

ANALISIS CHAID SEBAGAI ALAT BANTU STATISTIKA UNTUK SEGMENTASI PASAR (Studi Kasus pada Koperasi Syari'ah Al-Hidayah)

Yohanes Sondang Kunto

Staf Pengajar Fakultas Ekonomi, Universitas Kristen Petra, Surabaya

Siti Nurul Hasana

Peneliti Lepas – Alumnus Universitas Brawijaya, Malang

Abstrak: Segmentasi pasar adalah bahasan klasik di bidang pemasaran yang tidak pernah kehilangan daya tariknya. Dalam riset-riset segmentasi, ada dua metode pendekatan yaitu metode interdependensi dan dependensi. Statistika menyediakan banyak alat bantu untuk riset-riset segmentasi. Salah satu alat bantu statistika untuk riset segmentasi yang menggunakan pendekatan dependensi adalah analisis CHAID (*Chi-Squared Automatic Interaction Detection analysis*). Analisis CHAID akan menghasilkan diagram yang mirip dengan diagram pohon keputusan yang menyediakan informasi tentang derajat hubungan antara variabel dependen terhadap variabel independen dan informasi mengenai karakteristik segmen. Analisis CHAID mampu menangani variabel independen monotonik, bebas, dan mengambang. Pada kasus Koperasi Syari'ah Al-Hidayah, analisis CHAID digunakan untuk menentukan segmentasi nasabah berdasarkan status kredit nasabah sebagai variabel dependen dan data demografis nasabah sebagai variabel independen. Segmentasi nasabah dengan bantuan analisis CHAID dilakukan untuk mengidentifikasi segmen nasabah yang memiliki rasio kredit macet yang paling rendah. Berdasarkan hasil analisis CHAID didapatkan empat segmen. Segmen nasabah dengan karakteristik mempunyai penghasilan rata-rata sebesar $\leq$ Rp. 1.500.000,00, berjenis kelamin perempuan, serta berusia > 30 tahun adalah segmen yang memiliki rasio kredit macet yang paling rendah.

Kata kunci:

Abstract: Market segmentation is a classic topic in marketing which is never loss its attractiveness. In segmentation research, there is two approach methods which is interdependency method and dependency method. Statistics provides many tools for segmentation research. One of statistical tool for segmentation research which takes the dependency method as an approach is CHAID analysis (*Chi-Squared Automatic Interaction Detection analysis*). CHAID analysis would provide decision tree like diagram which provide information about degree of association from dependent variable to the independent variables and the information about segments characteristic. CHAID analysis could handle monotonic, free, and floating independent variables. In Koperasi Syari'ah Al-Hidayah case, CHAID analysis was used to determine the segment of its relationship based on the relationship credit status as the dependent variables and demographics data as independent variables. Relationship segmentation using CHAID analysis was conducted to identify relationship segment which has the lowest ratio of unpaid credit. Using CHAID analysis, four segments were identified. Relationship segment which has monthly income less than Rp 1.500.000,00, female, and over 30 years old is the lowest ratio of unpaid credit segment.

Keywords:

PENDAHULUAN

Salah satu bahasan klasik dalam bidang pemasaran adalah segmentasi pasar. Sekalipun tergolong ke dalam bahasan yang klasik, segmentasi pasar tidak pernah tidak pernah kehilangan daya tariknya. Hal ini dikarenakan segmentasi pasar adalah kunci utama untuk mencapai strategi pemasaran yang ideal.

Segmentasi pasar secara sederhana dapat diartikan sebagai proses untuk membagi suatu populasi individu ke dalam kelompok atau segmen yang lebih kecil yang berbeda karakteristiknya satu

dengan yang lain. Segmen-segmen yang terbentuk tersebut akan digunakan sebagai acuan untuk menentukan target pasar.

Menurut Kasali (1998), secara garis besar ada dua metode yang lazim dipakai dalam riset-riset segmentasi dewasa ini, yaitu metode interdependensi dan metode dependensi. Dalam metode interdependensi, segmentasi hanya dibentuk oleh variabel-variabel bebas. Dalam metode ini, variabel-variabel bebas tidak dikaitkan dengan variabel terikat yang harus dijelaskan atau diprediksikan. Tujuannya adalah menemukan kelompok-kelompok konsumen

(segmen-segmen) yang memiliki kesamaan respons terhadap variabel-variabel bebas tertentu, di mana variabel-variabel bebas tersebut bersifat sangat luas dan umumnya berbentuk pernyataan-pernyataan.

Berbeda dengan metode interdependensi, dalam metode dependensi, segmen-segmen dihasilkan karena ada hubungan antara variabel terikat (*dependent variable*) dengan sejumlah variabel bebas (*independent variable*). Teknik yang dibutuhkan di sini adalah sebuah pendekatan yang dapat menunjukkan variabel-variabel independen yang memiliki derajat hubungan atau pengaruh kuat terhadap variabel terikat, dan mengidentifikasi segmen-segmen yang paling berbeda menurut variabel-variabel tersebut.

Menurut Lehmann dan Eherler (2000), salah satu metode dependensi yang sering digunakan dalam segmentasi pasar adalah analisis CHAID (*Chi-Squared Automatic Interaction Detection analysis*).

Dalam penelitian ini, analisis CHAID akan diterapkan untuk menentukan segmentasi nasabah koperasi Syari'ah Al-Hidayah yang berlokasi di desa Druju kecamatan Sumbermanjing Wetan, Malang. Tujuan dari segmentasi ini adalah mengidentifikasi segmen nasabah potensial dengan harapan resiko kredit macet dapat diminimumkan. Data yang digunakan adalah status kredit dan data demografis dari nasabah tahun 2002-2003.

TINJAUAN PUSTAKA

Analisis CHAID

CHAID adalah singkatan dari *Chi-squared Automatic Interaction Detector*. CHAID pertama kali diperkenalkan dalam sebuah artikel berjudul "*An Exploratory Technique for Investigating Large Quantities of Categorical Data*" oleh Dr. G.V. Kass tahun 1980. Prosedurnya merupakan bagian dari teknik terdahulu yang dikenal dengan *Automatic Interaction Detector* (AID), dan menggunakan statistik *chi-square* sebagai alat utamanya.

