

**SELEKSI VARIABEL KUALITATIF
MELALUI *PROPORTIONAL REDUCTION IN UNCERTAINTY* (PRU)**

Achmad Zanbar Soleh , Peris Siregar, Resa Septiani Pontoh

Jurusan Statistika FMIPA Universitas Padjadjaran

Jl. Raya Bandung-Sumedang KM.21 Jatinangor-Sumedang

Email: a_zanbar_s@yahoo.com, echachicha@yahoo.com

ABSTRAK

Proportional Reduction In Uncertainty (PRU) akan digunakan untuk memilih sekelompok variabel kualitatif yang memiliki pengaruh paling dominan dengan pendekatan multinom multivariat. Perumusan PRU akan didasarkan pada konsep entropi shannon dan selanjutnya akan dihitung nilai informasi kumulatif dari setiap variabel dominan yang terpilih.

Kata kunci: PRU, Entropi Shannon, dan informasi kumulatif.

A. PENDAHULUAN

Kegiatan pengambilan sampel yang berbasis data kualitatif seperti data tentang jenis kelamin, jenis pekerjaan, asal daerah, pendidikan, dan lainnya dengan menggunakan angket sering kita jumpai dalam kehidupan sehari-hari. Angket digunakan untuk menjangkau informasi yang dibutuhkan dalam suatu penelitian terkait dengan karakteristik responden.

Penggunaan angket dalam bidang riset pemasaran biasanya digunakan untuk mengetahui perilaku, persepsi, ataupun prevalensi pelanggan terhadap suatu produk barang atau jasa baik itu terhadap produk baru ataupun produk yang telah dipasarkan.

Terkait hal tersebut, perlu kiranya dipelajari kebutuhan (*needs*), harapan (*expectation*), keinginan (*wants*), dan kepuasan (*satisfaction*) dari konsumen produk tersebut. Keempat faktor di atas akan dikemas dalam bentuk pertanyaan-pertanyaan (atribut-atribut) yang disebarakan kepada konsumen.

Salah satu kegiatan analisis statistika dalam memperlakukan setiap atribut yang ditanyakan adalah memilih sekelompok pertanyaan yang dominan mempengaruhi konsumen dalam menilai suatu produk. Hal ini diperlukan untuk menentukan rencana kegiatan yang harus dilakukan produsen dalam menjaga dan meningkatkan pangsa pasar.

Penelitian ini akan menjawab permasalahan bagaimana proses untuk menyeleksi atribut dominan dari sejumlah atribut yang ditanyakan melalui pendekatan *Proportional Reduction in Uncertainty* (PRU).

Maksud dari penelitian ini adalah menerapkan pendekatan *Proportional Reduction in Uncertainty* (PRU) dalam menyeleksi variabel kualitatif yang dominan dalam mempengaruhi topik penelitian, sedangkan tujuannya adalah mendapatkan alternatif formulasi dalam menentukan variabel dominan dari sejumlah variabel yang diamati secara multinomial multivariate melalui *Proportional Reduction in Uncertainty* (PRU).

Seleksi variabel kualitatif menjadi bagian penting dalam mendapatkan variabel dominan yang akan sangat berguna untuk membuat rencana dan strategi perbaikan terhadap suatu produk yang terkendala oleh dana, waktu, dan tenaga.

B. METODE PENELITIAN

Metode PRU dikembangkan oleh Shannon yang didasarkan pada konsep ketidakteraturan (entropi) jawaban dari data kualitatif. Entropi menurut Shannon dapat diinterpretasikan sebagai ukuran ketidakpastian dan dapat dirumuskan dalam bentuk peluang.

Pada kasus diskrit, entropi dapat dijelaskan sebagai berikut:

Misalkan suatu random experiment memiliki ruang sampel $\mathfrak{R} = \{a_1, a_2, \dots, a_n\}$ dengan distribusi peluang p_1, p_2, \dots, p_n dan $p_i = P(a_i)$ yang memenuhi $p_i \geq 0, 1 \leq i \leq n$ dan

$$\sum_{i=1}^n p_i = 1.$$

Selanjutnya entropi pada \mathfrak{R} adalah fungsi

$$H : \left\{ (p_1, p_2, \dots, p_n) \mid p_i \geq 0, 1 \leq i \leq n, \sum_{i=1}^n p_i = 1 \right\} \rightarrow [0, \log n]$$

Yang didefinisikan oleh $H(p_1, p_2, \dots, p_n) = -\sum_{k=1}^n p_k \cdot \log p_k$ dengan $p_k \cdot \log p_k = 0$ untuk $p_k = 0$. Dengan kata lain, entropi pada \mathfrak{R} dapat dituliskan $H(\mathfrak{R}) = H(p_1, p_2, \dots, p_n)$.

