

# KONSEP PEMAHAMAN CHI SQUARED AUTOMATIC INTERACTION DETECTION (CHAID Analysis) DISERTAI TUTORIAL PENERAPAN CONTOH KASUS DENGAN SPSS PADA SEGMENT PASAR KARTU KREDIT TERBARU YANG DIKELUARKAN OLEH BANK Z

CHAID ini dipergunakan pertama kali di Eropa mengingat pentingnya pengetahuan tentang segmen pasar untuk produk terbaru yang akan dikembangkan. Tentu, kalau kita sudah tahu yang mana yang jadi segmen pasar, maka kinerja pemasaran barang, promosi produk dan ketepatan sasaran produk akan memberikan hasil yang jauh lebih akurat.

Analisis CHAID ini pada dasarnya bersifat segmentatif yaitu memecah-mecah suatu problematik yang besar banget menjadi bagian-bagian yang kecil untuk mempermudah analisis segmen pasar. Tidak berlebihan jika ingin mengetahui siapa yang akan menjadi segmen kita adalah kunci mati untuk membuat target penjualan.

Selain bersifat segmentatif, CHAID ini juga bersifat prediktif sehingga dalam bidang ekonomi bisnis, analisis CHAID ini sangat diandalkan kepiawaiannya dalam membuat pemecahan untuk menentukan sebuah keputusan.

Ada memang sumber tertentu yang mengatakan bahwa analisis CHAID ini adalah bagian dari Data Mining Process. Jawabannya ya, memang betul seperti itu. Data-data yang ada akan dibangun lalu dipecah-pecah untuk pengambilan keputusan (decision making). Nah, sekarang bagaimana prinsip kerja analisis CHAID ini karena pada prinsipnya adalah dengan sistem pohon klasifikasi (Trees).

Adapun populasi yang kita amati akan disegmentasi ke dalam beberapa group berdasarkan sebuah variabel bebas yang paling kuat alias “number one” berasosiasi (ingat bukan berkorelasi yaaa) terhadap variabel terikatnya. Mengapa bias dikatakan berasosiasi? Ya karena kita menggunakan data yang berskala non metrik (nominal dan ordinal) untuk analisis CHAID ini.

Berikut ini contoh kasus misalnya, sebuah perusahaan mobile phone akan mengeluarkan produk terbarunya katakanlah HP dengan spesifikasi tertentu. Dimana sebelum produk ini mengempakkan sayapnya di pasaran, manajer pemasaran harus bisa membaca dulu bagaimana segmen pasar HP itu nanti. Karakteristik pembeli yang bagaimana yang akan membeli produk tersebut dan akan laris dimana produk HP itu nantinya.

Karakteristik misalnya dari sisi gender (segmen pasar kita laki-laki atau perempuan), pekerjaan (apakah yang bekerja di swasta, wiraswasta/businessman atau di

pemerintahan), status (apakah mereka yang sudah married atau masih lajang atau masih single parent), bisa juga dari karakteristik umur dan lain-lain sehingga hal tersebut perlu dianalisis. CHAID ini termasuk ke dalam keluarga besar statistik non parametrik yang juga bersifat eksploratif yaitu mengetahui variabel mana yang paling dominan dalam menjelaskan variabel respons (konsepnya mirip analisis faktor EFA) dan kita akhirnya mengetahui karakteristik dari objek/observasi yang kita teliti sehingga kita bisa memutuskan apakah akan memasarkan produk ke objek-objek sedemikian atau tidak. Contohnya, misal dalam analisis riset ekonomi perbankan dan dengan analisis CHAID ini nanti, sang manajer tahu karakteristik mana nasabah yang akan diberikan kredit pinjaman dan mana yang tidak (mirip dengan pola pikir analisis diskriminan).

Sehingga dapat dikatakan bahwa analisis CHAID ini adalah gabungan pola pikir analisis faktor EFA dengan analisis diskriminan

Berikut ini akan diberikan contoh kasusnya. Untuk penerapan, silahkan lihat pada data excel

Adapun kasus contoh sebagai berikut: Seorang manajer riset pemasaran ingin mengetahui bagaimana segmen pasar dari kartu kredit yang akan dikeluarkan oleh perusahaan perbankan tempat dia bekerja, katakanlah bank Z.