CHAID secara keseluruhan bekerja untuk menduga sebuah variabel tunggal, disebut sebagai variabel dependen, yang didasarkan pada sejumlah variabel-variabel yang lain, disebut sebagai variabel-variabel independen. CHAID merupakan suatu teknik iteratif yang menguji satu-persatu variabel independen yang digunakan dalam klasifikasi, dan menyusunnya berdasarkan pada tingkat signifikansi statistik *chi-square* terhadap variabel dependennya (Gallagher, 2000).

CHAID digunakan untuk membentuk segmentasi yang membagi sebuah sampel menjadi dua atau lebih kelompok yang berbeda berdasarkan sebuah

kriteria tertentu. Hal ini kemudian diteruskan dengan membagi kelompok-kelompok tersebut menjadi kelompok yang lebih kecil berdasarkan variabel variabel independen yang lain. Prosesnya berlanjut sampai tidak ditemukan lagi variabel independen – variabel independen yang signifikan secara statistik. Segmen-segmen yang dihasilkan akan bersifat saling lepas yang secara statistik akan memenuhi kriteria pokok segmentasi dasar (Bagozzi, 1994). Hasilnya juga akan memberikan peringkat pada variabel yang merupakan variabel independen paling signifikan sampai yang tidak signifikan.

CHAID memilih variabel-variabel variabel independennya atas dasar uji *chi-square* antara kategori variabel-variabel yang tersedia dengan kategori-kategori variabel dependennya (seperti yang terdapat pada statistika dasar bahwa uji *chi-square* merupakan uji non parametrik yang sesuai untuk menguji hubungan antar variabel yang berbentuk kategori) (Myers, 1996).

Pada dasarnya, dari beberapa definisi CHAID di atas dapat disimpulkan bahwa CHAID adalah sebuah metode untuk mengklasifikasikan data kategori di mana tujuan dari prosedurnya adalah untuk membagi rangkaian data menjadi subgrup-subgrup berdasarkan pada variabel dependennya (Lehmann dan Eherler, 2001).

Menurut Baron dan Phillips (Sharp et al., 2002), analisis CHAID dapat diringkas menjadi 3 elemen kunci, yaitu:

1. Uji signifikan *chi-square*, uji ini dilakukan untuk mengidentifikasi variabel independen yang paling signifikan dalam data.
2. Koreksi Bonferroni.
3. Sebuah algoritma yang digunakan untuk menggabungkan kategori-kategori variabel.

Variabel-Variabel dalam Analisis CHAID

Dalam analisis CHAID variabel yang digunakan dibedakan atas variabel terikat (variabel dependen) dan variabel bebas (variabel independen). Klasifikasi dalam CHAID dilakukan berdasarkan pada hubungan yang ada antara kedua variabel tersebut, oleh karena itu CHAID termasuk dalam metode dependensi dalam menentukan segmentasi.

Menurut Gallagher (2000), CHAID akan membedakan variabel variabel independennya menjadi tiga bentuk yang berbeda, yaitu:

1. **Monotonik:** kategori-kategori pada variabel ini dapat dikombinasikan atau digabungkan oleh CHAID hanya jika keduanya berdekatan satu sama lain, yaitu variabel-variabel yang kategori-

nya mengikuti urutan aslinya (data ordinal), contohnya: usia atau pendapatan.

2. **Bebas:** kategori-kategori pada variabel ini dapat dikombinasikan atau digabungkan walaupun keduanya berdekatan atau tidak satu sama lain (data nominal), contohnya: pekerjaan, kelompok etnik, dan area geografis.
3. **Mengambang (*floating*):** kategori-kategori pada variabel ini akan diperlakukan seperti monotonik kecuali untuk kategori yang terakhir (yaitu *missing value*), yang dapat berkombinasi dengan kategori manapun.

Deskripsi Matematis Analisis CHAID

Menurut Gallagher (2000), CHAID pada dasarnya merupakan sebuah proses 4 langkah yang iteratif.

- Pemeriksaan tiap variabel independen menggunakan uji *chi-square* untuk menentukan kategori mana yang nantinya signifikan untuk menunjukkan perbedaan dalam variabel dependen; dan mengumpulkan pula semua kategori yang tidak signifikan
- Penentuan variabel independen mana yang paling signifikan, yang terbaik untuk digunakan dalam membedakan variabel dependen berdasarkan nilai kesignifikan hasil uji yang dilakukan.
- Pembagian data menggunakan kategori variabel independen tersebut dengan peringkat yang paling signifikan.
- Untuk tiap tingkatan selanjutnya:
 - Pemeriksaan kategori variabel-variabel independen yang tersisa untuk menentukan peringkat yang paling signifikan dalam penentuan perbedaan variabel dependen selanjutnya, dan memisahkannya dengan yang tidak signifikan.
 - Penentuan variabel independen mana yang paling signifikan dan kemudian diteruskan lagi dengan pembagian datanya menggunakan variabel ini.
- Pengulangan langkah ke-4 untuk semua subgrup sampai teridentifikasi semua pembagian yang secara statistik telah signifikan.

