

**METODE IMPROVED CHAID (CHI-SQUARED AUTOMATIC INTERACTION DETECTION) PADA ANALISIS KREDIT MACET BMT (BAITUL MAL WA TAMWIL)**

Muhammad Muhajir

Program Studi Statistika FMIPA Universitas Islam Indonesia

Jalan Kaliurang KM 14,5, Sleman, Yogyakarta, 55584

E-mail: muhammad.muhajir.stat89@gmail.com

**ABSTRAK**

Penelitian ini bertujuan untuk menentukan segmentasi nasabah berdasarkan karakteristik status kreditnya pada BMT El Bummi 372. Segmentasi ini bertujuan untuk mengidentifikasi segmen nasabah potensial dengan harapan resiko kredit macet yang dapat diminimumkan. Penelitian ini menggunakan data sekunder, yang terdiri dari 945 nasabah yang terdiri dari 94 macet dan 851 tidak macet. Salah satu metode riset segmentasi yaitu *Improved CHAID (Chi-Squared Automatic Interaction Detection analysis)* yang merupakan perbaikan dari metode *CHAID* dengan mengelaborasi tabel kontingensinya dengan *Tschuprow's T*. *Improved CHAID* akan menghasilkan diagram yang mirip dengan diagram pohon keputusan yang menyediakan informasi tentang derajat hubungan antara variabel dependen terhadap variabel independen serta informasi mengenai karakteristik segmen. Pada kasus BMT El Bummi 372, didapatkan enam segmen. Segmen nasabah dengan karakteristik pendidikan (Diploma/S1/Profesi, S2/S3), usia  $\leq 50$  tahun, ukuran keluarga  $\leq 3$  orang dan penghasilan  $>Rp. 1.500.000,00$  adalah segmen yang memiliki rasio kredit macet yang paling rendah.

**Kata Kunci:** Segmentasi nasabah, Status Kredit, *Tschuprow's T*, *Improved CHAID*

**ABSTRACT**

*This research aims to determine customer segmentation based on characteristics of credit status customers in BMT EL Bummi 372. The aim of this segmentation to identify a potential customer by loan status stalled which can be minimized. This research used secondary data, which consist of 94 customers with unpaid status and 851 customers with paid status. One of method market segmentation is Improved CHAID (Chi-Squared Automatic Interaction Detection analysis) which improved contingency table by Tschuprow's T. Improved CHAID will produce the decision tree which provide information on degrees relations between variables dependent on variables independent and information about characteristic segments. In BMT El Bummi 372 case, there are six segments were identified. Relationship segment which has education (elementary school, junior high school, and senior high school), less than equal 50 years, family size less than equal 3 persons, and monthly income more than Rp 1.500.000,00 is the lowest ratio of unpaid credit segment.*

**Keyword:** Customer segmentation, Credit status, *Tschuprow's T*, *Improved CHAID*

**Pendahuluan**

Globalisasi di Indonesia, menyebabkan tingginya kompetisi pada Metode *Improved CHAID (chi-squared automatic interaction detection)* Pada Analisis Kredit Macet BMT (Baitul Mal wa Tamwil) Muhammad Muhajir

sektor perusahaan jasa seperti *Baitul Mal Wa Tamwil (BMT)*. Persaingan antar BMT yang semakin ketat untuk menarik

minat para konsumen atau mempertahankan nasabah dalam roda bisnis perekonomian merupakan salah satu cara dalam mempertahankan eksistensinya.

Setiap perusahaan jasa khususnya BMT, selalu berusaha untuk mencapai target dengan cara menarik nasabah sebanyak-banyaknya untuk menabung maupun pemberian pembiayaan (kredit) di BMT tersebut (Suhendi dkk, 2004). Pembiayaan yang diberikan kepada para nasabah tidak akan lepas dari resiko terjadinya pembiayaan bermasalah yang akhirnya dapat memengaruhi terhadap kinerja bank syariah ataupun lembaga keuangan syariah lainnya tersebut (Samti, 2011).

Dalam rangka meminimalkan resiko kredit tersebut, dapat dilakukan dengan cara memperhatikan faktor ketidaklayakan debitur melalui status demografi nasabahnya berdasarkan status kredit, sehingga diperlukan segmentasi pasar berdasarkan hubungan faktor-faktor tersebut (Kunto dan Hasana, 2006).

