDTV
43	[tanya] 50 in plasma	s50p	0,8895	[tanya] “merk_tv”?
	[jawab] toshiba	t50p	0,9500	[jawab] Toshiba t50p 50 in

				Plasma HDTV
44	[tanya] 50 in plasma	s50p	0,8895	[tanya] “merk_tv”?
	[jawab] lg	l50p	0,9500	[jawab] LG l50p 50 in Plasma HDTV
45	[tanya] 50 in plasma	s50p	0,8895	[tanya] “merk_tv”?
	[jawab] sony	so50p	0,9500	[jawab] Sony so50p 50 in Plasma HDTV

Dari data hasil pengujian pada tabel 4.1 dan tabel 4.2 menunjukkan bahwa dengan memberikan input kalimat permintaan yang memuat 2 nilai atribut produk, sistem memberikan dua macam tanggapan yang berbeda. Pada tabel 4.1 dapat dilihat bahwa sistem dapat menemukan data target dan menjawab pertanyaan pengguna, dengan melakukan inferensi berdasarkan 2 nilai atribut produk yang diinputkan oleh pengguna. Sedangkan pada tabel 4.2 dapat dilihat bahwa sistem belum dapat menemukan data target karena dari hasil inferensi menggunakan 2 nilai atribut produk, belum ada simpul di jaringan Bayes yang nilai probabilitasnya melebihi ambang batas yang telah ditetapkan yaitu sebesar 0,9. Tanggapan yang diberikan oleh sistem yaitu meminta kepada pengguna untuk memasukkan informasi tambahan berupa nilai atribut yang berkaitan dengan produk yang sedang dicari. Sistem dapat menemukan data target dan menjawab pertanyaan pengguna setelah mendapatkan 1 data tambahan dari pengguna dan melakukan inferensi berdasarkan 3 nilai atribut produk.

Rincian langkah-langkah yang dilakukan dapat dilihat pada penjelasan di bawah ini, yang dijelaskan dengan mengambil sampel sebuah kalimat permintaan sebagai berikut :

Kalimat permintaan : “saya ingin membeli tv toshiba 50 in”

Hasil ekstraksi kata kunci : tv, toshiba, 50 in

Hasil pencarian berdasarkan pencocokan kata kunci : 3 data

Data hasil pencarian selengkapnya dapat dilihat pada tabel 4.3 di bawah ini :

Tabel 4.3. Hasil pencarian berdasarkan pencocokan kata kunci pada percobaan 2

No	Data target hasil pencarian
1	Tv toshiba 50 in lcd
2	Tv toshiba 50 in led
3	Tv toshiba 50 in plasma

Dari tabel 4.3 di atas dapat dilihat bahwa pencarian langsung berdasarkan pencocokan kata kunci tidak berhasil menemukan data target karena himpunan kata kunci yang terdapat pada kalimat permintaan terlalu sedikit, sehingga ruang lingkup data yang dihasilkan menjadi luas. Oleh karena itu proses pencarian data target harus dilanjutkan dengan melakukan inferensi menggunakan jaringan Bayes semantik. Adapun tahapan yang dilakukan pada proses inferensi menggunakan jaringan Bayes semantik dapat ditunjukkan pada langkah-langkah di bawah ini.

Langkah pertama :

Pada jaringan Bayes, semua simpul yang berupa kata kunci yang terdapat pada kalimat permintaan pengguna diberikan nilai 1 sebagai berikut :

$$P(tv) = 1.0$$

$$P(toshiba | tv) = 1.0$$

$$P(50 in | tv) = 1.0$$

Langkah kedua :

Menghitung probabilitas semua simpul yang terhubung dengan simpul kata kunci “tv”, “toshiba”, dan “50 in” yaitu simpul “t50c”, “t50e”, dan “t50p”. Hasil pengujian selengkapnya dapat dilihat pada tabel 4.4 di bawah ini.

Tabel 4.4. Probabilitas simpul-simpul yang terhubung dengan kata kunci pada percobaan 2

Simpul S_i	Probabilitas S_i
t50c	0,8905
t50e	0,9115
t50p	0,8485

Secara manual, perhitungan probabilitas dari simpul-simpul yang ditunjukkan pada tabel 4.4 di atas dapat dijelaskan sebagai berikut :

$$P(t50c | toshiba, 50in, lcd) = 0,95$$

$$P(t50c | toshiba, 50in, lcd^c) = 0,6$$

$$P(t50c | toshiba, 50in^c, lcd) = 0,6$$

$$P(t50c | toshiba, 50in^c, lcd^c) = 0,3$$

$$P(t50c | toshiba^c, 50in, lcd) = 0,4$$

$$P(t50c | toshiba^c, 50in, lcd^c) = 0,2$$

$$P(t50c | toshiba^c, 50in^c, lcd) = 0,2$$

$$P(t50c | toshiba^c, 50in^c, lcd^c) = 0,05$$

$$P(toshiba | tv) = 1$$

$$P(toshiba^c | tv) = 0$$

$$P(50in | tv) = 1$$

$$P(50in^c | tv) = 0$$

$$P(lcd | tv) = 0,83$$

$$P(led | tv) = 0,89$$

$$P(plasma | tv) = 0,71$$

$$P(lcd^c | tv) = 0,17$$

$$P(led^c | tv) = 0,11$$

$$P(plasma^c | tv) = 0,29$$

Maka :

$$\begin{aligned} P(t50c) &= 0,95 \times p(\text{toshiba}|tv) \times p(50in|tv) \times p(lcd|tv) + \\ & 0,6 \times p(\text{toshiba}|tv) \times p(50in|tv) \times p(lcd^c|tv) + \\ & 0,6 \times p(\text{toshiba}|tv) \times p(50in^c|tv) \times p(lcd|tv) + \\ & 0,3 \times p(\text{toshiba}|tv) \times p(50in^c|tv) \times p(lcd^c|tv) + \\ & 0,6 \times p(\text{toshiba}^c|tv) \times p(50in|tv) \times p(lcd|tv) + \\ & 0,3 \times p(\text{toshiba}^c|tv) \times p(50in|tv) \times p(lcd^c|tv) + \\ & 0,3 \times p(\text{toshiba}^c|tv) \times p(50in^c|tv) \times p(lcd|tv) + \\ & 0,05 \times p(\text{toshiba}^c|tv) \times p(50in^c|tv) \times p(lcd^c|tv) \\ &= (0,95 \times 1 \times 1 \times 0,83) + (0,6 \times 1 \times 1 \times 0,17) + \\ & (0,6 \times 1 \times 0 \times 0,83) + (0,3 \times 1 \times 0 \times 0,17) + \\ & (0,6 \times 0 \times 1 \times 0,83) + (0,3 \times 0 \times 1 \times 0,17) + \\ & (0,3 \times 0 \times 0 \times 0,83) + (0,05 \times 0 \times 0 \times 0,17) \\ &= 0,8905 \end{aligned}$$

$$\begin{aligned}
P(t50e) &= 0,95 \times p(\text{toshiba}|\text{tv}) \times p(50\text{in}|\text{tv}) \times p(\text{led}|\text{tv}) + \\
& 0,6 \times p(\text{toshiba}|\text{tv}) \times p(50\text{in}|\text{tv}) \times p(\text{led}^c|\text{tv}) + \\
& 0,6 \times p(\text{toshiba}|\text{tv}) \times p(50\text{in}^c|\text{tv}) \times p(\text{led}|\text{tv}) + \\
& 0,3 \times p(\text{toshiba}|\text{tv}) \times p(50\text{in}^c|\text{tv}) \times p(\text{led}^c|\text{tv}) + \\
& 0,6 \times p(\text{toshiba}^c|\text{tv}) \times p(50\text{in}|\text{tv}) \times p(\text{led}|\text{tv}) + \\
& 0,3 \times p(\text{toshiba}^c|\text{tv}) \times p(50\text{in}|\text{tv}) \times p(\text{led}^c|\text{tv}) + \\
& 0,3 \times p(\text{toshiba}^c|\text{tv}) \times p(50\text{in}^c|\text{tv}) \times p(\text{led}|\text{tv}) + \\
& 0,05 \times p(\text{toshiba}^c|\text{tv}) \times p(50\text{in}^c|\text{tv}) \times p(\text{led}^c|\text{tv}) \\
&= (0,95 \times 1 \times 1 \times 0,89) + (0,6 \times 1 \times 1 \times 0,11) + \\
& (0,6 \times 1 \times 0 \times 0,89) + (0,3 \times 1 \times 0 \times 0,11) + \\
& (0,6 \times 0 \times 1 \times 0,89) + (0,3 \times 0 \times 1 \times 0,11) + \\
& (0,3 \times 0 \times 0 \times 0,89) + (0,05 \times 0 \times 0 \times 0,11) \\
&= 0,9115
\end{aligned}$$

$$\begin{aligned}
P(t50p) &= 0,95 \times p(\text{toshiba}|\text{tv}) \times p(50\text{in}|\text{tv}) \times p(\text{plasma}|\text{tv}) + \\
& 0,6 \times p(\text{toshiba}|\text{tv}) \times p(50\text{in}|\text{tv}) \times p(\text{plasma}^c|\text{tv}) + \\
& 0,6 \times p(\text{toshiba}|\text{tv}) \times p(50\text{in}^c|\text{tv}) \times p(\text{plasma}|\text{tv}) + \\
& 0,3 \times p(\text{toshiba}|\text{tv}) \times p(50\text{in}^c|\text{tv}) \times p(\text{plasma}^c|\text{tv}) + \\
& 0,6 \times p(\text{toshiba}^c|\text{tv}) \times p(50\text{in}|\text{tv}) \times p(\text{plasma}|\text{tv}) + \\
& 0,3 \times p(\text{toshiba}^c|\text{tv}) \times p(50\text{in}|\text{tv}) \times p(\text{plasma}^c|\text{tv}) + \\
& 0,3 \times p(\text{toshiba}^c|\text{tv}) \times p(50\text{in}^c|\text{tv}) \times p(\text{plasma}|\text{tv}) + \\
& 0,05 \times p(\text{toshiba}^c|\text{tv}) \times p(50\text{in}^c|\text{tv}) \times p(\text{plasma}^c|\text{tv}) \\
&= (0,95 \times 1 \times 1 \times 0,71) + (0,6 \times 1 \times 1 \times 0,29) + \\
& (0,6 \times 1 \times 0 \times 0,71) + (0,3 \times 1 \times 0 \times 0,29) + \\
& (0,6 \times 0 \times 1 \times 0,71) + (0,3 \times 0 \times 1 \times 0,29) + \\
& (0,3 \times 0 \times 0 \times 0,71) + (0,05 \times 0 \times 0 \times 0,29) \\
&= 0,8485
\end{aligned}$$

