

---

---

## ANALISIS KLASIFIKASI KEMISKINAN DI KOTA SEMARANG MENGUNAKAN ALGORITMA QUEST

Dwi Ispriyanti<sup>1</sup>, Alan Prahutama<sup>2</sup> dan Mustafid<sup>3</sup>

<sup>1,2,3</sup>Departemen Statistika, Fakultas Sains dan Matematika Universitas Diponegoro  
Alamat Email: dwiispriyanti@yahoo.com

### ABSTRAK

Kemiskinan menjadi suatu permasalahan bagi negara-negara berkembang termasuk Indonesia. Kemiskinan merupakan keadaan kondisi perekonomian suatu orang atau sekelompok orang yang tidak bisa memenuhi kebutuhan dasar seperti pendidikan, kesehatan, pangan, perumahan dan lainnya. Analisis klasifikasi kemiskinan merupakan bagian dari analisis kemiskinan yang mengategorikan rumah tangga atau kelompok kedalam kategori miskin dan tidak miskin. Pengkategorian tersebut didasarkan pada pengeluaran perkapita yang dibandingkan dengan nilai garis kemiskinan. Metode klasifikasi didalam statistika salah satunya pohon klasifikasi, yang meliputi antara lain Algoritma CART, QUEST, ID3, C45 dan lainnya. Algoritma QUEST merupakan pohon klasifikasi biner dengan prosedur pemilihan variabel penyekat/pemisah, penentuan titik sekat/pemisah serta proses pemberhentian. Pada penelitian ini hasil klasifikasi menggunakan algoritma QUEST dengan semua variabel prediktor diasumsikan skala rasio maka hasil klasifikasi yang didapat mempunyai akurasi 94.9%. Variabel prediktor yang mempengaruhi antara lain penerimaan beras miskin, jenis bahan bakar utama untuk memasak, jenis dinding rumah tinggal yang digunakan dan sumber utama air minum. Sedangkan hasil klasifikasi menggunakan algoritma QUEST dengan variabel prediktornya diasumsikan skala nominal juga menghasilkan akurasi 94.9%. Variabel prediktor yang mempengaruhi antara lain penerimaan beras miskin, jenis bahan bakar utama yang digunakan untuk memasak, bahan utam dinding rumah serta jumlah anggota rumah tangga.

**Kata Kunci:** QUEST, klasifikasi kemiskinan, Tingkat Akurasi

### PENDAHULUAN

Penurunan tingkat kemiskinan menjadi salah satu program utama Sustainable Development Goal's (SDG's) yang telah disepakati negara-negara di dunia. Kemiskinan juga menjadi salah satu permasalahan bagi negara-negara berkembang termasuk Indonesia. Menurut Undang-Undang No. 24 tahun 2004, kemiskinan merupakan kondisi perekonomian seseorang atau sekelompok orang yang tidak terpenuhinya hak-hak dasarnya untuk

mempertahankan dan mengembangkan kehidupan yang bermartabat. Kebutuhan dasar tersebut meliputi kebutuhan pangan, kesehatan, pendidikan, perumahan, pekerjaan, air bersih, pertanahan, sumber daya alam, lingkungan hidup. Kota Semarang merupakan ibu kota provinsi di Jawa Tengah yang menjadi pusat perekonomian. Analisis kemiskinan di suatu ibu kota provinsi menjadi sangat penting, karena bisa mengukur kemajuan wilayah lainnya. Menurut Badan Pusat

Statistika, berdasarkan pengeluaran per kapita tiap bulan, rumah tangga dapat dikategorikan menjadi rumah tangga miskin dan rumah tangga tidak miskin menggunakan garis kemiskinan.

Analisis kemiskinan pada paper ini merupakan analisis yang didasarkan pada ketepatan klasifikasi pengkategorian rumah tangga miskin dan tidak miskin yang ada di kota Semarang. Didalam analisis statistika yang termasuk kedalam metode klasifikasi antara lain Classification Regression Tree (CART), Algoritma ID3, Algoritma C45, dan Algoritma QUEST.

