



UNIVERSITAS INDONESIA

**PREDIKSI PREFERENSI PELANGGAN WARALABA
MAKANAN CEPAT SAJI DENGAN MENGGUNAKAN
PENDEKATAN DATA MINING**

SKRIPSI

**VINNY
0706275145**

**FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INDUSTRI
DEPOK
JUNI 2011**



UNIVERSITAS INDONESIA

**PREDIKSI PREFERENSI PELANGGAN WARALABA
MAKANAN CEPAT SAJI DENGAN MENGGUNAKAN
PENDEKATAN DATA MINING**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar
Sarjana Teknik**

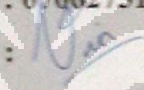
**VINNY
0706275145**

**FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INDUSTRI
DEPOK
JUNI 2011**

..

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri,
dan semua sumber baik yang dikutip maupun dirujuk
telah saya nyatakan dengan benar.

Nama : Vinny
NPM : 0706275145
Tanda Tangan : 
Tanggal : 9 Juni 2011

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :
Nama : Vinny
NPM : 0706275145
Program Studi : Teknik Industri
Judul Skripsi : Prediksi Preferensi Pelanggan Waralaba Makanan
Cepat Saji dengan Menggunakan Pendekatan Data
Mining

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Teknik pada Program Studi Teknik Industri, Fakultas Teknik, Universitas Indonesia

DEWAN PENGUJI

Pembimbing : Ir. Isti Surjandari Prajitno MT., MA., Ph.D. ()
Penguji : Ir. Erlinda Muslim, MEE. ()
Penguji : Arian Dhini S.T., M.T. ()
Penguji : Ir. Amar Rachman MEIM. ()

Ditetapkan di : Depok
Tanggal : 16 Juni 2011

UCAPAN TERIMA KASIH

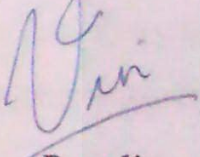
Puji syukur saya panjatkan kepada Sang Triratna Yang Maha Suci, karena berkah dan perlindungan-Nya, penulis bisa menyelesaikan skripsi ini tepat waktu dan tanpa kendala. Penulisan skripsi ini ditujukan untuk memenuhi salah satu syarat mendapatkan gelar Sarjana Teknik Departemen Teknik Industri pada Fakultas Teknik Universitas Indonesia. Penulis sangat menyadari bahwa tanpa adanya bimbingan dan bantuan dari berbagai pihak, akan sangat sulit bagi penulis untuk dapat menyelesaikan skripsi ini. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Ibu Ir. Isti Surjandari, Ph.D. selaku dosen pembimbing skripsi tersayang, yang selalu membimbing, mengarahkan, memotivasi, menyarankan, memperbaiki, dan membantu setiap langkah penyusunan skripsi ini.
2. Pak Noragraito, selaku GM Papa Ron's Pizza, Mbak Valen, dan Mas Egi yang telah memberikan kesempatan kepada penulis untuk melakukan penelitian di Papa Ron's Pizza, serta membantu proses penyebaran kuesioner ke gerai-gerai yang tersebar di JABODETABEK.
3. Stephen, sahabat sekaligus rekan kerja yang telah sangat memberikan perhatian serta membantu proses pengolahan data penelitian ini.
4. Dimas Adityamurthi dan Harry Bian Pramudia, selaku teman yang juga turut membantu peninjauan literatur serta proses pengolahan data penelitian ini.
5. Radita Tanaya, Devyana Fitri, dan Sartika Chandra, selaku teman yang membantu proses pengeditan kuesioner penelitian.
6. Teman-teman satu bimbingan Bu Isti : Dyah, Deddy, Adhi, Gina, Tama, dan Ratna, atas dukungan dan kerjasamanya pada proses pengerjaan skripsi ini.
7. Christine, Wilson Chandra, Vinna Christina, Fenny, Martina, Meilinda, Srijanti Chandra, dan Oscar Hammer Stein yang telah menjadi sahabat terbaik serta selalu menemani penulis menghadapi masa-masa sulit selama ini.
8. Teman-teman KMBUI BPH XVIII: Charles Tirta, Khanti Paramita, Yulianti, Prissilia, Lita Viani, Berry, Dini Lestari, Sandi Tannu Wijaya, dan teman-teman KMBUI angkatan 2007 lainnya yang telah menjadi sahabat, mendukung, serta telah 'membentuk' penulis menjadi orang yang lebih baik.

9. Cheryl, Ocha, Deborah, Melati, Monasisca, dan teman-teman TIUI angkatan 2007 atas persahabatan dan kerjasamanya yang luar biasa selama 4 tahun yang sangat mengesankan ini.
10. Keluarga besar tercinta, Mama, Papa, Eric Winata Gunawan, dan Viviyanti atas kasih sayang yang telah diberikan kepada penulis selama ini serta dukungan terhadap proses penyelesaian skripsi ini.
11. Terakhir, pihak lain yang tak bisa disebutkan namanya satu per satu di sini.

Akhir kata, penulis menyadari bahwa tulisan ini masih sangat jauh dari sempurna. Akan tetapi, penulis berharap agar kehadiran skripsi ini bisa bermanfaat dan berkontribusi bagi pengembangan ilmu pengetahuan.

Depok, 9 Juni 2011


Penulis

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Indonesia, saya yang bertanda tangan di bawah ini:

Nama : Vinny
NPM : 0706275145
Program Studi : Teknik Industri
Departemen : Teknik Industri
Fakultas : Teknik
Jenis karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul:

**Prediksi Preferensi Pelanggan Waralaba Makanan Cepat Saji dengan
Menggunakan Pendekatan Data Mining**

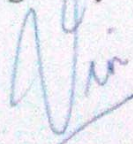
Berserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok

Pada tanggal : 9 Juni 2010

Yang menyatakan

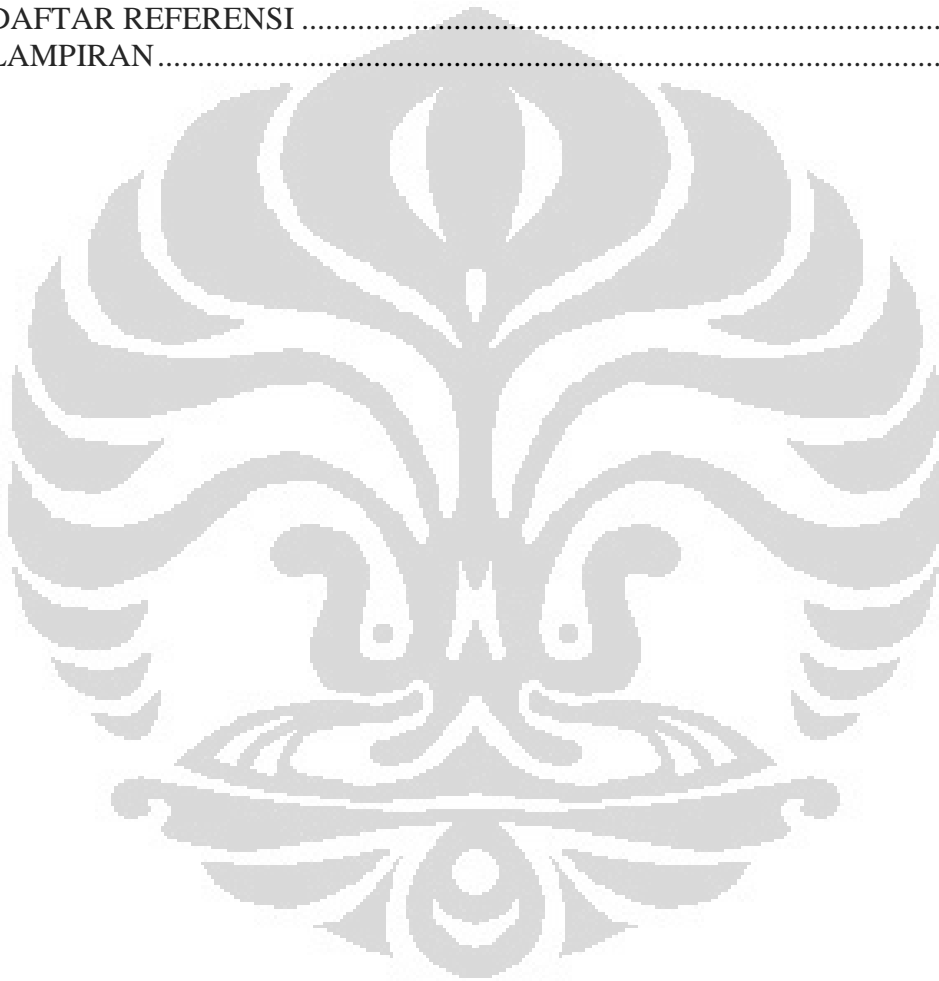


(Vinny)

DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS.....	ii
HALAMAN PENGESAHAN.....	iii
UCAPAN TERIMA KASIH.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	vi
ABSTRAK	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
DAFTAR LAMPIRAN.....	xiii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Diagram Keterkaitan Masalah	4
1.3 Perumusan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Batasan Penelitian.....	5
1.6 Metodologi Penelitian.....	5
1.7 Diagram Alir Metodologi Penelitian	7
1.8 Sistematika Penulisan	8
BAB 2 TINJAUAN LITERATUR.....	9
2.1 Jaringan Saraf Manusia.....	9
2.1.1 Neuron	10
2.1.2 Komunikasi antar Neuron.....	11
2.2 Transformasi Jaringan Saraf Manusia ke Jaringan Saraf Tiruan	13
2.3 Neural Network.....	14
2.4 Fungsi Aktivasi Sigmoid.....	21
2.5 Backpropagation	21
2.6 Gradient Descent Method	22
2.7 Backpropagation Rules	24
2.8 Contoh Backpropagation	24
2.9 Termination Criteria.....	26
2.10 Learning Rate.....	27
2.11 Momentum Term	28
2.12 Analisis Klaster.....	30
2.13 Analisis Klaster K-Means	31
2.14 Perkembangan Industri Fast Food	32
BAB 3 PENGUMPULAN DAN PENGOLAHAN DATA	34
3.1 Pengumpulan data.....	34
3.2 Penentuan Kriteria dan Atribut	35
3.3 Pengolahan Data	38
3.3.1 <i>K-Means Cluster Analysis</i>	38
3.3.2 <i>Neural Network Backpropagation</i>	40

3.4 Hasil Pengolahan Data.....	41
3.5 Hasil Data Demografi	42
BAB 4 ANALISIS.....	46
4.1 Analisis Neural Network.....	46
4.2 Analisis Data Demografi.....	57
BAB 5 KESIMPULAN DAN SARAN.....	59
5.1 Kesimpulan.....	59
5.2 Saran.....	60
DAFTAR REFERENSI	61
LAMPIRAN.....	62

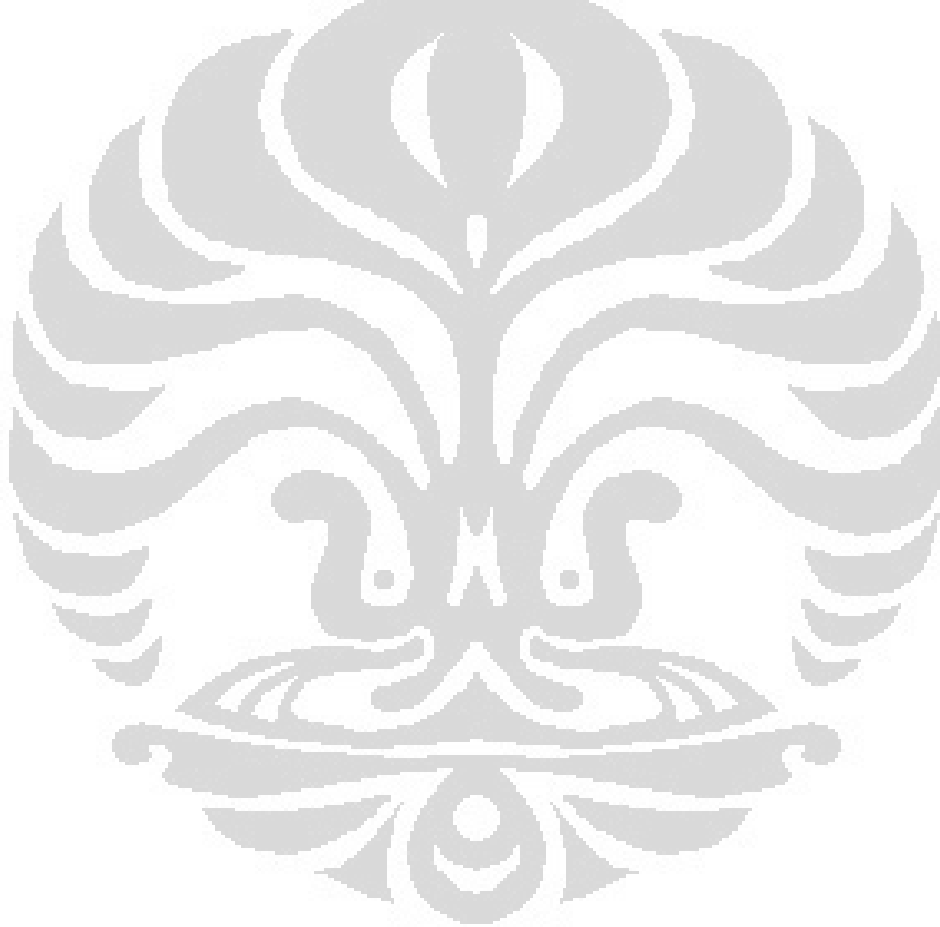


DAFTAR GAMBAR

Gambar 1.1	<i>Market Share Pizza</i> di Indonesia.....	3
Gambar 1.2	Diagram Keterkaitan Masalah	4
Gambar 1.3	Diagram Alir Metodologi Penelitian	7
Gambar 2.1	Struktur <i>Neuron</i> Vertebrata	10
Gambar 2.2	Sinapsis Kimiawi pada Saraf Manusia.....	12
Gambar 2.3	Analogi Peniruan Sistem Saraf Manusia.....	14
Gambar 2.4	Saraf Manusia dan <i>Artificial Neuron Model</i>	16
Gambar 2.5	<i>Feedforward Neural Network</i>	17
Gambar 2.6	<i>Recurrent Neural Network</i>	18
Gambar 2.7	Sebuah <i>Neural Network</i> Sederhana	18
Gambar 2.8	Grafik Fungsi <i>Sigmoid</i>	21
Gambar 2.9	Kurva SSE terhadap W_1	23
Gambar 2.10	Nilai η Akan Menghasilkan Algoritma Melampaui <i>Global Minimum</i>	27
Gambar 2.11	Nilai Momentum α Kecil.	29
Gambar 2.12	Nilai Momentum α Besar	29
Gambar 2.13	Pengelompokan pada <i>Cluster Analysis</i>	31
Gambar 3.1	Tampilan Layar Utama Matlab	40
Gambar 3.2	Tampilan Layar Editor Matlab	40
Gambar 3.3	Data Demografi Pekerjaan Responden.....	42
Gambar 3.4	Data Demografi Lokasi Tempat Tinggal	43
Gambar 3.5	Data Demografi Jenis Kelamin	43
Gambar 3.6	Data Demografi Pendapatan.....	44
Gambar 3.7	Data Demografi Frekuensi Makan Di Restoran	44
Gambar 3.8	Data Demografi Rata-Rata Pengeluaran Untuk Setiap Kali Makan di Restoran.....	45
Gambar 4.1	Recognition Rate Data Test.....	46
Gambar 4.2	Waktu <i>Training</i>	46
Gambar 4.3	Waktu <i>Testing</i>	47
Gambar 4.4	Representasi Grafik <i>Z-Score</i> Masing-Masing Kelas terhadap Atribut Yang Diberikan.....	50

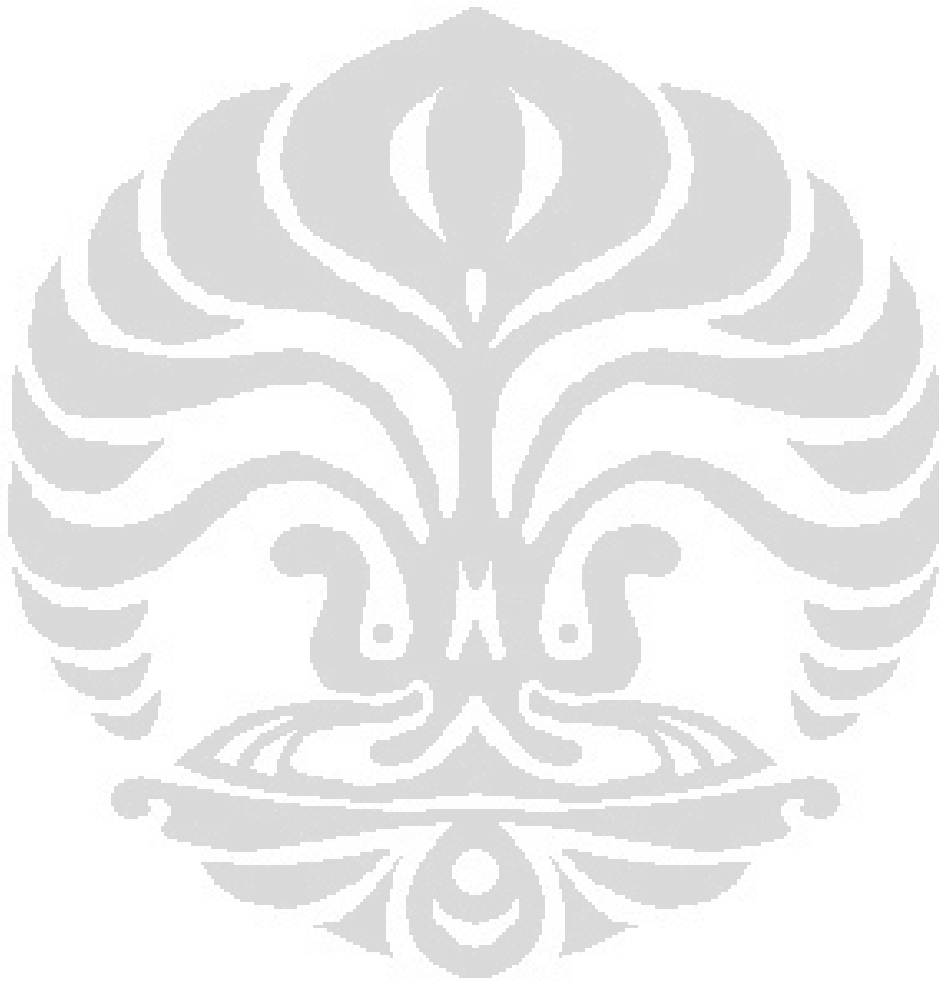
DAFTAR TABEL

Tabel 2.1	Perbandingan Saraf Manusia dengan <i>Neural Network</i>	14
Tabel 2.2	<i>Data Inputs Dan Initial Values</i> untuk Beban <i>Neural Networks</i>	19
Tabel 3.1	Analisis Reabilitas Cronbach's Alpha	38
Tabel 3.2	Data Deskriptif Standarisasi	39
Tabel 3.3	Hasil <i>Recognition Rate</i> dan <i>Time</i>	41
Tabel 4.1	<i>Final Cluster Centers</i>	47
Tabel 4.2	Peringkat Atribut yang Dipilih Oleh Kelas 1.....	52
Tabel 4.3	Peringkat Atribut yang Dipilih Oleh Kelas 2.....	53
Tabel 4.4	Peringkat Atribut yang Dipilih Oleh Kelas 3.....	55