Sang manajer melakukan riset dengan survei terkait kartu kredit baru tersebut dengan memanfaatkan variabel bebas penelitian: Income, Pekerjaan, Jenis Kelamin, Asal Daerah, Status dan terakhir adalah Kepemilikan Kartu Kredit. Nah untuk income terbagi atas income lebih besar dari 3 juta (1), income 1,5 s/d 3 juta (2) dan income kurang dari 1,5 juta (3). Pekerjaan hanya dibagi atas PNS (1) dan Non PNS (2) jenis kelamin jelaslah ada dua hehehe yaitu laki-laki (1) dan perempuan (2). Asal daerah ada Jawa (1) dan luar Jawa (2) Status terbagi atas married (1) dan single or single parent(2). Terakhir, kepemilikan kartu kredit dibagi atas sudah ada/punya (1) dan belum ada (2).

Selanjutnya untuk variabel terikatnya dinamakan keputusan lalu terbagi atas setuju akan memakai (1) dan masih ragu memakai (2).

**Langkah awalnya pengerjaan dalam SPSS:**

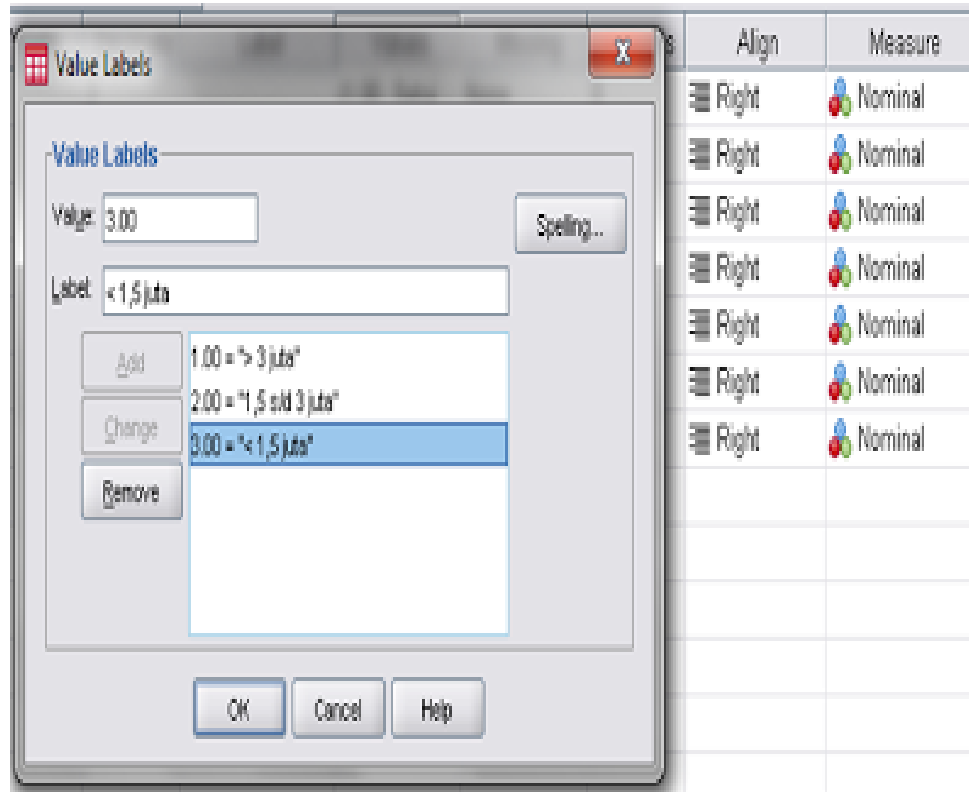
Berikut tampilan dalam Data View:

is chaid.sav [DataSet1] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

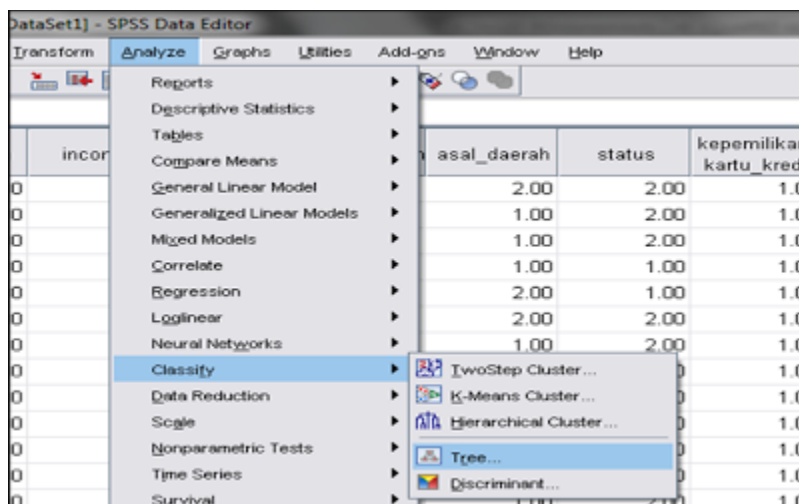
keputusan	income	pekerjaan	jenis_kelamin	asal_daerah	status	kepemilikan_kartu_kredit
1.00	2.00	2.00	1.00	2.00	2.00	1.00
1.00	1.00	2.00	1.00	1.00	2.00	1.00
1.00	1.00	2.00	1.00	1.00	2.00	1.00
1.00	1.00	2.00	1.00	1.00	1.00	1.00
1.00	2.00	2.00	1.00	2.00	1.00	1.00
1.00	1.00	1.00	1.00	2.00	2.00	1.00
1.00	1.00	2.00	1.00	1.00	2.00	1.00
1.00	1.00	2.00	1.00	1.00	2.00	1.00
1.00	2.00	2.00	1.00	1.00	2.00	1.00
1.00	3.00	1.00	2.00	2.00	2.00	1.00
1.00	2.00	2.00	2.00	2.00	1.00	1.00
1.00	1.00	2.00	1.00	1.00	2.00	1.00
1.00	1.00	2.00	1.00	1.00	2.00	1.00
1.00	2.00	1.00	1.00	2.00	2.00	1.00
1.00	3.00	2.00	1.00	2.00	2.00	2.00
1.00	3.00	2.00	2.00	1.00	2.00	2.00
1.00	2.00	2.00	2.00	1.00	2.00	1.00
1.00	2.00	2.00	1.00	1.00	2.00	1.00
1.00	3.00	2.00	1.00	2.00	2.00	1.00

Selanjutnya pada variabel view silahkan diatur valuenya masing-masing untuk variabel misalnya variabel income lebih besar dari 3 juta dibuat valuenya 1, income 1,5 s/d 3 juta dibuat 2 dan income kurang dari 1,5 juta dibuat 3. Berikut hasil gambarannya akan diperoleh:

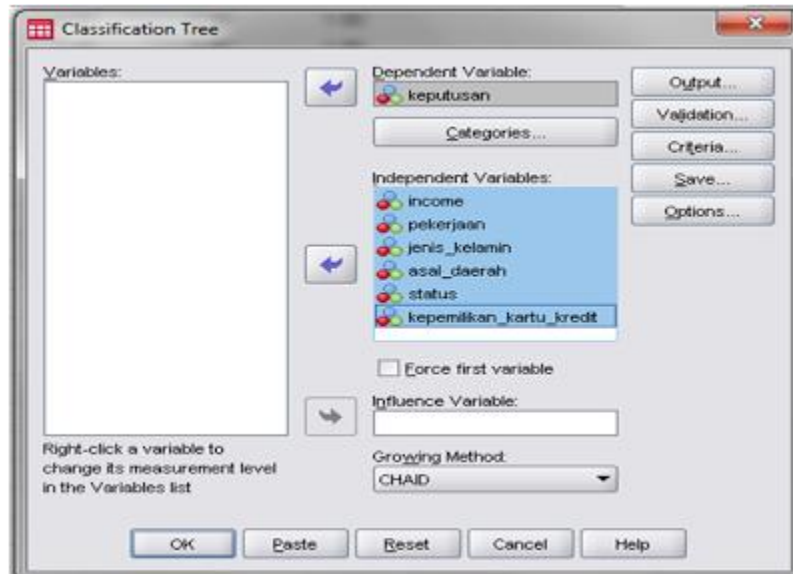


Selanjutnya atur juga measure (skala data) ke dalam skala data nominal (non metrik). Untuk variabel yang lain, silahkan diatur juga valuenya pada bagian Value Labels.

