

IMPLENTASI ALGORITMA CLASSIFICATION AND REGRESSION TREES (CART) DALAM KLASIFIKASI PENERIMA BANTUAN SOSIAL BERAS SEJAHTERA (Studi Kasus : Desa Kuala Sempang Kabupaten Bintan)

Angga Andarama, Nurfalinda, Nola Ritha

Anggaandaramah@gmail.com

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Maritim Raja Ali Haji

Abstract

Poverty is a problem that is being faced by national development in Indonesia. Social assistance is one way for the government to overcome the problem. In order to help analyze gaps in the community the author decided to conduct a study using the Classification And Regression Trees (CART) Algorithm, in this study poverty survey data for the Kuala Sangkang village residents in 2018 and data for RASTRA recipients in the Kuala Sangkang Village in 2018. Parameters used include KRT name, number of household members, floor area, floor type, wall type, drinking water source, defecation facility, lighting source, energy for cooking and PKH status. Based on the results of analysis and testing using the Classification and Regression method Trees (CART), the resulting accuracy of 76.47%.

Keywords: CART algorithm, social assistance, poverty, WEB

I. Pendahuluan

Kemiskinan merupakan persoalan paling mendasar yang menjadi pusat perhatian pemerintah di negara manapun, Kemiskinan merupakan gambaran dari perekonomian dinegara tersebut. suatu negara dikatakan miskin biasanya ditandai dengan tingkat pendapatan perkapita rendah, mempunyai tingkat pertumbuhan penduduk yang tinggi (lebih dari 2 persen per tahun) dan sebagian besar tenaga kerja bergerak di sektor pertanian dan terbelenggu dalam lingkaran kemiskinan (Munir, 2002).

Di Indonesia kemiskinan menjadi salah satu permasalahan yang menjadi fokus utama pemerintah, kemiskinan menjadi semakin kompleks karena kemiskinan tidak hanya berkaitan dengan rendahnya tingkat pendapatan, tetapi juga berkaitan dengan rendahnya tingkat pendidikan dan juga kesehatan, oleh karena itu tingkat kemiskinan biasanya menjadi tolak ukur tingkat kesejahteraan masyarakat. Dalam upaya penanggulangan setiap tahunnya pemerintah selalu mencanangkan upaya penanggulangan kemiskinan dengan membuat program-program dalam upaya mengurangi angka kemiskinan dan mensejahterakan masyarakat. Salah satu program penanggulangan kemiskinan yang dilakukan oleh pemerintah adalah program pemberian bantuan sosial beras sejahtera (BANSOS RASTRA), yaitu berupa bantuan pangan dalam bentuk beras (natura) yang diberikan oleh pemerintah untuk disalurkan setiap bulannya dengan tujuan untuk mengurangi beban pengeluaran keluarga tanpa dikenakan biaya tebus/harga. Beras yang disalurkan adalah beras berkualitas medium sejumlah 10 Kg.

Program pemberian bantuan sosial beras sejahtera (BANSOS RASTRA) juga dilaksanakan di Desa Kuala Sempang Kabupaten Bintan Kepulauan Riau. Dalam pelaksanaannya di masyarakat pemerintah setempat memberikan bantuan berdasarkan kriteria kemiskinan yang digunakan oleh

Badan Pusat Statistik (BPS), Beberapa contoh parameter yang digunakan diantaranya adalah luas lantai, jenis2 lantai, luas dinding, sumber penerangan, dll. Dalam rangka untuk membantu pemerintah Kabupaten Bintan untuk menganalisa kelayakan penerima RASTRA di masyarakat penulis memutuskan membuat suatu program yang dapat digunakan untuk mengklasifikasikan masyarakat yang akan menerima BANSOS tersebut.

Algoritma *Classification and Regression Trees (CART)* dipilih oleh peneliti untuk mengklasifikasikan penerima bantuan sosial beras sejahtera (BANSOS RASTRA) di Desa Kuala Sempang Kabupaten Bintan karena metode *CART* merupakan salah satu metode atau algoritma dari salah satu teknik eksplorasi data yaitu teknik pohon keputusan. *CART* terbilang sederhana namun merupakan metode yang kuat. *CART* bertujuan untuk mendapatkan suatu kelompok data yang akurat sebagai penciri dari suatu pengklasifikasian. Hal ini didasarkan pada penelitian yang sebelumnya dilakukan oleh Ritno, dkk (2018) yaitu “Implementasi Algoritma *Classification and Regression Trees (CART)* dalam Klasifikasi Ekonomi Keluarga pada Desa Dagang Kelambir Tg. Morawa”. Pada penelitian ini peneliti mengklasifikasikan ekonomi keluarga pada Dusun IV Desa Dagang Kelambir untuk mengetahui keluarga yang berhak menerima program bantuan dengan menggunakan Algoritma *Classification and Regression Trees (CART)*, hasil dari penelitian ini dari 45 sampel data yang digunakan dihasilkan 3 klasifikasi yaitu Tinggi, Sedang dan Rendah.