DAFTAR LAMPIRAN

Lampiran A	<i>Syntax</i> Matlab Pengambilan Data	62
Lampiran B	<i>Syntax</i> Matlab <i>Backpropagation</i>	63
Lampiran C	<i>Syntax</i> Matlab Pengolahan Data	64
Lampiran D	Tahapan Standarisasi Data.....	65
Lampiran E	Tahapan K-Means Cluster Analysis	67
Lampiran F	Proses Pembuatan Neural Network dan Pengolahan Data	70
Lampiran G	Data Primer	74
Lampiran H	Kuesioner	79



ABSTRAK

Nama : Vinny
Program Studi : Teknik Industri
Judul : Prediksi Preferensi Pelanggan Waralaba Makanan Cepat Saji dengan Menggunakan Pendekatan Data Mining

Masuknya waralaba makanan cepat saji asing, khususnya *pizza*, yang menawarkan kualitas yang baik dengan harga yang lebih murah, membuat waralaba domestik kalah bersaing. Oleh karena itu, penelitian ini bertujuan untuk mengetahui preferensi konsumen dan atribut yang paling mempengaruhi, akan kualitas salah satu waralaba domestik di Indonesia dengan pendekatan *data mining*. Adapun metode yang digunakan dalam penelitian ini adalah analisis *cluster* dan *neural network backpropagation*. Hasil penelitian menunjukkan bahwa *neural network* lebih cepat menghafalkan karakter-karakter responden yang dibagi dalam 3 kelas, dimana kelas 1 lebih memperhatikan makanan sehat, kenyamanan, dan pelayanan restoran, serta memiliki anggota yang paling banyak. Selain itu, ditemukan 15 atribut utama yang mendapatkan nilai preferensi yang baik dari responden-responden di kelas 1 ini. Dengan demikian, perusahaan bisa menyusun strategi pemasaran baru dalam rangka merebut pangsa pasar yang lebih besar.

Kata Kunci:

Neural network, cluster analysis, backpropagation, pizza, waralaba.

ABSTRACT

Name : Vinny
Study Program : Industrial Engineering
Judul : Predicting Consumer Preference of Fast Food Franchise:
A Data Mining Approach.

The entry of multinational fast food franchise, especially pizza franchise, which offers great quality with better price, made our domestic franchise lost the competition. Thus, this research was conducted to find out consumer preference and the most influential attribute of a domestic fast food franchise by using *data mining*. The methods used in the research were cluster analysis and neural network backpropagation. The result showed that neural network could be tested faster if the respondents were divided into 3 classes whose first class was filled with the people cared about healthy food, comfort, service, and had the most members. Besides, 15 main attributes gained better preference value from this class. With the last result, the company can build a new marketing strategy in order to grab a better market share.

Keywords:

Neural network, cluster analysis, backpropagation, pizza, franchise.

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Semakin meluasnya jangkauan dan jenis media informasi dalam satu dekade terakhir, membuat masyarakat bisa mendapatkan informasi dengan mudah. Informasi bukanlah ‘barang mewah’ lagi bagi masyarakat. Koran, majalah,

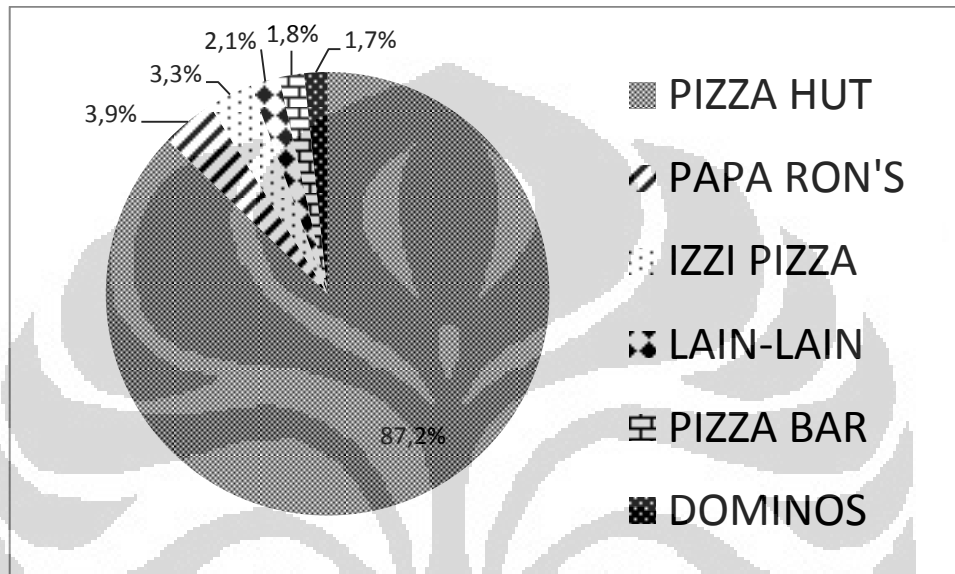
perusahaan dan mengarah pada perbaikan-perbaikan kontinu dalam produk atau proses (Surjandari, 2010, p. 1-2). Inovasi ini tentunya bukanlah hal sederhana yang bisa dilakukan tanpa aturan. Perusahaan harus terus mengembangkan kreativitasnya dalam rangka menyediakan jasa ataupun produk yang nantinya akan ditawarkan.

Adapun kreativitas inovasi ini bisa didasarkan pada intuisi perusahaan (*technology push*) ataupun membaca kebutuhan pasar (*market pull*). Perdebatan antara kelebihan dan kekurangan dari dua metode ini telah dimulai sejak tahun 1934 (Schumpeter, 1934 di dalam Astebro dan Dahlin, 2003). Akan tetapi, terdapat fakta yang menunjukkan bahwa peningkatan permintaan pasar sebenarnya disebabkan oleh permintaan konsumen yang semakin meningkat (Schmookler, 1962 di dalam Astebro dan Dahlin, 2003). Hasilnya, kebutuhan konsumen adalah *item* terpenting dalam sebuah proses inovasi (Freeman, 1982 di dalam Astebro dan Dahlin, 2003). Selain biaya yang lebih murah, perusahaan tidak harus lagi ‘meraba-raba’ kebutuhan pasar dan mengedukasi fungsi produk ke pasar nantinya.

, dan Fat Burger adalah beberapa contoh *brand* terkenal yang hidup saat ini. Kesuksesan usaha waralaba di atas, membuat para pelaku bisnis domestik ‘tergiur’ untuk turut ambil bagian dan Papa Ron’s Pizza adalah satu di antaranya.

Sejak tahun 2000, Papa Ron’s Pizza yang merupakan perusahaan waralaba makanan cepat saji domestik, telah memiliki 43 gerai yang tersebar di seluruh Indonesia. Tujuh belas gerai di antaranya berada di JABODETABEK (Jakarta, Bogor, Depok, Tangerang, dan Bekasi) dan sisanya berada di luar kota. Anang Sukandar, Ketua Asosiasi Franchise Indonesia (AFI), mengungkapkan bahwa

Papa Ron's Pizza merupakan restoran nasional yang mampu bersaing di pasar internasional. Akan tetapi, data menunjukkan bahwa *market share* yang didapatkan oleh perusahaan ini jauh lebih kecil dibandingkan dengan pesaing utama. Oleh karena itu, dibutuhkan penelitian lebih lanjut akan preferensi pasar dalam rangka mendongkrak persentase *market share* yang masih $\frac{1}{22}$ kali dari *market share* PIZZA HUT.



Gambar 1.1: Market Share Pizza di Indonesia

(Sumber: Media Referensi Indonesia,2009)

Selain sejumlah alasan di atas, proses penelitian kebutuhan pasar saat ini telah berkembang ke arah penggunaan *multiatribut* yang lebih bisa membaca kebutuhan konsumen yang banyak (Hayashi, Hsieh, & Setiono, 2009). Berbeda dengan *conjoint analysis* yang lebih direkomendasikan untuk penggunaan *atribut* yang sedikit dengan level yang bisa dikombinasikan dengan mudah (Helm, Manthey,

1.2 Diagram Keterkaitan Masalah

Adapun hubungan keterkaitan masalah yang dihadapi oleh Papa Ron's Pizza adalah sebagai berikut:

Gambar 1.2 Diagram Keterkaitan Masalah

1.3 Perumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan sebelumnya, masalah utama yang diangkat adalah minimnya pengetahuan akan preferensi masyarakat, terhadap waralaba makanan cepat saji di Indonesia, terutama *pizza*. Oleh karena itu, dibutuhkan penelitian preferensi ini dengan menggunakan metode *data mining* yang memungkinkan penyajian atribut yang banyak dan tetap *userfriendly*.

1.4 Tujuan Penelitian

Adapun tujuan penelitian yang ingin dicapai adalah mengetahui preferensi konsumen dan atribut yang paling mempengaruhi kualitas salah satu perusahaan waralaba makanan cepat saji di Indonesia.

1.5 Batasan Penelitian

Batasan dari penelitian ini adalah:

- Objek penelitian adalah Papa Ron's Pizza yang merupakan salah satu usaha waralaba di Indonesia yang telah memiliki sistem yang sudah stabil.
- Data yang diambil dalam penelitian ini adalah data primer yang didapatkan melalui penyebaran kuesioner kepada pelanggan Papa Ron's Pizza.
- Survey akan dilakukan pada cabang-cabang Papa Ron's Pizza yang tersebar di JABODETABEK.
- Responden berasal dari berbagai kalangan usia, jenis kelamin, dan pekerjaan, khususnya konsumen yang sedang mengunjungi gerai saat itu.
- Hasil dari penelitian ini akan diolah dengan menggunakan *tools neural network*, khususnya *backpropagation* dan *cluster analysis*.

1.6 Metodologi Penelitian

Penelitian ini dilakukan dengan berdasarkan pada tiga tahapan utama yang terbagi menjadi:

- Tahap Studi Pendahuluan

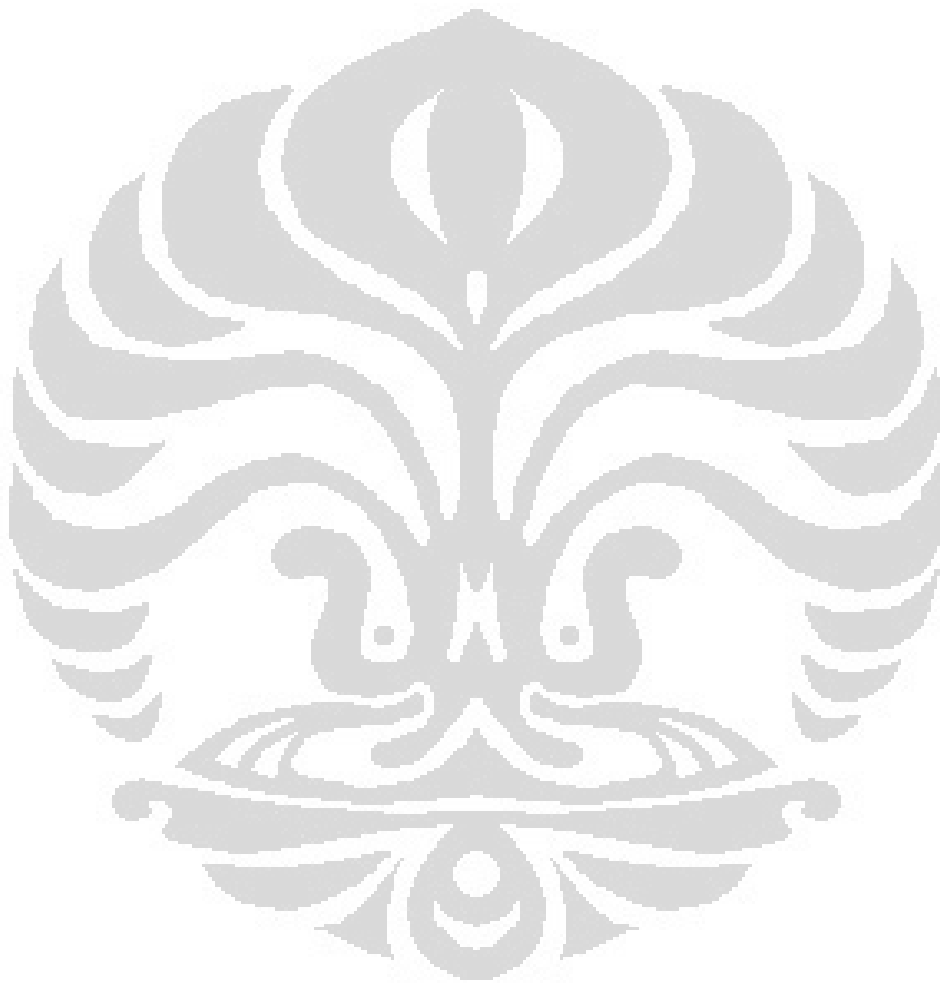
Pada bagian ini, terdapat sejumlah subtahapan yang harus dilalui, di antaranya:

- Merumuskan dan mendefinisikan masalah.
- Melakukan studi literatur pendukung penelitian.
- Menentukan data-data primer yang harus diambil.
- Merancang kuesioner penelitian.
- Tahap Pengambilan Data
Pada bagian ini, penelitian akan melalui tahapan seperti di bawah:
 - Menyebarkan kuesioner.
 - Merekap kuesioner yang telah diisi responden.
 - Memilah data menjadi *training set* dan *validation set*.
- Tahap Pengolahan Data
Pada bagian ini, hasil data yang diambil akan diolah dengan metode *neural network*, khususnya *backpropagation* dan *cluster analysis*.
- Tahap Penganalisisan
Pada bagian ini, hasil pengolahan data akan:
 - Dianalisis lebih jauh.
 - Disimpulkan sehingga tujuan dari penelitian ini bisa tercapai.

1.7 Diagram Alir Metodologi Penelitian

Berikut adalah diagram alir metodologi penelitian yang akan dilakukan:

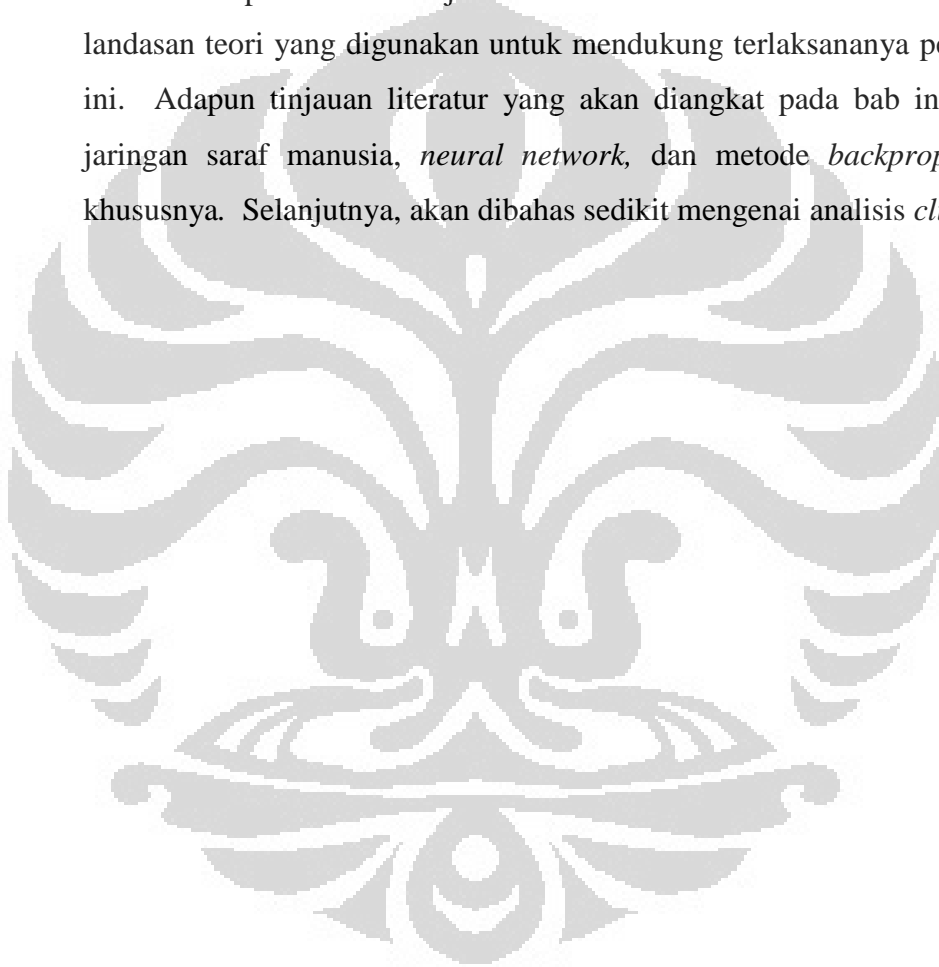
Gambar 1.3 Diagram Alir Metodologi Penelitian