Algoritma QUEST (*Quick, Unbiased, and Efficient Statistical Tree*) pertama kali dikenalkan tahun 1997 oleh Loh dan Shih. QUEST termasuk ke dalam pohon klasifikasi biner, yaitu pohon yang menghasilkan dua simpul setiap sekatnya. Pohon klasifikasi QUEST dibentuk dari pemilihan variabel penyekat, pemilihan titik sekat untuk variabel yang dipilih, dan proses pemberhentian [1].

Pemilihan variabel penyekat digunakan untuk menentukan simpul yang akan dibentuk. Variabel penyekat dapat dipilih dari variabel prediktor kontinu maupun kategorik. Variabel prediktor kontinu digunakan uji ANOVA F untuk menguji perbedaan rata-rata antar kelas pada variabel prediktor kontinu  $X$  atau digunakan uji Levene F untuk menguji kesamaan varian antar kelas. Sedangkan untuk variabel kategorik digunakan uji chi-kuadrat Pearson ( $\chi^2$ ) untuk menguji kebebasan antar variabel respon  $Y$  dan variabel prediktor kategorik  $X$  [2].

Pemilihan variabel penyekat pada QUEST adalah memilih nilai  $p$ -value terkecil dari hasil uji setiap variabel prediktor terhadap variabel respon dan

dibandingkan dengan koreksi Bonferroni. Koreksi Bonferroni diusulkan untuk menggunakan tingkat kepercayaan lebih kecil ketika beberapa uji dilakukan. Ketika  $n$  uji dilakukan maka masing-masing uji menggunakan tingkat kepercayaan  $\frac{\alpha}{n}$ , dengan  $\alpha$  adalah tingkat

## METODOLOGI PENELITIAN

Data yang digunakan dalam penelitian ini merupakan data Susenas 2016 di kota Semarang. Data yang diambil sebanyak 930 rumah tangga. Klasifikasi kemiskinan didasarkan pada pengeluaran per kapita (makanan dan nonmakanan). Jika pengeluaran per kapita selama sebulan kurang dari batas garis kemiskinan maka dikategorikan miskin. Jika ini melebihi dari nilai garis kemiskinan dikatakan tidak miskin. Nilai garis kemiskinan di kota Semarang pada tahun 2016 bernilai Rp 382.160.

Adapun variabel yang digunakan adalah sebagai berikut variabel dependen bernilai kategorik (1= miskin; 0=tidak miskin);  $X_1$ : Jenis Kelamin kepala RT (1=Laki-Laki; 2=perempuan);  $X_2$ : umur kepala RT;  $X_3$ : banyaknya anggota RT;  $X_4$ :Ijazah tertinggi kepala RT (1: Tidak ada ijazah; 2: SD; 3: SMP; 4:SMA; 5: D1/D2/D3; 6: S1; 7: S2/S3);  $X_5$ : Apakah kepala RT bekerja (1: Ya; 2=Tidak);  $X_6$ =lapangan usaha (0: Tidak bekerja; 1=pertanian; 2=pertambangan; 3=Industri pengolahan; 4: Listrik dan Gas; 5: bangunan; 6: perdagangan, hotel dan rumah makan; 7: Transportasi dan pergudangan; 8: Keuangan dan asuransi; 9: Jasa);  $X_7$ =status pekerjaan (0: Tidak bekerja; 1:berusaha sendiri; 2: berusaha dibantu buruh tidak tetap; 3: berusaha dibantu buruh tetap; 4: Buruh/karyawan/pegawai; 5: pekerja bebas; 6: pekerja keluarga);  $X_8$ =status kepemilikan rumah (1: milik sendiri; 2:

Kontrak/Sewa; 3: Bebas sewa; 4: Dinas; 5: Lainnya);  $X_9$ =bahan bangunan atap (1: Beton; 2: Genteng Keramik; 3: Genteng Metal; 4: Genteng Tanah Liat; 5: Asbes; 6: Seng; 7: Bambu; 8: Kayu/Sirap; 9: Jerami; 10: Lainnya);  $X_{10}$ =Bahan Utama dinding Rumah (1: Tembok; 2: kayu; 3: batang kayu; 4: Lainnya);  $X_{11}$ =Bahan Utama Lantai Rumah (1: Marmer/Granit; 2: Keramik; 3:Ubin/Tegel; 4; Semen/bata merah; 5: Tanah; 6: Lainnya);  $X_{12}$ = Penggunaan Fasilitas tempat buang air besar (1:Ada digunakan hanya ART sendiri; 2: Ada, digunakan bersama dengan ART lainnya; 3: Ada, di MCK Umum; 4: Tidak ada);  $X_{13}$ = Tempat pembuangan akhir tinja (1:tangki dengan dasar semen; 2: Tangki tanpa dasar semen; 3: IPAL; 4: Kolam/Sawah/Sungai/Danau; 5: Lubang Tanah; 6:Lainnya);  $X_{14}$ =Sumber Air Minum (1: Air kemasan bermerk; 2: Air isi ulang; 3:Leding Meteran; 4: Leding eceran; 5: sumur bor/pompa; 6:Sumur terlindung; 7: Sumur Tak terlindung; 8:Mata air terlindung; 9: Lainnya); ;  $X_{15}$ =Bahan bakar utama memasak (0: tidak memasak dirumah; 1:Listrik; 2: elpiji 5.5kg; 3: elpiji 12kg; 4:elpiji 3 kg; 5: minyak tanah; 6: kayu bakar);  $X_{16}$ =menerima beras miskin (1: Ya; 2: Tidak).

Analisis yang akan dilakukan meliputi dua skenario. Skenario pertama adalah dengan mengasumsikan variabel independen yang digunakan adalah berskala rasio. Skenario kedua adalah variabel independen yang digunakan merupakan skala nominal.

## HASIL DAN PEMBAHASAN

### Variabel independen diasumsikan berskala rasio

Pada klasifikasi menggunakan algoritma QUEST ini, mengacu pada

enam belas variabel yang digunakan kedalam model regresi logistik biner. Algoritma QUEST termasuk dalam kategori pohon klasifikasi biner. Algoritma Quest mampu mengatasi data dalam bentuk diskrit maupun kontinue untuk variabel independen yang digunakan. Pemisahan variabel menggunakan Anova F, Levene test, dan Chi-Square.

Hasil dari penerapan algoritma QUEST ke data kategori rumah tangga miskin dan tidak miskin berupa pohon klasifikasi dapat dilihat pada Lampiran 1. Dengan menggunakan algoritma QUEST, dihasilkan suatu pohon dengan aturan klasifikasi yang digunakan untuk memprediksi apakah rumah tangga tersebut dikategorikan miskin atau tidak. Dari pohon yang terbentuk, variabel yang penting dalam mengklasifikasikan rumah tangga miskin atau tidak antara lain penerimaan beras miskin ( $X_{16}$ ) Bahan bakar utama untuk memasak ( $X_{15}$ ), Sumber utama air untuk minum ( $X_{14}$ ), serta bahan utama dinding rumah tinggal ( $X_{10}$ ).

Dari pohon klasifikasi yang terbentuk terdapat sembilan simpul akhir, sehingga didapat aturan klasifikasi sebagai berikut:

1. Pernah menerima beras miskin  $\leq 1.5$  dan Bahan Bakar Utama untuk memasak adalah  $\leq 4.1$  maka rumah tangga tersebut akan dikategorikan tidak miskin. Artinya suatu rumah tangga yang cenderung untuk menerima beras miskin, akan tetapi bahan bakar yang digunakan untuk memasak menggunakan listrik, elpiji 3 kg atau 12 kg makan akan dikategorikan tidak miskin.
2. Pernah menerima beras miskin  $\leq 1.5$  dan Bahan Bakar Utama untuk memasak adalah  $\geq 4.1$  maka rumah

tangga tersebut akan dikategorikan tidak miskin ataupun miskin (peluangnya sama). Artinya suatu rumah tangga yang cenderung untuk menerima beras miskin, dengan bahan bakar yang digunakan untuk memasak menggunakan minyak, elpiji 3 kg, arang, bara/briket/lainnya