Dari tabel 4.4 maupun hasil perhitungan secara manual di atas dapat dilihat bahwa nilai probabilitas tertinggi dimiliki oleh simpul “t50e” yaitu sebesar 0,9115. Karena nilai probabilitas dari simpul tersebut telah melebihi nilai ambang batas yang ditetapkan yaitu sebesar 0,9 maka dapat dikatakan bahwa sistem telah berhasil menemukan data target yang dicari. Selanjutnya sistem akan membaca data-data lengkap dari simpul “t50e”, dan menampilkan informasi produk target ke pengguna sebagai berikut :

Produk target :

“Toshiba t50e 50 in LED HDTV”

4.3. Percobaan 3

Percobaan dilakukan terhadap produk kelompok pertama yaitu produk “tv” yang memiliki tiga buah atribut yaitu “merk_tv”, “ukuran” layar, dan jenis “layar”. Data masukan berupa kalimat permintaan (*query*) dengan berbagai kombinasi yang didalamnya memuat kata kunci yang merupakan nilai data (*value*) dari 1 buah atribut produk. Hasil pengujian selengkapnya disajikan pada tabel 4.5 di bawah ini :

Tabel 4.5. Hasil pengujian dengan input awal memuat satu atribut yaitu jenis “layar”

No	Input dari Pengguna	Probabilitas Tertinggi		Tanggapan dari Sistem
		Simpul	P	
1	[tanya] tv lcd	s32c	0,8562	[tanya] “merk_tv” yang anda inginkan?
	[jawab] samsung	s32c	0,9150	[jawab] Samsung s32c 32 in LCD HDTV
2	[tanya] tv lcd	s32c	0,8562	[tanya] “merk_tv” ?
	[jawab] lg	l32c	0,9150	[jawab] LG l32c 32 in LCD HDTV
3	[tanya] tv lcd	s32c	0,8562	[tanya] “merks_tv” ?
	[jawab] sharp	sh32c	0,9150	[jawab] Sharp sh32c 32 in LCD HDTV
4	[tanya] tv lcd	s32c	0,8562	[tanya] “merk_tv” ?
	[jawab] sony	so32c	0,9150	[jawab] Sony so32c 32 in LCD HDTV
5	[tanya] tv lcd	s32c	0,8562	[tanya] “merk_tv” ?
	[jawab] toshiba	t32c	0,9150	[jawab] Toshiba t32c 32 in LCD HDTV
6	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” yang anda inginkan?
	[jawab] 19 in	s19c	0,8895	[tanya] “merk_tv” ?
	[jawab] samsung	s32c	0,9500	[jawab] Samsung s19c 19 in LCD HDTV
7	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 19 in	s19c	0,8895	[tanya] “merk_tv” ?
	[jawab] lg	l19c	0,9500	[jawab] LG l19c 19 in LCD HDTV
8	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 19 in	s19c	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so19c	0,9500	[jawab] Sony so19c 19 in LCD

				HDTV
9	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 32 in	s32c	0,8895	[tanya] “merk_tv” ?
	[jawab] lg	l32c	0,9500	[jawab] LG l32c 32 in LCD HDTV
10	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 32 in	s32c	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so32c	0,9500	[jawab] Sony so32c 32 in LCD HDTV
11	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 32 in	s32c	0,8895	[tanya] “merk_tv” ?
	[jawab] toshiba	t32c	0,9500	[jawab] Toshiba t32c 32 in LCD HDTV
12	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 40 in	s40c	0,8895	[tanya] “merk_tv” ?
	[jawab] sharp	sh40c	0,9500	[jawab] Sharp sh40c 40 in LCD HDTV
14	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 40 in	s40c	0,8895	[tanya] “merk_tv” ?
	[jawab] toshiba	t40c	0,9500	[jawab] Toshiba t40c 40 in LCD HDTV
15	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 40 in	s40c	0,8895	[tanya] “merk_tv” ?
	[jawab] samsung	s40c	0,9500	[jawab] Samsung s40c 40 in LCD HDTV
16	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 46 in	s46c	0,8895	[tanya] “merk_tv” ?
	[jawab] sharp	sh46c	0,9500	[jawab] Sharp sh46c 46 in LCD HDTV
17	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 46 in	s46c	0,8895	[tanya] “merk_tv” ?
	[jawab] toshiba	t46c	0,9500	[jawab] Toshiba t46c 46 in LCD HDTV
18	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 50 in	s50c	0,8895	[tanya] “merk_tv” ?
	[jawab] lg	l50c	0,9500	[jawab] LG l50c 50 in LCD HDTV
19	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 50 in	s50c	0,8895	[tanya] “merk_tv” ?
	[jawab] sharp	sh50c	0,9500	[jawab] Sharp sh50c 50 in LCD HDTV
20	[tanya] tv lcd	s32c	0,8562	[tanya] “ukuran” ?
	[jawab] 50 in	s50c	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so50c	0,9500	[jawab] Sony so50c 50 in LCD HDTV
21	[tanya] tv led	s32e	0,8562	[tanya] “merk_tv” ?

	[jawab] samsung	s32e	0,9150	[jawab] Samsung s32e 32 in LED HDTV
22	[tanya] tv led	s32e	0,8562	[tanya] “merk_tv” ?
	[jawab] lg	l32e	0,9150	[jawab] LG l32e 32 in LED HDTV
23	[tanya] tv led	s32e	0,8562	[tanya] “merk_tv” ?
	[jawab] sharp	sh32e	0,9150	[jawab] Sharp sh32e in LED HDTV
24	[tanya] tv led	s32e	0,8562	[tanya] “merk_tv” ?
	[jawab] sony	so32e	0,9150	[jawab] Sony so32e 32 in LED HDTV
25	[tanya] tv led	s32e	0,8562	[tanya] “merk_tv” ?
	[jawab] toshiba	t32e	0,9150	[jawab] Toshiba t32e 32 in LED HDTV
26	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 19 in	s19e	0,8895	[tanya] “merk_tv” ?
	[jawab] lg	l19e	0,9500	[jawab] “LG 119e 19 in LED HDTV
27	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 19 in	s19e	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so19e	0,9500	[jawab] Sony so19e 19 in LED HDTV
28	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 19 in	s19e	0,8895	“merk_tv” ?
	[jawab] toshiba	t19e	0,9500	[jawab] Toshiba t19e 19 in LED HDTV
29	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 32 in	s32e	0,8895	[tanya] “merk_tv” ?
	[jawab] sharp	sh32e	0,9500	[jawab] Sharp sh32e 32 in LED HDTV
30	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 32 in	s32e	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so32e	0,9500	[jawab] Sony so32e 32 in LED HDTV
31	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 32 in	s32e	0,8895	[tanya] “merk_tv” ?
	[jawab] toshiba	t32e	0,9500	[jawab] Toshiba t32e 32 in LED HDTV
32	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 40 in	s40e	0,8895	[tanya] “merk_tv” ?
	[jawab] samsung	s40e	0,9500	[jawab] Samsung s40e 40 in LED HDTV
33	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 40 in	s40e	0,8895	[tanya] “merk_tv” ?
	[jawab] lg	l40e	0,9500	[jawab] LG 140e 40 in LED HDTV

34	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 40 in	s40e	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so40e	0,9500	[jawab] Sony so40e 40 in LED HDTV
35	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 46 in	s46e	0,8895	[tanya] “merk_tv” ?
	[jawab] lg	l46e	0,9500	[jawab] LG l46e 46 in LED HDTV
36	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 46 in	s46e	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so46e	0,9500	[jawab] Sony so46e 46 in LED HDTV
37	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 46 in	s46e	0,8895	[tanya] “merk_tv” ?
	[jawab] toshiba	t46e	0,9500	[jawab] Toshiba t46e 46 in LED HDTV
38	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 50 in	s50e	0,8895	[tanya] “merk_tv” ?
	[jawab] sharp	sh50e	0,9500	[jawab] Sharp sh50e 50 in LED HDTV
39	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 50 in	s50e	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so50e	0,9500	[jawab] Sony so50e 50 in LED HDTV
40	[tanya] tv led	s32e	0,8562	[tanya] “ukuran” ?
	[jawab] 50 in	s50e	0,8895	[tanya] “merk_tv” ?
	[jawab] toshiba	t50e	0,9500	[jawab] Toshiba t50e 50 in LED HDTV
41	[tanya] tv plasma	s32p	0,8562	[tanya] “merk_tv” ?
	[jawab] samsung	s32p	0,9150	[jawab] Samsung s32p 32 in Plasma HDTV
42	[tanya] tv plasma	s32p	0,8562	[tanya] “merk_tv” ?
	[jawab] lg	l32p	0,9150	[jawab] LG l32p 32 in Plasma HDTV
43	[tanya] tv plasma	s32p	0,8562	[tanya] “merk_tv” ?
	[jawab] sharp	sh32p	0,9150	[jawab] Sharp sh32p 32 in Plasma HDTV
44	[tanya] tv plasma	s32p	0,8562	[tanya] “merk_tv” ?
	[jawab] sony	so32p	0,9150	[jawab] Sony so32p 32 in Plasma HDTV
45	[tanya] tv plasma	s32p	0,8562	[tanya] “merk_tv” ?
	[jawab] toshiba	t32p	0,9150	[jawab] Toshiba t32p 32 in Plasma HDTV
46	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 19 in	s19p	0,8895	[tanya] “merk_tv” ?
	[jawab] lg	l19p	0,9500	[jawab] LG l19p 19 in Plasma