1.8 Sistematika Penulisan

Penelitian ini dirangkum ke dalam suatu bentuk penulisan sistematis, yang terdiri atas:

- Bab 1 merupakan Bab Pendahuluan. Bab ini berisikan penjelasan mengenai latar belakang permasalahan, perumusan masalah, tujuan penelitian, batasan penelitian, metodologi penelitian, serta sistematika penulisan.
- Bab 2 merupakan Bab Tinjauan Literatur. Bab ini berisikan landasan-landasan teori yang digunakan untuk mendukung terlaksananya penelitian ini. Adapun tinjauan literatur yang akan diangkat pada bab ini adalah jaringan saraf manusia, *neural network*, dan metode *backpropagation* khususnya. Selanjutnya, akan dibahas sedikit mengenai analisis *cluster*



BAB 2 TINJAUAN LITERATUR

2.1 Jaringan Saraf Manusia

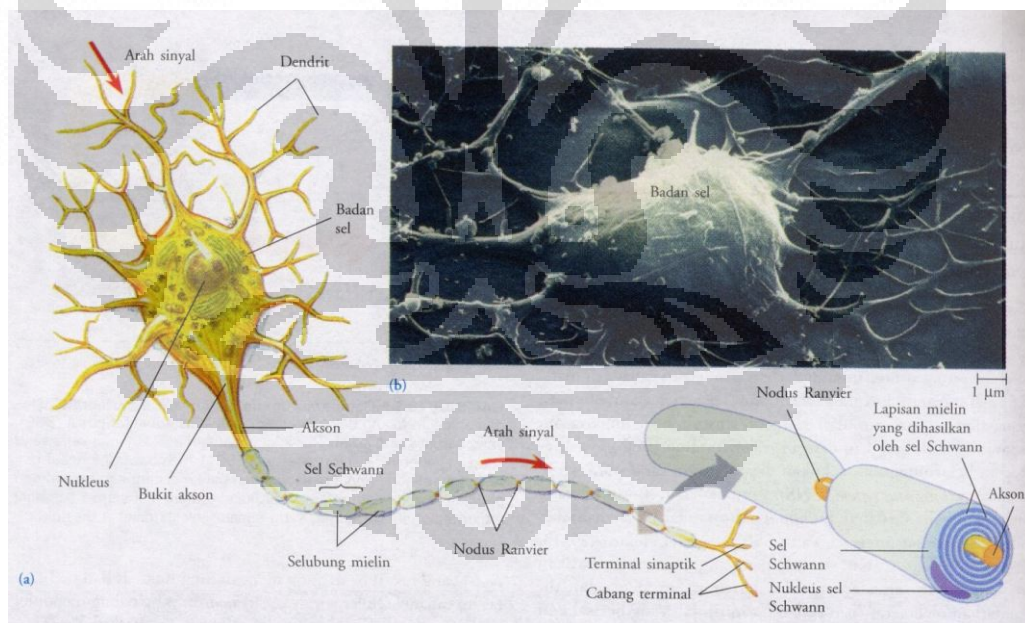
Microprocessor yang ada di komputer adalah satu di antara sekian banyak produk yang mengimplementasikan cara kerja otak. *Microprocessor* terbuat dari *chip* komputer rumit yang membuatnya bisa melakukan tugas yang kompleks. Akan tetapi, sistem saraf yang tersusun atas *neuron* hidup, melakukan tugas yang jauh lebih banyak dan lebih kompleks ketika membaca atau memahami kata-kata dari tulisan ini. Sistem saraf manusia bisa jadi merupakan kumpulan materi yang terorganisasi paling rumit di dunia. Satu sentimeter kubik otak manusia bisa mengandung lebih dari 50 juta sel saraf, yang masing-masing bisa berkomunikasi dengan ribuan *neuron* lain dalam jaringan kerja pengolahan informasi. Hal ini membuat komputer yang paling rumit di dunia sekalipun terlihat primitif. Jalur-jalur saraf ini mengontrol setiap persepsi dan pergerakan kita dan membuat kita mampu belajar, mengingat, berpikir, dan menjadi sadar akan diri kita sendiri dan lingkungan sekitar kita (Campbell et al, 2004).

Secara umum, sistem saraf mempunyai tiga fungsi yang saling tumpang tindih: *input* sensoris, integrasi, dan *output* motoris. *Input* adalah penghantaran atau konduksi sinyal dari reseptor sensoris, misalnya sel-sel pendeteksi cahaya di mata, ke pusat integrasi. Integrasi adalah proses penerjemahan informasi yang berasal dari stimulus reseptor sensoris oleh lingkungan, kemudian dihubungkan dengan respons tubuh yang sesuai. Sebagian besar integrasi dilakukan dalam sistem saraf pusat (SSP atau *central nervous system*, CNS), yaitu otak dan sumsum tulang belakang (pada vertebrata). *Output* motoris adalah penghantaran sinyal dari pusat integrasi, yaitu SSP, ke sel-sel efektor (*effector cells*), sel-sel otot atau sel kelenjar yang mengaktualisasikan respons tubuh terhadap stimulus tersebut. Sinyal tersebut dihantarkan oleh saraf (*nerve*), berkas mirip tali yang berasal dari penjurulan *neuron* yang terbungkus dengan ketat dalam jaringan ikat. Saraf yang menghubungkan sinyal motoris dan sensoris antara sistem saraf pusat dan bagian tubuh lain secara bersamaan disebut sebagai sistem saraf tepi (SST atau *peripheral nervous system*, PNS). Dari reseptor ke efektor, informasi

dikomunikasikan dalam satu saraf dari satu *neuron* ke *neuron* berikutnya melalui kombinasi sinyal listrik dan sinyal kimiawi. (Campbell et al, 2004)

2.1.1 Neuron

Neuron adalah unit fungsional sistem saraf yang dikhususkan untuk menghantarkan dan mengirimkan sinyal dalam tubuh dari satu lokasi ke lokasi yang lain. Meskipun terdapat banyak jenis *neuron* yang berbeda dalam hal struktur dan fungsinya, sebagian besar *neuron* mempunyai beberapa ciri yang sama. Sebuah *neuron* mempunyai badan sel (*cell body*) yang relatif lebih besar dan mengandung nukleus dan berbagai ragam organel seluler lainnya. Ciri *neuron* yang paling menonjol adalah penjurulan yang mirip serat, yang disebut *prosesus*, sehingga sel mampu mencapai jarak yang jauh untuk menghantarkan pesan. Ada dua jenis penjurulan *neural* yang umum: *dendrit*, yang mengirimkan sinyal dari ujungnya ke seluruh bagian *neuron*; dan *akson*, yang menghantarkan pesan ke ujung *neuron*.



Gambar 2.1 Struktur *Neuron* Vertebrata

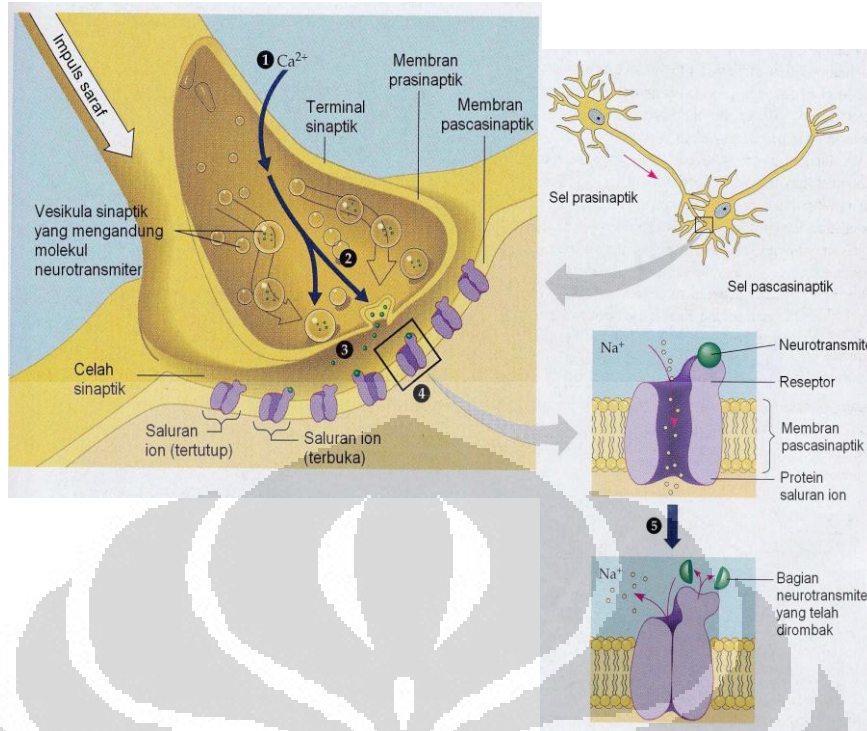
(Sumber: Campbell et al, 2004)

Dendrit pada kebanyakan *neuron*, seperti *neuron* pada Gambar 2.1, mempunyai banyak percabangan (nama itu diturunkan dari Bahasa Yunani *dendron* yang berarti pohon). Dengan demikian, *dendrit* adalah adaptasi

struktural yang meningkatkan luas permukaan *neuron*, tempat *neuron* itu menerima *input* dari *neuron* lain atau reseptor sensoris. Banyak *neuron* mempunyai sebuah akson tunggal, yang mungkin sangat panjang. Pada kenyataannya, saraf psikiatik pada kaki, mempunyai akson yang memanjang dari bagian bawah sumsum tulang belakang sampai ke otot betis dan telapak kaki, suatu jarak yang bisa sampai satu meter atau lebih. Bukit akson (*axon hillock*) adalah daerah pada badan sel tempat akson bercabang. Pada daerah ini impuls dihantarkan ke akson umumnya dibangkitkan. Banyak akson dalam sistem saraf vertebrata terbungkus oleh lapisan insulasi yang disebut selubung mielin (*Myelin sheath*), dan dibentuk oleh sel-sel pendukung. Pada sistem saraf tepi, sel-sel pendukung disebut sel Schwann (*Schwann cell*) yang menyusun selubung mielin, sedangkan dalam sistem saraf pusat, sel-sel pendukung yang disebut *oligodendrosit* menghasilkan selubung mielin. Akson bisa bercabang, dan masing-masing cabang bisa mencapai ratusan hingga ribuan ujung-ujung khusus yang disebut terminal sinaptik (*synaptic terminal*), yang mengirimkan sinyal ke sel lain dengan melepaskan *messenger* kimiawi yang disebut *neurotransmitter*. Lokasi kontak antara terminal sinaptik dan sel target (baik *neuron* lain ataupun sel efektor, misalnya sel otot) disebut sinapsis (*synapse*). Dengan demikian, sinapsis adalah persambungan di mana satu *neuron* berkomunikasi dengan *neuron* lain dalam satu jalur neural, atau di mana sebuah *neuron* berkomunikasi dengan sebuah sel otot atau sel kelenjar (Campbell et al, 2004).

2.1.2 Komunikasi antar Neuron

Sinapsis adalah persambungan unik yang mengontrol komunikasi antara satu *neuron* dengan sel-sel lain. Sinapsis ditemukan di antara dua *neuron*, yaitu antara reseptor sensoris dengan *neuron* sensoris, antara *neuron* motoris dengan sel otot yang dikontrolnya, dan antara *neuron* dengan sel kelenjar. Sel yang menghantarkan sinyal itu disebut sel prasinaptik (*presynaptic cell*), dan sel yang menerimanya disebut sel pascasinaptik (*postsynaptic cell*).



Gambar 2.2 Sinapsis Kimiawi pada Saraf Manusia

(Sumber: Campbell et al, 2004)

Sinapsis terdiri atas dua jenis, di antaranya:

1. Sinapsis Listrik

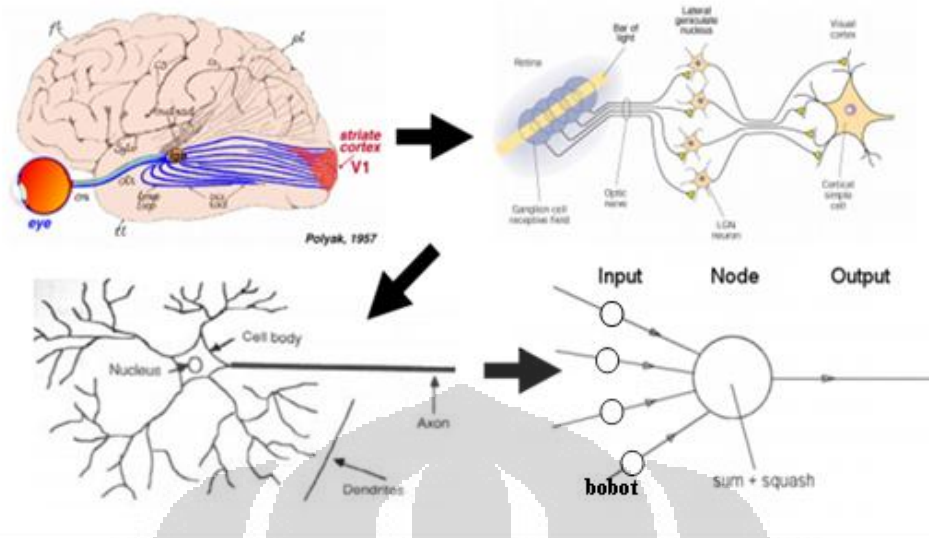
Sinapsis listrik memungkinkan potensial aksi merambat secara langsung dari satu sel prasinaptik ke sel pascasinaptik. Sel-sel itu dihubungkan dengan persambungan longgar, yaitu saluran antarsel yang mengalirkan ion potensial aksi lokal agar mengalir antar-*neuron*. Ujung akson raksasa pada udang dan *crustacea* lain biasanya saling berhubungan dan dikopel dengan sinapsis listrik. Hal ini memungkinkan impuls merambat dari satu *neuron* ke *neuron* lain tanpa penundaan dan tanpa kehilangan kekuatan sinyal. Sinapsis listrik dalam sistem saraf pusat vertebrata menyelaraskan aktivitas *neuron* yang bertanggung jawab atas sejumlah pergerakan cepat dan khas. Contohnya, sinapsis listrik pada otak yang membuat beberapa jenis ikan mampu mengibaskan ekornya dengan sangat cepat ketika melarikan diri dari pemangsa. Akan tetapi, sinapsis kimiawi jauh lebih umum dibandingkan dengan sinapsis listrik pada vertebrata dan sebagian besar invertebrata.

2. Sinapsis Kimiawi

Pada sinapsis kimiawi, sebuah celah sempit, atau celah sinaptik, memisahkan sel prasinaptik dari sel pascasinaptik. Adanya celah tersebut menyebabkan sel-sel tidak dapat dihubungkan secara elektrik, dan potensial aksi yang terjadi pada sel prasinaptik tidak dapat dirambatkan secara langsung ke membran sel pascasinaptik. Solusinya, suatu rangkaian kejadian mengubah sinyal listrik potensial aksi yang tiba di terminal sinaptik itu diubah menjadi sinyal kimiawi yang mengalir melewati sinapsis, dimana sinyal kimiawi ini nantinya akan diubah kembali menjadi sinyal listrik pada sel pascasinaptik. Terdapat banyak kantung yang disebut vesikula sinaptik (*synaptic vesicle*) di dalam sitoplasma ujung akson prasinaptik. Masing-masing vesikula mengandung ribuan molekul *neurotransmitter*, zat yang dibebaskan sebagai *messenger* antarsel ke dalam celah sinaptik. Kebanyakan *neurotransmitter* yang ditemukan dalam sistem saraf hewan adalah berbeda. Sebagian besar *neuron* hanya mensekresikan satu jenis *neurotransmitter*. Akan tetapi, sebuah *neuron* tunggal bisa menerima sinyal kimiawi dari berbagai *neuron* yang mensekresikan *neurotransmitter* yang berbeda-beda dari terminal sinaptiknya (Campbell et al, 2004)

2.2 Transformasi Jaringan Saraf Manusia ke Jaringan Saraf Tiruan

Sistem kerja saraf manusia ini akhirnya diadopsi oleh seorang *neurophysiologist* Warren McCulloch dan *logician* Walter Pitts untuk membuat jaringan saraf tiruan (*neural network*) pada tahun 1943. Akan tetapi, teknologi yang tersedia saat itu belum memungkinkan mereka berbuat lebih jauh. *Neural network* ini dibuat dengan meniru komponen pembangun dari jaringan saraf manusia dan mekanisme dari jaringan saraf otak manusia. *Neuron* yang dibuat saat itu tidak bisa di-*train*. Akan tetapi, mereka merepresentasikan sejumlah fungsi logika. Model inilah yang nantinya menjadi fondasi untuk perkembangan *neural network* selanjutnya. Komponen *neural network* ini adalah masukan, *neuron*, bobot, masukan, dan keluaran. Berikut adalah gambaran dan perbandingan antara jaringan saraf manusia dengan *neural network*.



Gambar 2.3 Analogi Peniruan Sistem Saraf Manusia

(Sumber: Stephen, 2004)

Tabel 2.1 Perbandingan Saraf Manusia dengan *Neural Network*

Jaringan Saraf Manusia	Neural Network
Dendrit	Masukan
Badan Sel	<i>Neuron</i>
Akson	Keluaran
Neurotransmitter	Bobot
Jumlah <i>neuron</i> banyak	Jumlah <i>neuron</i> sedikit

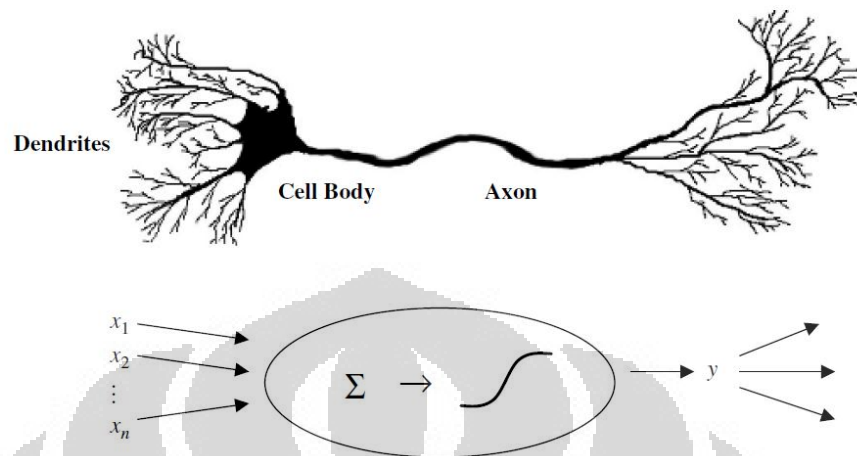
Selain sejumlah perbedaan di Tabel 2.1, terdapat pula kesamaan antara jaringan saraf manusia dengan *neural network*. Sebuah *neuron* tunggal bisa menerima sinyal kimiawi dari berbagai *neuron* yang mensekresikan *neurotransmitter* yang berbeda-beda dari terminal sinaptiknya. Demikian pula dengan *neural network*, yang bisa menerima masukan dari berbagai *neuron* lain.