Sedangkan Magidson dalam Bagozzi (1994), menerangkan bahwa langkah-langkah analisis CHAID secara garis besar dapat dibagi menjadi tiga tahap, yaitu Penggabungan, Pemisahan dan Penghentian. Tahap-tahap tersebut dapat dijabarkan sebagai berikut :

Tahap 1: Penggabungan

Untuk tiap variabel independen, $X_1, X_2, \dots, X_k$

- Bentuk tabel kontingensi dua arah dengan variabel dependennya.
- Hitung statistik *chi-square* untuk setiap pasang kategori yang dapat dipilih untuk digabung menjadi satu, untuk menguji kebebasannya dalam sebuah sub tabel kontingensi $2 \times J$ yang dibentuk oleh sepasang kategori tersebut dengan variabel dependennya yang mempunyai sebanyak J kategori.
- Untuk masing-masing nilai *chi-square* berpasangan, hitung *p-value* berpasangan bersamaan. Di antara pasangan-pasangan yang tidak signifikan, gabungkan sebuah pasangan kategori yang paling mirip (yaitu pasangan yang mempunyai nilai *chi-square* berpasangan terkecil) menjadi sebuah kategori tunggal, dan kemudian dilanjutkan ke langkah nomor 4. Tetapi apabila semua pasangan kategori yang tersisa adalah signifikan, lanjutkan ke langkah nomor 5.
- Untuk suatu kategori gabungan yang terdiri dari 3 kategori atau lebih, ujilah untuk melihat apakah suatu kategori variabel independen seharusnya dipisah dengan menguji kesignifikan antara kategori tersebut dengan kategori yang lain dalam satu kategori gabungan. Jika didapat nilai *chi-square* yang signifikan, pisahkan kategori tersebut dengan yang lain. Jika lebih dari satu kategori yang bisa dipilih untuk dipisah, pisahkan salah satu yang mempunyai nilai *chi-square* tertinggi. Kemudian kembali ke langkah nomor 3.
- Dengan cara memilih, gabungkan suatu kategori yang mempunyai sedikit pengamatan yang tidak sesuai dengan kategori lain yang paling mirip, seperti yang diukur oleh nilai *chi-square* berpasangan yang terkecil.
- Hitung *p-value* terkoreksi Bonferroni didasarkan pada tabel yang telah digabung.

Tahap 2: Pemisahan

- Pilihlah variabel independen terbaik, yaitu variabel independen dengan nilai *p-value* yang terendah, dan kemudian melakukan pembagian kelompok dengan variabel independen ini (yaitu gunakan masing-masing kategori-kategori variabel independen tersebut, yang telah digabung secara optimal, untuk menentukan sub pembagian dari kelompok induk menjadi sub kelompok yang baru). Jika tidak ada variabel independen dengan nilai *p-value* yang signifikan, jangan memulai pembagian kelompok tersebut.

Tahap 3: Penghentian

- Kembali ke langkah nomor 1 untuk menganalisis sub kelompok berikutnya. Hentikan ketika semua sub kelompok telah dianalisis dan juga telah berisi

pengamatan-pengamatan dengan jumlah yang terlalu sedikit.

Bagian-Bagian Utama dari Analisis CHAID

Uji *Chi-Square* (Khi-kuadrat, χ^2)

Teknik uji ini memungkinkan kita untuk mengetahui independensi antara dua variabel pada tiap levelnya. Misal suatu variabel pertama memiliki r kategori dan variabel kedua memiliki c kategori maka n_{ij} adalah pengamatan pada variabel pertama di level i dan variabel kedua di level j , secara umum tabel disajikan sebagai berikut:

Tabel 1. Struktur Data Uji *Chi-Square*

Kolom Baris	1	2	...	c	Total
1	n_{11}	n_{12}	...	n_{1c}	$n_{1\cdot}$
2	n_{21}	n_{22}	...	n_{2c}	$n_{2\cdot}$
.	.	.	...	.	.
.	.	.	...	.	.
.	.	.	...	.	.
r	n_{r1}	n_{r2}	...	n_{rc}	$n_{r\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$	...	$n_{\cdot c}$	n

Tabel 2. Probabilitas Sel

Kolom Baris	1	2	...	c	Total
1	p_{11}	p_{12}	...	p_{1c}	$P_{1\cdot}$
2	p_{21}	p_{22}	...	p_{2c}	$P_{2\cdot}$
.	.	.	...	.	.
.	.	.	...	.	.
.	.	.	...	.	.
r	p_{r1}	p_{r2}	...	p_{rc}	$P_{r\cdot}$
Total	$P_{\cdot 1}$	$P_{\cdot 2}$	...	$P_{\cdot c}$	P

di mana

p_{ij} adalah probabilitas kejadian irisan antara baris i dan kolom j

$P_{i\cdot}$ adalah probabilitas total pada baris ke- i

$P_{\cdot j}$ adalah probabilitas total kolom ke- j

Hipotesis pada pengujian *chi-square* adalah:

H_0 : $p_{ij} = p_{i\cdot}p_{\cdot j}$ (tidak terdapat hubungan antara baris dan kolom (bebas))

H_1 : $p_{ij} \neq p_{i\cdot}p_{\cdot j}$ (terdapat hubungan antara baris dan kolom (tidak bebas))

Sedangkan statistik ujinya adalah:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - E_{ij})^2}{E_{ij}} \quad \text{di mana}$$

$$E_{ij} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{n} \quad (1)$$

di mana

n_{ij} = banyaknya pengamatan pada baris ke- i dan kolom ke- j

E_{ij} = nilai harapan pengamatan pada baris ke- i dan kolom ke- j

$n_{i\cdot}$ = total banyaknya pengamatan pada baris ke- i

$n_{\cdot j}$ = total banyaknya pengamatan pada baris ke- j

n = total banyaknya responden

Keputusan yang diambil dari uji *chi-square* ini adalah

H_0 ditolak jika nilai $\chi^2_{\text{hit}} > \chi^2_{\text{tabel}}$ atau $p\text{-value} < \alpha$.

CHAID menggunakan statistik *chi-square* dalam dua cara. Yang pertama, statistik *chi-square* digunakan untuk menentukan apakah kategori-kategori dalam sebuah variabel independen bersifat seragam dan bisa digabungkan menjadi satu. Yang kedua, ketika semua variabel independen sudah diringkas menjadi bentuk yang signifikan dan tidak mungkin digabung lagi, kemudian statistik *chi-square* digunakan untuk menentukan variabel independen mana yang paling signifikan untuk membagi atau membedakan kategori-kategori dalam variabel dependen (Gallagher, 2000).