				HDTV
47	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 19 in	s19p	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so19p	0,9500	[jawab] Sony so19p 19 in Plasma HDTV
48	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 19 in	s19p	0,8895	[tanya] “merk_tv” ?
	[jawab] toshiba	t19p	0,9500	[jawab] Toshiba t19p 19 in Plasma HDTV
49	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 32 in	s32p	0,8895	[tanya] “merk_tv” ?
	[jawab] samsung	s32p	0,9500	[jawab] Samsung s32p 32 in Plasma HDTV
50	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 32 in	s32p	0,8895	[tanya] “merk_tv” ?
	[jawab] lg	l32p	0,9500	[jawab] LG l32p 32 in Plasma HDTV
51	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 32 in	s32p	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so32p	0,9500	[jawab] Sony so32p 32 in Plasma HDTV
52	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 40 in	s40p	0,8895	[tanya] “merk_tv” ?
	[jawab] lg	l40p	0,9500	[jawab] LG l40p 40 in Plasma HDTV
53	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 40 in	s40p	0,8895	[tanya] “merk_tv” ?
	[jawab] sharp	sh40p	0,9500	[jawab] Sharp sh40p 40 in Plasma HDTV
54	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 40 in	s40p	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so40p	0,9500	[jawab] Sony so40p 40 in Plasma HDTV
55	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 46 in	s46p	0,8895	[tanya] “merk_tv” ?
	[jawab] sharp	sh46p	0,9500	[jawab] Sharp sh46p 46 in Plasma HDTV
56	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 46 in	s46p	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so46p	0,9500	[jawab] Sony so46p 46 in Plasma HDTV
57	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 46 in	s46p	0,8895	[tanya] “merk_tv” ?
	[jawab] toshiba	t46p	0,9500	[jawab] Toshiba t46p 46 in Plasma HDTV
58	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?

	[jawab] 50 in	s50p	0,8895	[tanya] “merk_tv” ?
	[jawab] samsung	s50p	0,9500	[jawab] Samsung s50p 50 in Plasma HDTV
59	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 50 in	s50p	0,8895	[tanya] “merk_tv” ?
	[jawab] sony	so50p	0,9500	[jawab] Sony so50p 50 in Plasma HDTV
60	[tanya] tv plasma	s32p	0,8562	[tanya] “ukuran” ?
	[jawab] 50 in	s50p	0,8895	[tanya] “merk_tv” ?
	[jawab] toshiba	t50p	0,9500	[jawab] Toshiba t50p 50 in Plasma HDTV

Dari data hasil pengujian pada tabel 4.5 menunjukkan bahwa dengan memberikan input kalimat permintaan yang hanya memuat 1 nilai atribut produk, sistem memberikan dua macam tanggapan yang berbeda. Pada beberapa kondisi input, sistem dapat menemukan data target dan menjawab pertanyaan pengguna setelah meminta 1 data tambahan kepada pengguna dan melakukan inferensi berdasarkan 2 nilai atribut produk yang diinputkan oleh pengguna. Dan pada beberapa kondisi input yang lainnya, sistem dapat menemukan data target dan menjawab pertanyaan pengguna, setelah meminta 2 data tambahan kepada pengguna dan melakukan inferensi berdasarkan 3 nilai atribut produk.

Rincian langkah-langkah yang dilakukan dapat dilihat pada penjelasan di bawah ini, yang dijelaskan dengan mengambil sampel sebuah kalimat permintaan sebagai berikut :

Kalimat permintaan : “saya ingin membeli tv samsung”

Hasil ekstraksi kata kunci : tv, samsung

Hasil pencarian berdasarkan pencocokan kata kunci : 15 data

Data hasil pencarian selengkapnya dapat dilihat pada tabel 4.6 di bawah ini :

Tabel 4.6. Hasil pencarian berdasarkan pencocokan kata kunci pada percobaan 3

No	Data target hasil pencarian
1	Tv samsung 19 in lcd
2	Tv samsung 19 in led
3	Tv samsung 19 in plasma
4	Tv samsung 32 in lcd
5	Tv samsung 32 in led
6	Tv samsung 32 in plasma
7	Tv samsung 40 in lcd
8	Tv samsung 40 in led
9	Tv samsung 40 in plasma
10	Tv samsung 46 in lcd
11	Tv samsung 46 in led
12	Tv samsung 46 in plasma
13	Tv samsung 50 in lcd
14	Tv samsung 50 in led
15	Tv samsung 50 in plasma

Dari tabel 4.6 di atas dapat dilihat bahwa pencarian langsung berdasarkan pencocokan kata kunci tidak berhasil menemukan data target karena himpunan kata kunci yang terdapat pada kalimat permintaan terlalu sedikit, sehingga ruang lingkup data yang dihasilkan menjadi luas. Oleh karena itu proses pencarian data target harus dilanjutkan dengan melakukan inferensi menggunakan jaringan Bayes semantik. Adapun tahapan yang dilakukan pada proses inferensi menggunakan jaringan Bayes semantik dapat ditunjukkan pada langkah-langkah di bawah ini.

Langkah pertama :

Pada jaringan Bayes, semua simpul yang berupa kata kunci yang terdapat pada kalimat permintaan pengguna diberikan nilai 1 sebagai berikut :

$$P(tv) = 1.0$$

$$P(samsung | tv) = 1.0$$

Langkah kedua :

Menghitung probabilitas semua simpul yang terhubung dengan simpul kata kunci “tv” dan “samsung”. Hasil pengujian selengkapnya dapat dilihat pada tabel 4.7 di bawah ini.

Tabel 4.7. Probabilitas simpul-simpul yang terhubung dengan kata kunci pada percobaan 3

Simpul S_i	Probabilitas S_i
s19c	0,81196
s32c	0,85635
s40c	0,79830
s46c	0,78805
s50c	0,79147
s19e	0,83227
s32e	0,87705
s40e	0,81849
s46e	0,80815
s50e	0,81160
s19p	0,77134
s32p	0,81495
s40p	0,75792
s46p	0,74785
s50p	0,75121

Dari tabel 4.7 di atas dapat dilihat bahwa nilai probabilitas tertinggi dimiliki oleh simpul “s32e” yaitu sebesar 0,87705, tetapi nilai probabilitas tersebut masih lebih rendah dari ambang batas yang telah ditetapkan yaitu sebesar 0,9. Dengan demikian, sampai dengan tahap ini sistem masih belum dapat menemukan data target yang diinginkan, dan akan dilanjutkan dengan proses pengumpulan informasi tambahan melalui jaringan semantik.

Langkah ketiga :

Sistem akan mencari informasi yang berkaitan dengan kata kunci “tv” dan kata kunci “samsung” pada jaringan semantik, yaitu berupa simpul-simpul objek yang memiliki hubungan O-A (Objek-Atribut) atau A-V (Atribut-Value). Hasil pengujian dapat dilihat pada tabel 4.8 di bawah ini.

Tabel 4.8. Simpul jaringan semantik yang memiliki hubungan dengan objek “tv” dan “samsung” pada percobaan 3

Objek1	Relasi	Objek2
Tv	Has-a	Merk_tv
Tv	Has-a	Ukuran
Tv	Has-a	Layar
Merk_tv	Has-a	Samsung

Dari tabel 4.8 di atas dapat dilihat bahwa objek “tv” memiliki hubungan O-A dengan objek “merk_tv”, “ukuran”, dan “layar”. Atau dengan kata lain dapat dijelaskan bahwa objek “tv” memiliki tiga buah atribut yaitu “merk_tv”, “ukuran” dan “layar”. Sedangkan objek “samsung” memiliki hubungan A-V dengan objek “merk_tv”, atau dengan kata lain dapat dijelaskan bahwa atribut “merk_tv” memiliki nilai data (*value*) “samsung”.

Berdasarkan data-data di atas, maka sistem akan meminta informasi tambahan kepada pengguna melalui kalimat permintaan yang baru, yaitu berupa nilai data (*value*) dari atribut “ukuran” atau atribut “layar”. Dalam percobaan ini sistem memilih atribut “ukuran” yang memiliki beberapa nilai data (*value*) seperti ditunjukkan pada tabel 4.9 di bawah ini.

Tabel 4.9. Nilai data (*value*) dari atribut “ukuran” pada percobaan 3

Objek1	Relasi	Objek2
Ukuran	Has-a	19 in
Ukuran	Has-a	32 in
Ukuran	Has-a	40 in

Ukuran	Has-a	46 in
Ukuran	Has-a	50 in

Setelah mengidentifikasi nilai-nilai data (*value*) dari atribut “ukuran”, maka selanjutnya sistem akan mengajukan kalimat permintaan kepada pengguna sebagai berikut :

[Sistem] : “ukuran” yang diinginkan? 19 in, 32 in, 40 in, 46 in, 50 in?