Dari beberapa metode segmentasi pasar (nasabah) yang ada, penulis tertarik menggunakan metode *Improved CHAID* (*Chi-Squared Automatic Interaction Detection*). Metode ini dipilih karena menghasilkan improvisasi tabel

kontingensi *Tschuprow's T* pada algoritma *CHAID*, sehingga hasil segmentasinya akan lebih baik. Metode ini digunakan pada segmentasi nasabah, dan diharapkan dapat menemukan aturan klasifikasi yang dapat digunakan untuk memprediksi potensi seorang nasabah baru dengan pinjaman berstatus macet.

### Tujuan Penelitian

Penelitian ini bertujuan untuk mengidentifikasi segmen nasabah potensial dengan harapan resiko kredit macet dapat diminimumkan berdasarkan aturan klasifikasi yang sesuai.

### Metode Penelitian

Untuk melakukan segmentasi nasabah yang potensial dengan cara mengimplementasikannya dengan metode klasifikasi pohon *Improve CHAID*. Sistem tersebut memiliki tiga buah komponen utama yaitu :

- i. Variabel dataset nasabah BMT El Bummi 372 (*input*)
- ii. Proses signifikansi variabel dengan uji *Tschuprow's T* untuk pembentukan segmen nasabah
- iii. Proses pembuatan diagram pohon *Improve CHAID* dengan aturan "*Top-down stopping rule*"

Data dan *software* yang digunakan dalam penelitian ini adalah

data sekunder, yaitu data Pengguna Pembiayaan BMT El Bummi 372 Patuk, Gunung Kidul tahun 2011-2013 yang akan diolah pada software SIPINA. Data yang diambil hanyalah data nasabah yang melakukan peminjaman selama tahun 2011-2013. Data tersebut kemudian dipergunakan sebagai variabel dalam melakukan analisis *Improved CHAID*. Variabel dalam penelitian ini terdiri dari variabel dependen dan independen sebagai berikut (Kunto dan Hasana, 2006):

**a) Variabel Dependen**

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

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

**b) Variabel Independen**

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

1. Usia

Usia nasabah dibedakan sebagai berikut:

- Usia Remaja ( $\leq 23$  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 keamanan diraih.

- Usia Lanjut ( $\geq 51$  tahun)

Usia di mana nasabah mulai menikmati hari tuanya.

2. 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/Profesi
- S2/S3

3. Penghasilan rata-rata keluarga  
Penghasilan rata-rata keluarga nasabah per bulan BMT 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
- $\geq$  Rp. 2.000.000,00

4. Ukuran keluarga

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

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

### Rancangan Penelitian

#### Algoritma *Improved CHAID*

Secara umum ada dua tahap utama dalam algoritma *Improved CHAID* sebagai berikut (Belaid, 2001):

#### 1) Memisahkan dataset menjadi training dan *testing data*

Pada tahap ini dataset dibagi menjadi dua antara lain training data yaitu data yang akan dijadikan input dalam proses pembentukan pohon klasifikasi untuk menemukan aturan klasifikasi. *Testing data* yaitu data yang akan digunakan dalam pengujian aturan klasifikasi yang telah terbentuk.

#### 2) Melakukan Proses Analisis *CHAID*

Ada lima langkah analisis *CHAID* dengan *Tschuprow's T* sebagai berikut (Kass, 1980):

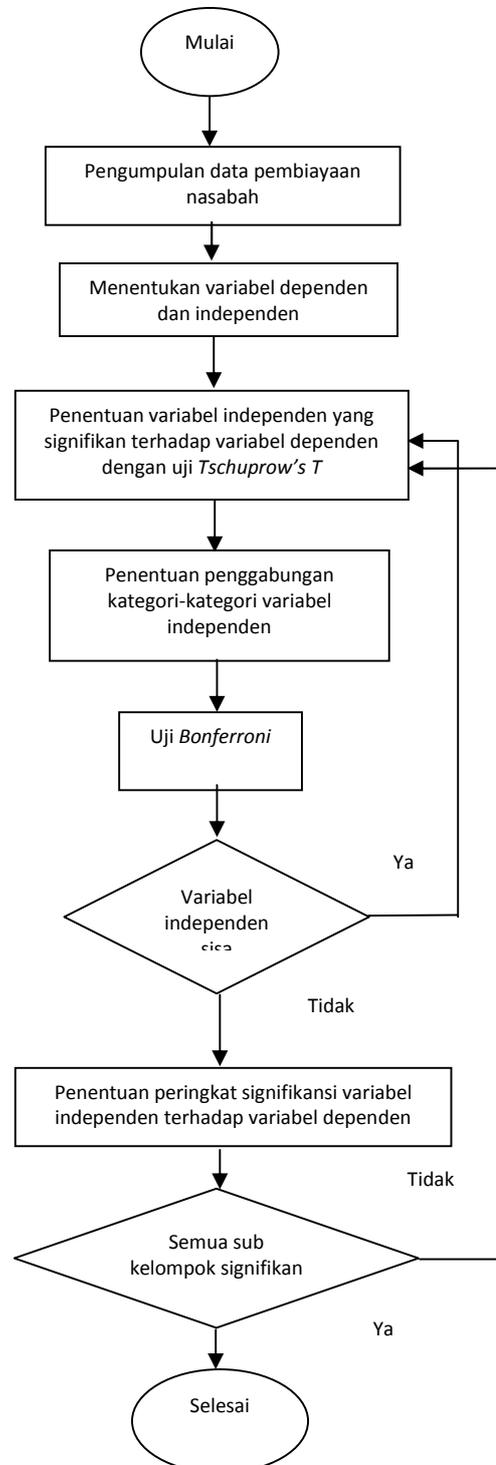
Untuk tiap variabel independen,  $X_1, X_2, \dots, X_k$  dan variabel dependen  $Y$