2.3 *Neural Network*

“Sebuah jaringan saraf adalah sebuah prosesor yang terdistribusi paralel dan mempunyai kecenderungan untuk menyimpan pengetahuan yang didapatkannya dari ‘pengalaman’ sebelumnya dan membuatnya tetap tersedia untuk digunakan. Hal ini menyerupai kerja otak dalam dua hal yaitu:

1. Pengetahuan diperoleh dari jaringan melalui suatu proses belajar

2. Kekuatan hubungan antar sel saraf yang dikenal dengan bobot sinapsis digunakan untuk menyimpan pengetahuan (Haykin, 1994).



Gambar 2.4 Saraf Manusia dan *Artificial Neuron Model*

(Sumber: Larose, 2005)

Input (x_i) dikumpulkan dari *neuron* sebelumnya (atau dari data set) dan dikombinasikan melalui fungsi kombinasi seperti penjumlahan (Σ). Selanjutnya, hal ini akan menjadi masukan ke dalam sebuah fungsi aktivasi untuk memproduksi respons *output* (y), yang mana akan dihubungkan dengan *neurons* selanjutnya.

Adapun kelebihan dari penggunaan *neural network* adalah keadaannya yang *robust* terhadap *noisy data*. Dikarenakan *network* memiliki banyak *nodes* (*neuron* tiruan), dengan beban yang dihubungkan dengan masing-masing koneksi, *network* bisa mempelajari untuk bekerja di antara contoh-contoh data set yang tidak informatif (keliru). Sebaliknya, penggunaan *neural network* cenderung memberikan interpretasi yang buram terhadap manusia dan membutuhkan waktu *training* yang lebih panjang (Larose, 2005).

Jenis pembelajaran *neural network* terbagi menjadi dua, yaitu *supervised learning* dan *unsupervised learning*. Berikut keterangannya:

1. *Supervised Learning*

Pada metode ini, proses pembelajaran *neural network* diarahkan untuk memberikan target terhadap keluaran yang dihasilkan. Salah satu contohnya *Backpropagation*.

2. *Unsupervised Learning*

Metode ini tidak memberikan pengarahan kepada *neural network* terhadap target keluaran dan contohnya adalah *Self Organizing Map (SOM)*.

Pada umumnya, sebuah *neural network* terdiri atas 3 lapisan utama sebagai berikut:

1. Lapisan masukan (*input layer*)

Lapisan ini hanya bertugas untuk meneruskan masukan dan tidak akan melakukan komputasi apapun. Jumlah *node* pada lapisan masukan sama dengan jumlah ciri atau atribut pada pola yang akan dikenali.

2. Lapisan tersembunyi (*hidden layer*)

Lapisan ini melakukan komputasi. Jumlah *node* pada layer ini ditentukan oleh si pengguna. Jumlah yang semakin banyak akan meningkatkan fleksibilitas dan kekuatan dari jaringan untuk mengidentifikasi pola yang kompleks. Akan tetapi, jumlah yang terlalu banyak akan menyebabkan *overfitting*, yang menyebabkan proses penguatan *training set* akan mengeneralisasi *validation set*. Jika hal ini terjadi, maka hal yang seharusnya dilakukan adalah mengurangi jumlah *node* pada *hidden layer*.

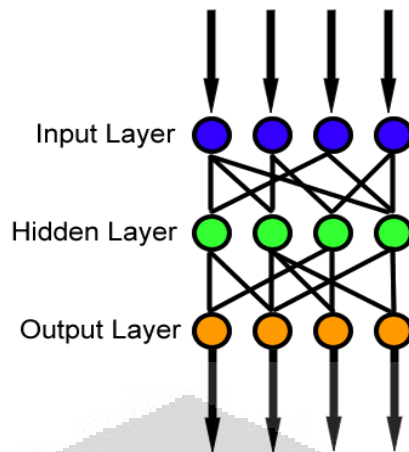
3. Lapisan keluaran (*output layer*)

Lapisan ini juga melakukan komputasi. Jumlah *node* pada lapisan keluaran sama dengan kelas pola dan tergantung dari jumlah klasifikasinya.

Dengan berdasarkan pada arsitektur pola koneksi antar-*neuron* dari lapisan-lapisan ini, *neural network* bisa dibedakan menjadi 2 jenis, yaitu:

1. Struktur *feedforward*

Struktur ini adalah yang paling pertama dan sederhana dari *neural network*. Pada struktur ini, informasi bergerak hanya melalui satu arah, dari *node input* menuju *node tersembunyi*, dan terakhir menuju *node output*. Tidak ada siklus ataupun *loop* pada jaringan ini.

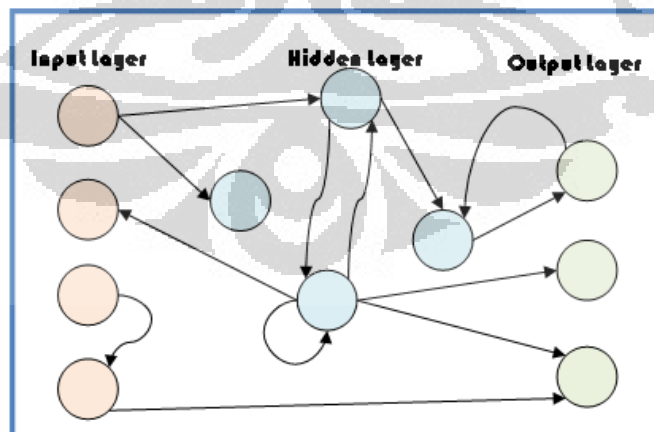


Gambar 2.5 *Feedforward Neural Network*

(Sumber: Stephen, 2010)

2. Struktur *recurrent (feedback)*

Berbeda dengan *feedforward*, *neural network* jenis ini bisa memiliki *signal* yang berjalan pada kedua arah dengan membuat *loop* pada jaringan. *Feedback network* bisa sangat *powerful* dan menyulitkan. Jaringan ini dinamis. Pernyataan mereka akan terus berubah sampai mereka mencapai keadaan ekuilibrium. Mereka akan tetap berada di titik ekuilibrium sampai terjadi perubahan *input* dan sebuah titik ekuilibrium baru ditemukan.

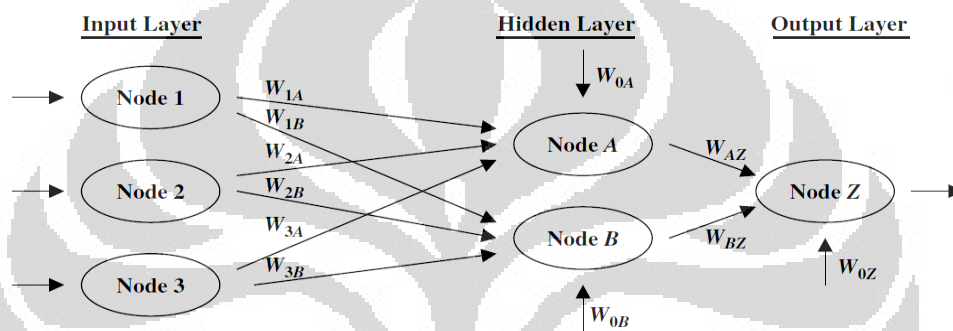


Gambar 2.6 *Recurrent Neural Network*

(Sumber: Stephen, 2010)

Selain itu, sebuah *neural network* juga ditentukan oleh fungsi aktivasi yang memberikan ciri tertentu. Sebuah fungsi aktivasi sebaiknya memiliki beberapa karakteristik, yaitu kontinu, dapat diturunkan, dan monoton naik.

Secara umum, *neural network* biasanya memiliki 3 *layer* seperti yang telah dijelaskan sebelumnya. Akan tetapi, bisa saja terdiri atas satu atau lebih *hidden layer*. *Neural network* terkoneksi secara keseluruhan. Dengan kata lain, setiap *node* pada *layer* sebelumnya akan terhubung dengan *layer* berikutnya. Setiap hubungan antara *nodes* memiliki beban (misalnya W_{1A}). Pada umumnya, beban ini secara acak bernilai antara 0 dan 1.



Gambar 2.7 Sebuah *Neural Network* Sederhana

(Sumber: Larose, 2004)

Input layer akan menerima *inputs* dari data set, misalnya nilai atribut dan secara sederhana akan melewati nilai-nilai ini ke *hidden layer*, tanpa proses apapun. Oleh karena itu, *nodes* pada *input layer* tidak akan berbagi detail struktur *node* seperti yang dilakukan oleh *hidden layer* dan *output layer*.

Selanjutnya, Tabel 2.2 akan menginvestigasi struktur dari *hidden layer nodes* dan *output layer nodes* dengan menggunakan *sampel data* yang disediakan.

Tabel 2.2 *Data Inputs* dan *Initial Values* Beban *Neural Networks*

$x_0 = 1.0$	$W_{0A} = 0.5$	$W_{0B} = 0.7$	$W_{0Z} = 0.5$
$x_1 = 0.4$	$W_{1A} = 0.6$	$W_{1B} = 0.9$	$W_{AZ} = 0.9$
$x_2 = 0.2$	$W_{2A} = 0.8$	$W_{2B} = 0.8$	$W_{BZ} = 0.9$
$x_3 = 0.7$	$W_{3A} = 0.6$	$W_{3B} = 0.4$	

(Sumber: Larose, 2004)

Sebuah *combination function* (biasanya *summation*, Σ) menghasilkan kombinasi linear dari *node inputs* dan beban hubungan pada sebuah *single scalar value*, yang disebut sebagai *net*. Oleh karena itu, untuk *node j*,

$$\text{net}_j = \sum_i W_{ij} x_{ij} = W_{0j} \cdot x_{0j} + W_{1j} \cdot x_{1j} + \dots + W_{lj} x_{lj} \quad (2.1)$$

dimana x_{ij} merepresentasikan *input i* ke *node j*, W_{ij} merepresentasikan *weights* yang terasosiasi dengan *input i* ke *node j*, dan ada $I + 1$ *input* ke *node j*. Perlu dicatat bahwa x_1, x_2, \dots, x_I merepresentasikan *inputs* dari *upstream nodes*, yang mana x_0 merepresentasikan *input* yang konstan. Analog dengan faktor yang konstan pada model regresi, yang secara unik bernilai $x_{0j} = 1$. Oleh karena itu, setiap *hidden layer* ataupun *output layer node j* memiliki ekstra-*input* yang sebanding dengan beban $W_{0j} x_{0j} = W_{0j}$.

Sebagai contohnya, pada *node A* kita akan mendapatkan:

$$\begin{aligned} \text{net}_A &= \sum W_{iA} x_{iA} = W_{0A}(1) + W_{1A} x_{1A} + W_{2A} x_{2A} + W_{3A} x_{3A} \\ &= 0.5 + 0.6(0.4) + 0.8(0.2) + 0.6(0.7) = 1.32 \end{aligned} \quad (2.2)$$

Pada *node A*, fungsi kombinasi $\text{net}_A = 1,32$ ini akan digunakan sebagai *input* pada sebuah fungsi aktivasi. Pada *neuron* biologis, sinyal dikirimkan di antara *neuron* ketika kombinasi *input* ke *neuron* berikutnya melewati *threshold* tertentu, dan *neuron* tersebut akan terangsang. Hal ini adalah tindakan *nonlinear* dikarenakan respons rangsang tidak perlu berhubungan secara linear dengan *increment* dari *sigmoid* pada *node Z*, dengan hasil:

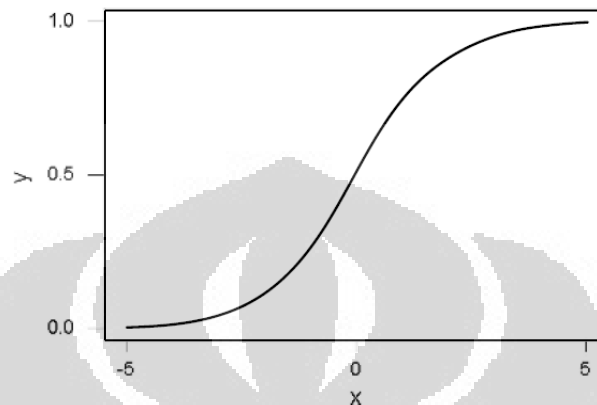
$$f(\text{net}_Z) = \frac{1}{1 + e^{-1,9461}} = 0,875$$

Nilai ini akan merepresentasikan prediksi nilai dari target untuk observasi yang pertama.

2.4 Fungsi Aktivasi Sigmoid

Penggunaan fungsi aktivasi *sigmoid* memang paling banyak digunakan. Hal ini disebabkan oleh kelakuan *kombinasi* linear, *curvilinear* dan hampir konstan, tergantung dari nilai input. Gambar 2.8 menunjukkan grafik dari fungsi *sigmoid*, dimana $y = f(x) = 1/(1 + e^{-x})$, untuk $-5 < x < 5$. Selain itu, kebanyakan

input dari x bernilai $-1 < x < 1$, perilaku dari $f(x)$ ini hampir linear. Ketika nilai *input* bergerak menjauh dari bagian tengah grafik, maka $f(x)$ berubah menjadi konstan. Fungsi *sigmoid* sering dikenal sebagai *squashing function*, karena ia mengambil nilai real dari *input* dan menjadikannya sebagai *output* antara 0 dan 1.



Gambar 2.8 Grafik Fungsi *Sigmoid*

(Sumber: Larose, 2004)

2.5 Backpropagation

Pertanyaan yang muncul selanjutnya adalah bagaimana cara sebuah *neural network* belajar. *Neural network* merepresentasikan sebuah metode *supervised* yang membutuhkan sebuah *training set* dari sebuah data yang lengkap, termasuk variabel target. *Neural network* dengan menggunakan algoritma *backpropagation* mempunyai bentuk arsitektur dengan struktur *feedforward* yang terdiri atas 3 lapisan, yaitu *input*, *hidden*, dan *output*. Masing-masing lapisan terdiri atas sejumlah *neuron* yang banyaknya tergantung pada aplikasi penggunaan algoritma ini.

Setiap observasi dari *training set* ini akan diproses melalui *network*, dan sebuah *output* akan dihasilkan melalui *output node* (mengasumsikan bahwa kita hanya memiliki satu *output node*). Nilai *output* ini akan dibandingkan dengan nilai aktual target yang diinginkan. Prediksi *error* ini analog dengan bentuk *residual* pada model regresi. Untuk mengukur seberapa baiknya prediksi *output* yang sesuai dengan nilai target, kebanyakan *neural network* menggunakan *sum of squared errors* seperti ini:

$$SSE = \sum_{records} \sum_{output\ nodes} (actual - output)^2 \quad (2.7)$$

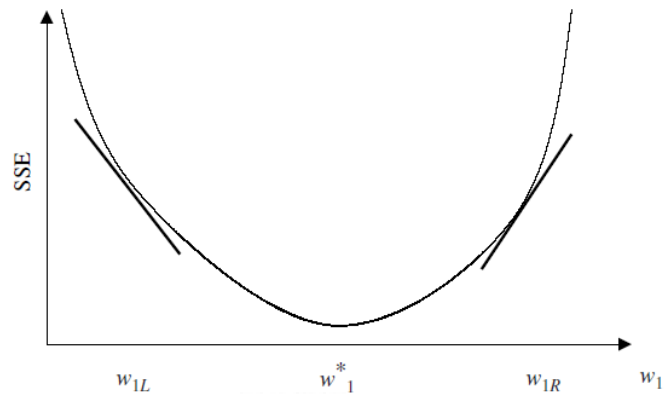
dimana kuadrat dari prediksi *errors* dijumlahkan secara keseluruhan dari semua *output nodes* dan semua data yang sedang di-*train*. Masalah utama dari *backpropagation* adalah untuk mengkonstruksi beban set model yang meminimumkan SSE. Pada bagian ini, *weight* dianalogikan sebagai parameter dari model regresi. Nilai sebenarnya dari tiap beban yang akan meminimumkan SSE yang tidak diketahui, dan untuk itulah perlu diestimasi dari data yang ada. Akan tetapi, dikarenakan bentuk nonlinear dari fungsi *sigmoid* yang mempermeasi jaringan, maka tidak ada solusi yang bisa digunakan untuk meminimumkan SSE.

2.6 Gradient Descent Method

Untuk menyelesaikan proses optimisasi, *gradient-descent method* bisa digunakan untuk membantu dalam mencari *set of weights* yang akan meminimumkan SSE. Seandainya kita memiliki set (vektor) dari m *weights* $\mathbf{w} = w_0, w_1, w_2, \dots, w_m$ pada model *neural network* yang ingin mencari nilai dari setiap *weight* yang juga akan meminimumkan SSE. Kita bisa menggunakan *gradient descent method*, yang akan memberikan arahan bahwa kita harus menyesuaikan *weights* untuk menurunkan SSE. Gradien dari SSE terhadap vektor pada *weights* \mathbf{w} pada vektor derivatif adalah:

$$\nabla \text{SSE}(\mathbf{w}) = \left[\frac{\partial \text{SSE}}{\partial w_0}, \frac{\partial \text{SSE}}{\partial w_1}, \dots, \frac{\partial \text{SSE}}{\partial w_m} \right] \quad (2.8)$$

Untuk mengilustrasikan bagaimana kerja dari *gradient descent works*, bisa dilihat kasus dimana hanya ada *single weight* w_1 . Pada gambar 2.9, yang memplot *error* dari SSE terhadap *range* dari nilai untuk w_1 . Selanjutnya akan dipilih nilai dari w_1 yang akan meminimumkan SSE. Nilai optimal dari beban w_1 diindikasikan sebagai w_1^* . Selanjutnya, akan dikembangkan *rule* yang akan membantu kami untuk memindahkan nilai w_1 yang sekarang ke nilai optimal w_1^* sebagai berikut: $w_{\text{new}} = w_{\text{current}} + \Delta w_{\text{current}}$, dimana $\Delta w_{\text{current}}$ adalah perubahan lokasi dari w .