3. Pernah menerima beras miskin  $\geq 1.5$  dan sumber utama untuk minum  $\leq 3.1$  maka akan dikategorikan rumah tangga tidak miskin. Artinya suatu rumah tangga yang cenderung tidak menerima beras miskin dan bahan utama sumber air minum adalah air kemasan bermerk, air isi ulang dan air ledeng meteran maka rumah tangga tersebut dikategorikan masuk kedalam rumah tangga tidak miskin.
4. Pernah menerima beras miskin  $\geq 1.5$ ; sumber utama untuk minum  $> 3.1$ ; serta bahan utama dinding rumah tinggal adalah  $\leq 1.8$  maka rumah tangga tersebut dikategorikan tidak miskin. Artinya suatu rumah tangga cenderung tidak menerima beras miskin, sedangkan sumber air minum yang digunakan ledeng eceran, sumur bor, sumur terlindung/tak terlindung, mata air tak terlindung/terlindung, air hujan. Sedangkan dinding utama rumah tinggal berbahan dasar tembok maka rumah tangga tersebut dikategorikan tidak miskin. Artinya sumber utama air minum walaupun menggunakan sumur, tetapi bukan faktor penentu utama rumah tangga tersebut dikategorikan miskin.
5. Pernah menerima beras miskin  $\geq 1.5$ ; sumber utama untuk minum  $> 3.1$ ; serta bahan utama dinding rumah tinggal adalah  $> 1.8$  maka rumah tangga tersebut dikategorikan tidak miskin. Artinya suatu rumah tangga cenderung tidak menerima beras

miskin, sedangkan sumber air minum yang digunakan ledeng eceran, sumur bor, sumur terlindung/tak terlindung, mata air tak terlindung/terlindung, air hujan. Sedangkan dinding utama rumah tinggal berbahan dasar kayu atau bambu maka rumah tangga tersebut dikategorikan tidak miskin.

Tingkat akurasi yang dihasilkan adalah sebagai berikut:

**Tabel 1.** Tingkat akurasi klasifikasi kemiskinan di kota Semarang menggunakan Algoritma QUEST

Variabel	Prediksi	
	Tidak Miskin	Miskin
Tidak Miskin	883	0
Miskin	47	0

Tabel 1 menunjukkan nilai akurasi klasifikasi kemiskinan di kota Semarang menggunakan algoritma QUEST mencapai 94.9% yang dikategorikan sangat tinggi. Akan tetapi perlu dilihat kategori antara rumah tangga miskin dan tidak miskin perbandingannya cukup tinggi yaitu 47 banding 883. Jika dilihat hasil akurasinya, kategori rumah tangga tidak miskin sangat tepat 100% diklasifikasikan kedalam rumah tangga tidak miskin. Akan tetapi untuk kategori rumah tangga miskin, semuanya diklasifikasikan kedalam rumah tangga tidak miskin, sehingga ketepatannya 0%. Hal ini dapat disimpulkan bahwa, Algoritma QUEST akan baik digunakan untuk menginterpretasikan rumah tangga tidak miskin. Algoritma QUEST tidak bisa menginterpretasikan rumah tangga miskin.

**Variabel independen diasumsikan berskala nominal**

Lampiran 2 menunjukkan pohon klasifikasi menggunakan Algoritma QUEST untuk variabel independen berskala nominal. Berdasarkan gambar

tersebut faktor-faktor yang menjadi penentu klasifikasi rumah tangga miskin di kota Semarang antara lain penerimaan beras miskin, jenis bahan bakar utama yang digunakan untuk memasak, bahan utama dinding rumah serta jumlah anggota rumah tangga. Berdasarkan hasil analisis dapat dibuat beberapa penjabaran sebagai berikut:

1. Rumah tangga yang menerima beras miskin 14.5% dikategorikan miskin, 85.5% dikatehorikan tidak miskin.
2. Rumah tangga yang tidak menerima beras miskin 98% dikategorikan rumah tangga tidak miskin, 2% dikategorikan rumah tangga miskin.
3. Rumah tangga yang tidak menerima beras miskin, jenis bahan bakar utama untuk memasak adalah minyak tanah, 66.7% dikategorikan tidak miskin dan 33.3% dikategorikan miskin.
4. Rumah tangga yang tidak menerima beras miskin, jenis bahan bakar utama untuk memasak adalah selain minyak tanah, 98.1% dikategorikan tidak miskin dan 1.9% dikategorikan miskin
5. Rumah tangga yang tidak menerima beras miskin, jenis bahan bakar utama untuk memasak adalah selain minyak tanah, dan bahan utama dinding rumah tempat tinggal, berbahan dasar tembok maka 98.5% dikategorikan tidak miskin sementara 1.5% dikategorikan miskin.
6. Rumah tangga yang tidak menerima beras miskin, jenis bahan bakar utama untuk memasak adalah selain minyak tanah, dan bahan utama dinding rumah tempat tinggal, berbahan dasar plesteran anyaman bamboo atau anyaman bamboo maka 90.6% dikategorikan tidak miskin, 9.4% dikategorikan miskin.

7. Rumah tangga yang tidak menerima beras miskin, jenis bahan bakar utama untuk memasak adalah selain minyak tanah, dan bahan utama dinding rumah tempat tinggal, berbahan dasar tembok dan banyaknya anggota rumah tangga kurang dari 8 sampai 9 orang maka 98.6% dikatergorikan tidak miskin dan 1.4 dikategorikan miskin
8. Rumah tangga yang tidak menerima beras miskin, jenis bahan bakar utama untuk memasak adalah selain minyak tanah, dan bahan utama dinding rumah tempat tinggal, berbahan dasar tembok dan banyaknya anggota rumah tangga lebih dari dari 8 sampai 9 orang maka 66.7 dikategorikan tidak miskin dan 33.3% dikategorikan miskin.

#### KESIMPULAN

Berdasarkan hasil analisis yang didapatkan dapat disimpulkan bahwa

1. Kategori rumah tangga miskin dibandingkan dengan rumah tangga tidak miskin sangat sedikit sekali yaitu 47: 883.
2. Variabel prediktor diasumsikan skala rasio maka variabel yang mempengaruhi antara lain penerimaan beras miskin, jenis bahan bakar utama untuk memasak, jenin dinding rumah tinggal yang digunakan dan sumber utama air minum.
3. Sedangkan hasil klasifikasi menggunakan algoritma QUEST dengan variabel prediktornya diasumsikan skala nominal, Variabel prediktor yang mempengaruhi antara lain penerimaan beras miskin, jenis bahan bakar utama yang digunakan untuk memasak, bahan utaam dinding

---

---

rumah serta jumlah anggota rumah tangga.

4. Nilai akurasi yang dihasilkan dengan variabel independen diasumsikan berskala nominal ataupun rasio mempunyai nilai akurasi sebesar 94.9%.