1. Bentuk tabel kontingensi dua arah dengan variabel dependennya berdasarkan *Tschuprow's T*.
2. Hitung statistik *Tschuprow's T* untuk setiap pasang kategori yang dapat dipilih untuk digabung menjadi satu, untuk menguji kebebasannya dalam sebuah sub tabel kontingensi 2 x 2 atau minimal dua kelompok yang menunjukkan bahwa jumlah kelompok adalah persis sama dengan jumlah nilai kelas.
3. Hitung *p-value* terkoreksi *Bonferroni* didasarkan pada tabel

yang telah digabung(Sharp dkk, 2002).

4. 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.
5. 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.
6. Lihat tingkat akurasi dari aturan yang telah terbentuk dari hasil pohon klasifikasi. Klasifikasi dengan tingkat keakuratan  $\geq 80\%$  yang dianjurkan (baik) digunakan dalam pengklasifian (Ramdhany dkk, 2006).

Alur kerja segmentasi nasabah dengan metode *Improved CHAID* digambarkan pada Gambar 1.



Gambar 1. Rancangan Penelitian

## Pembahasan

Pada bab ini akan dibahas aplikasi algoritma *Improved CHAID* dalam penentuan segmentasi nasabah yang mempunyai dataset berskala ordinal. Metode ini digunakan pada segmentasi nasabah, dan diharapkan dapat menemukan aturan klasifikasi yang dapat digunakan untuk memprediksi potensi seorang nasabah baru dengan pinjaman berstatus macet

### Demografi Nasabah

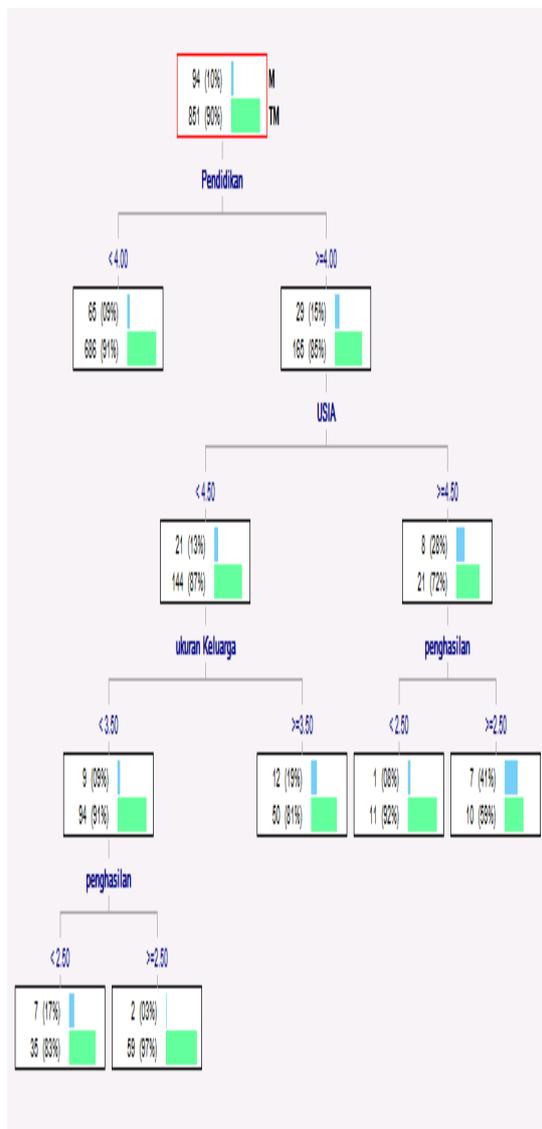
Data yang diambil hanyalah data nasabah yang melakukan peminjaman selama tahun 2011-2013. Jumlah total nasabah pada kurun waktu tiga tahun tersebut adalah sebanyak 945 nasabah. Data tersebut kemudian dipergunakan sebagai variabel dalam melakukan analisis *Improved CHAID*. Deskripsi data peminjam BMT El Bummi 372 yang melakukan peminjaman dapat dilihat pada Tabel 1.