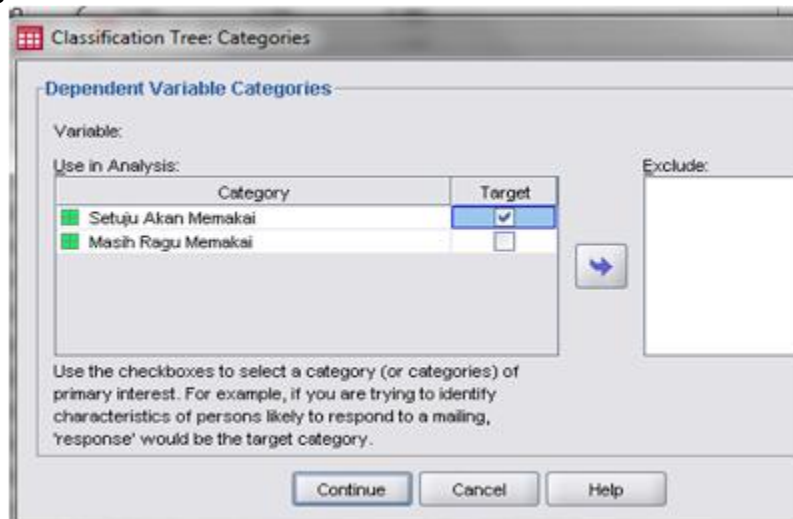
Kemudian klik Analyze, Classify dan pilih Tree sehingga diperoleh gambar sebagai berikut:



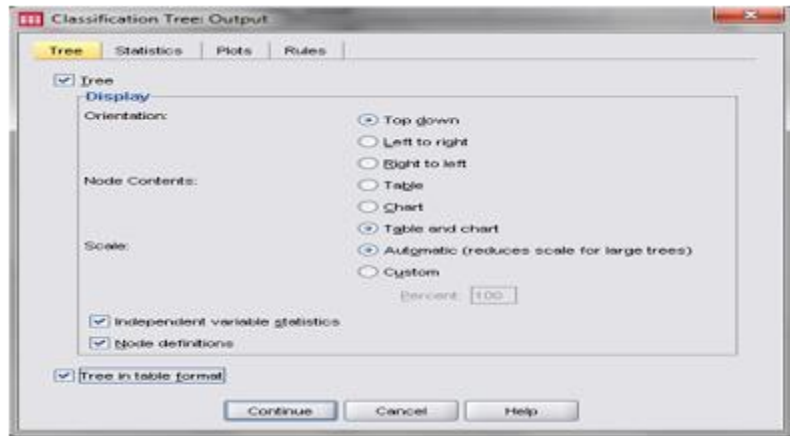
Masukkan variabel keputusan ke dalam bagian Dependent Variable sedangkan keenam variabel bebas yang dipakai (mulai dari income sampai kepemilikan kartu kredit) silahkan masukkan ke bagian Independent Variables. Berikut ilustrasinya.



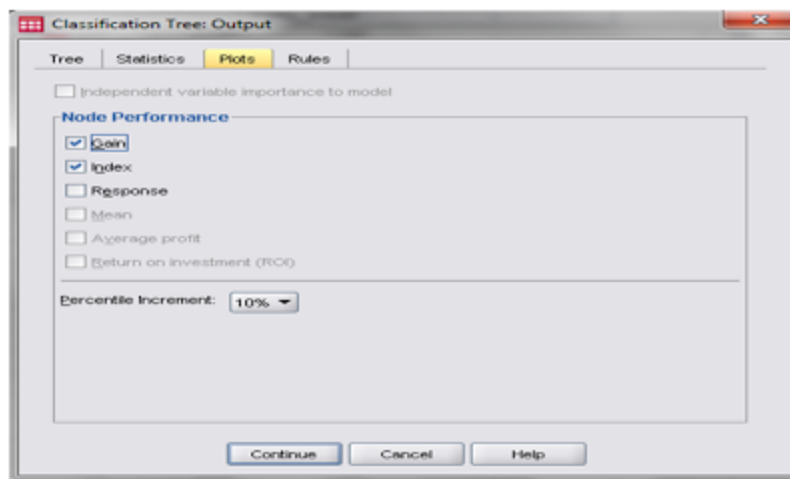
Selanjutnya klik pada Categories dan cek list pada kategori setuju akan memakai karena itu yang akan kita lihat