Koreksi Bonferroni (*Bonferroni Correction*)

Koreksi Bonferroni adalah suatu proses koreksi yang digunakan ketika beberapa uji statistik untuk kebebasan atau ketidakbebasan dilakukan secara bersamaan (Sharp et al., 2002). Koreksi Bonferroni biasanya digunakan dalam perbandingan berganda.

Ketika terdapat sebanyak M uji perbandingan yang sudah dikatakan bebas satu sama lain, peluang untuk melakukan kesalahan tipe 1 atau α (dalam satu atau lebih uji-uji tersebut), akan sama dengan 1 dikurangi peluang untuk tidak melakukan kesalahan tipe 1 dalam uji-uji tersebut, di mana nilainya akan lebih besar dari α yang telah ditentukan. Secara umum, hal tersebut dapat dirumuskan sebagai berikut (Bagozzi, 1994):

$$1 - (1 - \alpha)^M > \alpha \quad (2)$$

di mana

M = pengali Bonferroni

α = salah tipe 1

Pengali Bonferroni untuk masing-masing tipe variabel variabel independen adalah berbeda. Gallagher (2000) menyebutkan bahwa pengali Bonferroni untuk masing-masing jenis variabel variabel independen adalah sebagai berikut:

1. variabel independen Monotonik

$$M = \binom{c-1}{r-1} \quad (3)$$

• variabel independen Bebas

$$M = \sum_{i=0}^{r-1} (-1)^i \frac{(r-i)^c}{i!(r-i)!} \quad (4)$$

- variabel independen Mengambang (*Floating*)

$$M = \binom{c-2}{r-2} + r \binom{c-2}{r-1} \quad (5)$$

di mana

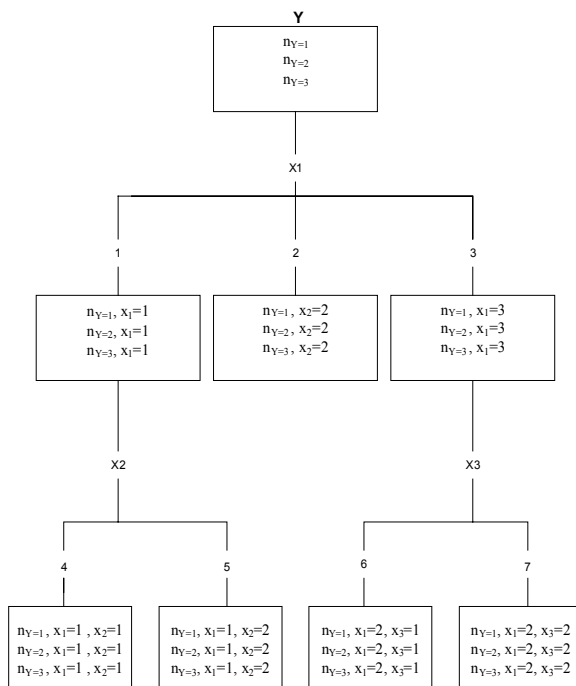
M = pengali Bonferroni

c = kategori variabel dependen

r = kategori variabel independen

Diagram Pohon (*Tree Diagram*)

Hasil pembentukan segmen dalam CHAID akan ditampilkan dalam sebuah diagram pohon. Secara umum diagram pohon dari CHAID adalah sebagai berikut (Lehmann dan Eherler, 2001):



Gambar 1. Diagram Pohon dalam Analisis CHAID

Diagram pohon CHAID mengikuti aturan “dari atas ke bawah” (*Top-down stopping rule*), di mana diagram pohon disusun mulai dari kelompok induk, berlanjut di bawahnya sub kelompok yang berturut-turut dari hasil pembagian kelompok induk berdasarkan kriteria tertentu (Myers, 1996). Tiap-tiap node dari diagram pohon ini menggambarkan sub kelompok dari sampel yang diteliti. Setiap node akan berisi keseluruhan sampel dan frekuensi absolut n_i untuk tiap kategori yang disusun di atasnya. Pada pohon klasifikasi CHAID terdapat istilah kedalaman

(*depth*) yang berarti banyaknya tingkatan node-node sub kelompok sampai ke bawah pada node sub kelompok yang terakhir. Pada kedalaman pertama, sampel dibagi oleh X_1 sebagai variabel independen terbaik untuk variabel dependen berdasarkan uji chi-square. Tiap node berisi informasi tentang frekuensi variabel Y, sebagai variabel dependen, yang merupakan bagian dari sub kelompok yang dihasilkan berdasarkan kategori yang disebutkan (X_1). Pada kedalaman ke-2 (node X_2 dan X_3) merupakan pembagian dari X_1 (untuk node ke-1 dan ke-3). Dengan cara yang sama, sampel selanjutnya dibagi oleh variabel penjelas yang lain, yaitu X_2 dan X_3 , dan selanjutnya menjadi sub kelompok pada node ke-4, 5, 6, dan 7 (Lehmann dan Eherler, 2001).

Dari masing-masing node tersebut juga ditampilkan persentase responden untuk tiap-tiap kategori dari variabel dependen, dan juga ditunjukkan jumlah total responden untuk masing-masing node (Myers, 1996).

Secara ringkas, Bagozzi (1994) menyatakan bahwa, diagram pohon, yang merupakan inti dari analisis CHAID, akan berisi:

1. Simbol yang menerangkan tentang kategori tertentu (atau kategori-kategori yang telah digabungkan).
2. Sebuah ringkasan data dari variabel dependen dalam kelompok tersebut (misalnya persentase respon).
3. Ukuran sampel untuk kelompok tersebut, atau biasa dilambangkan dengan “n”.

