

# PENELUSURAN KARAKTERISTIK PERILAKU KONSUMEN DENGAN METODE AUTOMATIC INTERACTION DETECTION (AID)

**Agus Rusgiyono**

Staf Pengajar Prodi Statistika Jurusan Matematika FMIPA UNDIP

## **Abstract**

AID methods used to see relation between respons variable with a number of variable of predictor and also see related pattern between predictors. In AID procedures data disjointed in two group determined by variable of predictor most explaining of difference values of respons variables, and this recurring process so that yielded to with refer to splits in data. Every split yield new data sub which its variable values is mutually and exclusive exhaustive. Its end results with refer toing crotch is so-called tree of AID

**Keywords** : split, tree of AID

## **1. Pendahuluan**

Konsumen dalam mempertimbangkan untuk menggunakan barang atau jasa selalu mempertimbangkan dua faktor berkaitan dengan kualitas jasa dan harga barang atau jasa. Berkaitan dengan jasa di samping konsumen menilai kualitas dari segi kesenjangan jasa yakni perbedaan antara apa yang diharapkan pelanggan akan mereka terima dan persepsi mereka terhadap jasa yang benar-benar diserahkan maka penilaian kualitas jasa juga melibatkan lima dimensi yang luas sebagai kriteria yakni kehandalan, keberwujudan, daya tanggap, jaminan dan empati<sup>[4]</sup>.

Data-data yang terkait dengan perilaku konsumen ini diperoleh dengan metode survey dengan memakai skala linkert 1-5 pada setiap pertanyaan pada masing-masing dimensi kepuasan pelanggan maupun harga barang dan jasa, dan variabel respon biner berupa keputusan untuk jadi membeli atau tidak<sup>[5]</sup>.

Untuk melihat hubungan antara variabel respon berupa keputusan untuk membeli barang atau jasa dalam kaitannya dengan variabel penjelas berupa kualitas barang atau jasa dan harga barang atau jasa telah banyak dikembangkan metode statistika untuk menanganinya, baik dengan metode statistika deskriptif maupun dengan statistika inferensial. Dalam penelitian data survey lebih sering digunakan pendekatan eksplorasi data seperti statistika deskriptif. Bila tujuannya untuk melihat hubungan antar variabel maka statistika deskriptif memiliki banyak kekurangan. Di lain pihak penggunaan statistika inferensial memerlukan banyak hal yang harus dikerjakan. Sehingga dalam menangani data survey diperlukan analisis data yang dapat menjembatani kekurangan kedua metode tersebut.

Di dalam pustaka<sup>[1]</sup>, disebutkan salah satu alternatif yang dapat ditawarkan di sini adalah metode Automatic Interaction Detection (AID). Dalam prosedur AID, data dipisah dalam dua kelompok yang ditentukan oleh variabel prediktor yang paling menjelaskan perbedaan nilai-nilai variabel respon, dan proses ini berulang sehingga dihasilkan serangkaian split-split dalam data. Tiap split menghasilkan sub data baru yang nilai-nilai variabel responnya bersifat mutually exclusive dan exhaustive. Hasil akhirnya adalah serangkaian cabang pohon yang disebut pohon AID. Sehingga dapat dipandang metode AID ini cocok untuk tujuan melihat hubungan antar variabel walaupun jumlah variabelnya banyak, tanpa harus melakukan banyak hal seperti pada statistika Inferensial.

## 2. Rumusan Masalah

Menjelaskan langkah kerja metode Automatic Interaction Detection dalam melihat karakteristik perilaku konsumen khususnya keinginannya untuk membeli barang sebagai variabel respon yang tergantung pada variabel penjelas khususnya kualitas barang/jasa.

## 3. Pembahasan

Perilaku konsumen barang atau jasa seringkali berbeda pada berbagai faktor yang berkaitan dengan nilai kualitas barang atau jasa, harga, cara pemasaran dan sebagainya. Keinginan konsumen untuk membeli suatu barang di samping didorong oleh kebutuhan akan barang tersebut juga terpengaruh pada tingkat pendapatan yang dipunyainya. Opini semacam ini menarik untuk ditelusuri karena sering tidak diketahui mana yang lebih kuat atau mana yang menyebabkan suatu faktor tereliminasi karena dominannya faktor yang lain <sup>[2]</sup>.

Automatic Interaction detection (AID) diperkenalkan pertamakalinya tahun 1964 di Institute for Social Research, University of Michigan oleh Morgan dan Sonquist <sup>[1]</sup>.