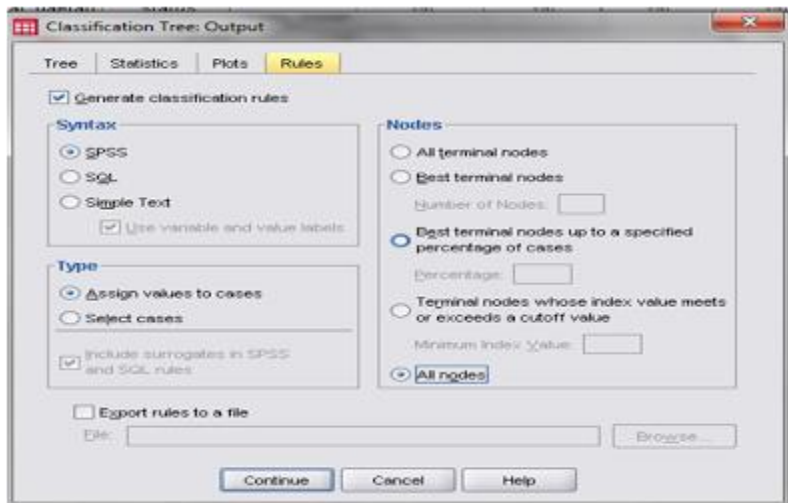
Lalu sekarang klik pada Output. Nah pada tree note content pilih yang tree and tables biar nanti ada pohon klasifikasi dan ada tabelnya juga (tabel response). Cek list juga pada Tree in table format.



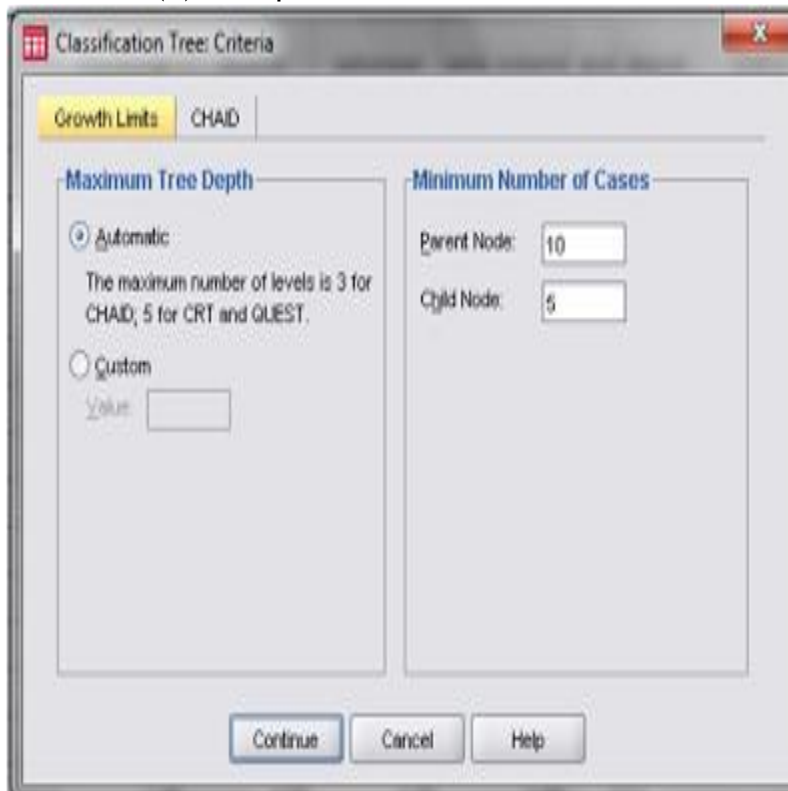
Selanjutnya pada Plots, silahkan berikan tanda centang (✓) pada Gain dan Index saja. Adapun gambar sebagai berikut:



Kemudian pada Rules aktifkan Generate Classification Rules dan pada Nodes silahkan pilih All Nodes (untuk menampilkan seluruh hierarki dalam pohon klasifikasi)



Lalu klik Continue dan kembali ke tampilan jendela Classification Tree. Sekarang, klik pada bagian Criteria dan pada Minimum Number of Cases atur aja untuk node parents (10) dan untuk note child (5). Lalu pilih Continue dan OK.