[Pengguna] : 19 in

Hasil ekstraksi kata kunci tambahan : 19 in

Langkah keempat :

Setelah mendapatkan data kata kunci tambahan dari pengguna, sistem akan mengupdate simpul kata kunci “19 in” pada jaringan Bayes, dengan memberikan nilai 1 sebagai berikut :

$$P(19 \text{ in} | tv) = 1.0$$

Kemudian sistem akan menghitung ulang probabilitas semua simpul yang terhubung dengan simpul kata kunci “tv”, “samsung”, dan “19 in”. Hasil pengujian selengkapnya dapat dilihat pada tabel 4.10 di bawah ini.

Tabel 4.10. Probabilitas simpul-simpul yang terhubung dengan kata kunci awal dan kata kunci tambahan pada percobaan 3

Simpul S_i	Probabilitas S_i
s19c	0,8905
S19e	0,9115
S19p	0,8485

Dari tabel 4.10 di atas dapat dilihat bahwa nilai probabilitas tertinggi dimiliki oleh simpul “s19e” yaitu 0,9115. Karena nilai probabilitas dari simpul tersebut telah melebihi nilai ambang batas yang ditetapkan yaitu sebesar 0,9 maka dapat dikatakan bahwa sistem telah berhasil menemukan data target yang dicari.

Selanjutnya sistem akan membaca data-data lengkap dari simpul “s19e”, dan menampilkan informasi produk target ke pengguna sebagai berikut :

Produk target :

“Samsung s19e 19 in LED HDTV”

4.4. Percobaan 4

Percobaan dilakukan terhadap produk kelompok kedua yaitu produk “handycam” yang memiliki empat buah atribut yaitu “merk_cam”, “media” penyimpanan, optical “zoom”, dan jenis “sensor”. Data masukan berupa kalimat permintaan (*query*) dengan berbagai kombinasi yang didalamnya memuat kata kunci yang merupakan nilai data (*value*) dari 2 buah atribut produk. Hasil pengujian selengkapnya disajikan pada tabel 4.11 di bawah ini :

Tabel 4.11. Hasil pengujian dengan input awal memuat dua atribut yaitu “merk” dan “media” penyimpan

No	Input dari Pengguna	Probabilitas Tertinggi		Tanggapan dari Sistem
		Simpul	P	
1	[tanya] handycam sony minidv	somd40m	0,8473	[tanya] “zoom” yang anda inginkan?
	[jawab] 10x	somd10m	0,9150	[jawab] Sony somd10m Minidv Handycam with 10x Zoom
2	[tanya] handycam sony minidv	somd40m	0,8473	[tanya] “zoom”?
	[jawab] 20x	Somd20m	0,9150	[jawab] Sony somd20m Minidv Handycam with 20x Zoom
3	[tanya] handycam sony minidv	somd40m	0,8473	[tanya] “zoom”?
	[jawab] 40x	somd40m	0,9150	[jawab] Sony somd40m Minidv Handycam with 40x Zoom
4	[tanya] handycam sony hdd	sohd40m	0,8473	[tanya] “zoom”?
	[jawab] 10x	Sohd10m	0,9150	[jawab] Sony sohd10m Hard Disk Handycam with 10x Zoom
5	[tanya] handycam	sohd40m	0,8473	[tanya] “zoom”?

	sony hdd				
	[jawab] 20x	Sohd20m	0,9150		[jawab] Sony sohd20m Hard Disk Handycam with 20x Zoom
6	[tanya] handycam sony hdd	somd40m	0,8473		[tanya] "zoom"?
	[jawab] 40x	somd40m	0,9150		[jawab] Sony sohd40m Hard Disk Handycam with 40x Zoom
7	[tanya] handycam sony dvd	sodv40m	0,8473		[tanya] "zoom"?
	[jawab] 10x	Sodv10m	0,9150		[jawab] Sony sodv10m DVD Handycam with 10x Zoom
8	[tanya] handycam sony dvd	sodv40m	0,8473		[tanya] "zoom"?
	[jawab] 10x	Sodv20m	0,9150		[jawab] Sony sodv20m DVD Handycam with 20x Zoom
9	[tanya] handycam sony dvd	sodv40m	0,8473		[tanya] "zoom"?
	[jawab] 10x	sodv40m	0,9150		[jawab] Sony sodv40m DVD Handycam with 40x Zoom
10	[tanya] handycam canon minidv	cmd40m	0,8473		[tanya] "zoom"?
	[jawab] 10x	cmd10m	0,9150		[jawab] Canon cmd10m Minidv Handycam with 10x Zoom
11	[tanya] handycam canon minidv	cmd40m	0,8473		[tanya] "zoom"?
	[jawab] 20x	cmd20m	0,9150		[jawab] Canon cmd20m Minidv Handycam with 20x Zoom
12	[tanya] handycam canon minidv	cmd40m	0,8473		[tanya] "zoom"?
	[jawab] 40x	cmd40m	0,9150		[jawab] Canon cmd40m Minidv Handycam with 40x Zoom
13	[tanya] handycam canon hdd	chd40m	0,8473		[tanya] "zoom"?
	[jawab] 10x	chd10m	0,9150		[jawab] Canon chd10m Hard Disk Handycam with 10x Zoom
14	[tanya] handycam canon hdd	chd40m	0,8473		[tanya] "zoom"?
	[jawab] 20x	chd20m	0,9150		[jawab] Canon chd20m Hard Disk Handycam with 20x Zoom

15	[tanya] handycam canon hdd	chd40m	0,8473	[tanya] “zoom”?
	[jawab] 40x	chd40m	0,9150	[jawab] Canon chd40m Hard Disk Handycam with 40x Zoom
16	[tanya] handycam canon dvd	cdv40m	0,8473	[tanya] “zoom”?
	[jawab] 10x	cdv10m	0,9150	[jawab] Canon cdv10m DVD Handycam with 10x Zoom
17	[tanya] handycam canon dvd	cdv40m	0,8473	[tanya] “zoom”?
	[jawab] 20x	cdv20m	0,9150	[jawab] Canon cdv20m DVD Handycam with 20x Zoom
18	[tanya] handycam canon dvd	cdv40m	0,8473	[tanya] “zoom”?
	[jawab] 40x	cdv40m	0,9150	[jawab] Canon cdv40m DVD Handycam with 40x Zoom
19	[tanya] handycam jvc sdcad	jsd40m	0,8473	[tanya] “zoom”?
	[jawab] 10x	jsd10m	0,9150	[jawab] JVC jsd10m Sdcad Handycam with 10x Zoom
20	[tanya] handycam jvc sdcad	jsd40m	0,8473	[tanya] “zoom”?
	[jawab] 20x	jsd20m	0,9150	[jawab] JVC jsd20m Sdcad Handycam with 20x Zoom
21	[tanya] handycam jvc sdcad	jsd40m	0,8473	[tanya] “zoom”?
	[jawab] 40x	jsd40m	0,9150	[jawab] JVC jsd40m Sdcad Handycam with 40x Zoom
22	[tanya] handycam jvc minidv	jmd40m	0,8473	[tanya] “zoom”?
	[jawab] 10x	jmd10m	0,9150	[jawab] JVC jmd10m Minidv Handycam with 10x Zoom
23	[tanya] handycam jvc minidv	jmd40m	0,8473	[tanya] “zoom”?
	[jawab] 20x	Jmd20m	0,9150	[jawab] JVC jmd20m Minidv Handycam with 20x Zoom
24	[tanya] handycam jvc minidv	jmd40m	0,8473	[tanya] “zoom”?
	[jawab] 40x	Jmd40m	0,9150	[jawab] JVC jmd40m Minidv Handycam with 40x Zoom
25	[tanya] handycam jvc hdd	Jhd40m	0,8473	[tanya] “zoom”?
	[jawab] 10x	Jhd10m	0,9150	[jawab] JVC jhd10m Hard Disk Handycam with 10x Zoom

26	[tanya] handycam jvc hdd	Jhd40m	0,8473	[tanya] “zoom”?
	[jawab] 20x	Jhd20m	0,9150	[jawab] JVC jhd20m Hard Disk Handycam with 20x Zoom
27	[tanya] handycam jvc hdd	Jhd40m	0,8473	[tanya] “zoom”?
	[jawab] 40x	Jhd40m	0,9150	[jawab] JVC jhd40m Hard Disk Handycam with 40x Zoom

Dari data hasil pengujian pada tabel 4.11 menunjukkan bahwa dengan memberikan input kalimat permintaan yang memuat 2 nilai atribut produk, sistem dapat menemukan data target dan menjawab pertanyaan pengguna setelah meminta 1 data tambahan kepada pengguna dan melakukan inferensi berdasarkan 3 nilai atribut produk.

Rincian langkah-langkah yang dilakukan dapat dilihat pada penjelasan di bawah ini, yang dijelaskan dengan mengambil sampel sebuah kalimat permintaan sebagai berikut :

Kalimat permintaan : “saya mencari handycam sony dengan media penyimpan harddisk”

Hasi ekstraksi kata kunci : handycam, sony

Pada saat sistem melakukan ekstraksi dan identifikasi kata kunci, kata “harddisk” tidak akan teridentifikasi sebagai kata kunci karena kata tersebut tidak didefinisikan di tabel kata kunci, tetapi kata harddisk memiliki sinonim atau padanan kata yaitu “hdd” yang ada di tabel kata kunci.