Metode AID bertujuan menjelaskan hubungan antara variabel respon dengan variabel prediktor. Pertama-tama keseluruhan sampel dibelah menjadi dua sub kelompok yang ditentukan oleh variabel prediktor yang paling menjelaskan perbedaan nilai-nilai variabel respon dan proses ini berulang untuk setiap sub kelompok selanjutnya yang telah terbentuk. Kriteria pengukuran pemisahan sub kelompok pada variabel respon digunakan fungsi jarak. Strategi dasarnya adalah menguji masing masing variabel prediktor sedemikian sehingga menyebabkan pembelahan biner menghasilkan sub kelompok-sub kelompok yang memaksimalkan fungsi jarak terpilih. Proses berlanjut dengan pembelahan sub kelompok berturut-turut yang diperoleh dari tahap sebelumnya atau tahap yang menggunakan kriteria yang sama untuk memutuskan prediktor mana yang akan digunakan. Hasil dari proses pembelahan biner adalah pohon hirarki (bertingkat) yang disebut pohon AID.

Pemisahan pengamatan ke dalam dua kelompok pada metode AID ini biasa disebut sebagai pembelahan biner. Pembelahan biner yang dilakukan oleh suatu variabel prediktor adalah suatu cara pembelahan  $N$  pengamatan menjadi dua kelompok berdasarkan nilai-nilai yang dimiliki variabel prediktor tersebut.

Dalam setiap tahap, pembagian sebuah sub kelompok menjadi dua bagian berdasarkan oleh salah satu variabel prediktor. Pembelahan yang dilakukan berdasarkan variabel prediktor terpilih adalah yang meminimalkan jumlah kuadrat sisa dari variabel respon yang diperoleh pada saat kelompok dibelah menjadi dua bagian. Hal ini ekuivalen dengan memaksimalkan jumlah kuadrat antar kelompok. Kemudian dari dua sub kelompok yang dihasilkan, yang mempunyai jumlah kuadrat kesalahan maksimum dipisah lagi berdasarkan variabel lainnya dan seterusnya hingga akhirnya diperoleh kelompok-kelompok pengamatan yang mempunyai ciri variabel respon dan variabel prediktor tertentu yang berkaitan. Kriteria jarak ini digunakan untuk memaksimalkan kekuatan penjelasan dari pembelahan biner di setiap tahap, seperti halnya pada analisis variansi.

### Uji Signifikansi dalam AID

Misalkan kelompok dalam pembahasan ini yaitu keputusan untuk membeli barang yang digolongkan dalam dua klasifikasi yakni jadi membeli dan tidak jadi membeli. Terhadap masing-masing golongan diamati pendapatannya per bulan memiliki mean dan variansi  $\bar{X}$  dan  $S^2$ , dan variabel prediktor berupa variabel nilai persepsi kualitas barang/jasa dan nilai persepsi harga barang/jasa yang memiliki  $c$  kategori, dengan  $N_i, \bar{X}_i$  masing – masing

adalah banyaknya pengamatan dan rata-rata variabel ke  $- i$  dan  $f_i = \frac{N_i}{N}$  adalah proporsi observasi dalam kategori ke  $- i$ .

Setelah kelompok membelah menjadi  $M_1$  dan  $\bar{Y}_1$  yaitu jumlah observasi dan rata-rata pengamatan sub kelompok pertama dan  $M_2$  dan  $\bar{Y}_2$  yaitu jumlah observasi dan rata-rata pengamatan sub kelompok kedua.

Proporsi jumlah kuadrat yang dijelaskan oleh pembelahan split, diberikan oleh :

$$P = \frac{\text{jumlah kuadrat antar kelompok}}{\text{jumlah kuadrat kelompok}} = \frac{M_1 M_2 (\bar{Y}_2 - \bar{Y}_1)^2}{N^2 S^2}$$

Selanjutnya diperiksa distribusi dari :

$$K = NP = \frac{M_1 M_2 (\bar{Y}_2 - \bar{Y}_1)^2}{NS^2} = \max\{K_i\} = \max\{NP_i\}$$

di mana  $P_i$  adalah proporsi varian yang dijelaskan oleh metode pembelahan dari kelompok ke- $i$ .