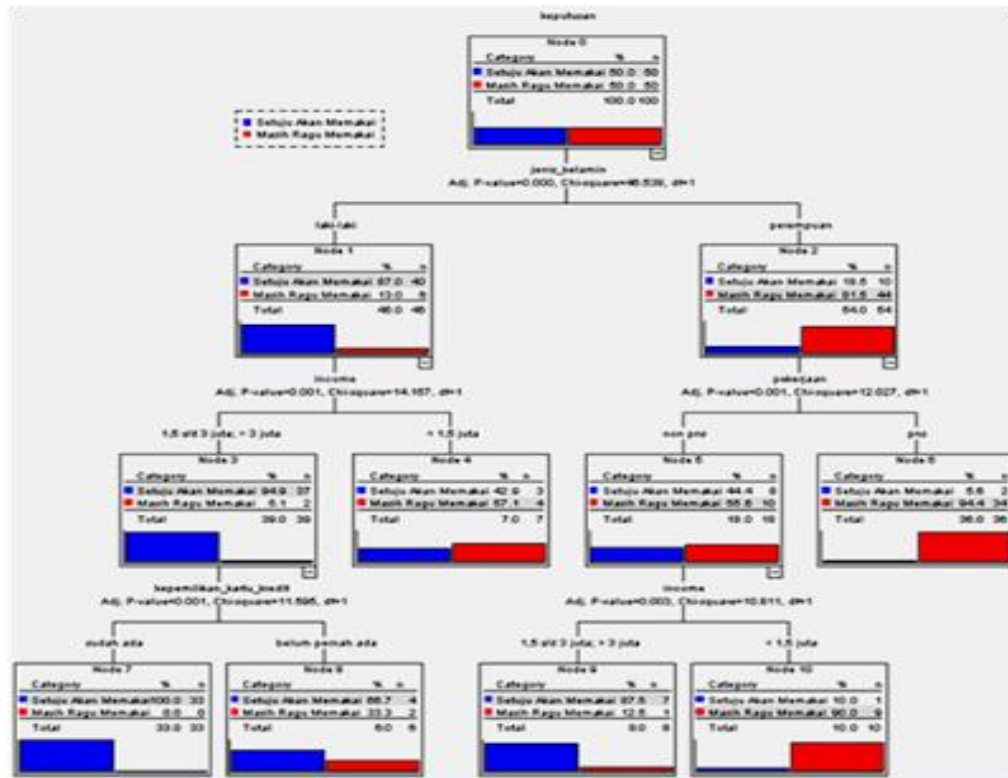
Adapun untuk output dapat dilihat sebagai berikut:

Specifications	Growing Method	CHAID
	Dependent Variable	keputusan
	Independent Variables	income, pekerjaan, jenis_kelamin, asal_daerah, status, kepemilikan_kartu_kredit
	Validation	None
	Maximum Tree Depth	3
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	jenis_kelamin, income, kepemilikan_kartu_kredit, pekerjaan
	Number of Nodes	11
	Number of Terminal Nodes	6
	Depth	3

Langkah awal bahwa dalam analisis CHAID kita akan berkenalan dengan tiga jenis node yaitu node indukan (parent node), node anakan (child node) dan node final (terminal node). Dalam output model Summary terlihat bahwa ada 11 node yang terbentuk dan untuk kepentingan analisis adalah fokus kepada 6 terminal node.

Tingkat kedalaman (Depth) analisis CHAID sejumlah 3 yaitu jumlah hierarki pada pohon klasifikasi berdasarkan jenis kelamin, hierarki berdasarkan income dan pekerjaan serta hierarki berdasarkan kepemilikan kartu kredit dan income.