METODE PENELITIAN

Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder, yaitu data Pengguna Pembiayaan Koperasi Syari'ah Al-Hidayah Druju, Sumbermanjing Wetan Tahun 2002-2003. Data yang diambil hanyalah data nasabah yang melakukan peminjaman selama tahun 2002-2003. Data tersebut kemudian dipergunakan sebagai variabel dalam melakukan analisis CHAID. Variabel dalam penelitian ini terdiri dari variabel dependen dan independen sebagai berikut:

a) Variabel Dependen

Variabel dependen yang digunakan dalam penelitian ini adalah pengguna pembiayaan koperasi yang dibedakan menjadi dua kategori, yaitu:

1. Anggota dengan pinjaman yang berstatus macet (M)
2. Anggota dengan pinjaman yang berstatus tidak macet (TM)

b) Variabel Independen

Variabel independen dalam penelitian ini diambil dari variabel demografis data pengguna pembiayaan koperasi. Variabel tersebut terdiri dari 5 komponen demografis, yaitu:

1. Jenis kelamin

Untuk mengetahui karakteristik nasabah berdasarkan jenis kelamin yang dominan, maka jenis kelamin konsumen dibedakan menjadi:

- Laki-laki
- Perempuan

2. Usia

Usia nasabah dibedakan sebagai berikut:

- Usia Remaja (≤ 23 tahun)

Usia remaja merupakan usia transisi yang pada umumnya sangat mudah dipengaruhi oleh faktor-faktor eksternal.

- Usia 24 – 30 tahun

Usia ini merupakan usia kedewasaan nasabah serta biasanya usia di mana nasabah mulai merintis karir dalam kehidupannya.

- Usia 31 – 40 tahun

Usia yang merupakan usia peningkatan karir dan kematangan dalam bersikap.

- Usia 41 – 50 tahun

Usia di mana pada masa ini biasanya menjadi masa-masa kejayaan bagi nasabah dan masa di mana kemapanan diraih.

- Usia Lanjut (≥ 51)

Usia di mana nasabah mulai menikmati hari tuanya.

3. Pendidikan terakhir

Nasabah dapat dikelompokkan menurut tingkat pendidikan yang telah dicapai. Dalam hal ini tingkat pendidikan nasabah akan dibagi menjadi:

- SD
- SMP/ sederajat
- SMU/ sederajat
- Diploma
- S1/S2/S3

4. Pekerjaan

nasabah koperasi memiliki berbagai macam pekerjaan. Jenis pekerjaan nasabah dibagi menjadi:

- PNS/Pegawai Pemerintah/TNI/BUMN
- Karyawan Swasta
- Wiraswasta / Dagang
- Petani dan Peternak
- Lainnya

5. Penghasilan rata-rata keluarga

Berdasarkan keterangan dari tempat penelitian dilakukan, penghasilan rata-rata keluarga nasabah per bulan koperasi ini dibagi menjadi:

- $\leq$ Rp. 750.000,00

- Rp. 750.000,00 – Rp. 1.500.000,00
- Rp. 1.500.000,00 – Rp. 2.000.000,00
- Rp. 2.000.000,00

6. Ukuran keluarga

Ukuran keluarga adalah jumlah orang atau individu dalam satu rumah tempat nasabah tinggal. Variabel ini dibedakan atas:

- 1 orang
- 2 orang
- orang
- 4 orang
- 5 orang atau lebih

Data tersebut di atas selanjutnya akan dianalisis menggunakan software SPSS 13.0.

Metode Analisis Data

Langkah-langkah dalam melakukan analisis CHAID secara garis besar adalah sebagai berikut:

1. Memasukkan semua data berdasarkan kategori yang ditentukan sebagai berikut:

Tabel 3. Tabulasi Kategori Tiap Variabel

Variabel Dependen		Kategori
1. Status Pinjaman		
1.	Macet	0
2.	Tidak macet	1
Variabel Independen		Kategori
- Jenis Kelamin		
1.	Laki-laki	1
2.	Perempuan	2
- Usia		
1.	Usia Remaja (≤ 23 tahun)	1
2.	Usia 24 – 30 tahun	2
3.	Usia 31 – 40 tahun	3
4.	Usia 41 – 50 tahun	4
5.	Usia Lanjut (≥ 51)	5
- Pendidikan		
1.	SD	1
2.	SMP / sederajat	2
3.	SMU / sederajat	3
4.	Diploma	4
5.	S1 / S2 / S3	5
- Pekerjaan		
1.	PNS / Pegawai Pemerintah / TNI / BUMN	1
2.	Karyawan Swasta	2
3.	Wiraswasta / Dagang	3
4.	Petani / Peternak	4
5.	Lainnya	5
- Penghasilan		
1.	$\leq$ Rp. 750.000,00	1
2.	$>$ Rp. 750.000,00 – Rp. 1.500.000,00	2
3.	$>$ Rp. 1.500.000,00 – Rp. 2.000.000,00	3
4.	$>$ Rp. 2.000.000,00	4
- Ukuran Keluarga		

- | | | |
|----|--------------------|---|
| 1. | 1 orang | 1 |
| 2. | 2 orang | 2 |
| 3. | 3 orang | 3 |
| 4. | 4 orang | 4 |
| 5. | 5 orang atau lebih | 5 |
- Menentukan terlebih dahulu semua skala variabel yang akan digunakan dengan tepat dan benar.
 - Menentukan kategori target dari kategori-kategori variabel dependen. Hal ini dilakukan untuk memunculkan beberapa grafik lain sebagai informasi lebih lanjut dalam data yang ada. Kategori target yang dipergunakan bisa salah satu atau semua kategori yang ada pada variabel dependen.
 - Selanjutnya akan dilanjutkan dengan proses matematis analisis CHAID sesuai prosedur pada sub bab 2.1.2. Proses ini akan menerapkan 3 langkah analisis CHAID, yaitu langkah Penggabungan, Pemisahan, dan Pemberhentian. Dalam langkah Penggabungan akan mulai diterapkan uji *chi-square* dan pengali Bonferroni sebagai pengoreksinya. Pada langkah Penggabungan sebagian besar proses akan menggunakan uji *chi-square* saja. Kemudian dilakukan iterasi pada kedua langkah tersebut, dan proses iterasi akan berhenti apabila sudah tidak ada lagi variabel independen yang tersisa untuk diuji hubungannya dengan variabel dependen, atau juga apabila terbentuknya node pada diagram pohon telah memenuhi batasan yang ditentukan oleh peneliti. Proses ini disebut dengan proses Pemberhentian.
 - Menentukan segmentasi nasabah dengan menginterpretasikan diagram pohon CHAID
 - Menentukan target pasar berdasarkan hasil segmentasi nasabah yang sudah terbentuk.