Gambar 2.9 Kurva SSE Terhadap w_1

(Sumber: Larose, 2004)

Sekarang, seandainya *current weight*, w_{current} dekat dengan w_{1L} , maka kita harus meningkatkan *current weight* ke nilai w_1^* . Di sisi lain, jika w_{current} dekat dengan w_{1R} , kita akan menurunkan nilai tersebut untuk membuatnya lebih dekat dengan nilai optimal dari w_1^* . Sekarang, nilai derivatif dari $\partial \text{SSE} / \partial w_1$ adalah *slope* dari kurva SSE pada w_1 . Untuk nilai dari w_1 yang dekat dengan w_{1R} , maka ini adalah *slope* yang positif, begitupula sebaliknya. Selanjutnya, proses penyesuaian dari w_{current} ini adalah nilai negatif dari tanda derivatif terhadap SSE pada w_{current} sehingga bernilai $-(\partial \text{SSE} / \partial w_{\text{current}})$.

Selanjutnya, hal yang harus diperhatikan adalah, seberapa jauh proses penyesuaian ini akan berlanjut. Jika kurva tergolong kurva yang lancip, besar penyesuaian akan lebih besar dibandingkan jika kurva berbentuk datar. Pada akhirnya, hasilnya akan dikalikan dengan konstanta η (eta), yang disebut sebagai *learning rate*, dimana nilainya berkisar antara 0 hingga 1. Hasilnya, bentuk dari $\Delta w_{\text{current}} = -\eta (\partial \text{SSE} / \partial w_{\text{current}})$, yang berarti perubahan dari *current weight* senilai dengan negatif dari konstanta dikalikan dengan fungsi *slope of error* dari w_{current} .

2.7 Backpropagation Rules

Dengan menggunakan fungsi aktivasi *sigmoid* dan *gradient descent*, Mitchell menyatakan *rules* dari *backpropagation* adalah sebagai berikut:

$$w_{ij,\text{new}} = w_{ij,\text{current}} + \Delta w_{ij} \quad \text{dimana} \quad \Delta w_{ij} = \eta \delta_j x_{ij} \quad (2.9)$$

δ_j merepresentasikan *responsibility* dari *error* tertentu pada *node j*. *Error responsibility* ini dihitung dengan menggunakan *partial derivative* dari fungsi *sigmoid* terhadap net_j , dan menjadikannya bentuk sebagai berikut:

$$\delta_j = \begin{cases} \text{output}_j(1 - \text{output}_j)(\text{actual}_j - \text{output}_j) & \text{for output layer nodes} \\ \text{output}_j(1 - \text{output}_j) \sum_{\text{downstream}} W_{jk} \delta_j & \text{for hidden layer nodes} \end{cases} \quad (2.10)$$

dimana $\sum_{\text{downstream}} W_{jk} \delta_j$ merupakan *weighted sum* dari *error responsibilities* untuk *nodes downstream* dari *hidden layer node* tertentu.

2.8 Contoh Backpropagation

Kembali ke contoh soal sebelumnya, *output* yang dihasilkan adalah 0,875. Diasumsikan bahwa nilai aktual dari atribut target adalah 0,8 dan *learning rate* dari $\eta=0,1$. Maka *prediction error* yang dihasilkan adalah $0,8 - 0,875 = -0,075$ dan kita bisa mengaplikasikan *rules* ini untuk mengilustrasikan bagaimana algoritma *backpropagation* bekerja untuk menyesuaikan *weights* dengan menunjukkan *responsibility error* pada berbagai *nodes*. Walaupun sangat mungkin untuk meng-*update weight* setelah semua *records* telah dibaca, *neural network* menggunakan *stochastic (online) backpropagation*, yang meng-*update weights* dari setiap *record*.

$$\delta_A = \text{output}_A(1 - \text{output}_A) \sum_{\text{downstream}} W_{jk} \delta_j$$

Node yang berada pada bagian *downstream* dari *node A* hanyalah *node Z*. *Weight*

$$\Delta W_{2A} = \eta \delta_A x_2 = 0.1(-0.00123)(0.2) = -0.0000246.$$

$$W_{2A,\text{new}} = W_{2A,\text{current}} + \Delta W_{2A} = 0.8 - 0.0000246 = 0.7999754.$$

Untuk *weight* W_{3A} , maka:

$$\Delta W_{3A} = \eta \delta_A x_3 = 0.1(-0.00123)(0.7) = -0.0000861. \quad (2.19)$$

$$W_{3A,\text{new}} = W_{3A,\text{current}} + \Delta W_{3A} = 0.6 - 0.0000861 = 0.5999139.$$

Pada akhirnya, bisa ditentukan W_{0A} , sebagai berikut:

$$\Delta W_{0A} = \eta \delta_A (1) = 0.1(-0.00123) = -0.000123. \quad (2.20)$$

$$W_{0A,\text{new}} = W_{0A,\text{current}} + \Delta W_{0A} = 0.5 - 0.000123 = 0.499877$$

Dengan cara yang sama dicari *weights* W_{0B} , W_{1B} , W_{2B} , dan W_{3B} .

Dengan adanya nilai-nilai ini, maka proses perbandingan akan dimulai. *Network* akan mengkalkulasi nilai prediksi untuk variabel target. Kemudian, akan dibandingkan nilai *output* hasil prediksi dengan target yang sebenarnya. Kemudian disaring kesalahan prediksi dari seluruh jaringan, dan disesuaikan lagi *weights* yang dihasilkan hingga didapatkan *prediction error* yang terkecil.

2.9 Termination Criteria

Algoritma *neural network* akan terus memproses kerja melalui *training data set*, *record* demi *record*, menyesuaikan *weights* secara konstan untuk mengurangi *prediction error*. Hal ini akan mengakibatkan *pass* data yang sangat banyak sebelum *termination criterion* ditemukan. Jika waktu *training* menjadi hal yang penting, maka boleh di-*set* jumlah *pass* yang diinginkan, atau jumlah *realtime* yang memudahkan algoritma bekerja. Hal ini dinamakan sebagai *termination criteria*. Walaupun demikian, hal yang didapatkan pada *training time* yang sempit adalah menurunnya kualitas model.

Oleh karena itu, penggunaan *termination criterion* yang menyesuaikan dengan SSE pada *training data* telah dikurangi pada batasan tertentu. Akan tetapi, karena kefleksibelannya, *neural network* rentan akan *overfitting*, ketika ‘menghafalkan’ pola yang istimewa pada *training set* dan tidak memperlihatkan generalisasi pada data yang tak terlihat.

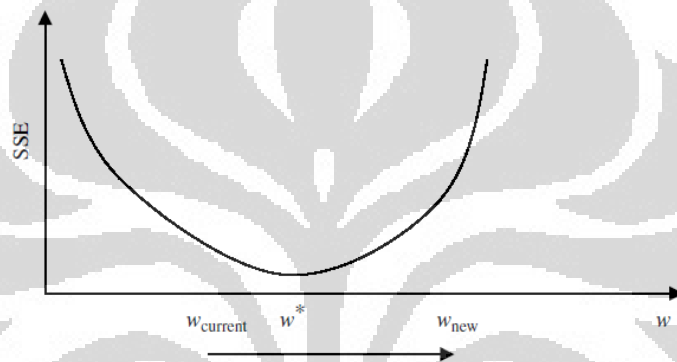
Oleh karena itu, sebuah *neural network* pada umumnya harus mengimplementasi prosedur *cross validation*, sebagai berikut :

1. Memelihara bagian dari data *set original* sebagai *holdout validation set*,
2. Men-*train neural network* dengan sejumlah *training data*.
3. Mengaplikasikan *weight* yang dipelajari dari *training data* pada *validation data*.
4. Mengatur dua *set weights*, yaitu “*current*” *set* yang didapatkan dari *training data*, dan satu lagi “*best*” *set* yang didapatkan dari SSE terendah pada *validation data*.
5. Ketika *current set of weights* telah memiliki SSE yang lebih baik dari *best set of weights*, maka terminasikan algoritma.

Telepas dari penggunaan *stopping criterion*, *neural network* tidak menjamin untuk mendapatkan solusi terbaik (*global minimum*) pada SSE. Lebih sering, algoritma menjadi *stuck* pada bagian *local minimum*, yang merepresentasikan sebuah solusi yang baik, bukan yang terbaik. Secara teknis, hal ini telah menghasilkan masalah yang tidak dapat diatasi:

2.10 Learning Rate

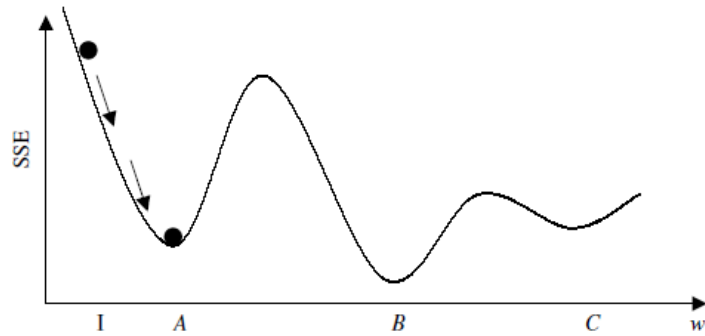
Sebuah *learning rate* η , $0 < \eta < 1$ dipilih untuk membantu *neural network* mendapatkan *global minimum* untuk SSE.



Gambar 2.10 Nilai η Menghasilkan Algoritma Melampaui *Global Minimum*

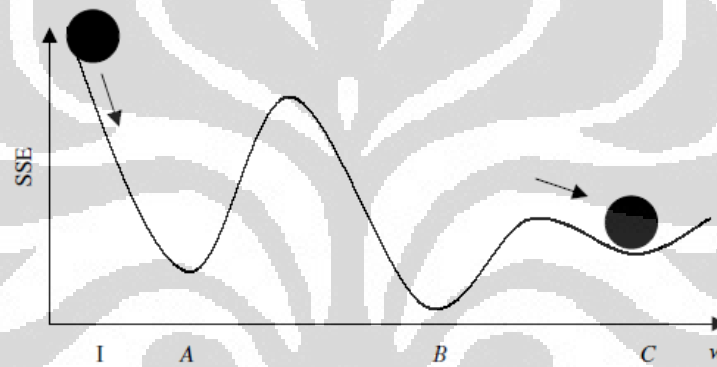
(Sumber : Larose, 2004)

Ketika *learning rate* terlalu kecil, penyesuaian beban akan menjadi terlalu kecil. Dengan demikian, jika η terlalu kecil, maka *network* akan menjalani waktu yang pada B. Pada Gambar 2.11, dimisalkan nilai *momentum term* α , disimbolkan sebagai sebuah bola dengan massa yang kecil. Jika kita menjatuhkan bola ini, maka ia akan berhenti di lereng yang pertama dan tak akan mencapai titik B. Kemudian pada Gambar 2.12, dimisalkan nilai *momentum term* α bernilai besar, disimbolkan sebagai sebuah bola dengan massa yang besar. Dengan demikian, ia akan melewati titik B dan menuju titik C. Oleh karena itu, nilai α harus dilakukan secara tepat untuk menghindari kedua hal ini.



Gambar 2.11 Nilai Momentum α Kecil

(Sumber : Larose, 2004)

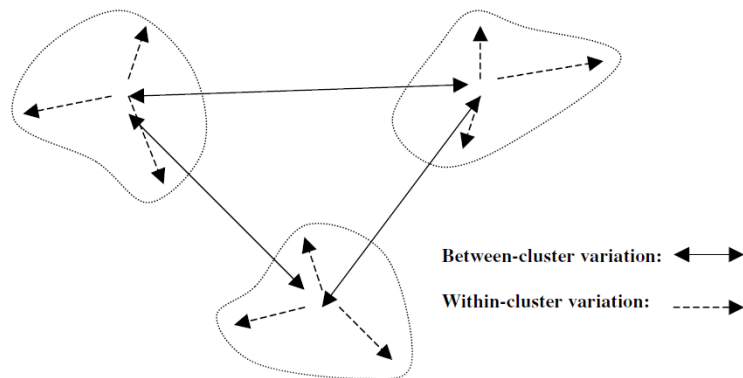


Gambar 2.12 Nilai Momentum α Besar

(Sumber : Larose, 2004)

2.12 Analisis Kluster

Adapun fungsi dari analisis kluster adalah mengelompokkan data, hasil observasi, ataupun kasus ke dalam beberapa kelas yang berisi objek yang sama. Metode *cluster* ini sedikit berbeda dengan metode klasifikasi karena tidak ada



Gambar 2.13. Pengelompokan pada *cluster analysis*.

(Sumber : Larose, 2004)

2.13 Analisis Kluster *K-Means*

Algoritma *K-means clustering* adalah algoritma yang paling mudah dan efektif digunakan untuk menemukan kluster pada data. Cara kerja algoritma adalah sebagai berikut:

1. Menentukan berapa kluster k dari data yang harus dipartisikan.
2. Menentukan secara acak data k yang dijadikan sebagai pusat kluster awal.
3. Untuk setiap data, dicari titik tengah kluster sehingga setiap *data set*, memiliki titik tengahnya.
4. Untuk setiap kluster k , dicari *cluster centroid*, dan meng-*update* lokasi dari setiap titik tengah kluster ke nilai baru dari *centroid*.
5. Diulangi terus tahapan 3-5 hingga sistem konvergen atau berakhir.

Untuk tahapan nomor 3, jarak *euclidean* adalah yang paling cocok, walaupun kriteria lain juga bisa digunakan. Untuk melakukan tahapan 4, langkah-langkah yang harus dilalui adalah sebagai berikut:

- Misalkan kita memiliki n data points $(a_1, b_1, c_1), (a_2, b_2, c_2), \dots (a_n, b_n, c_n)$.
- *Centroids* dari titik-titik ini adalah berada di gravitasi dari poin-poin ini yang berlokasi pada $(\sum \frac{a_i}{n}, \sum \frac{b_i}{n}, \sum \frac{c_i}{n})$
- Dengan demikian, titik-titik $(1,1,1), (1,2,1), (1,3,1),$ dan $(2,1,1)$ akan memiliki *centroid*

$$\left(\frac{1+1+1+2}{4}, \frac{1+2+3+1}{4}, \frac{1+1+1+1}{4}\right) = (1.25, 1.75, 1.00) \quad (2.25)$$

Algoritma ini akan terus berjalan hingga titik *centroid* tidak berubah lagi. Dengan kata lain, algoritma ini akan terus dijalankan hingga kluster C_1, C_2, \dots, C_k , dimiliki setiap *cluster center* yang berada pada kluster.

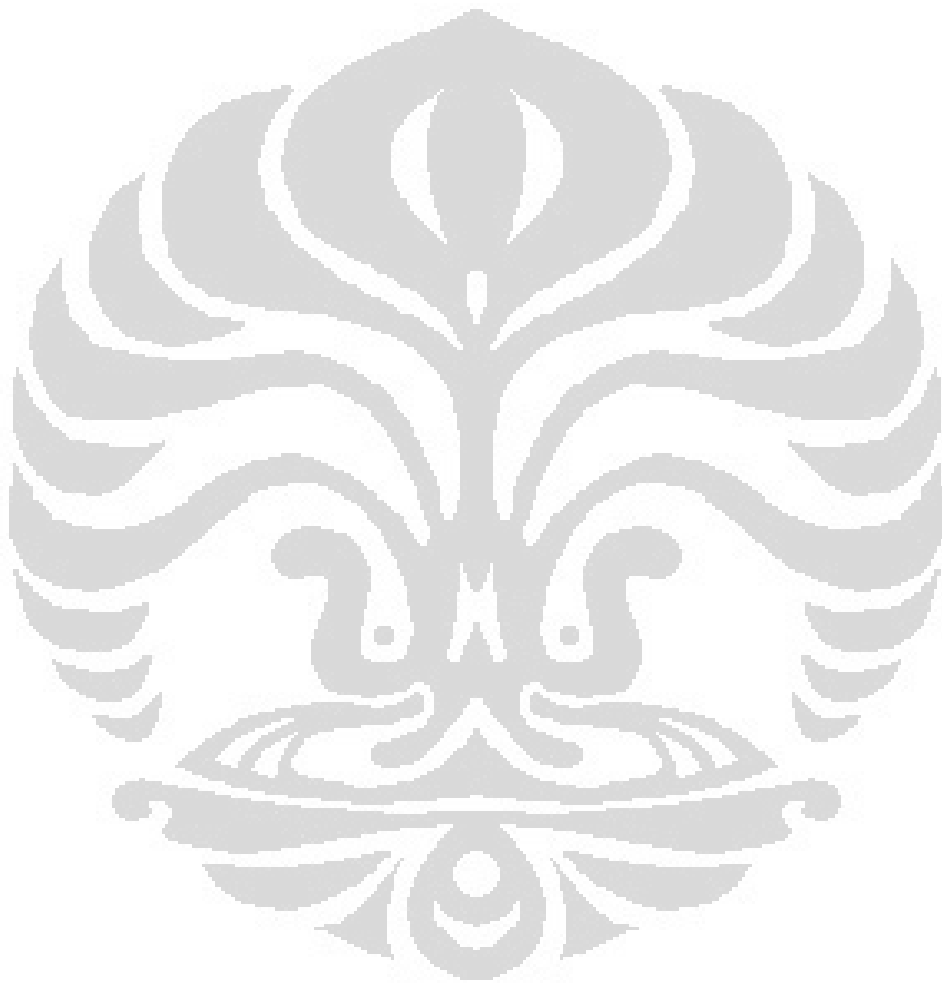
2.14 Perkembangan Industri *Fast Food*

Kata *fast food* sendiri berarti makanan yang disajikan dengan cepat. Kebanyakan dari makanan ini dibungkus dengan kemasan yang lebih dikenal dengan 'take away'. Dengan kehadiran sistem *franchise* dan kepopuleran dari makanan ini, membuat restoran *fast food* dan gerai-gerainya tersebar hampir di seluruh dunia dan konsumen yang mencakup semua umur.