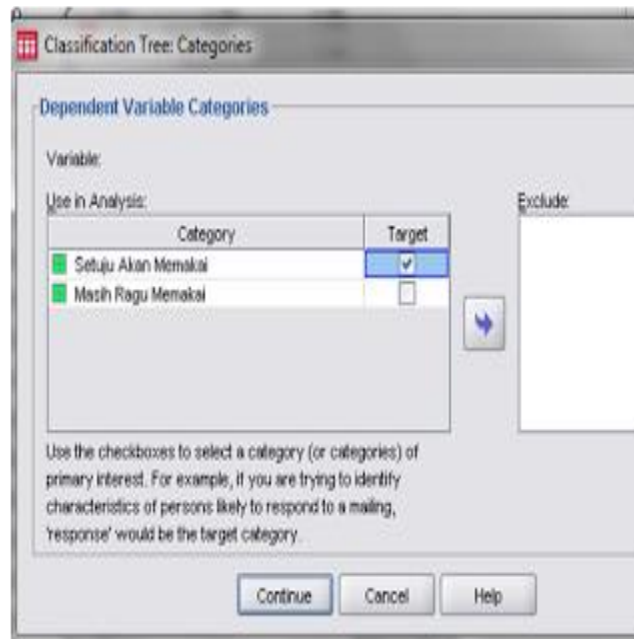
Lebih jelasnya dapat dilihat output treesnya berikut ini:



Dari gambar terlihat bahwa untuk node 0 kan node paling awal tuuh , jadi tidak memiliki parent node. Sedangkan untuk node 1 dan 2 baru punya parent node yaitu **node 0** (node induk).Selanjutnya, untuk node 3 dan 4 punya parent node yaitu **node 1**. Next, untuk node 5 dan 6 juga punya parent node yaitu **node 2**hehehe.. Selanjutnya, untuk node 7 dan 8 punya parent node yaitu **node 3** dan terakhir neeh sob untuk node 9 dan 10 punya parent node yaitu **node 5**.

Sedangkan child note, dari parent node. Dari gambar maka diperoleh 5 node yang jadi parent node dan 10 child node (node 1 sampai 10

Pada output dapat kita melihat terminal nodenya seperti gambar sbb:



Sehingga dari gambar tersebut akan dapat kita tentukan target segmen pasar dan tentu yang kita pilih sebagai target adalah kategori setuju akan memakai. Sehingga untuk informasi terminal node target dapat kita lihat pada Target Category: Setuju Akan Memakai.

## Target Category: Setuju Akan Memakai

Gains for Nodes

Node	Node		Gain		Response	Index
	N	Percent	N	Percent		
7	33	33.0%	33	66.0%	100.0%	200.0%
9	8	8.0%	7	14.0%	87.5%	175.0%
8	6	6.0%	4	8.0%	66.7%	133.3%
4	7	7.0%	3	6.0%	42.9%	85.7%
10	10	10.0%	1	2.0%	10.0%	20.0%
6	36	36.0%	2	4.0%	5.6%	11.1%

Growing Method: CHAID  
 Dependent Variable: keputusan

Sejalan dengan output pada Model Summary bahwa ada 6 buah terminal node yaitu node 7, 9, 8, 4, 10 dan 6. Hal ini dapat dilihat node terminal sudah langsung sudah langsung terurut dari yang terbesar berdasarkan response dan indexnya. Pertanyaannya apakah keenam terminal node ini yang menjadi segmen pasar produk kartu kredit bank Z?

Segmen pasar produk adapun Ke-enam terminal node tersebut bukan merupakan Inget terminal node itu ibarat kata berisikan kumpulan-kumpulan kandidat alias calon segmen pasar produk kartu kredit keluaran Bank Z. Sehingga untuk memilih segmen pasar dari kandidat segmen pada terminal node silahkan akan dapat dilihat nilai Indexnya. Jika nilai index lebih besar dari 100%, maka itulah yang akan jadi segmen pasar produk kartu kreditnya.

Selanjutnya akan kita lihat bahwa hanya ada 3 node terpilih yaitu node 7, node 9 dan node 8 yang menjadi segmen/pemakai kartu kredit itu dan segmen terbesar adalah node 7. Perhatikan kembali outp

ut trees di atas. Segmen terbesar adalah pria dengan income/penghasilan lebih dari sama dengan 1,5 juta rupiah dan sudah memiliki kartu kredit. Selanjutnya terbesar kedua adalah wanita dengan pekerjaan non PNS dengan income/penghasilan lebih dari sama dengan 1,5 juta rupiah. Selanjutnya segmen terakhir adalah pria dengan penghasilan lebih dari sama dengan 1,5 juta rupiah dan belum pernah memiliki kartu kredit.