Untuk lebih memperjelas langkah matematis analisis CHAID pada metode penelitian, disajikan skema metode penelitian pada Gambar 2.

HASIL PENELITIAN

Analisis CHAID

Segmentasi yang dihasilkan oleh analisis CHAID dengan dibantu software SPSS 13.0 pada nasabah yang melakukan peminjaman di koperasi Al-Hidayah tahun 2002–2003 dapat dilihat dari diagram pohon klasifikasi CHAID (CHAID *classification tree*) seperti pada Gambar 3.

Diagram pohon hasil analisis CHAID pada Gambar 3, menerangkan bahwa pada node teratas diketahui jumlah total nasabah yang melakukan peminjaman adalah 209 nasabah, terdiri dari 19 nasabah (9,1%) dengan status kreditnya macet dan

190 nasabah (90,9%) dengan status kredit yang tidak macet.

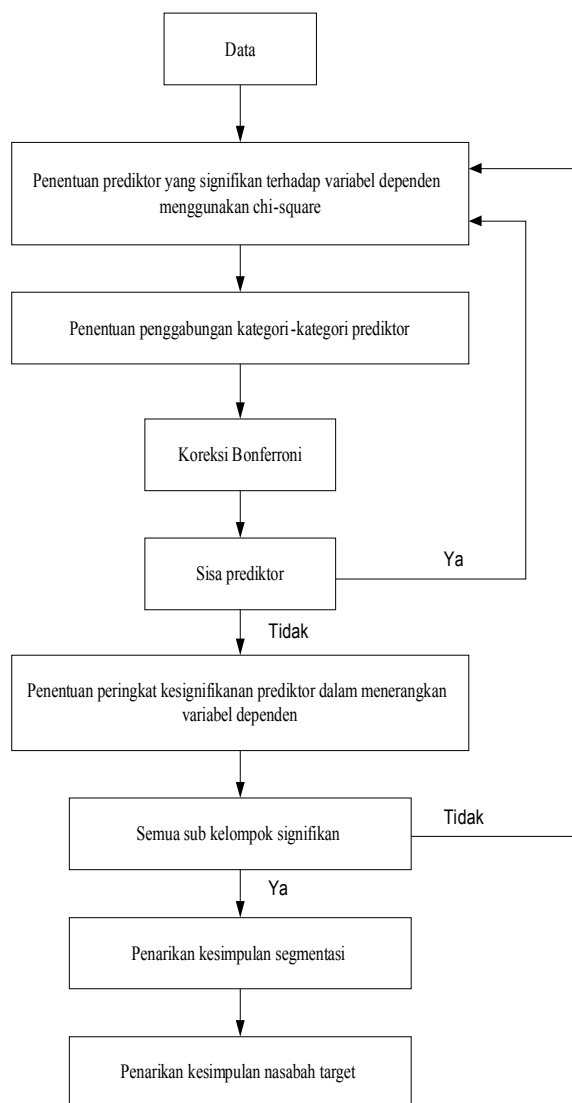
Tahap pertama dalam analisis CHAID adalah tahap Penggabungan. Dalam penelitian ini, variabel penghasilan dibagi menjadi 4 kategori, yaitu:

- $\leq$ Rp. 750.000,00
- $>$ Rp. 750.000,00 – Rp. 1.500.000,00
- $>$ Rp. 1.500.000,00 – Rp. 2.000.000,00
- $>$ Rp. 2.000.000,00

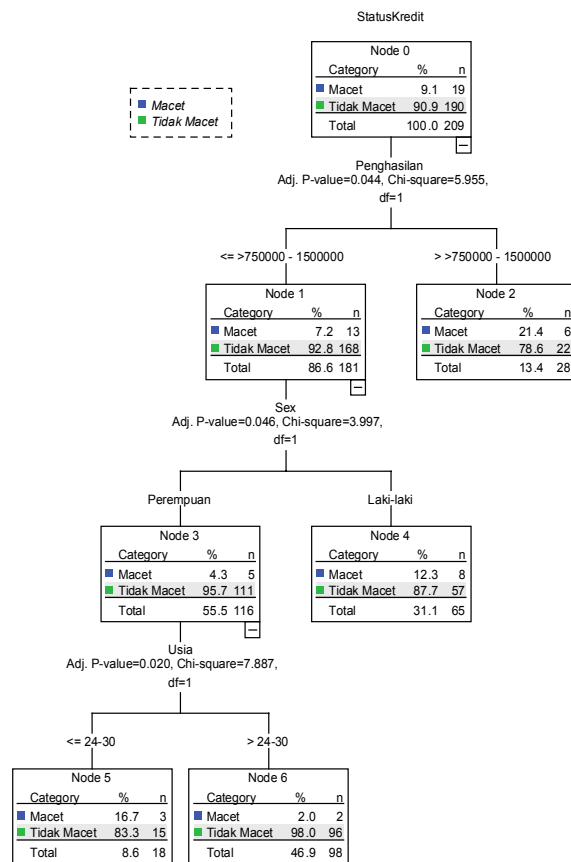
Setelah melalui analisis CHAID, variabel ini kemudian diringkas menjadi 2 kategori, seperti yang dapat dilihat pada diagram pohon kedalaman yang ke-1, yaitu:

- Kategori $\leq$ Rp. 750.000,00 digabung dengan kategori $>$ Rp. 750.000,00 – Rp. 1.500.000,00

Kategori $>$ Rp. 1.500.000,00 – Rp. 2.000.000,00 digabung dengan kategori $>$ Rp. 2.000.000,00



Gambar 2. Skema Metode Penelitian



Gambar 3. Diagram Pohon Analisis CHAID

Hal ini berarti bahwa kategori ke-1 dan ke-2 memenuhi syarat kesignifikanan chi-square untuk bisa digabung menjadi satu kategori tunggal. Begitu pula untuk kategori ke-3 dan ke-4 yang juga digabung menjadi satu kategori tunggal berdasarkan analisis CHAID yang ada.