Konsep dari *fast food* ini digeneralisasi dari perkembangan penduduk yang tinggi. Akan tetapi, penelitian menunjukkan bahwa tren ini memiliki akarnya, misalnya kehadiran *stands* di Roma kuno dan toko-toko mie di negara-negara Asia. Pada masa itu, roti yang rata, *falafel*, dan makanan hangat yang baru disajikan sangat menarik perhatian pelewat jalan. Kios-kios ini menyajikan makanan untuk masyarakat dengan strata yang tidak terlalu tinggi. Sekarang, tren ini menjadi bagian dari bentuk hidup yang *ultra-modern*. Sejarah dari Romawi regional lainnya. Makanan ringan seperti *sandwiches* adalah hasil eksperimen dari industri ini.

Kebanyakan variasi *fast food* berbagi kesamaan dalam hal makanan dan budaya. Industri ini sekarang tumbuh subur di dunia internasional dengan adanya kehadiran *niche chains*. Perkembangannya mulai dari alternatif menu yang lebih sehat ke pelayanan konvensional, telah menghasilkan sistem seperti *portable foods* yang bisa diambil sendiri oleh pelanggan. Pada kebanyakan gerai dan *drive-ins*, pelanggan bisa melihat cara penyajiannya sehingga bisa mengkonfirmasi dan mempromosikan standar higienis yang dimiliki. Konsep makan dan pergi ini tidak hanya mengeliminasi kebutuhan peralatan makan, tetapi juga membuat para pelanggan duduk menikmati makanan yang merupakan ciri khas dari tradisi budaya atau etnis tertentu. Sekarang, *fast food industry* telah beroperasi pada toko yang nyaman dan mewah, dan semuanya bersaing untuk

menstandarisasi proses produksi untuk menghasilkan makanan yang murah dan lezat (Borade, 2010).



BAB 3

PENGUMPULAN DAN PENGOLAHAN DATA

3.1 Pengumpulan data

Data-data yang digunakan untuk penelitian ini adalah data primer yang didapatkan dengan menyebarkan kuesioner ke pelanggan Papa Ron's Pizza yang tersebar di JABODETABEK. Kuesioner yang disebarkan berjumlah 600 buah. Jumlah responden yang berhasil didapatkan adalah 346 responden, dengan data yang bisa dipakai adalah 323 data.

Pada pendekatan *data mining*, khususnya metode *neural network* tidak terdapat ketentuan mengenai berapa jumlah data yang harus diambil. Semakin banyak data yang di-*train*, akan membuat *neural network* yang dibuat semakin 'cerdas' dalam mengenali karakter objek. Oleh karena itu, jumlah data yang semakin banyak akan menghasilkan *model neural network* yang semakin baik.

Adapun kuesioner yang dibuat untuk penelitian ini, menggunakan skala *likert* dengan skala 1-5 (1 menunjukkan sangat tidak suka, dan 5 menunjukkan sangat suka). Kuesioner terdiri atas 8 halaman dengan bentuk seperti yang di lampiran. Kuesioner diberikan kepada pelanggan pada saat mereka telah selesai memesan menu makanan dan menunggu menu disajikan. Rata-rata waktu pengisian adalah 10-15 menit sehingga pelanggan bisa dengan leluasa mengisi kuesioner karena Papa Ron's Pizza memiliki *service time* sekitar 15 menit.

Didirikan sejak tahun 2000, saat ini Papa Ron's Pizza sudah membuka 43 gerai yang tersebar di seluruh Indonesia dan mendapatkan banyak prestasi. Selain mendapatkan penghargaan sebagai "The Best Tasting Pizza" versi harian Jakarta Post, juga mendapatkan peringkat ketiga terbaik terhadap kepuasan pelanggan versi majalah SWA. Papa Ron's Pizza adalah satu-satunya restoran Pizza yang merupakan merek lokal di antara merek waralaba internasional lainnya.

Adapun visi dan misi dari perusahaan ini adalah:

- Visi :
"agar dapat dikenal oleh pelanggan dan franchisee sebagai salah satu dari 3 perusahaan pizza terbesar di asia"

- Misi :
 - “untuk menjaga kesempurnaan standar mutu kualitas internasional dalam produk dan pelayanan dengan memenuhi permintaan rasa, kenikmatan, kenyamanan dan nilai khusus bagi pelanggan”
 - kesempurnaan mutu dari pelayanan kualitas internasional
 - jaminan kepuasan pelanggan

3.2 Penentuan Kriteria dan Atribut

Kelebihan dari metode pendekatan *data mining* adalah keleluasaan untuk menggunakan sebanyak mungkin atribut yang diinginkan. Pada penelitian ini, 4 kriteria utama yang dijadikan sebagai objek penelitian adalah sebagai berikut:

1. Keadaan restoran

Kriteria ini tentu sangat penting untuk diteliti lebih lanjut. Meski produk utama yang dijual oleh restoran adalah makanan, hal-hal seperti suasana, letak, dan keadaan yang berada di dalam restoran juga menjadi faktor tersendiri yang menentukan keputusan pelanggan untuk kembali. Pada kriteria ini, ditentukan sepuluh buah atribut yang akan dipertimbangkan oleh responden sebagai preferensi. Berikut adalah atribut yang ditentukan:

1. Bersih dan rapi
2. Mudah mencari tempat duduk
3. Restoran yang sudah familiar/sudah dikenal baik
4. Berada di lokasi yang mudah dijangkau
5. Suasana restoran yang nyaman dan menyenangkan
6. Cocok untuk membawa anggota keluarga
7. Memiliki kamar mandi dan tempat cuci tangan yang bersih
8. Restoran yang nyaman untuk anak kecil
9. Memiliki dekorasi yang menarik perhatian (atraktif)
10. Memiliki tempat duduk yang nyaman

2. Menu makanan

Kriteria ini haruslah dipertimbangkan lebih lanjut, terlebih untuk restoran seperti Papa Ron's Pizza yang memiliki pasar yang telah matang. Meski menu yang disajikan hampir sama dengan restoran *fast food* lainnya, harus

diketahui karakter spesifik menu makanan yang membuat pelanggan untuk Menawarkan pizza dengan topping yang menarik

11. Menawarkan menu yang telah disesuaikan dengan selera lokal
12. Menawarkan makanan Italia lain yang unik dan lezat
13. Menawarkan *dessert* yang dingin
14. Menawarkan *dessert* yang hangat

3. Jenis pelayanan

Jenis pelayanan juga merupakan sebuah faktor yang perlu ditelaah lebih lanjut, terutama karena budaya Indonesia yang masih kental akan adat timur. Selain alasan ini, Papa Ron's Pizza juga harus mengetahui jenis layanan apakah yang harus ia miliki sebagai sebuah kelebihan yang diinginkan konsumen. Berikut adalah atribut yang ditentukan:

15. Menyediakan pelayanan yang cepat dan efisien
16. *Delivery order* yang mudah
17. Menawarkan diskon dan kupon makan
18. Pelayanan dari karyawan yang terampil
19. Menyediakan fasilitas internet (Wifi)

4. Bentuk pemasaran

Strategi pemasaran menjadi kelebihan tersendiri bagi setiap perusahaan. Oleh karena itu, faktor ini menjadi penting untuk dipertanyakan lebih lanjut. Berikut adalah atribut-atribut yang ditentukan:

20. Telah memiliki pengalaman dan citra yang baik
21. Mengesankan seperti restoran bertaraf internasional
22. Sering memasang iklan di media
23. Sering menawarkan undangan makan gratis
24. Turut berpartisipasi dalam acara penting/besar
25. Menawarkan diskon untuk pembayaran menggunakan kartu kredit
26. Mengadakan lomba makan dengan hadiah menarik
27. Mengadakan program makan sepuasnya.

Setelah semua kriteria dan atribut ditentukan dengan baik, responden diarahkan untuk menentukan seberapa besar ekspektasi mereka terhadap atribut yang ada dengan menggunakan skala *likert*.

Selain mendapatkan data-data preferensi tersebut, pelanggan juga diminta untuk mencantumkan data demografi yang terdiri atas:

1. Pekerjaan
2. Lokasi tempat tinggal
3. Jenis kelamin
4. Pendapatan rata-rata perbulan (atau uang saku bagi yang belum bekerja)
5. Frekuensi makan di restoran perbulan (baik pribadi maupun bersama keluarga)
6. Rata-rata pengeluaran yang digunakan untuk setiap kali makan di restoran cepat saji.

Data-data demografi ini ditujukan untuk mengetahui segmen pasar seperti apa yang sebaiknya dibidik oleh Papa Ron's Pizza untuk dijadikan sebagai pasar utama. Membentuk *niche* market tentunya akan menjadi sebuah solusi untuk pasar yang matang ini.

3.3 Pengolahan Data

Sebelum diolah, sebaiknya data dianalisis terlebih dahulu validitas (mengukur apa yang diukur) dan reliabilitasnya (kekonsistenan pengukuran). Untuk validitas, digunakan *validity content*, yaitu kuesioner diperlihatkan kepada pihak yang berkompeten. Untuk reliabilitas, digunakan parameter Cronbach's Alpha:

Tabel 3.1. Analisis Reabilitas Cronbach's Alpha

Reliability Statistics	
Cronbach's Alpha	N of Items
.900	38

Terlihat bahwa nilai Cronbach's Alpha lebih besar dari 0,7 dan data dinyatakan *reliable*.

3.3.1 K-Means Cluster Analysis

Pada bagian ini, responden ini akan dikelompokkan menjadi beberapa kelas terlebih dahulu. Hal ini ditujukan untuk mengelompokkan responden yang karakternya hampir mirip. Basis data yang dikelompokkan adalah hasil preferensi responden terhadap skala *likert* yang diberikan pada masing-masing atribut. Adapun tahapan yang harus dilalui pada analisis *cluster* adalah:

1. Menilai perlu tidaknya melakukan transformasi data

Pada data yang ada, terlihat isi data sangatlah bervariasi dalam satuan. Terdapat atribut yang cenderung sangat disukai, dan atribut yang mendapatkan respon yang tidak begitu disukai. Perbedaan yang mencolok seperti ini akan menyebabkan bias dalam analisis *cluster*. Oleh karena itu, data asli harus ditransformasikan (standarisasi) sebelum bisa dianalisis. Dengan demikian, perlu dilakukan transformasi terhadap variabel yang relevan ke bentuk *z-score*. Adapun tahapan-tahapan proses standarisasi bisa dilihat di Lampiran D. Berikut adalah output yang menginformasikan data deskriptif dari variabel-variabel, yang akan digunakan saat penafsiran hasil *cluster* :

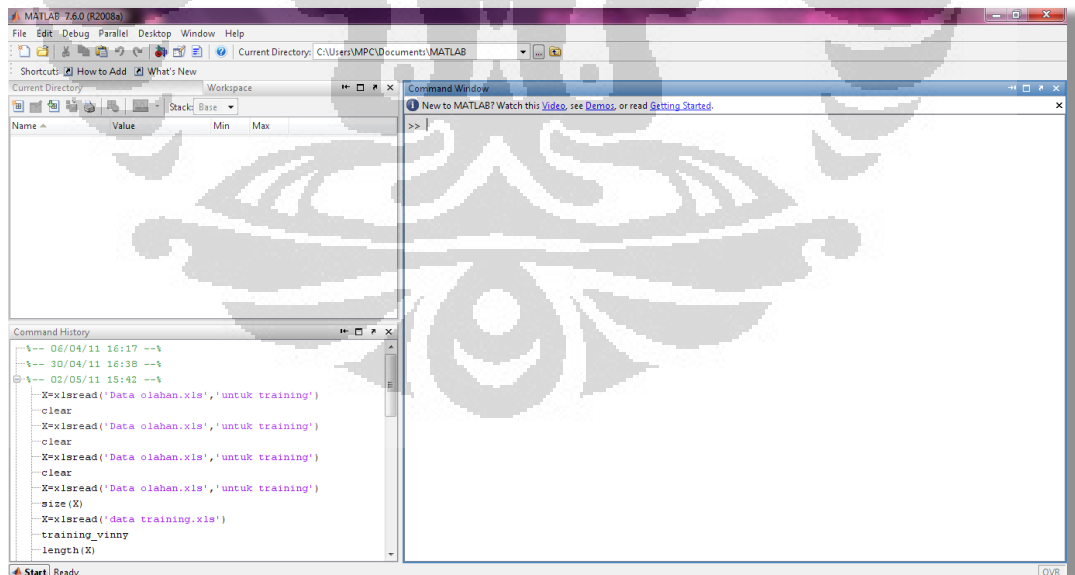
Tabel 3.2 Data Deskriptif Standarisasi

2. Analisis Cluster

Setelah data distandarisasi, proses dilanjutkan dengan memilih metode pembentukan *cluster*, yakni metode *non-hierarchical cluster* atau *K-Means Cluster*. Hasil nilai standarisasi ini akan dikelompokkan menjadi 2 kelas terlebih dahulu. Selanjutnya, proses pembagian kelas ini dilanjutkan hingga 5 kelas dan keanggotaan kelas masing-masing responden disimpan untuk dilanjutkan ke proses *training neural network*.

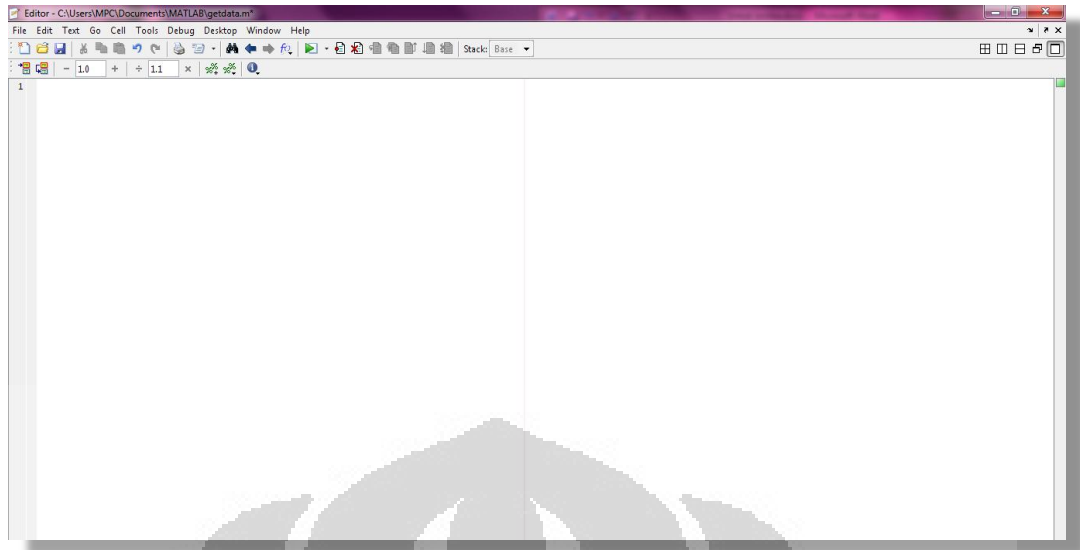
3.3.2 Neural Network Backpropagation

Setelah responden-responden yang mirip dikelompokkan, maka tahap selanjutnya adalah memasukkan data ke dalam model *neural network backpropagation*. Sebelum memulai proses pemrograman, akan dibahas terlebih dahulu mengenai software *Matlab 7.6.0 (R2008a)*. Berikut adalah tampilan layar utama dari software ini:



Gambar 3.1 Tampilan Layar Utama Matlab

Selain itu, terdapat layar lain yang terbuka setelah layar ini terbuka, yaitu layar **editor**. Layar ini berfungsi sebagai tempat memasukkan perintah atau kode program yang akan dijalankan nantinya. Berikut adalah tampilan layar **editor**:



Gambar 3.2 Tampilan Layar **Editor** Matlab

Untuk lebih jelasnya, berikut adalah sederetan tahapan yang dilalui dari membuat program hingga data berhasil di-*train*:

1. Proses Pembuatan Program.

Pada penelitian ini, terdapat 3 buah perintah utama untuk mengolah keseluruhan data, yaitu program pengambilan data, perintah *backpropagation*, dan program untuk mengolah data. Keseluruhan program ini ditujukan untuk men-*train* dan men-*test* data yang ada.

2. Proses *Training*

Proses *training* ini ditujukan untuk membuat *neural network* mengetahui karakter dari tiap kelas dengan cara ‘mengenali’ dan ‘menghafalkan’ nilai-nilai *z-score* dari masing-masing responden pada setiap kelas.

3. Proses *Testing*

Selanjutnya, proses *testing* ditujukan untuk ‘menguji’ *neural network* apakah ia bisa mengenal data-data responden baru yang terdapat pada masing-masing kelas. Untuk lebih jelasnya, ketiga tahapan ini bisa dilihat di Lampiran F.

3.4 Hasil Pengolahan Data

Setelah proses *training* dan *testing* dilakukan, berikut adalah hasil dari *recognition rate* dan *time* yang dihasilkan dari pembagian sejumlah kelas:

Tabel 3.3 Hasil *Recognition Rate* dan *Time*.