Sekarang dari ketiga segmen pasar (target penjualan), manajer pemasaran bank Z ingin tahu berapa perolehan yang bisa diperoleh dari ketiga segmen tersebut. Adapun dapat kita perhatikan output Gain for Nodes pada bagian Percent Gain seperti pada table sebagai berikut:

## Target Category: Setuju Akan Memakai

Gains for Nodes

Node	Node		Gain		Response	Index
	N	Percent	N	Percent		
7	33	33.0%	33	66.0%	100.0%	200.0%
9	8	8.0%	7	14.0%	87.5%	175.0%
8	6	6.0%	4	8.0%	66.7%	133.3%
4	7	7.0%	3	6.0%	42.9%	85.7%
10	10	10.0%	1	2.0%	10.0%	20.0%
6	36	36.0%	2	4.0%	5.6%	11.1%

Growing Method: CHAID  
Dependent Variable: keputusan

Perhatikan bahwa untuk segmen pertama, produk kartu kredit terbaru keluaran bank Z akan mampu meraup pasar sampai level 66%. Artinya jika kita arahkan pemasaran baik melalui media promosi (periklanan), kunjungan dan purposive targetted system (langsung menawarkan kepada karakteristik seperti segmen pertama yaitu pria dengan income/penghasilan lebih dari sama dengan 1,5 juta rupiah dan sudah memiliki kartu kredit), maka sebesar 66% dari target segmen pertama akan berhasil diraup.

Segmen pasar kedua memang pihak bank Z akan memperoleh Gain yang tidak sebesar segmen pertama. Kita akan sanggup meraup sampai level 14% untuk populasi pada segmen kedua (wanita dengan pekerjaan non PNS dengan income/penghasilan lebih dari sama dengan 1,5 juta rupiah).

Selanjutnya, kita akan bisa meraup pasar sampai ke level 8% untuk segmen pasar ketiga (pria dengan penghasilan lebih dari sama dengan 1,5 juta rupiah dan belum pernah memiliki kartu kredit). Hal inilah yang menjadi target pasar kartu kredit Bank Z.

Dengan mengetahui siapa segmen pasar kita maka akan sangat efisien dalam pengelolaan biaya untuk melakukan strategi pemasaran. Sudah jelas bahwa dari segmen pertama (segmen terbesar), kita memperoleh/meraup pasar sampai ke level 66%, maka sangat disarankan promosi produk kartu kredit ini lebih diarahkan kepada segmen tersebut. Dalam hal ini tidak masalah untuk segmen ini yaitu pria dengan

pekerjaan apakah dia PNS atau non PNS (swasta atau wiraswasta/pebisnis) asalkan penghasilannya lebih dari sama dengan 1,5 juta rupiah dan sudah pernah memiliki kartu kredit, maka segmen ini akan sangat menjanjikan.

Promosi yang lebih gencar juga perlu diperhatikan untuk segmen kedua yaitu populasi wanita dengan income lebih dari 1,5 juta rupiah dengan pekerjaan non PNS (swasta atau wiraswasta/pebisnis). Biaya promosi lebih baik dialokasikan dengan kunjungan ke pabrik-pabrik/perusahaan-perusahaan swasta dan mendatangi target pasar seperti pada karakteristik segmen kedua. Karena gain yang diperoleh mencapai 14%. Begitu pula halnya untuk segmen yang ketiga (yang terakhir).

**Risk**

Estimate	Std. Error
.090	.029

Growing Method: CHAID  
Dependent Variable: keputusan

**Classification**

Observed	Predicted		
	Setuju Akan Memakai	Masih Ragu Memakai	Percent Correct
Setuju Akan Memakai	44	6	88.0%
Masih Ragu Memakai	3	47	94.0%
Overall Percentage	47.0%	53.0%	91.0%

Growing Method: CHAID  
Dependent Variable: keputusan

Kita lihat bahwa standar errornya sangat kecil yaitu 0,029 dengan risiko kesalahan yang diperkirakan adalah 9%. Sama halnya dengan analisis diskriminan, kita bisa lihat kesahihan/keakuratan hasil penelitian dengan CHAID Analysis untuk kasus identifikasi segmen-segmen pasar kartu kredit terbaru keluaran bank Z adalah  $91/100 = 91\%$  (masuk dalam kategori sangat akurat).