Fungsi H atau entropi ini harus memenuhi sifat-sifat berikut:

1. $H(p_1, p_2, \dots, p_n) = -\sum_{k=1}^n p_k \cdot \log p_k = 0$ jika dan hanya jika $p_i = 0$ untuk suatu i dan $p_j = 0$ yang berlaku apabila $i \neq j; 1 \leq i \leq n$.

2. Untuk suatu distribusi peluang $p_1, p_2, \dots, p_n; \sum_{i=1}^n p_i = 1$ akan berlaku hubungan

$$\text{berikut ini: } H(p_1, p_2, \dots, p_n) \leq H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right)$$

Taksiran entropi yang melibatkan dua variabel kualitatif dapat diperoleh melalui tabel kontingensi. Secara umum apabila yang terlibat lebih dari dua variabel kualitatif, maka taksiran entropi dapat diperoleh dari tabel kontingensi multiple.

Selanjutnya rumusan PRU akan didefinisikan sebagai berikut. Misalkan X_i dan X_j adalah variabel kualitatif dengan himpunan modalitas masing-masing $\{x_{i1}, x_{i2}, \dots, x_{ik}\}$ dan $\{x_{j1}, x_{j2}, \dots, x_{jk}\}$. Proporsi X_j yang diterangkan oleh X_i dapat diukur melalui: $PRU(X_j \setminus X_i) = \frac{H(X_j) - H(X_j \setminus X_i)}{H(X_j)}$ dengan $H(X_i)$ dan $H(X_j)$ menyatakan entropi marginal dan $H(X_j \setminus X_i)$ menyatakan entropi bersyarat.

Misalkan X_1, X_2, \dots, X_p adalah variabel-variabel kualitatif maka tahapan memilih k buah variabel dominan melalui PRU yang dikembangkan oleh Shannon akan dijelaskan sebagai berikut:

1. Buat matrik M ukuran $p \times p$ dengan $M_{ij} = PRU(X_i \setminus X_j)$
2. Hitung jumlah setiap baris
3. Pilih variabel yang memiliki jumlah baris paling maksimum
4. Hapus baris dan kolom yang memiliki jumlah baris paling maksimum
5. Kembali ke tahap 2.
6. Proses berhenti pada saat menyisakan 2 baris.

Untuk mengetahui kualitas dari setiap variabel dominan yang terpilih maka digunakan informasi kumulatif (CI) yang dirumuskan sebagai berikut:

$$CI(k) = CI(k-1) + [1 - CI(k-1)] \frac{SPRU(k)}{(n-k+1)}$$

Dengan $CI(0) = 0$, $SPRU(k)$ menyatakan jumlah PRU dari variabel terpilih pada tahap ke-k.

Dengan demikian, proses pemilihan faktor-faktor dominan melalui metode PRU telah dapat diaplikasikan melalui data ordinal yang diperoleh melalui penyebaran angket.

C. HASIL PENELITIAN DAN PEMBAHASAN

Aplikasi pemilihan faktor-faktor dominan melalui metode PRU dilakukan untuk mengukur kepuasan nasabah suatu bank terhadap pelayanan yang diberikan oleh karyawannya. Penelitian ini menggunakan data yang diperoleh melalui penyebaran angket kepada 22 orang responden yang mengisi secara lengkap setiap item pertanyaan yang diajukan. Terdapat 21 variabel digunakan untuk mengukur tingkat kepuasan nasabah bank yakni:

1. Kesederhanaan penggunaan rekening (K1)
2. Kompetensi pegawai (K2)

3. Pengurusan faktur oleh bank (K3)
4. Kualitas penyambutan (K4)
5. Kerendahan hati pegawai (K5)
6. Bank terorganisir dengan baik (K6)
7. Kecepatan transfer (K7)
8. Kemudahan mendapatkan buku cek (K8)
9. Pemeliharaan hubungan personal (K9)
10. Kemudahan mendapat informasi (K10)
11. Kemudahan transfer melalui rekening (K11)
12. Pegawai yang bersahabat (K12)
13. Keseriusan bank menyelesaikan masalah (K13)
14. Kesiapan pegawai melaksanakan tugasnya (K14)
15. Keamanan rekening (K15)
16. Bank terpercaya (K16)
17. Frekuensi penarikan rekening (K17)
18. Kecepatan pelayanan (K18)
19. Keamanan di bank (K19)
20. Keteraturan pengambilan rekening (K20)
21. Kesederhanaan operasi secara keseluruhan (K21)

Tabel di bawah ini menampilkan hasil perhitungan PRU dan kumulatif informasi (CI) dari data kepuasan pelayanan perbankan.

Tahap	Item terpilih	SPRU	IC (%)
1	K16	3,28992	16,45
2	K18	2,78588	28,70
3	K8	2,56204	38,85
4	K4	2,17148	46,66
5	K21	2,00645	53,35
6	K1	1,84291	59,08
7	K6	1,77581	64,27
8	K7	1,72720	69,02
9	K10	1,48485	72,85
10	K14	1,45307	76,44

11	K13	1,38764	79,71
12	K3	1,35524	82,77
13	K15	1,25656	85,48
14	K20	1,12565	87,82
15	K11	1,09089	90,03
16	K19	1,04415	92,11
17	K5	1,01964	94,12
18	K17	1,00864	96,10
19	K9	1,00136	98,05
20	K12	1,00000	100

Terpilihnya K16; bank terpercaya pada tahap 1 proses pemilihan variabel dominan memberikan informasi bahwa nasabah bank menjadikan k16 (bank terpercaya) sebagai prioritas utama yang mempengaruhi pelayanan perbankan dengan menyumbang informasi sebesar 16,45%. Dengan demikian hal pertama yang penting diperhatikan oleh suatu bank adalah menanamkan kepercayaan terhadap nasabah dalam meningkatkan pelayanan mereka.

Faktor dominan kedua yang mempengaruhi pelayanan bank kepada nasabah adalah K18 (kecepatan pelayanan) dengan sumbangan informasi sebesar 12,25%. Terkait hal tersebut, bank diharuskan melatih dan meningkatkan kemampuan para pegawainya untuk dapat bekerja secara cepat dan terampil. Melalui proses yang sama, maka faktor dominan terakhir adalah K12 (pegawai yang bersahabat) dengan sumbangan informasi sebesar 1,95%.

D. KESIMPULAN DAN SARAN

PRU memberikan rumusan pemilihan variabel dominan dengan memperhatikan ketidakpastian jawaban yang diberikan oleh responden. Nilai SPRU berbanding lurus dengan besarnya sumbangan informasi yang diberikan oleh variabel tersebut. Variabel dengan nilai SPRU paling besar adalah variabel yang mendapatkan prioritas utama untuk diperhatikan dan juga memberikan sumbangan informasi terbesar terhadap permasalahan yang dihadapi.

Penelitian ini hanya mempelajari proses menghitung nilai SPRU dengan asumsi bahwa setiap sel dalam tabel kontingensi terisi. Sedangkan pada kenyataan sehari-hari banyak ditemukan kasus dengan nilai sel tabel kontingensi bernilai nol atau kosong. Untuk itu perlu adanya penelitian lanjutan terkait penentuan nilai derajat kebebasan jika ditemukan sel bernilai nol atau kosong pada tabel kontingensi.

E. DAFTAR PUSTAKA

1. Agresti, A., 1996. An Introduction To Categorical Data Analysis, A Wiley-Interscience Publication.
2. Budhi, W.S., 1994. Kalkulus Multi Variabel dan Pemakaiannya. Diktat Kuliah MA-212, Jurusan Matematika ITB.
3. De Falguerolles, A.J.S., 1993, Un Modele Graphique pour la Selection de variables Qualitatives, Revue de Statistiques Appliquees, XLI(2), 23-41.
4. Hogg. R.V. and Craig, A.T., 1995. Introduction to Mathematical Statistic, Macmillan Publishing Co., Inc. New York.
5. Klir, G.J. and Folger, T.A., 1988. Fuzzy Set, Uncertainty, and Information. Prentice Hall, Englewood Cliffs, New Jersey 07632.
6. Reynold, H.T., 1977. Analysis Of Nominal Data. Sage Publication, Inc.