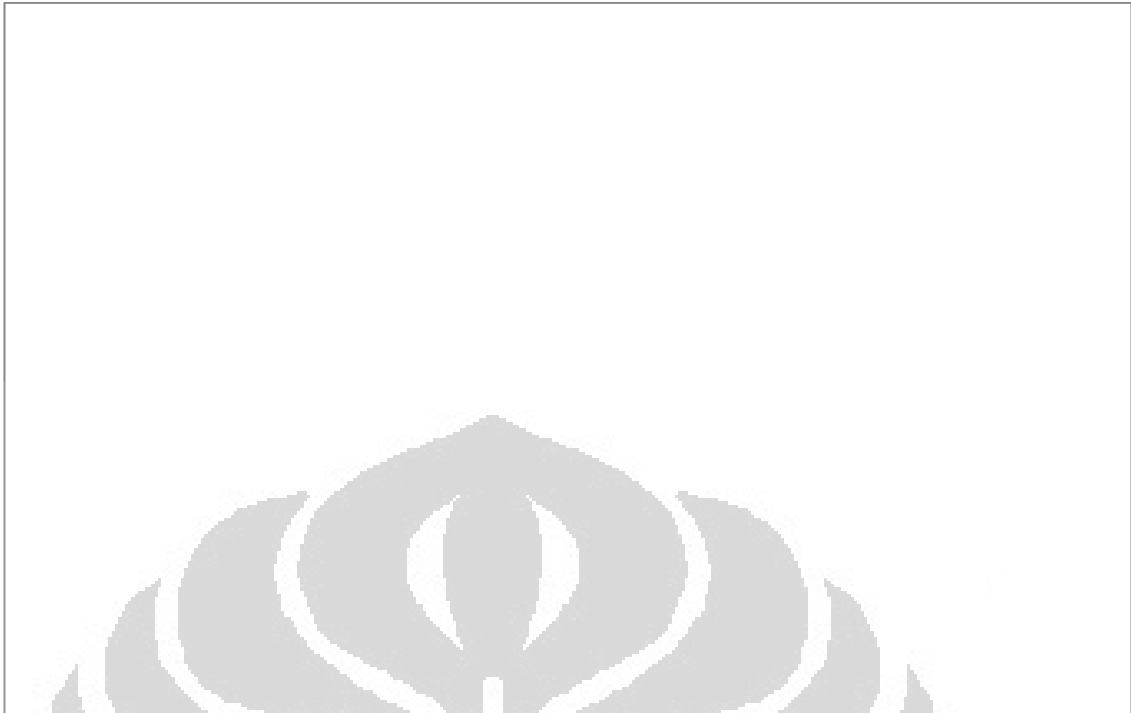
Recognition rate data train ini menunjukkan persentase kemampuan dari model *neural network* untuk ‘menghafalkan’ karakter pada kelas-kelas yang diberikan. Di sisi lain, *recognition rate data test* menunjukkan persentase kebenaran dari model untuk ‘mengenali’ keanggotaan dari data *z-score* responden yang diberikan/dites kepadanya. Bisa dilihat bahwa persentase *recognition rate data test* menurun seiring bertambahnya kelas.

Selanjutnya, *training time* adalah waktu yang dibutuhkan oleh model untuk ‘menghafal’ karakter kelas, sedangkan *testing time* adalah waktu yang dibutuhkan oleh model untuk ‘mengenali’ keanggotaan dari data *z-score* responden yang diberikan/dites kepadanya. *Training time* cenderung meningkat seiring bertambahnya kelas, sedangkan untuk *testing time* pembagian tiga kelas menunjukkan waktu yang terbaik. Kemudian, *error total* ini menunjukkan kesalahan dari model secara keseluruhan.

3.5 Hasil Data Demografi

Pada bagian ini, akan dibahas lebih lanjut mengenai data-data demografi yang didapatkan dari pembagian kuesioner. Berikut adalah presentasi dari data demografi yang didapatkan:

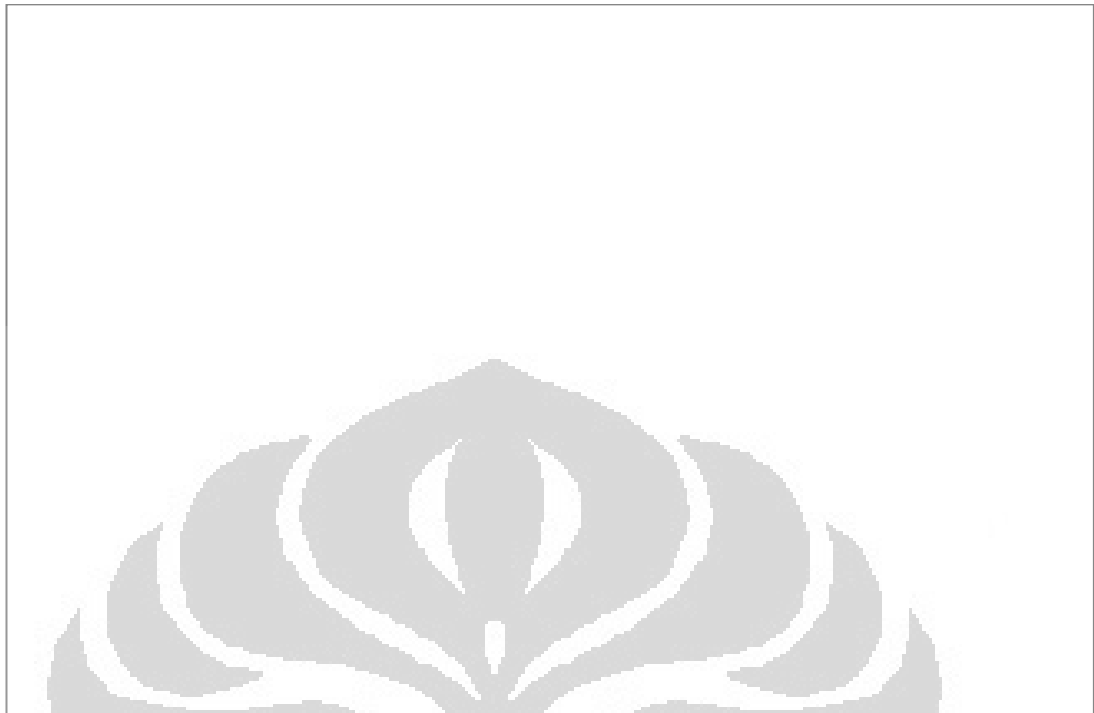
- Pekerjaan



Gambar 3.3 Data Demografi Pekerjaan Responden

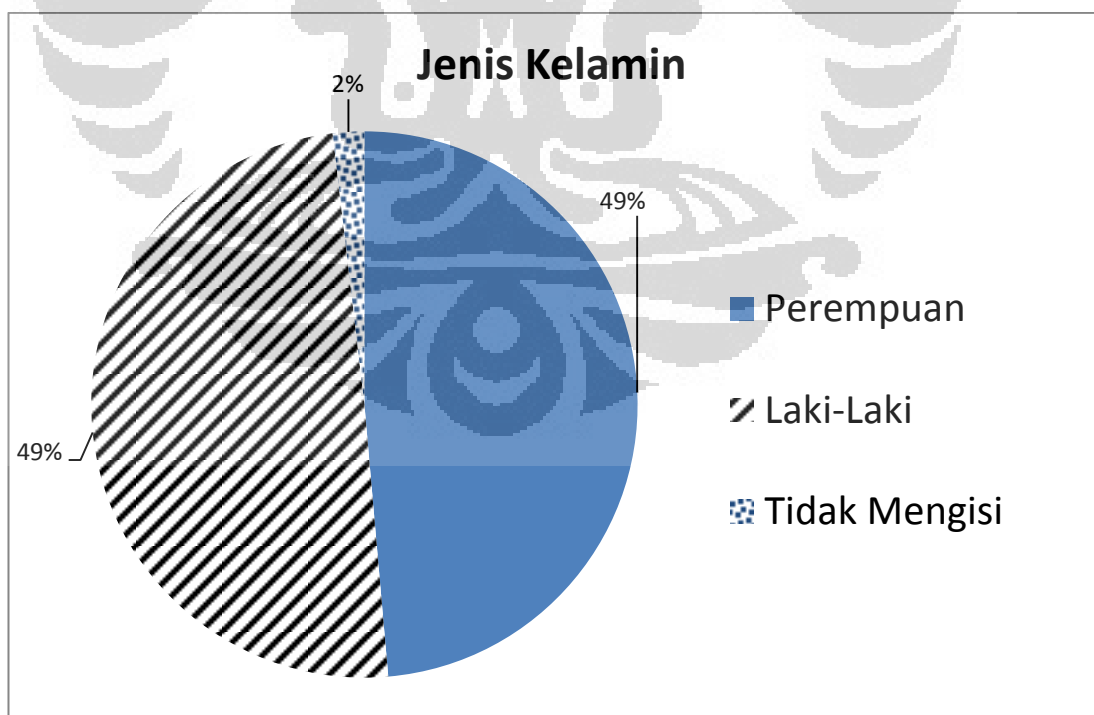


- Lokasi tempat tinggal responden saat ini.



Gambar 3.4 Data Demografi Lokasi Tempat Tinggal

- Jenis Kelamin



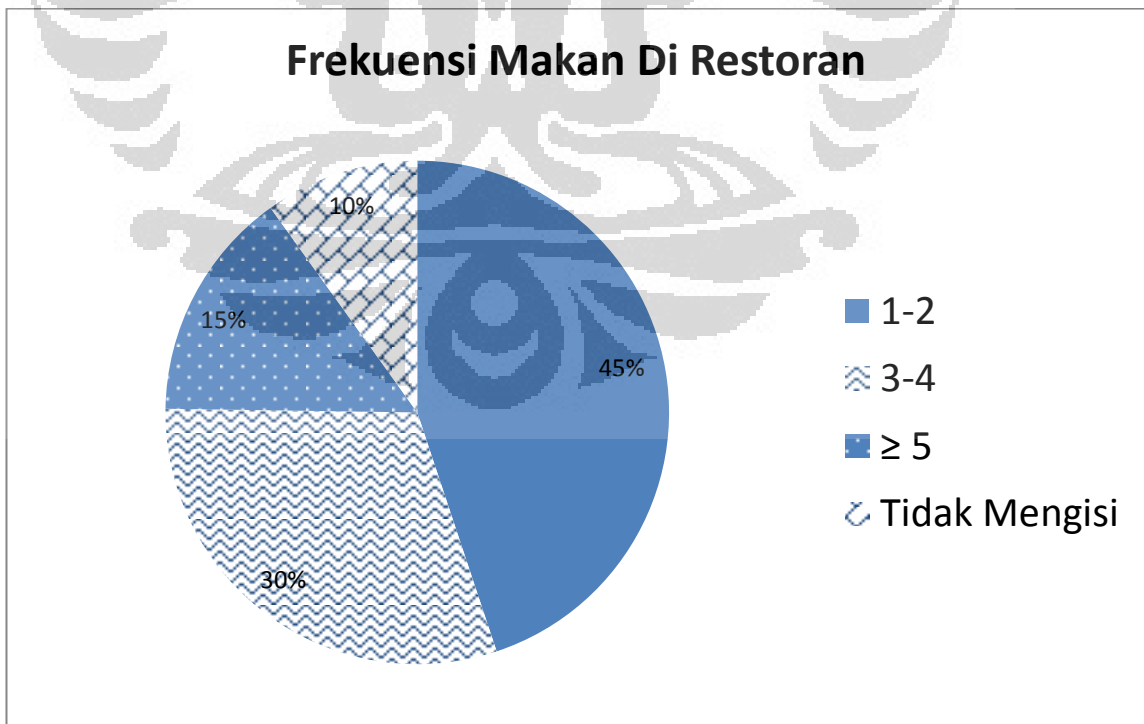
Gambar 3.5 Data Demografi Jenis Kelamin

- Pendapatan rata-rata perbulan (atau uang saku bagi yang belum bekerja)



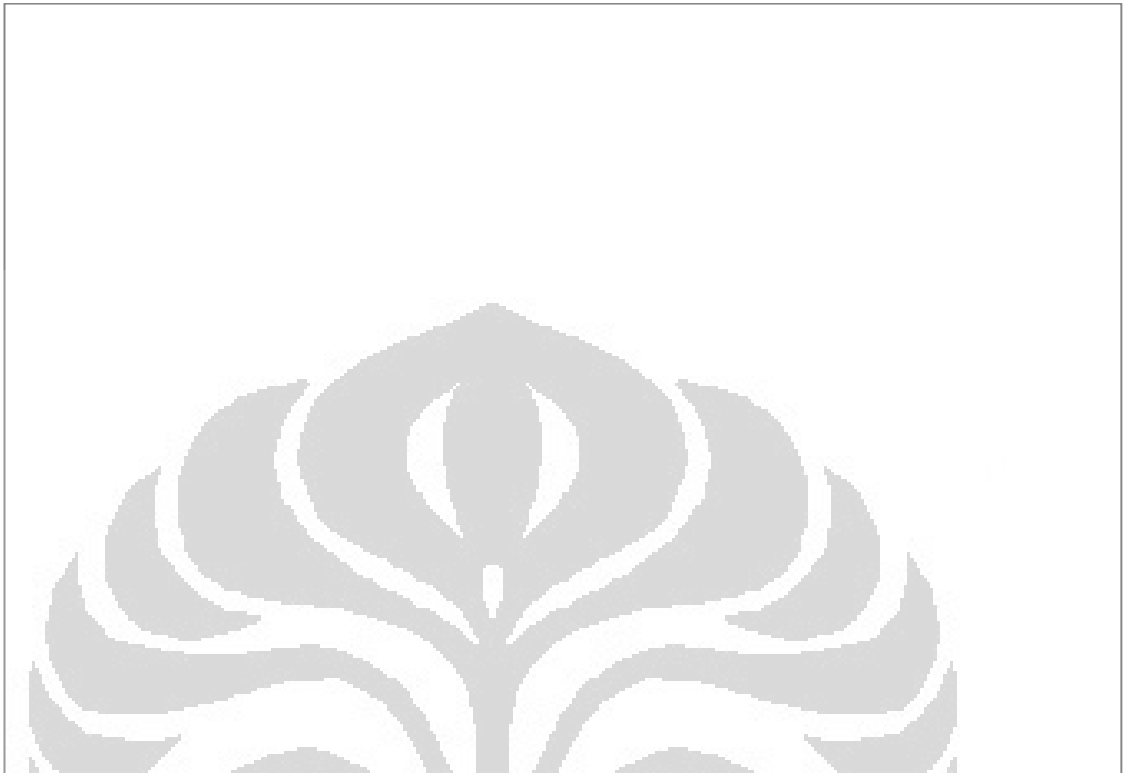
Gambar 3.6 Data Demografi Pendapatan

- Frekuensi makan di restoran perbulan (baik pribadi maupun bersama keluarga)



Gambar 3.7 Data Demografi Frekuensi Makan di Restoran

- Rata-rata pengeluaran yang digunakan untuk setiap kali makan di restoran cepat saji.

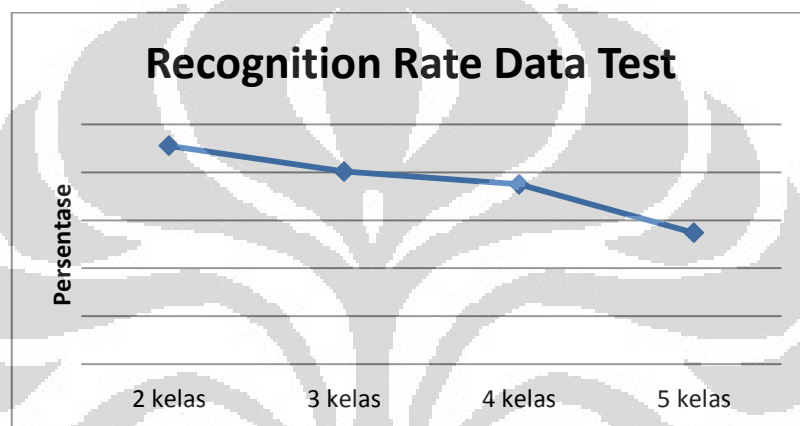


Gambar 3.8 Data Demografi Rata-Rata Pengeluaran untuk Setiap Kali Makan di Restoran

ANALISIS

4.1 Analisis Neural Network

Dari hasil pengolahan data pada bab sebelumnya, terlihat dengan jelas bahwa semakin banyak pembagian kelas, maka semakin kecil pula jumlah anggota di dalam satu kelas. Anggota yang semakin sedikit ini akan menurunkan kemampuan *neural network* untuk bisa mengenali keanggotaan responden dengan baik. Hal ini terlihat dari persentase *recognition rate data test* yang dihasilkan oleh model:

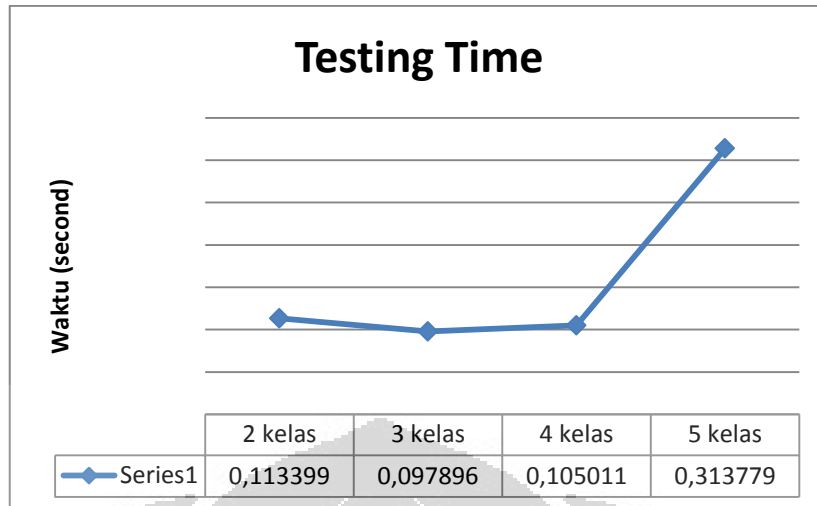


Gambar 4.1 *Recognition Rate Data Test*

Selanjutnya, waktu yang dibutuhkan oleh *neural network* untuk menghafalkan karakter anggota dalam satu kelas juga meningkat seiring bertambahnya jumlah kelas.

Gambar 4.2 *Waktu Training*

Akan tetapi, pada proses *testing*, waktu yang dibutuhkan model untuk mengenali karakter anggota dari 3 kelas bernilai lebih kecil dibandingkan kelas lain.