Dengan melakukan pengecekan sinonim, maka hasil akhir dari ekstraksi kata kunci yaitu : handycam, sony, hdd

Hasil pencarian berdasarkan pencocokan kata kunci : 9 data

Data hasil pencarian selengkapnya dapat dilihat pada tabel 4.12 di bawah ini :

Tabel 4.12. Hasil pencarian berdasarkan pencocokan kata kunci pada percobaan 4

No	Data target hasil pencarian
1	Handycam sony hdd 10x ccd
2	Handycam sony hdd 10x cmos
3	Handycam sony hdd 10x 3ccd
4	Handycam sony hdd 20x ccd
5	Handycam sony hdd 20x cmos
6	Handycam sony hdd 20x 3ccd
7	Handycam sony hdd 40x ccd
8	Handycam sony hdd 40x cmos
9	Handycam sony hdd 40x 3ccd

Dari tabel 4.12 di atas dapat dilihat bahwa pencarian berdasarkan pencocokan kata kunci tidak berhasil menemukan data target karena himpunan kata kunci yang terdapat pada kalimat permintaan terlalu sedikit, sehingga ruang lingkup data yang dihasilkan menjadi luas. Oleh karena itu proses pencarian data target harus dilanjutkan dengan melakukan inferensi menggunakan jaringan Bayes semantik. Adapun tahapan yang dilakukan pada proses inferensi menggunakan jaringan Bayes semantik dapat ditunjukkan pada langkah-langkah di bawah ini.

Langkah pertama :

Pada jaringan Bayes, semua simpul yang berupa kata kunci yang terdapat pada kalimat permintaan pengguna diberikan nilai 1 sebagai berikut :

$$P(\text{handycam}) = 1.0$$

$$P(\text{sony} \mid \text{tv}, \text{handycam}) = 1.0$$

$$P(\text{hdd} \mid \text{handycam}) = 1.0$$

Langkah kedua :

Menghitung probabilitas semua simpul yang terhubung dengan simpul kata kunci “handycam”, “sony”, dan “hdd”. Hasil pengujian selengkapnya dapat dilihat pada tabel 4.13 di bawah ini.

Tabel 4.13. Probabilitas simpul-simpul yang terhubung dengan kata kunci pada percobaan 4

Simpul S_i	Probabilitas S_i
Sohd10c	0,80420
Sohd10m	0,83370
Sohd103	0,78650
Sohd20c	0,80706
Sohd20m	0,83641
Sohd203	0,78945
Sohd40c	0,81850
Sohd40m	0,84725
Sohd403	0,70938

Dari tabel 4.13 di atas dapat dilihat bahwa nilai probabilitas tertinggi dimiliki oleh simpul “sohd40m” yaitu sebesar 0,84725, tetapi nilai probabilitas tersebut masih lebih rendah dari ambang batas yang telah ditetapkan yaitu sebesar 0,9. Dengan demikian, sampai dengan tahap ini sistem masih belum dapat menemukan data target yang diinginkan, dan akan dilanjutkan dengan proses pengumpulan informasi tambahan melalui jaringan semantik.

Langkah ketiga :

Sistem akan mencari informasi yang berkaitan dengan kata kunci “handycam” dan kata kunci “sony”, “hdd” pada jaringan semantik, yaitu berupa simpul-simpul objek yang memiliki hubungan O-A (Objek-Atribut) atau A-V (Atribut-Value).

Hasil pengujian dapat dilihat pada tabel 4.14 di bawah ini.

Tabel 4.14. Simpul jaringan semantik yang memiliki hubungan dengan objek “handycam” dan “sony”, dan “hdd” pada percobaan 4

Objek1	Relasi	Objek2
Handycam	Has-a	Merk_cam
Handycam	Has-a	Media

Handycam	Has-a	Zoom
Handycam	Has-a	Sensor
Media	Has-a	Hdd
Merk_cam	Has-a	Sony

Dari tabel 4.14 di atas dapat dilihat bahwa objek “handycam” memiliki hubungan O-A dengan objek “merk_cam”, “media”, “zoom” dan “sensor”. Atau dengan kata lain dapat dijelaskan bahwa objek “handycam” memiliki empat buah atribut yaitu “merk_cam”, “media”, “zoom”, dan “sensor”. Sedangkan objek “sony” memiliki hubungan A-V dengan objek “merk_cam”, atau dengan kata lain dapat dijelaskan bahwa atribut “merk_cam” memiliki nilai data (*value*) “sony”. Dan objek “hdd” memiliki hubungan A-V dengan objek “media”, atau dengan kata lain atribut “media” memiliki nilai data (*value*) “hdd”.

Berdasarkan data-data di atas, maka sistem akan meminta informasi tambahan kepada pengguna yang berupa nilai data dari atribut “zoom” atau atribut “sensor”. Dalam percobaan ini sistem memilih atribut “zoom” yang memiliki beberapa nilai data (*value*) seperti ditunjukkan pada tabel 4.15 di bawah ini.

Tabel 4.15. Nilai data (*value*) dari atribut “zoom” pada percobaan 4

Objek1	Relasi	Objek2
Zoom	Has-a	10x
Zoom	Has-a	20x
Zoom	Has-a	40x

Setelah mendapatkan nilai-nilai data (*value*) dari atribut “zoom”, selanjutnya sistem akan mengajukan kalimat permintaan kepada pengguna sebagai berikut :

[Sistem] : “zoom” yang diinginkan? 10x, 20x, 40x?

[Pengguna] : 40x

Hasil ekstraksi kata kunci tambahan : 40x

Langkah keempat :

Setelah mendapatkan data kata kunci tambahan dari pengguna, sistem akan mengupdate simpul kata kunci “40x” pada jaringan Bayes, dengan memberikan nilai 1 sebagai berikut :

$$P(40x | handycam) = 1.0$$

Kemudian sistem akan menghitung ulang probabilitas semua simpul yang terhubung dengan simpul kata kunci “handycam”, “sony”, “hdd”, “40x”. Hasil pengujian selengkapnya dapat dilihat pada tabel 4.16 di bawah ini.

Tabel 4.16. Probabilitas simpul-simpul yang terhubung dengan kata kunci awal dan kata kunci tambahan pada percobaan 4

Simpul S_i	Probabilitas S_i
sohd40c	0,8900
sohd40m	0,9150
sohd403	0,8750

Dari tabel 4.16 di atas dapat dilihat bahwa nilai probabilitas tertinggi dimiliki oleh simpul “sohd40m” yaitu 0,9150. Karena nilai probabilitas dari simpul tersebut telah melebihi nilai ambang batas yang ditetapkan yaitu sebesar 0,9 maka dapat dikatakan bahwa sistem telah berhasil menemukan data target yang dicari. Selanjutnya sistem akan membaca data-data lengkap dari simpul “sohd40m”, dan menampilkan informasi produk target ke pengguna sebagai berikut :

Produk target :

“Sony sohd40m Hard Disk Handycam with 40x Optical Zoom”

4.5. Percobaan 5

Percobaan dilakukan terhadap produk kelompok kedua yaitu produk “handycam” yang memiliki empat buah atribut yaitu “merk_cam”, “media” penyimpanan, optical “zoom”, dan jenis “sensor”. Data masukan berupa kalimat permintaan (*query*) dengan berbagai kombinasi yang didalamnya memuat kata

kunci yang merupakan nilai data (*value*) dari 1 buah produk. Hasil pengujian selengkapnya disajikan pada tabel 4.17 di bawah ini :

Tabel 4.17. Hasil pengujian dengan input awal memuat satu atribut yaitu “zoom”

No	Input dari Pengguna	Probabilitas Tertinggi		Tanggapan dari Sistem
		Simpul	P	
1	[tanya]handycam 10x zoom	Sosd10m	0,8447	[tanya] ”merk” yang anda inginkan?
	[jawab] jvc	Jsd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] dvd	Jdv10m	0,9150	[jawab] JVC jdv10m DVD Handycam with 10x Zoom
2	[tanya]handycam 10x zoom	Sosd10m	0,8447	[tanya] ”merk”?
	[jawab] jvc	Jsd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] sdcad	jsd10m	0,9150	[jawab] JVC jsd10m DVD Handycam with 10x Zoom
3	[tanya]handycam 10x zoom	Sosd10m	0,8447	[tanya] ”merk”?
	[jawab] jvc	Jsd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] minidv	jmd10m+	0,9150	[jawab] JVC jmd10m Minidv Handycam with 10x Zoom
4	[tanya]handycam 10x zoom	Sosd10m	0,8447	[tanya] ”merk”?
	[jawab] jvc	Jsd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] hdd	jhd10m	0,9150	[jawab] JVC jhd10m Hard Disk Handycam with 10x Zoom
5	[tanya]handycam 10x zoom	Sosd10m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] hdd	sohd10m	0,9150	[jawab] Sony sohd10m Hard Disk Handycam with 10x Zoom
6	[tanya]handycam 10x zoom	Sosd10m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] minidv	somd10m	0,9150	[jawab] Sony somd10m

				Minidv Handycam with 10x Zoom
7	[tanya]handycam 10x zoom	Sosd10m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] sdcard	sosd10m	0,9150	[jawab] Sony sosd10m Handycam with 10x Zoom
8	[tanya]handycam 10x zoom	Sosd10m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] dvd	sodv10m	0,9150	[jawab] Sony sodv10m DVD Handycam with 10x Zoom
9	[tanya]handycam 10x zoom	sosd10m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] hdd	chd10m	0,9150	[jawab] Canon chd10m Hard Disk Handycam with 10x Zoom
10	[tanya]handycam 10x zoom	Sosd10m	0,8447	[tanya] ”merk”?
	[jawab] Canon	csd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] minidv	cmd10m	0,9150	[jawab] Sony cmd10m Minidv Handycam with 10x Zoom
11	[tanya]handycam 10x zoom	sosd10m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] sdcard	csd10m	0,9150	[jawab] Canon csd10m Handycam with 10x Zoom
12	[tanya]handycam 10x zoom	sosd10m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd10m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] dvd	cdv10m	0,9150	[jawab] Canon cdv10m DVD Handycam with 10x Zoom
13	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] sdcard	csd20m	0,9150	[jawab] Canon csd20m Handycam with 20x Zoom