Berdasar Tabel 1 terlihat, bahwa nasabah peminjam kredit terbanyak berdasarkan variabelnya yaitu usia antara 41-50 tahun sebanyak 359 orang, pendidikan terakhir SMP/Sederajat sebanyak 308 orang, penghasilan antara Rp. 750.000,00 – Rp. 1.500.000,00, dan ukuran keluarga 2 orang sebanyak 340 orang.

Tabel 1. Deskriptif Data

Variabel Dependen	Frekuensi	Prosentase (%)
<b>Status Kredit</b>		
1. Macet	94	9.95
2. Tidak macet	851	90.05
<b>Variabel Independen</b>		
<b>Usia</b>		
1. Usia Remaja ( $\leq 23$ tahun)	43	4.55
2. Usia 24 – 30 tahun	129	13.65
3. Usia 31 – 40 tahun	269	28.47
4. Usia 41 – 50 tahun	359	37.99
5. Usia Lanjut ( $\geq 51$ tahun)	145	15.34
<b>Pendidikan</b>		
1. SD	161	17.04
2. SMP/ sederajat	308	32.59
3. SMU/ sederajat	282	29.84
4. Diploma/S1/Profesi	167	17.67
5. S2/S3	27	2.86
<b>Penghasilan</b>		
1. $\leq$ Rp. 750.000,00	342	36.19
2. Rp. 750.000,00 – Rp. 1.500.000,00	296	31.32
3. Rp. 1.500.000,00 – Rp. 2.000.000,00	199	21.06
4. $>$ Rp. 2.000.000,00	108	11.43
<b>Ukuran Keluarga</b>		
1. 1 Orang	52	5.50
2. 2 Orang	340	35.98
3. 3 Orang	185	19.58
4. 4 Orang	183	19.37
5. 5 Orang atau lebih	185	19.58

**Aplikasi Algoritma *Improved CHAID***



Gambar 2. Hasil Pohon Klasifikasi *Improved CHAID*

Dari pohon analisis *Improved CHAID* pada Gambar 2 diketahui bahwa pada node teratas diketahui jumlah total nasabah 945 nasabah, terdiri dari 94 nasabah (10%) dengan status kreditnya macet dan 851 nasabah (90%) dengan status kredit yang tidak macet. Dari pohon klasifikasi *Improved CHAID*

terbentuk enam segmen yang berbeda, sebagai berikut:

1. **Jika** nasabah dengan pendidikan <4 (SD, SMP, SMA), **maka** dari data diperoleh *Tidak Macet* sebanyak 686 nasabah dan *Macet* sebanyak 65 nasabah. Sehingga nasabah pada kelompok ini diduga akan mengalami kemacetan.
2. **Jika** nasabah dengan pendidikan  $\geq 4$  (Diploma/S1/Profesi, S2/S3), usia < 4.5 ( $\leq 50$  tahun), ukuran keluarga < 3.5 ( $\leq 3$  orang) dan penghasilan < 2.5 ( $\leq$  Rp. 1.500.000,00), **maka** dari data diperoleh *Tidak Macet* sebanyak 35 nasabah dan *Macet* sebanyak 7 nasabah. Sehingga nasabah pada kelompok ini diduga akan mengalami kemacetan.
3. **Jika** nasabah dengan pendidikan  $\geq 4$  (Diploma/S1/Profesi, S2/S3), usia < 4.5 ( $\leq 50$  tahun), ukuran keluarga < 3.5 ( $\leq 3$  orang) dan penghasilan  $\geq 2.5$  ( $>$ Rp. 1.500.000,00), **maka** dari data diperoleh *Tidak Macet* sebanyak 59 nasabah dan *Macet* sebanyak 2 nasabah. Sehingga nasabah pada kelompok ini diduga tidak akan mengalami kemacetan.
4. **Jika** nasabah dengan pendidikan  $\geq 4$  (Diploma/S1/Profesi, S2/S3), usia < 4.5 ( $\leq 50$  tahun) dan ukuran keluarga  $\geq 3.5$  ( $\geq 4$  orang), **maka** dari data diperoleh *Tidak Macet* sebanyak 50

nasabah dan *Macet* sebanyak 12 nasabah. Sehingga nasabah pada kelompok ini diduga akan mengalami kemacetan.