Gambar 4.3 Waktu *Testing*

Hal ini menunjukkan bahwa karakter anggota kelas pada pembagian 3 kelas memang sama. Oleh karena itu, *neural network* bisa dengan cepat mengenali keanggotaan dari masing-masing responden yang ada.

Selanjutnya, hal yang perlu ditelaah lebih lanjut adalah mencari tahu karakter seperti apa yang dimiliki pada masing-masing kelas. Perusahaan nantinya bisa menentukan strategi pemasaran yang tepat untuk responden yang berada pada kelas terbesar dalam rangka meraup *market share* yang lebih besar. Dengan demikian, hal yang harus ditelaah lebih lanjut adalah menganalisis *final cluster centers* pada analisis *clustering*.

Tabel 4.1 *Final Cluster Centers*

Tabel 4.1 *Final Cluster Centers (Lanjutan)*

Output ini merupakan akhir dari proses *clustering*. Angka pada Tabel 4.1 mengacu pada angka z , dengan ketentuan:

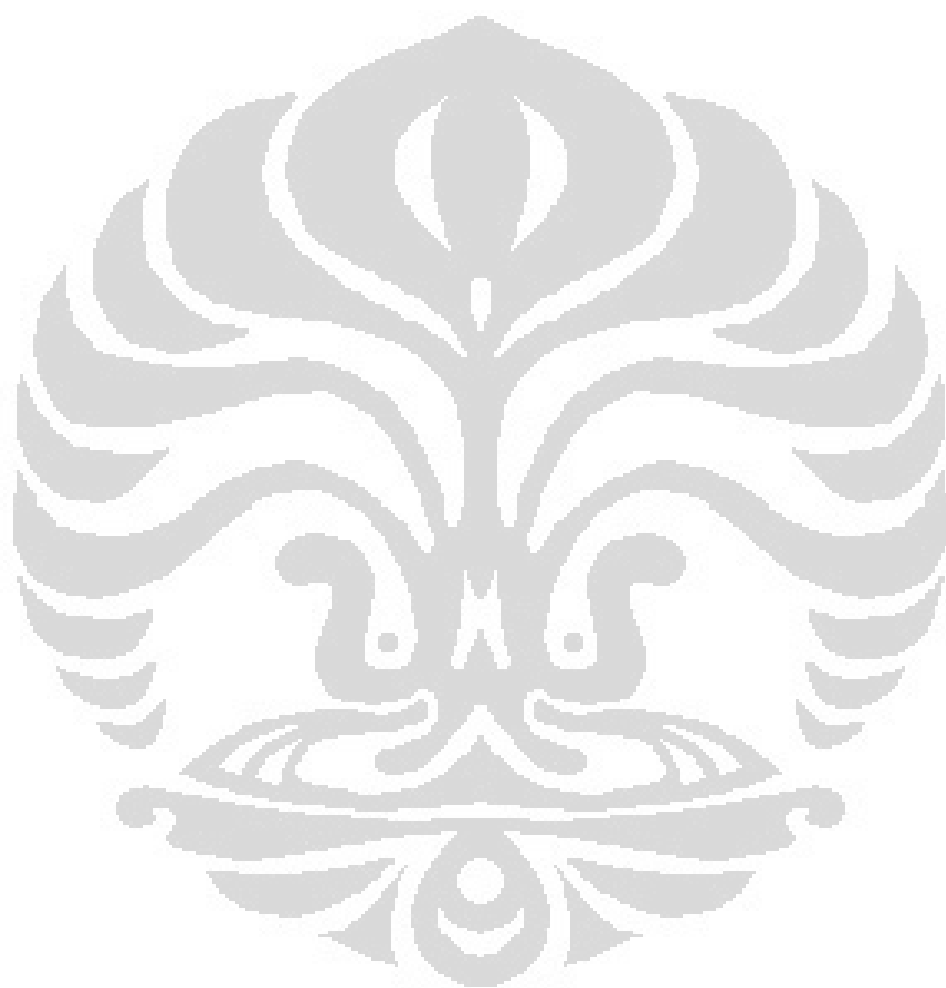
- Angka negatif berarti data di bawah rata-rata total
- Angka positif berarti data di atas rata-rata total.

Sebagai contoh, angka 0,28494 pada atribut pertama pada kelas pertama menyatakan rata-rata nilai yang diberikan responden pada kelas pertama terhadap atribut yang pertama adalah:

$$X = \mu + z \cdot \sigma \quad (4.1)$$

Maka rata-rata nilai yang diberikan responden terhadap atribut pertama di kelas 1 (dengan mengacu pada Tabel 3.2) adalah $4,4489 + 0,28494 (0,72212) = 4,654$. Untuk melakukan proses penganalisan yang lebih mendalam, perlu dilakukan penghitungan nilai rata-rata preferensi yang diberikan oleh masing-masing kelas. Hasil ini diperlihatkan pada Tabel 4.2, Tabel 4.3, dan Tabel 4.4.

Selanjutnya, untuk lebih mempermudah proses penganalisan, maka akan lebih baik jika nilai z -score dipresentasikan ke dalam sebuah grafik. Berikut adalah hasil z -score yang telah direpresentasikan ke dalam grafik:

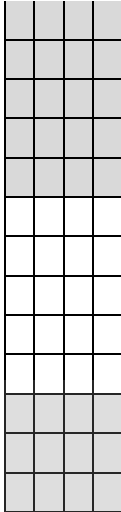


Berikut adalah analisis umum dari masing-masing kelas:

- Kelas 1 (Kelas Penggemar Suasana Menyenangkan dan Makanan Sehat)
Kelas ini memiliki anggota yang paling banyak dibandingkan dengan dua kelas yang lain. Jika dilihat dari grafik, kelas ini lebih cenderung memiliki nilai *z-score* yang paling besar pada atribut-atribut yang tergolong pada kriteria diberikan nilai rata-rata yang paling tinggi, kelompok ini memberikan nilai rata-rata preferensi tertinggi untuk makanan yang sehat dan nilai rata-rata terendah untuk menu makanan yang terus berganti sesuai jadwal makan, serta pizza yang tipis. Di sisi lain, golongan ini cenderung kurang menyukai program lomba makan dengan hadiah menarik dan

Tabel 4.2 Peringkat Atribut yang Dipilih Oleh Kelas 1

Tabel 4.2 Peringkat Atribut yang Dipilih Oleh Kelas 1 (Lanjutan)



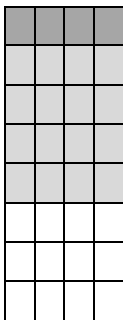
Karena kelas 1 memiliki jumlah responden paling banyak, maka hendaknya perhatian lebih diberikan kepada karakter kelas 1 sehingga perusahaan bisa berkesempatan untuk mendapatkan *market share* yang lebih baik. Dengan adanya data di Tabel 4.2, perusahaan bisa lebih memprioritaskan kebutuhan-kebutuhan apa saja yang harus dipenuhi terlebih

i terhadap 38 atribut yang diberikan:

Tabel 4.3 Peringkat Atribut yang Dipilih Oleh Kelas 2

Peringkat	Atribut	Keterangan Atribut	Nilai Rata-Rata
1	38	Mengadakan program makan sepuasnya.	4,025298
2	37	Mengadakan lomba makan dengan hadiah yang menarik	3,987339
3	29	Pelayanan dari karyawan yang terampil	3,987314
4	36	Menawarkan diskon untuk pembayaran dengan menggunakan kartu kredit	3,974666
5	19	Menawarkan makanan yang sehat	3,911426
6	14	Menawarkan menu yang beraneka ragam	3,835408

Tabel 4.3 Peringkat Atribut yang Dipilih Oleh Kelas 2 (Lanjutan)



Tabel 4.4 Peringkat Atribut yang Dipilih Oleh Kelas 3 (Lanjutan)

Peringkat	Atribut	Keterangan Atribut	Nilai Rata-Rata
34	16	Menyediakan minuman hangat yang beraneka ragam	4,238113
35	25	Menawarkan <i>dessert</i> yang hangat	4,133301
36	37	Mengadakan lomba makan dengan hadiah yang menarik	4,104758
37	20	Menawarkan pizza yang lebih tipis	4,076233
38	33	Sering memasang iklan di media	4,057095

4.2 Analisis Data Demografi

Selain analisis *cluster* yang dilakukan sebelumnya, berikut adalah analisis data demografi yang bisa menjelaskan segmen pasar pelanggan Papa Ron's Pizza:

1. Pekerjaan

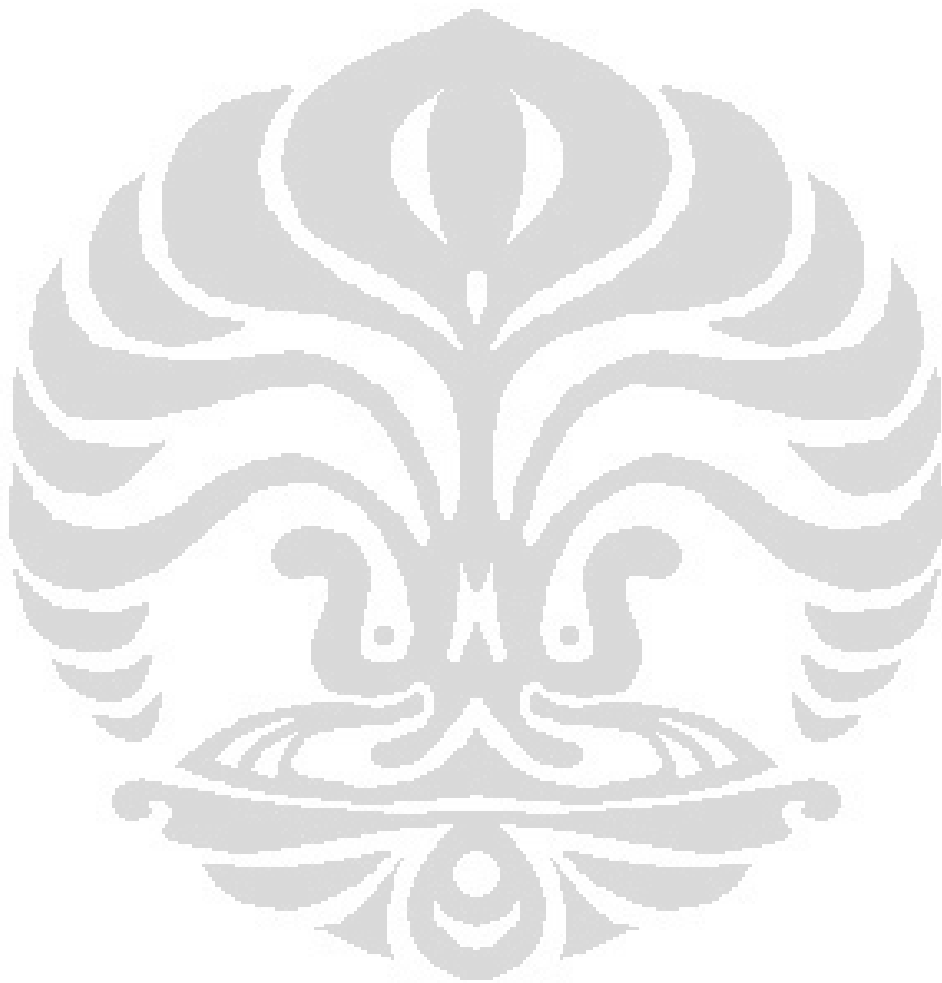
Dari 323 responden yang didapatkan, 45% di antaranya adalah karyawan. Oleh karena itu, perusahaan bisa menjadikannya sebagai target pasar utama. Pendirian gerai di daerah yang berdekatan dengan pusat perkantoran bisa menjadi sebuah strategi. Selain itu, perusahaan juga harus

Pada data demografi ini, tidak terdapat informasi signifikan yang bisa dijadikan sebagai strategi pemasaran.

2. Pendapatan rata-rata perbulan (atau uang saku bagi yang belum bekerja)

Dari data yang didapatkan dari pelanggan Papa Ron's Pizza, didapatkan bahwa sekitar 26% memiliki pendapatan Rp.2.501.000-Rp. 5.000.000. dan 22% memiliki pendapatan Rp. 1.301.000 – Rp 2.500.000,-. Dengan

demikian, informasi ini bisa menjadi masukan bagi perusahaan bahwa perusahaan hendaknya membidik target pasar menengah ke bawah.



3. Frekuensi makan di restoran perbulan
45% responden menjawab bahwa mereka hanya makan sekitar 1-2 kali di restoran per bulan, baik pribadi maupun bersama keluarga. Informasi ini memperkuat informasi bahwa target pasar perusahaan memang golongan masyarakat menengah ke bawah.
4. Rata-rata pengeluaran yang digunakan untuk setiap kali makan di restoran cepat saji.

besar. Selain itu, penelitian lanjutan untuk mempertajam karakter *niche market* bisa dijadikan sebagai solusi bagi perusahaan. Dengan demikian, perusahaan bisa benar-benar tahu karakter pasar ini dan menyediakan kebutuhan yang tepat untuk memenuhinya.



BAB 5 KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian, model yang dihasilkan, memiliki kemampuan mengenal yang berbeda jika jumlah pembagian kelas berbeda. Semakin banyak jumlah kelas yang dihasilkan, kemampuan model melakukan *training* data akan menurun. Akan tetapi, tidak begitu halnya pada proses *testing*.

Proses *testing* yang dilakukan oleh model *neural network* terhadap data yang ada, menunjukkan bahwa model bisa mengenal kelas yang berjumlah tiga dengan waktu yang paling singkat. Hal ini menunjukkan bahwa karakter anggota masing-masing kelas adalah sama sehingga model bisa dengan cepat mengenali keanggotaan responden. Selain itu, jumlah keanggotaan masing-masing kelas bisa membantu perusahaan untuk memprioritaskan preferensi konsumen.

Kelas yang memiliki anggota terbanyak adalah kelas satu dengan anggota 139 orang. Perusahaan hendaknya memberikan perhatian lebih terhadap karakter kelas satu. Selanjutnya, pada penggunaan analisis *cluster*, terlihat bahwa kelas 1 adalah kelompok yang cenderung memilih hidup sehat dan sangat memperhatikan kebersihan, kenyamanan, serta pelayanan dari sebuah restoran. Kelas ini memberikan rata-rata nilai preferensi lebih dari empat kepada 15 atribut. Dengan demikian, perusahaan bisa mempertajam karakter *niche market*.

Di sisi lain, data demografi yang didapatkan, bisa dijadikan strategi pendukung. Dari data ini, bisa dilihat bahwa perusahaan ini memiliki pelanggan yang lebih banyak berstatus sebagai karyawan dan golongan menengah ke bawah. Perusahaan harus menelaah lebih lanjut, keinginan dan karakter pasar ini. Selanjutnya, data demografi juga menunjukkan bahwa responden lebih banyak berasal dari Jakarta Utara dan Bekasi sehingga perusahaan bisa memperluas jaringan di kedua lokasi ini.

Sebagai penutup, model yang dibuat telah memungkinkan penggunaan atribut yang banyak untuk mengetahui berapa pembagian kelas yang sebaiknya dilakukan. Kemudian, model ini nantinya bisa digunakan untuk memprediksi responden-responden berikutnya jika diinginkan. Selanjutnya dengan bantuan

cluster analysis, perusahaan bisa mengetahui karakter yang lebih jauh dari masing-masing kelas sehingga karakter *niche market* bisa diketahui.

5.2 Saran

Model *neural network* yang dijadikan sebagai sarana penelitian adalah model *feedforward* yang hanya terdiri atas sebuah *hidden layer*. Penggunaan model *recurrent* yang memungkinkan *signal* berjalan pada kedua arah dengan membuat *loop* pada jaringan, bisa dijadikan sebagai penelitian lanjutan. Hal ini tentunya bertujuan untuk menciptakan model yang lebih *powerful* dan tentunya lebih optimum.

Di sisi lain, penambahan jumlah *hidden layer* pada model bisa juga menjadi lanjutan dari penelitian ini. Penambahan jumlah *hidden layer* ini tentunya akan menambah kompleksitas model. Hal ini akan meningkatkan kemampuan model untuk menjalankan sejumlah perintah yang dikehendaki oleh pengguna nantinya. Jika kedua saran ini dilakukan, maka proses perbandingan bisa dilakukan dalam rangka menciptakan model yang paling *powerful* untuk mengenal pasar.

DAFTAR REFERENSI

- Astebro, Thomas, dan Dahlin, Kristina. (2003). *Opportunity Knocks. Economic Journal*.
- Borade, Gaynor. (2010). *History of The Fast Food Industry*. Mei 4, 2011. <http://www.buzzle.com/articles/history-of-the-fast-food-industry.html>
- Campbell, N.A, Reece, J.B, & Mitchell, L.G. (2004). Biologi (Rahayu Lestari , Penerjemah). Jakarta: PT Penerbit Erlangga.
- Engelbrecht, Andries P. (2007). *Computational Inteligence* (2nd Ed.). New York: A John Wiley & Sons, Inc., Publication.
- Hayashi, Y. , Hsieh, M. H., dan Setiono, R. (2009). Predicting Consumer Preference for Fast-Food Franchises: A Data Mining Approach. *Journal of the Operational Research Society*, 60, 1221-1229.
- Haykin, S. (1994) *Neural Networks: A Comprehensive Foundation*, Prentice Hall.
- Helm, R. , Manthey, L., Scholl, A., & Steiner M. (2005). Solving Multiattribute Design Problems with Analytic Hierarchy Process and Conjoint Analysis: An Empirical Comparison. *European Journal of Operational Research*, 164, 760-777.
- Larose, Daniel T. (2004). *Discovering Knowledge in Data*. New York: A John Wiley & Sons, Inc., Publication.
- Papa Ron's Pizza. (2011). *Company History and Profile*. Maret 8, 2011. <http://www.paparonspizza.com/corporate/company-history-profile/>
- Stephen. (2011, Januari). *Penerapan Algoritma Genetika Pada Backpropagation Untuk Mengenal Bunga Iris dan Wajah*. Paper presented at Seminar of Electrical Engineering, Universitas Indonesia.
- Surjandari, Isti. (2010). *Conjoint Analysis: Konsep dan Aplikasi*. Jakarta: Penerbit Universitas Trisakti.