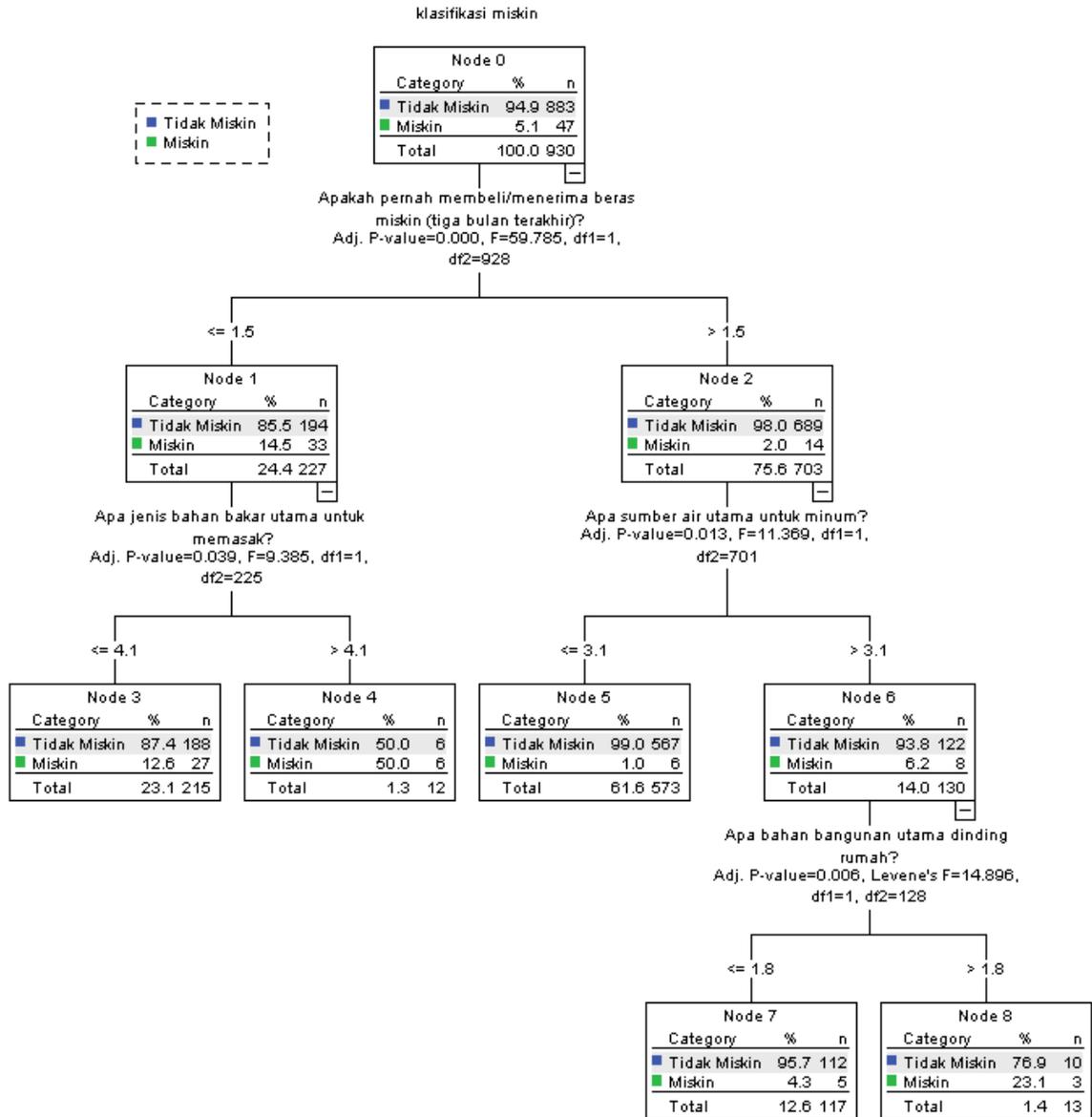
#### UCAPAN TERIMA KASIH

Penelitian ini disupport oleh Universitas Diponegoro melalui skim penelitian Riset Penerapan dan Pengembangan (RPP) dengan nomor kontrak 474-26/UN7.P4.3/PP/2018

#### DAFTAR PUSTAKA

1. Berry, M. J. A. dan Linoff, G. S. 2004. *Data Mining Techniques*. Wiley Publishing, Inc: Indiana.
2. Grabczewski, K. 2014. *Meta-Learning in Decision Tree Induction*. Springer International Publishing, Switzerland.
3. Kim, H. dan Loh, W.-Y. 2001. Classification Trees with Unbiased Multiway Splits. *Am. Statist. Assoc.* 96; 590-604.
4. Larose, D. T. 2005. *Discovering Knowledge in Data "An Introduction to Data Mining"*. John Wiley dan Sons, Inc: New Jersey.
5. Loh, W.-Y. dan Shih, Y.-S. 1997. Split Selection Methods for Classification Trees, *Statistica Sinica* 7; 815-840.
6. Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I. dan de Mendonca, A. 2011. Data mining methods in the prediction of Dementia. *BMC Research Notes*; 4:299.
7. Rokach, L. dan Maimon, O. 2008. *Data Mining with Decision Trees "Theory and Application"*. World Scientific Publishing Co. Pte. Ltd. : USA.

**Lampiran 1. Hasil Pohon Klasifikasi menggunakan Algoritma QUEST dengan variabel independen berskala rasio**



**Lampiran 2. Hasil Pohon Klasifikasi menggunakan ALgoritma QUEST dengan variabel independen berskala Nominal**