Di bawah hipotesis  $H_0$  bahwa variabel prediktor tidak berhubungan dengan variabel respon, setiap alokasi dari pengamatan pada kategori yang berbeda memberikan  $N_i$  yang sama kemungkinannya. Dengan asumsi ini, distribusi probabilitas dari  $X_i$  dalam kategori

ke-  $i$  dengan  $N_i$  pengamatan memiliki mean  $\bar{X}_i$  dan varian  $\frac{(1-f_i)S^2}{N_i}$ , di mana  $f_i$  adalah

proporsi pengamatan dalam kategori ke- $i$ .

$$\text{Misalkan } G_i = \sqrt{\frac{f_i}{1-f_i}}$$

$$\text{dan transformasi } \mathcal{G}_i = \frac{G_i (X_i - \bar{X}) \sqrt{N}}{S} \quad i = 1, 2, 3, \dots, c-1$$

digunakan, maka distribusi probabilitas dari  $\mathcal{G}_i$  memberikan proporsi pengamatan  $f_i$  dalam setiap kategori adalah  $h(\{\mathcal{G}_i\} | \{f_i\}) \sim N(0, \sum \mathcal{G}_i)$ .

$$\text{Misalkan dalam hal ini } c=2, \text{ maka hanya ada satu } \mathcal{G}_i \text{ yaitu } \mathcal{G}_1 = \frac{G_1 (X_1 - \bar{X}) \sqrt{N}}{S}$$

sehingga memiliki pendekatan distribusi normal baku .

$$\text{Statistik uji } K \text{ adalah kuadrat absolute dari } \mathcal{G}_1, \quad K = \mathcal{G}_1^2 \text{ akibatnya } K \sim \chi_1^2$$

Setelah dibuktikan bahwa untuk prosedur pembelahan kelompok, statistik uji  $K$  mengikuti distribusi Chi-kuadrat, maka selanjutnya nilai kritis dari  $K$  didasarkan pada nilai kritis dari distribusi  $\chi_{1,\alpha}^2$ . Jika  $K$  lebih besar dari nilai kritis tersebut maka split dapat dilakukan. Sebaliknya jika  $K$  kurang dari nilai kritis tersebut maka split dihentikan<sup>[3]</sup>.

Aplikasi untuk sampel kecil, dapat dilihat jika hanya ada dua kategori dan variabel respon dikotomi, masalah ini dapat direduksi sehingga terbentuk tabel kontingensi  $2 \times 2$  dengan statistic uji  $\chi^2$ . Aproximasi ini dapat diterima untuk ukuran sampel di atas 40. Jika terdapat tiga kategori dari variabel prediktor dan banyaknya observasi sama untuk kategori pertama dan ketiga, maka distribusi permutasi exact dari  $K$  dapat diperoleh dalam kasus "monotonic" pada variabel respon dikotomis dengan banyaknya observasi sama pada setiap bagian dikotomi, untuk total observasi  $N$ , misalkan  $n_1$  obervasi yang berada dalam

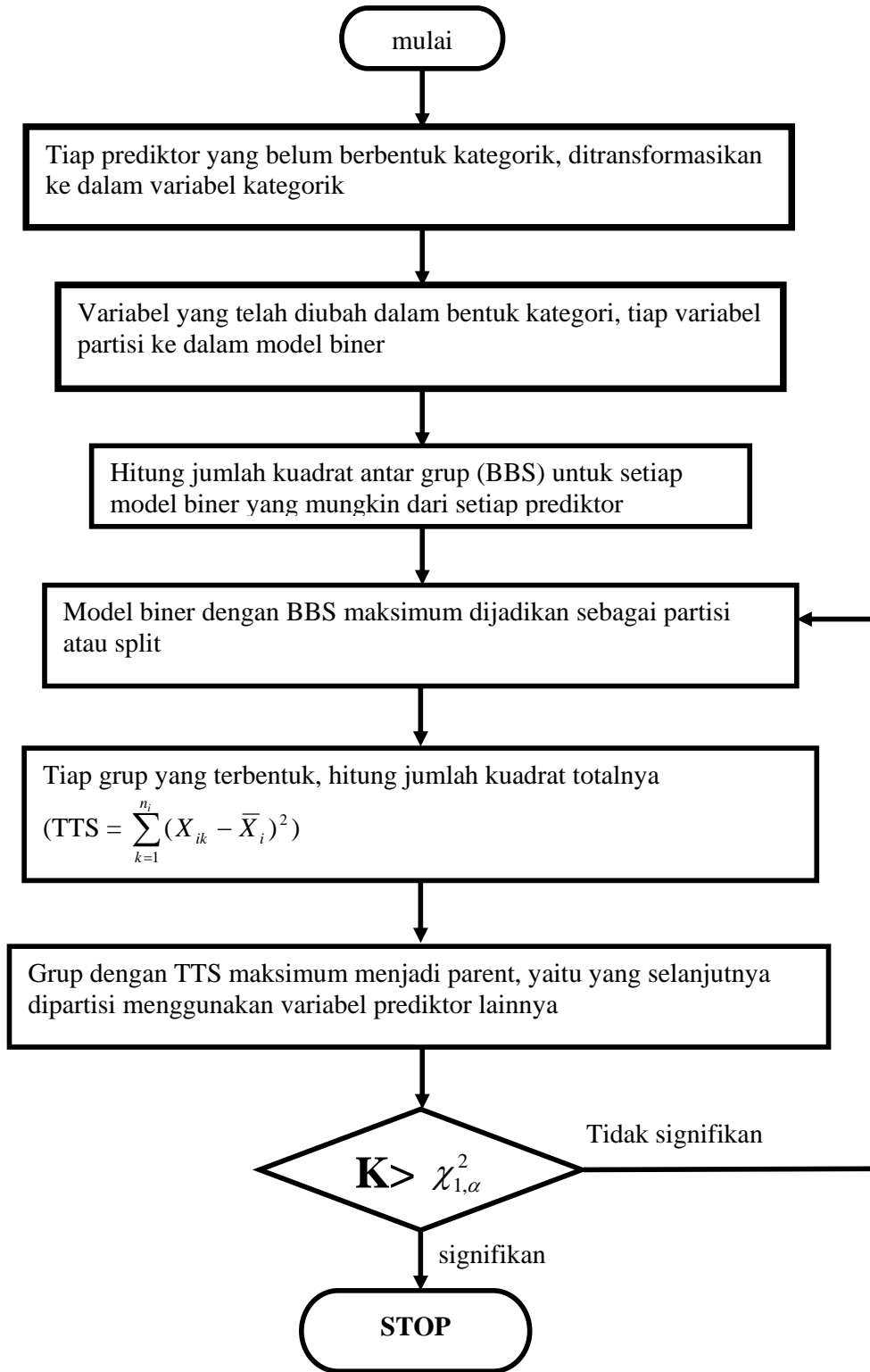
kategori pertama dan  $n_3$  banyaknya observasi yang berada pada kategori ketiga dengan  $n_1 = n_3$ .