Penggabungan kategori juga terjadi pada variabel usia. Dalam penelitian ini, variabel usia dibagi menjadi 5 kategori, yaitu:

- Usia Remaja (≤ 23 tahun)
- Usia 24 – 30 tahun
- Usia 31 – 40 tahun
- Usia 41 – 50 tahun
- Usia Lanjut (≥ 51)

Selanjutnya variabel ini oleh CHAID diringkas menjadi 2 kategori, seperti yang terlihat pada diagram pohon kedalaman yang ke-3 yaitu:

1. Kategori usia ≤ 23 tahun digabung dengan kategori usia 24 – 30 tahun
2. Kategori usia 31 – 50 tahun digabung dengan kategori usia ≥ 51 tahun

Hal ini berarti bahwa kategori ke-1 dan ke-2 memenuhi syarat kesignifikanan chi-square untuk bisa digabung menjadi satu kategori tunggal. Begitu pula untuk kategori ke-3, ke-4, dan ke-5 yang juga digabung menjadi satu kategori tunggal berdasarkan analisis CHAID yang ada.

Dari Tabel Model Summary pada Lampiran 2, dapat diketahui bahwa dalam penelitian ini sebenarnya ada 6 variabel bebas. Kemudian hasil analisis CHAID menunjukkan bahwa hanya ada 3 variabel bebas yang signifikan terhadap variabel terikatnya, yaitu variabel penghasilan, jenis kelamin, dan usia. Hal ini juga dapat dilihat dari hasil diagram pohon CHAID, bahwa pohon klasifikasi tersebut mempunyai 3 kedalaman, di mana variabel penghasilan membagi status kredit pada kedalaman ke-1, kemudian variabel jenis kelamin pada kedalaman ke-2, dan variabel usia pada kedalaman ke-3. Sehingga ada tiga variabel independen yang tersisa dan tidak dianggap mempunyai hubungan dengan variabel dependen, yaitu variabel pendidikan, pekerjaan, dan jumlah keluarga.

Nilai *p-value* dan nilai uji *chi-square* dari masing-masing variabel independen yang dianggap mempunyai hubungan dengan variabel dependennya dapat diringkas dalam tabel berikut:

Tabel 4. Nilai *P-value* dan *Chi-square* Variabel Dependen dalam Diagram Pohon

Variabel	Nilai <i>p-value</i>	Nilai <i>chi-square</i>
Penghasilan	0,044	5,955
Jenis Kelamin	0,046	3,997
Usia	0,020	7,887

di mana nilai *p-value* dalam tabel di atas merupakan nilai *p-value* setelah dikoreksi oleh pengali Bonferroni.

Dari Tabel 4. tersebut dapat diketahui bahwa apabila dilakukan pengambilan keputusan berdasarkan nilai *p-value*, di mana ketiga nilai tersebut kurang dari $\alpha=0,05$, yaitu 0,044; 0,046; dan 0,020, maka dapat ditarik kesimpulan bahwa keputusan uji chi-square adalah tolak H_0 . Hal ini berarti bahwa terdapat hubungan antara ketiga variabel tersebut dengan variabel dependen, yaitu status kredit.

Pada Gambar 3 dapat dilihat bahwa variabel penghasilan adalah variabel independen terbaik yang digunakan untuk membagi dan menerangkan variabel status kredit sebagai variabel dependen. Kemudian variabel jenis kelamin merupakan variabel independen yang signifikan untuk membagi kategori pada node ke-1 (kategori penghasilan rata-rata $\leq$ Rp. 750.000,00 dan kategori penghasilan rata-rata $>$ Rp.

750.000,00 – Rp. 1.500.000,00) pada variabel penghasilan. Dan variabel usia merupakan variabel independen yang signifikan dengan kategori pada node ke-3 (kategori jenis kelamin perempuan) pada variabel jenis kelamin.

Pohon klasifikasi CHAID pada Gambar 3 menerangkan bahwa nasabah koperasi Syari'ah Al-Hidayah yang melakukan pinjaman pada tahun 2002 – 2003 dibagi menjadi 4 segmen, yaitu:

Tabel 5. Segmentasi Nasabah Hasil Analisis CHAID

Segmen	Karakteristik
Ke-1	Nasabah dengan keluarga yang mempunyai penghasilan rata-rata sebesar $\leq$ Rp. 1.500.000,00, berjenis kelamin perempuan, serta berusia $\leq$ 30 tahun.
Ke-2	Nasabah dengan keluarga yang mempunyai penghasilan rata-rata sebesar $\leq$ Rp. 1.500.000,00, berjenis kelamin perempuan, serta berusia $>$ 30 tahun.
Ke-3	Nasabah dengan keluarga yang mempunyai penghasilan rata-rata sebesar $\leq$ Rp. 1.500.000,00, dan berjenis kelamin laki-laki.
Ke-4	Nasabah dengan keluarga yang mempunyai penghasilan rata-rata sebesar $>$ Rp. 1.500.000,00.