14	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] dvd	cdv20m	0,9150	[jawab] Canon cdv20m DVD Handycam with 20x Zoom
15	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] minidv	cmd20m	0,9150	[jawab] Canon cmd20m Minidv Handycam with 20x Zoom
16	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] hdd	chd20m	0,9150	[jawab] Canon chd20m Hard Disk Handycam with 20x Zoom
17	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] minidv	somd20m	0,9150	[jawab] Sony somd20m Minidv Handycam with 20x Zoom
18	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] hdd	sohd20m	0,9150	[jawab] Sony sohd20m Hard Disk Handycam with 20x Zoom
19	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] dvd	sodv20m	0,9150	[jawab] Sony sodv20m DVD Handycam with 20x Zoom
20	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] sdcard	sosd20m	0,9150	[jawab] Sony sosd20m Handycam with 20x Zoom

21	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] jvc	jsd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] minidv	jmd20m	0,9150	[jawab] JVC jmd20m Minidv Handycam with 20x Zoom
22	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] jvc	jsd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] hdd	jhd20m	0,9150	[jawab] JVC jhd20m Hard Disk Handycam with 20x Zoom
23	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] jvc	jsd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] dvd	jdvd20m	0,9150	[jawab] JVC jdvd20m DVD Handycam with 20x Zoom
24	[tanya]handycam 20x zoom	sosd20m	0,8447	[tanya] ”merk”?
	[jawab] jvc	jsd20m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] sdcad	jsd20m	0,9150	[jawab] JVC jsd20m Handycam with 20x Zoom
25	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] dvd	sodv40m	0,9150	[jawab] Sony sodv40m DVD Handycam with 40x Zoom
26	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] sdcad	sosd40m	0,9150	[jawab] Sony sosd40m Handycam with 40x Zoom
27	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] sony	sosd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] hdd	sohd40m	0,9150	[jawab] Sony sohd40m Hard Disk Handycam with 40x Zoom
28	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?

	[jawab] sony	sosd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] minidv	somd40m	0,9150	[jawab] Sony somd40m Minidv Handycam with 40x Zoom
29	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] sdcard	csd40m	0,9150	[jawab] Canon csd40m Handycam with 40x Zoom
30	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
31	[jawab] dvd	cmd40m	0,9150	[jawab] Canon cmd40m Minidv Handycam with 40x Zoom
	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
32	[jawab] dvd	cdv40m	0,9150	[jawab] Canon cdv40m DVD Handycam with 40x Zoom
	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] canon	csd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
33	[jawab] hdd	chd40m	0,9150	[jawab] Canon chd40m Hard Disk Handycam with 40x Zoom
	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] jvc	jsd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
34	[jawab] sdcard	jsd40m	0,9150	[jawab] JVC jsd40m Handycam with 40x Zoom
	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] jvc	jsd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
35	[jawab] minidv	jmd40m	0,9150	[jawab] JVC jmd40m Minidv Handycam with 40x Zoom
	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] jvc	jsd40m	0,8771	[tanya] “media”?

	[jawab] dvd	jdv40m	0,9150	[jawab] JVC jdv40m DVD Handycam with 40x Zoom
36	[tanya]handycam 40x zoom	sosd40m	0,8447	[tanya] ”merk”?
	[jawab] jvc	jsd40m	0,8771	[tanya] “media” penyimpan yang anda inginkan?
	[jawab] hdd	jhd40m	0,9150	[jawab] JVC jhd40m Hard Disk Handycam with 40x Zoom

Dari data hasil pengujian pada tabel 4.17 menunjukkan bahwa dengan memberikan input kalimat permintaan yang hanya memuat 1 nilai atribut produk, sistem dapat menemukan data target dan menjawab pertanyaan pengguna setelah meminta 2 data tambahan kepada pengguna dan melakukan inferensi berdasarkan 3 nilai atribut produk.

Rincian langkah-langkah yang dilakukan dapat dilihat pada penjelasan di bawah ini, yang dijelaskan dengan mengambil sampel sebuah kalimat permintaan sebagai berikut :

Kalimat permintaan : “saya mau membeli handycam sony”

Hasil ekstraksi kata kunci : handycam, sony

Hasil pencarian berdasarkan pencocokan kata kunci : 36 data

Data hasil pencarian selengkapnya dapat dilihat pada tabel 4.18 di bawah ini :

Tabel 4.18. Hasil pencarian berdasarkan pencocokan kata kunci pada percobaan 5

No	Data target hasil pencarian
1	Handycam sony sdcard 10x ccd
2	Handycam sony sdcard 10x cmos
3	Handycam sony sdcard 10x 3ccd
4	Handycam sony sdcard 20x ccd
5	Handycam sony sdcard 20x cmos
6	Handycam sony sdcard 20x 3ccd
7	Handycam sony sdcard 40x ccd
8	Handycam sony sdcard 40x cmos
9	Handycam sony sdcard 40x 3ccd

Dari tabel 4.18 di atas dapat dilihat bahwa pencarian berdasarkan pencocokan kata kunci tidak berhasil menemukan data target karena himpunan kata kunci yang terdapat pada kalimat permintaan terlalu sedikit, sehingga ruang lingkup data yang dihasilkan menjadi luas. Oleh karena itu proses pencarian data target harus dilanjutkan dengan melakukan inferensi menggunakan jaringan Bayes semantik. Adapun tahapan yang dilakukan pada proses inferensi menggunakan jaringan Bayes semantik dapat ditunjukkan pada langkah-langkah di bawah ini.

Langkah pertama :

Pada jaringan Bayes, semua simpul yang berupa kata kunci yang terdapat pada kalimat permintaan pengguna diberikan nilai 1 sebagai berikut :

$$P(\text{handycam}) = 1.0$$

$$P(\text{sony} \mid \text{tv, handycam}) = 1.0$$

Langkah kedua :

Menghitung probabilitas semua simpul yang terhubung dengan simpul kata kunci “handycam”, dan “sony”. Hasil pengujian selengkapnya dapat dilihat pada tabel 4.19 di bawah ini.

Tabel 4.19. Probabilitas simpul-simpul yang terhubung dengan kata kunci pada percobaan 5

Simpul S_i	Probabilitas S_i
sosd10c	0,76139
sosd10m	0,79152
sosd103	0,74331
somd10c	0,73387
somd10m	0,76440
somd103	0,71555
sohd10c	0,71246
sohd10m	0,74331
sohd103	0,69395

sodv10c	0,71246
sodv10m	0,74331
Sodv103	0,69395
Sosd20c	0,76434
sosd20m	0,79437
Sosd203	0,74632
Somd20c	0,73688
Somd20m	0,76734
Somd203	0,71860
Sohd20c	0,71552
Sohd20m	0,74632
Sohd203	0,69704
Sodv20c	0,71552
Sodv20m	0,74632
Sodv203	0,69704
Sosd40c	0,77615
Sosd40m	0,80578
Sosd403	0,75838
Somd40c	0,74893
Somd40m	0,77911
Somd403	0,73081
Sohd40c	0,72775
Sohd40m	0,75838
Sohd403	0,70938
Sodv40c	0,72775
Sodv40m	0,75838
Sodv403	0,70938

Dari tabel 4.19 di atas dapat dilihat bahwa nilai probabilitas tertinggi dimiliki oleh simpul “sosd40m” yaitu sebesar 0,80578, tetapi nilai probabilitas tersebut masih lebih rendah dari ambang batas yang telah ditetapkan yaitu sebesar

0,9. Dengan demikian, sampai dengan tahap ini sistem masih belum dapat menemukan data target yang diinginkan, dan akan dilanjutkan dengan proses pengumpulan informasi tambahan melalui jaringan semantik.

Langkah ketiga :

Sistem akan mencari informasi yang berkaitan dengan kata kunci “handycam” dan kata kunci “sony” pada jaringan semantik, yaitu berupa simpul-simpul objek yang memiliki hubungan O-A (Objek-Atribut) atau A-V (Atribut-Value). Hasil pengujian dapat di lihat pada tabel 4.20 di bawah ini.

Tabel 4.20. Simpul yang memiliki hubungan dengan objek “handycam” dan “sony” pada percobaan 5

Objek1	Relasi	Objek2
Handycam	Has-a	Merk_cam
Handycam	Has-a	Media
Handycam	Has-a	Zoom
Handycam	Has-a	Sensor
Merk_cam	Has-a	Sony

Dari tabel 4.20 di atas dapat dilihat bahwa objek “handycam” memiliki hubungan O-A dengan objek “merk_cam”, “media”, “zoom” dan “sensor”. Atau dengan kata lain dapat dijelaskan bahwa objek “handycam” memiliki empat buah atribut yaitu “merk_cam”, “media”, “zoom”, dan “sensor”. Sedangkan objek “sony” memiliki hubungan A-V dengan objek “merk_cam”, atau dengan kata lain dapat dijelaskan bahwa atribut “merk_cam” memiliki nilai data (*value*) “sony”.

Berdasarkan data-data di atas, maka sistem akan meminta informasi tambahan kepada pengguna yang berupa nilai data dari atribut “media”, “zoom” atau “sensor”. Dalam percobaan ini sistem memilih atribut “media” yang memiliki beberapa nilai data (*value*) seperti ditunjukkan pada tabel 4.21 di bawah ini.

Tabel 4.21. Nilai data (*value*) dari atribut “media” pada percobaan 5

Objek1	Relasi	Objek2
Media	Has-a	Sdcard
Media	Has-a	Minidv
Media	Has-a	Hdd
Media	Has-a	Dvd

Setelah mendapatkan nilai-nilai data (*value*) dari atribut “media”, selanjutnya sistem akan mengajukan kalimat permintaan kepada pengguna sebagai berikut :

[Sistem] : “media” penyimpanan yang diinginkan? Sdcard, Minidv,Hdd,Dvd?