5. **Jika** nasabah dengan pendidikan  $\geq 4$  (Diploma/S1/Profesi, S2/S3), usia  $\geq 4.5$  ( $\geq 51$  tahun) dan penghasilan  $\geq 2.5$  ( $>Rp. 1.500.000,00$ ), maka dari data diperoleh *Tidak Macet* sebanyak 11 nasabah dan *Macet* sebanyak 1 nasabah. Sehingga nasabah pada kelompok ini diduga akan mengalami kemacetan.
6. **Jika** nasabah dengan pendidikan  $\geq 4$  (Diploma/S1/Profesi, S2/S3), usia  $\geq 4.5$  ( $\geq 51$  tahun) dan penghasilan  $< 2.5$  ( $\leq Rp. 1.500.000,00$ ), maka dari data diperoleh *Tidak Macet* sebanyak 11 nasabah dan *Macet* sebanyak 1 nasabah. Sehingga nasabah pada kelompok ini diduga akan mengalami kemacetan.

### Tingkat Akurasi Metode

Pada tahap akurasi sistem perhitungan analisis *Improved CHAID*. Berikut tabel tingkat akurasi *output Improved CHAID*

Tabel 3. Tingkat Akurasi analisis *Improved CHAID*

<i>Output</i>	Akurasi (%)
<i>Improved CHAID</i>	90.5

Tabel 3 menunjukkan hasil tingkat akurasi analisis *Improved CHAID* sebesar 90.5 %. Hal ini berarti metode ini sudah cukup baik, bila digunakan dalam pengklasifikasian calon nasabah peminjam kredit.

### Kesimpulan.

BMT El Bummi 372 sebaiknya memprioritaskan calon nasabah dengan karakteristik nasabah menurut metode *Improved CHAID* yaitu nasabah dengan pendidikan (Diploma/S1/Profesi, S2/S3), usia  $\leq 50$  tahun, ukuran keluarga  $\leq 3$  orang dan penghasilan  $>Rp. 1.500.000,00$ , maka diduga tidak akan mengalami kemacetan. Selain itu BMT sebaiknya memperhatikan (selektif) terhadap calon nasabah dengan karakteristik peminjamnya

### Pustaka

- Belaid, A. 2001. *Data Mining in Document Structure Retroconversion*. ACM Transactions on Computational Logic, Vol. 2, No. 3.
- Kass, G., 1980, *An Exploratory Technique for Investigating Large Quantities of Categorical Data*, Applied Statistics, Vol. 29, No. 2, 119-127.
- Kunto, Y.S. dan Hasana, S.N. 2006. *Analisis CHAID sebagai Alat Bantu Statistika untuk Segmentasi Pasar*. Jurnal Manajemen, Vol.1 No.2. Universitas Kristen Petra, Surabaya. [http://fportfolio.petra.ac.id/user\\_files/05-011/MAR06010205.pdf](http://fportfolio.petra.ac.id/user_files/05-011/MAR06010205.pdf). Tanggal akses 13 September 2015.
- Ramdiany, D.N., Kustiyo, A., Handharyani, E., dan Buono, A., 2006, *Diagnosis Gangguan Sistem*

*Urinari pada Anjing dan Kucing Menggunakan VFI 5*, Institut Pertanian Bogor.

Samti, A.M., 2011. *Faktor-faktor yang Mempengaruhi Pengembal Kredit Bermasalah oleh Debitur Gerai Kredit Verena Bogor*. Skripsi S-1, FEM IPB, Bogor.

Sharp, A., Romaniuk, J., and, Cierpicki, S., 2002, *The Performance Of Segmentation Variables : A Comparative Study*, <http://anzmac.info/conference/199>

8/Cd\_rom/Sharp222.pdf, Tanggal akses : 15 September 2015.

Suhendi, H., Praja J.S., Abdullah H.N., Sadrah, H., Ridwan A.H., Solehudin, E., Januari, Y., Yusuf, D.K., 2004, *BMT dan Bank Islam*, Bandung : Pustaka Bani Quraisy.