Dari keempat segmen yang terbentuk, dapat ditabulasi jumlah masing-masing nasabah yang macet dan yang tidak macet dalam tabel sebagai berikut:

Tabel 6. Tabulasi Jumlah Nasabah Macet dan Tidak macet

Segmen ke-	Jumlah Macet	Jumlah Tidak Macet
1	3	15
2	2	96
3	8	57
4	6	22

Pada kasus yang dibahas dalam penelitian ini, apabila ingin menentukan nasabah target yang akan dijadikan acuan untuk menentukan nasabah baru yang dapat memperkecil pinjaman yang macet, maka akan dipilih segmen-segmen dengan jumlah nasabah yang pinjaman macetnya terkecil, serta mempunyai jumlah nasabah dengan pinjaman tidak macet terbesar. Dari tabulasi hasil analisis CHAID di atas diketahui bahwa segmen yang memenuhi untuk dijadikan acuan dalam membidik nasabah target adalah pada segmen ke 2 (node ke-6 pada diagram pohon). Nasabah pada segmen ke-2 ini mempunyai karakteristik sebagai nasabah dengan keluarga yang

mempunyai penghasilan rata-rata sebesar $\leq$ Rp. 1.500.000,00, berjenis kelamin perempuan, serta berusia $>$ 30 tahun. Segmen ke-2 ini menunjukkan bahwa tingkat pendidikan dan jenis pekerjaan tidak menjadi patokan lancar atau tidaknya peminjaman yang dilakukan oleh nasabah. Nasabah perempuan dengan waktu luang dan pekerjaan yang sedikit akan lebih mempunyai banyak kesempatan untuk dapat teratur melakukan segala kewajiban terhadap koperasi dibandingkan apabila nasabah tersebut laki-laki dengan tanggung jawab pekerjaan sebagai kepala keluarga, walaupun hal ini hanya didukung dengan tingkat pendapatan yang tergolong menengah ke bawah. Kedewasaan dan kematangan berpikir pada tingkat usia $>$ 30 tahun juga akan dapat lebih meningkatkan kesadaran akan tanggung jawab kepada koperasi. Jadi di masa yang akan datang Koperasi Syari'ah Al-Hidayah dapat memprioritaskan calon nasabah dengan karakteristik seperti pada segmen ke-2 agar selanjutnya jumlah pinjaman dengan status macet dapat diminimalisir.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan uraian-uraian di atas, maka dapat ditarik kesimpulan bahwa:

1. Analisis CHAID adalah salah satu alat statistika yang dapat digunakan untuk segmentasi pasar dengan pendekatan metode dependensi.
2. Analisis CHAID memiliki kemampuan untuk menganalisis variabel independen monotonik, bebas, dan mengambang.
3. Pada kasus Koperasi Syari'ah Al-Hidayah, analisis CHAID mendapatkan empat segmen yang berbeda, yaitu:
 - a. Nasabah dengan keluarga yang mempunyai penghasilan rata-rata sebesar $\leq$ Rp. 1.500.000,00, berjenis kelamin perempuan, serta berusia $\leq$ 30 tahun.
 - b. Nasabah dengan keluarga yang mempunyai penghasilan rata-rata sebesar $\leq$ Rp. 1.500.000,00, berjenis kelamin perempuan, serta berusia $>$ 30 tahun.
 - c. Nasabah dengan keluarga yang mempunyai penghasilan rata-rata sebesar $\leq$ Rp. 1.500.000,00, dan berjenis kelamin laki-laki.
 - d. Nasabah dengan keluarga yang mempunyai penghasilan rata-rata sebesar $>$ Rp. 1.500.000,00.
4. Untuk meminimumkan jumlah pinjaman dengan status macet di masa mendatang maka Koperasi Syari'ah Al-Hidayah sebaiknya memprioritaskan calon nasabah dengan karakteristik nasabah

dengan keluarga yang mempunyai penghasilan rata-rata sebesar $\leq$ Rp. 1.500.000,00, berjenis kelamin perempuan, serta berusia > 30 tahun. Hal ini karena dari hasil analisis CHAID, segmen ke-2 (node ke-6 pada diagram pohon CHAID) memiliki resiko kredit macet yang paling rendah yaitu sebesar 2% (2/98) dan besar segmen 46,89% (98/209) dari total nasabah.

Usman, H. dan Akbar, R.P.S. 1995. *Pengantar Statistika*. Bumi Aksara. Jakarta.

Saran

Saran yang bisa diberikan untuk penelitian selanjutnya tentang analisis CHAID adalah perlu pengkajian lebih lanjut dari segi reliabilitas analisis CHAID dan model prediksinya.

DAFTAR PUSTAKA

- Assauri, S. 2002. *Manajemen Pemasaran: Dasar, Konsep dan Strategi*. Edisi Ketujuh. PT. Raja Grafindo Persada. Jakarta.
- Bagozzi, R.P. 1994. *Advanced Methods of Marketing Research*. Blackwell Publishers Ltd., Oxford.
- Gallagher, C.A., 2000. *An Iterative Approach to Classification Analysis*. www.casact.org/library/ratemaking/90dp237.pdf. Tanggal akses : 18 Desember 2005.
- Kasali, R. 1998. *Membidik Pasar Indonesia: Segmentasi, Targeting, Positioning*. Gramedia. Jakarta.
- Kotler, P. 1997. *Marketing Management*. Jilid 1. 9th edition. Alih bahasa: Hendra Teguh dan Ronny A. Rusli. Prentice-Hall Inc., New Jersey.
- Kotler, P. dan Armstrong, G. 1997. *Principles of Marketing*. 7th Edition. Alih bahasa: Alexander Sindoro. Prentice-Hall Inc., New Jersey.
- Lehmann, T. dan Eherler, D. 2001. *Responder Profiling with CHAID and Dependency Analysis*. www.informatik.uni-freiburg.de/~ml/ecmlpkdd/WS-Proceedings/w10/lehmann.pdf. Tanggal akses: 12 Desember 2005.
- Myers, J.H. 1996. *Segmentation and Positioning for Strategic Marketing Decisions*. American Marketing Association. Chicago.
- Sharp, A., J. Romaniuk dan S. Cierpicki. 2002. *The Performance Of Segmentation Variables: A Comparative Study*. http://130.195.95.71:8081/www/ANZMAC1998/Cd_rom/Sharp222.pdf. Tanggal akses: 15 Pebruari 2006.
- Umar, H. 2000. *Riset Pemasaran dan Perilaku Konsumen*. Gramedia. Jakarta.