[Pengguna] : sdcard

Hasil ekstraksi kata kunci tambahan : sdcard

Langkah keempat :

Setelah mendapatkan data kata kunci tambahan dari pengguna, sistem akan mengupdate simpul kata kunci “sdcard” pada jaringan Bayes, dengan memberikan nilai 1 sebagai berikut :

$$P(\text{sdcard} \mid \text{handycam}) = 1.0$$

Kemudian sistem akan menghitung ulang probabilitas semua simpul yang terhubung dengan simpul kata kunci “handycam”, “sony”, “sdcard”. Hasil pengujian selengkapnya dapat dilihat pada tabel 4.22 di bawah ini.

Tabel 4.22. Probabilitas simpul-simpul yang terhubung dengan kata kunci awal dan kata kunci tambahan pada percobaan 5

Simpul S_i	Probabilitas S_i
sosd10c	0,80420
sosd10m	0,83370
sosd103	0,78650
sosd20c	0,80706
sosd20m	0,83641

sosd203	0,78945
sosd40c	0,81850
sosd40m	0,84725
sosd403	0,80125

Dari tabel 4.22 di atas dapat dilihat bahwa nilai probabilitas tertinggi dimiliki oleh simpul “sosd40m” yaitu 0,84725, tetapi nilai probabilitas tersebut masih lebih rendah dari ambang batas yang telah ditetapkan yaitu sebesar 0,9. Dengan demikian, sampai dengan tahap ini sistem masih belum dapat menemukan data target yang diinginkan, dan akan dilanjutkan dengan proses pengumpulan informasi tambahan melalui jaringan semantik.

Langkah kelima :

Sistem akan mencari informasi yang berkaitan dengan kata kunci “handycam”, “sony”, dan “sdcard” pada jaringan semantik, yaitu berupa simpul-simpul objek yang memiliki hubungan O-A (Objek-Atribut) atau A-V (Atribut-Value). Hasil pengujian dapat di lihat pada tabel 4.23 di bawah ini.

Tabel 4.23. Simpul jaringan semantik yang memiliki hubungan dengan objek “handycam”, “sony”, dan “sdcard” pada percobaan 5

Objek1	Relasi	Objek2
Handycam	Has-a	Merk_cam
Handycam	Has-a	Media
Handycam	Has-a	Zoom
Handycam	Has-a	Sensor
Merk_cam	Has-a	Sony
Media	Has-a	sdcard

Dari tabel 4.23 di atas dapat dilihat bahwa objek “handycam” memiliki hubungan O-A dengan objek “merk_cam”, “media”, “zoom” dan “sensor”. Atau dengan kata lain dapat dijelaskan bahwa objek “handycam” memiliki empat buah

atribut yaitu “merk_cam”, “media”, “zoom”, dan “sensor”. Sedangkan objek “sony” memiliki hubungan A-V dengan objek “merk_cam”, atau dengan kata lain dapat dijelaskan bahwa atribut “merk_cam” memiliki nilai data (*value*) “sony”. Dan objek “sdcard” memiliki hubungan A-V dengan objek “media”, atau dengan kata lain dapat dijelaskan bahwa atribut “media” memiliki nilai data (*value*) “sdcard”.

Berdasarkan data-data di atas, maka sistem akan meminta informasi tambahan kepada pengguna yang berupa nilai data dari atribut “zoom” atau “sensor”. Dalam percobaan ini sistem memilih atribut “zoom” yang memiliki beberapa nilai data (*value*) seperti ditunjukkan pada tabel 4.24 di bawah ini.

Tabel 4.24. Nilai data (*value*) dari atribut “zoom” pada percobaan 5

Objek1	Relasi	Objek2
Zoom	Has-a	10x
Zoom	Has-a	20x
Zoom	Has-a	40x

Setelah mendapatkan nilai-nilai data (*value*) dari atribut “sensor”, selanjutnya sistem akan mengajukan kalimat permintaan kepada pengguna sebagai berikut :

[Sistem] : “zoom” yang diinginkan? 10x, 20x, 40x?

[Pengguna] : 40x

Hasil ekstraksi kata kunci tambahan : 40x

Langkah keenam :

Setelah mendapatkan data kata kunci tambahan dari pengguna, sistem akan mengupdate simpul kata kunci “40x” pada jaringan Bayes, dengan memberikan nilai 1 sebagai berikut :

$$P(40x | handycam) = 1.0$$

Kemudian sistem akan menghitung ulang probabilitas semua simpul yang terhubung dengan simpul kata kunci “handycam”, “sony”, “sdcard”, ”40x”. Hasil pengujian selengkapnya dapat dilihat pada tabel 4.25 di bawah ini.

Tabel 4.25. Probabilitas simpul-simpul yang terhubung dengan kata kunci awal dan kata kunci tambahan kedua pada percobaan 5

Simpul S_i	Probabilitas S_i
Sosd40c	0,8900
Sosd40m	0,9150
Sosd403	0,8750

Dari tabel 4.25 di atas dapat dilihat bahwa nilai probabilitas tertinggi dimiliki oleh simpul “sosd40m” yaitu sebesar 0,9150. Karena nilai probabilitas dari simpul tersebut telah melebihi nilai ambang batas yang ditetapkan yaitu sebesar 0,9 maka dapat dikatakan bahwa sistem telah berhasil menemukan data target yang dicari. Selanjutnya sistem akan membaca data-data lengkap dari simpul “sosd40m”, dan menampilkan informasi produk target ke pengguna sebagai berikut :

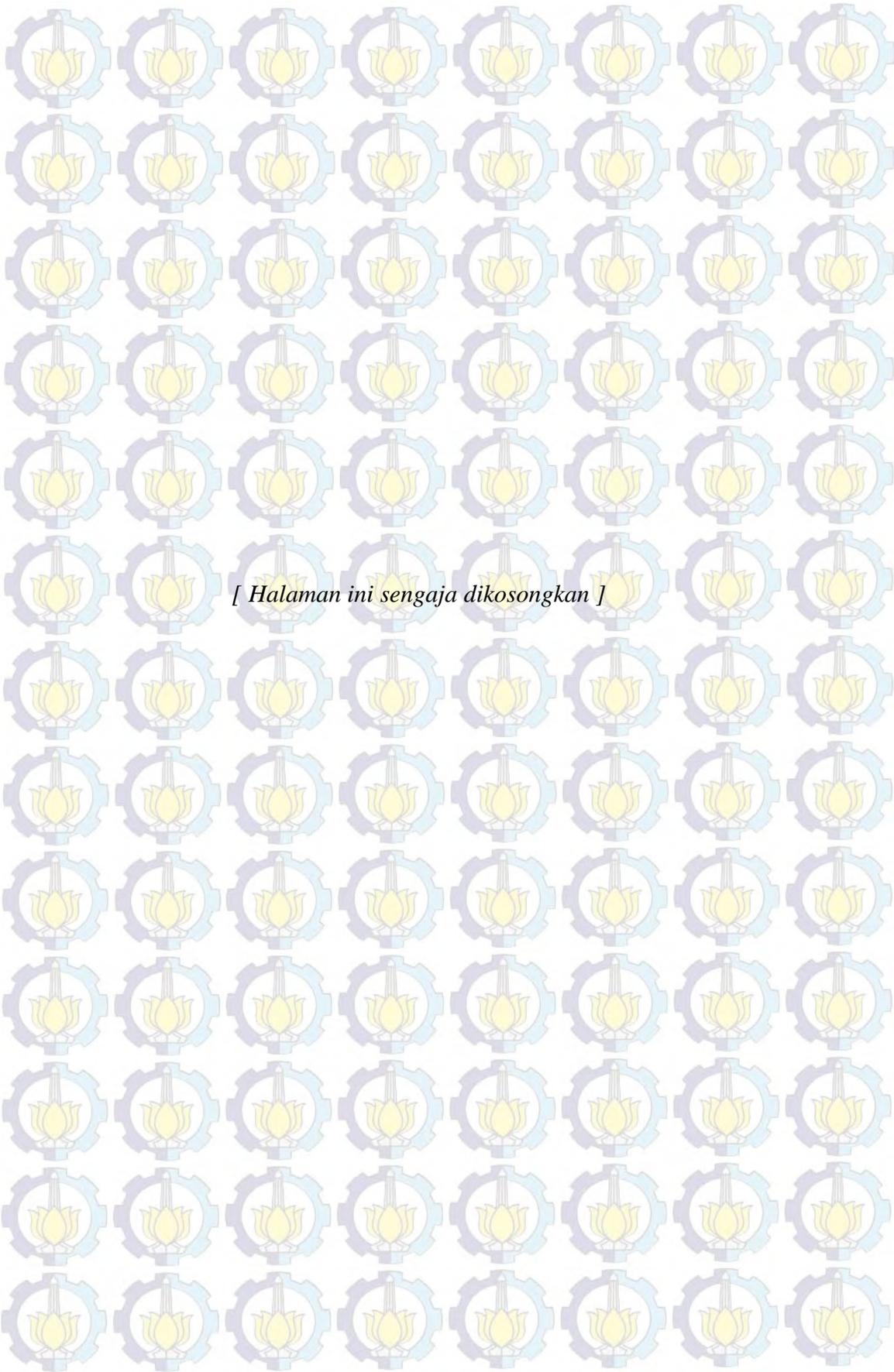
Produk target :

“Sony sosd40m Handycam with 40x Optical Zoom”

Selain data-data hasil uji coba yang telah ditunjukkan pada tabel hasil percobaan 1 hingga percobaan 5 di atas, secara keseluruhan uji coba yang telah dilakukan untuk mengukur unjuk kerja sistem dilakukan dengan memberikan masukan yang berupa kombinasi kata kunci dengan jumlah dan urutan yang berbeda-beda. Uji coba dilakukan sebanyak 910 kali dengan rincian untuk produk TV yang memiliki 3 buah atribut diuji coba sebanyak 370 kali, dan untuk produk Handycam yang memiliki 4 buah atribut diuji coba sebanyak 540 kali. Hasil selengkapnya dapat dilihat pada tabel 4.26 :