Perhatikan proses “random walk” dari (0,0) ke (N,0) dengan langkah positif dan negatif sebagaimana suatu observasi berada dalam sebuah klas dikotomi atau kelas yang lain, dapat diperlihatkan bahwa distribusi K bergantung pada posisi pada partikel pada waktu  $n_1$  dan waktu  $N - n_3$ . Dengan menggunakan prinsip refleksi dengan cermat dapat ditunjukkan bahwa :

$$\begin{aligned}
 & P\{K = Ni^2 / (Nn_1 - n_1^2)\} \\
 &= \frac{n_1!n_2!n_3!}{N!} \binom{\frac{1}{2}N}{\frac{1}{2}(n_1+i)} \binom{\frac{1}{2}N}{\frac{1}{2}(n_1-i)} x \left\{ \sum_{k=2}^{N-2} \binom{\frac{1}{2}(N-n_1-i)}{\frac{1}{2}(n_2-2i+k)} \binom{\frac{1}{2}(N-n_1+i)}{\frac{1}{2}(n_2+2i-k)} \right. \\
 & \left. + \sum_{k=0}^2 \binom{\frac{1}{2}(N-n_1-i)}{\frac{1}{2}(n_2-2i+k)} \binom{\frac{1}{2}(N-n_1+i)}{\frac{1}{2}(n_2+2i-k)} \right\}, \text{ jika } i = \pm 1, \pm 2, \dots, \pm n_1 \\
 &= 0
 \end{aligned}$$

*jika tidak maka akan berada pada pihak yang lain*

**Langkah kerja**



**Gambar 1.** Langkah Kerja Metode AID

#### 4. Kesimpulan

Setelah terbentuk pohon AID maka diperoleh 3 informasi:

1. Pengelompokan pengamatan variabel respon berupa keputusan konsumen untuk membeli barang atau jasa dalam sub kelompok yang relatif homogen menurut dimensi variabel penjelas kualitas barang/jasa dan harga.
2. Interaksi dan struktur keterkaitan antar dimensi variabel penjelas yakni dimensi kualitas barang/jasa dan harga.
3. Variabel penjelas yakni kualitas barang /jasa dan harga yang telah tereduksi akibat adanya dominasi dimensi variabel penjelas yang satu terhadap yang lain.

#### DAFTAR PUSTAKA

1. Fielding, A. and O' Muirheartaigh, C.A., *Binary Segmentasi in Survey Analysis, with Particular Referece to AID*, Mc Graw Hill, 1976.
2. Fitzroy, P.T., *Analytical Method for Marketing Management*, Mc Graw Hill, 1976.
3. Kass, G.V., Significance Testing in Automatic Interaction Detection, *Appl. Statist*, 1975, Vol. 24, No. 2.
4. Lovelock, C., *Principles of Service Marketing and Management*, Prentice – Hall, Inc., 1999.
5. Umar, H., *Metode Riset Bisnis*, PT.Gramedia Pustaka Utama, 2002.