Tabel 4.26. Rekapitulasi hasil ujicoba unjuk kerja sistem dalam menemukan data target berdasarkan jumlah input kata kunci

No	Jenis Produk	Jumlah Kata Kunci yang Diinputkan	Keberhasilan menemukan data target	
			Jumlah	Prosentase
1	TV, memiliki 3 atribut	2	145	39%
		3	225	61%
2	Handycam, memiliki 4 atribut	3	216	40%
		4	324	60%



[Halaman ini sengaja dikosongkan]

BAB 5 KESIMPULAN DAN SARAN

5.1. Kesimpulan

Dari ujicoba yang telah dilakukan pada penelitian ini didapatkan beberapa kesimpulan sebagai berikut :

1. Jaringan Bayes semantik dapat digunakan untuk membangun sistem temu kembali informasi yang dapat berinteraksi dengan pengguna untuk mengumpulkan tambahan informasi yang diperlukan untuk proses inferensi.
2. Dengan menggunakan data masukan berupa kalimat permintaan (*query*) yang memuat kata kunci berupa nomor seri atau model dari sebuah produk elektronik, maka data target dapat ditemukan dengan melakukan pencarian berdasarkan pencocokan kata kunci, tanpa harus melakukan inferensi menggunakan jaringan Bayes semantik. Hal ini dikarenakan nomor seri atau model dari sebuah produk elektronik bersifat unik, artinya setiap produk memiliki nomor seri atau model yang berbeda.
3. Inferensi menggunakan jaringan Bayes dapat menemukan data target jika pada kalimat permintaan memuat nilai data atribut produk dengan jumlah yang mencukupi, yaitu :
 - a. Untuk menemukan data target produk TV yang memiliki identitas 3 buah atribut, 39% hasil uji coba membutuhkan input 2 kata kunci, dan 61% hasil ujicoba membutuhkan input 3 kata kunci.
 - b. Dan untuk menemukan data target produk Handycam yang memiliki identitas 4 buah atribut, 40% hasil uji coba membutuhkan input 3 kata kunci, dan 60% hasil ujicoba membutuhkan input 4 kata kunci.
4. Jika inferensi menggunakan jaringan Bayes belum dapat menemukan data target karena nilai data atribut yang teridentifikasi pada kalimat permintaan pengguna belum mencukupi atau terlalu sedikit, sistem secara otomatis dapat meminta tambahan informasi yang berkaitan erat dengan data target yang sedang dicari dengan menggunakan hubungan semantik objek tersebut pada jaringan semantik.

5.2. Saran dan Penelitian Selanjutnya

Berikut ini adalah beberapa saran dan rencana penelitian selanjutnya yang dapat dilakukan untuk menyempurnakan sistem yang telah dibuat :

1. Menambah banyaknya atribut yang merupakan data-data identitas yang dimiliki oleh data target, yaitu lebih dari 4 buah atribut.
2. Menyempurnakan pemilihan kata kunci yang digunakan pada sistem agar dapat memberikan hasil yang lebih optimal, yaitu kombinasi kata kunci yang memiliki hubungan penting dengan data target, dan kata kunci yang populer atau banyak digunakan oleh pengguna untuk melakukan pencarian.

BAB 5 KESIMPULAN DAN SARAN

5.1. Kesimpulan

Dari ujicoba yang telah dilakukan pada penelitian ini didapatkan beberapa kesimpulan sebagai berikut :

1. Jaringan Bayes semantik dapat digunakan untuk membangun sistem temu kembali informasi yang dapat berinteraksi dengan pengguna untuk mengumpulkan tambahan informasi yang diperlukan untuk proses inferensi.
2. Dengan menggunakan data masukan berupa kalimat permintaan (*query*) yang memuat kata kunci berupa nomor seri atau model dari sebuah produk elektronik, maka data target dapat ditemukan dengan melakukan pencarian berdasarkan pencocokan kata kunci, tanpa harus melakukan inferensi menggunakan jaringan Bayes semantik. Hal ini dikarenakan nomor seri atau model dari sebuah produk elektronik bersifat unik, artinya setiap produk memiliki nomor seri atau model yang berbeda.
3. Inferensi menggunakan jaringan Bayes dapat menemukan data target jika pada kalimat permintaan memuat nilai data atribut produk dengan jumlah yang mencukupi, yaitu :
 - a. Untuk menemukan data target produk TV yang memiliki identitas 3 buah atribut, 39% hasil uji coba membutuhkan input 2 kata kunci, dan 61% hasil ujicoba membutuhkan input 3 kata kunci.
 - b. Dan untuk menemukan data target produk Handycam yang memiliki identitas 4 buah atribut, 40% hasil uji coba membutuhkan input 3 kata kunci, dan 60% hasil ujicoba membutuhkan input 4 kata kunci.
4. Jika inferensi menggunakan jaringan Bayes belum dapat menemukan data target karena nilai data atribut yang teridentifikasi pada kalimat permintaan pengguna belum mencukupi atau terlalu sedikit, sistem secara otomatis dapat meminta tambahan informasi yang berkaitan erat dengan data target yang sedang dicari dengan menggunakan hubungan semantik objek tersebut pada jaringan semantik.

5.2. Saran dan Penelitian Selanjutnya

Berikut ini adalah beberapa saran dan rencana penelitian selanjutnya yang dapat dilakukan untuk menyempurnakan sistem yang telah dibuat :

1. Menambah banyaknya atribut yang merupakan data-data identitas yang dimiliki oleh data target, yaitu lebih dari 4 buah atribut.
2. Menyempurnakan pemilihan kata kunci yang digunakan pada sistem agar dapat memberikan hasil yang lebih optimal, yaitu kombinasi kata kunci yang memiliki hubungan penting dengan data target, dan kata kunci yang populer atau banyak digunakan oleh pengguna untuk melakukan pencarian.

DAFTAR PUSTAKA

Acid, S., de Campos, L., Ferna'ndez-Luna, J., and huete, J. (2003), "An Information Retrieval Model Based On Simple Bayesian Networks". *International Journal of Intelligent Systems*, 18(2), 251-265.

Calado, P., da Silva, A., Laender, A., Riberio-Neto, B., and Vieira, R. (2004). "A Bayesian Network Approach To Searching Web Databases Through Keyword-Based Queries". *Information Processing and Management*, 40, 773-790.

Danenberg, Stewart A. (1998), "Semantic Network Designs for Courseware", Hillsdale, New Jersey: Lawrence Erlbaum Associates Publishers, 1998

Hasugian, Jonner (2009), *Dasar-Dasar Ilmu Perpustakaan dan Informasi*, USU Press, Medan.

Hong, J.-H. And Cho, S.-B. (2003), "A Two-Stage Bayesian Network For Effective Development Of Conversational Agent", *LNCS*, Vol. 2690, pp. 1-8).

Horvitz, E., Breese, J., Heckerman, D., Hovel, D., and Rommels, K. (1998), "The Lumiere Project: Bayesian User Modeling For Inferring The Goals And Needs Of Software Users", *In Proceedings of the 14th Conference Uncertainty in Artificial Intelligence*, pp. 256-265.

Kim, K.-M., Hong, J.-H., and Cho, S.-B. (2007), "A Semantic Bayesian Network Approach to Retrieving Information With Intelligent Conversational Agents", *Information Processing and Management*, Vol. 43, No. 1, hal. 225-236.

Manning, C.D., Raghavan, P., and Schütze, H. (2008), *Introduction to Information Retrieval*, Cambridge University Press, 1, 2008.

Meadow, Charles T. (1992), *Text Information Retrieval Systems*, Academic Press, London.

Pearl, J. (1998), *Probabilistic Reasoning In Intelligent Systems: Networks Of Plausible Inference*, Morgan Kaufmann Publishers.

Perugini, S. And Ramakrishnan, N.(2003), “Personalizing Web Sites With Mixed-Initiative Interaction”, *IEEE IT Professional*, 5(2), 9-15.

Ribeiro-neto, B., and Muntz, R. (1996), “ A Belief Network Model For IR”. In *Proc. Of The 19th Annual Int. ACM SIGIR Conf. On Research And Development In Information Retrieval*, pp. 253-260, ACM Press.

Samodra, J., Sumpeno, S., Hariadi, M. (2009), “Klasifikasi Dokumen Teks Berbahasa Indonesia dengan Menggunakan *Naive Bayes*”, *Seminar Nasional Electrical, Informatic and It's Education (SNEIE) 2009*, Malang, Indonesia.

Tuttle, H. And Croft, W. (1990), “Inference Networks For Document Retrieval”. In *Proc. Of The 13th Int. Conf. On Research And Development In Information Retrieval*, pp. 1-24.

BIOGRAFI PENULIS



Joko Samodra, lahir di kota Malang pada tanggal 12 Januari 1973. Penulis menamatkan pendidikan Strata 1 di Institut Sains dan Teknologi Palapa Malang pada tahun 1996 dengan gelar Sarjana Komputer. Saat ini penulis sedang berusaha menamatkan pendidikan Strata 2 di Institut Teknologi Sepuluh Nopember (ITS) Surabaya di Jurusan Teknik Elektro bidang keahlian

Jaringan Cerdas Multimedia konsentrasi Teknologi Permainan. Penulis bertempat tinggal di Jl.Saxofone No. 48 Malang, dan bekerja sebagai tenaga pengajar di Progam Studi Animasi Jurusan Seni dan Desain Universitas Negeri Malang.