II. Metode Penelitian

2.1 Data Mining

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam database. proses yang menggunakan tehnik statistik, matematika, kecerdasan buatan, dan mechine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Turban dkk. 2005)

2.2 Metode *Classification and Regression Trees (CART)*

CART (Classification and Regression Trees) merupakan salah satu metode atau algoritma dari salah satu teknik eksplorasi data yaitu teknik pohon keputusan. *CART* terbilang sederhana namun merupakan metode yang kuat. *CART* bertujuan untuk mendapatkan suatu kelompok data yang akurat sebagai penciri dari suatu pengklasifikasian, selain itu *CART* digunakan untuk menggambarkan hubungan antara variabel respon (variabel dependen atau tak bebas) dengan satu atau lebih variabel prediktor (variabel independen atau bebas).

Model pohon yang dihasilkan bergantung pada skala variabel respon, jika variabel respon data berbentuk kontinu maka model pohon yang dihasilkan adalah *Regression Trees* (pohon regresi) sedangkan bila variabel respon mempunyai skala kategorik maka pohon yang dihasilkan adalah *Classification Trees* (pohon klasifikasi) (Breiman dkk. 1993). Tujuan utama *CART* adalah untuk mendapatkan suatu kelompok data yang akurat sebagai penciri dari suatu pengklasifikasian (Timofeev, 2004). Pohon klasifikasi merupakan metode penyekatan data secara berulang (rekursif) dan secara biner (*binary recursive partitioning*), karena selalu membagi kumpulan data menjadi dua sekatan. Setiap sekatan data dinyatakan sebagai node dalam pohon yang terbentuk. Pembentukan pohon klasifikasi dilakukan melalui penyekatan gugus data dengan sederetan penyekat biner sampai dihasilkan node akhir

2.3 Proses Algoritma *Classification and Regression Trees (CART)*

Ada beberapa langkah-langkah yang terdapat didalam algoritma *CART* (Widodo dkk. 2014) adalah sebagai berikut :

- a. Menyusun Calon Cabang, Tahapan pertama dalam algoritma *CART* yaitu dengan menyusun calon cabang (*candidate split*), penyusunan ini dilakukan terhadap seluruh variabel prediktor secara lengkap (*exhaustive*).

- b. Menghitung Nilai PL dan PR Untuk menghitung nilai PL dan PR digunakan persamaan sbb :

$$PL = \frac{\text{jumlah catatan pada calon cabang kiri } tL}{\text{jumlah catatan pada data latihan}}$$

$$PR = \frac{\text{jumlah catatan pada calon cabang kanan } tR}{\text{jumlah catatan pada data latihan}}$$

- c. Menghitung Nilai $\phi (s|t)$ goodness Dari hasil perhitungan nilai PL dan PR bisa dilanjutkan dengan menghitung nilai Menghitung Nilai $\phi (s|t)$ goodness, Dimulai dengan menghitung nilai $P(j|tL)$ dan $P(j|tR)$ dengan menggunakan persamaan diatas

$$P(j|tL) = \frac{\text{jumlah catatan pada kategori } j \text{ pada calon cabang kiri } tL}{\text{jumlah catatan pada data latihan}}$$

$$P(j|tR) = \frac{\text{jumlah catatan pada kategori } j \text{ pada calon cabang kanan } tR}{\text{jumlah catatan pada data latihan}}$$

Lalu dilanjutkan dengan menghitung nilai $\phi (s|t)$ dengan menggunakan persamaan berikutnya

$$Q (s|t) = \sum_{j=i} \text{jumlah kategori}$$

$$| P(j|tL) - P(j|tR) | (5)$$

$$\phi (s|t) = 2PLPR (s|t) (6)$$

- d. Setelah semua nilai goodness dihitung dari seriap cabang, Langkah selanjutnya dari algoritma ini adalah menentukan calon cabang manakah yang akan benar-benar dijadikan cabang. Hal ini dilakukan dengan memilih calon cabang manakah yang akan yang memiliki nilai kesesuaian terbesar.

Keterangan

tL = Calon cabang kiri dari noktah keputusan t

tR = Calon cabang kanan dari noktah keputusan t

PL = Rasio jumlah record pada calon cabang kiri dengan jumlah record training

PR = Rasio jumlah record pada calon cabang kanan dengan jumlah record training

$P(j|tL)$ = Rasio jumlah record berkategori j pada calon cabang kiri

$P(j|tR)$ = Rasio jumlah record berkategori j pada calon cabang kanan

2.4 Bantuan Sosial Beras Sejahtera (BANSOS RASTRA)

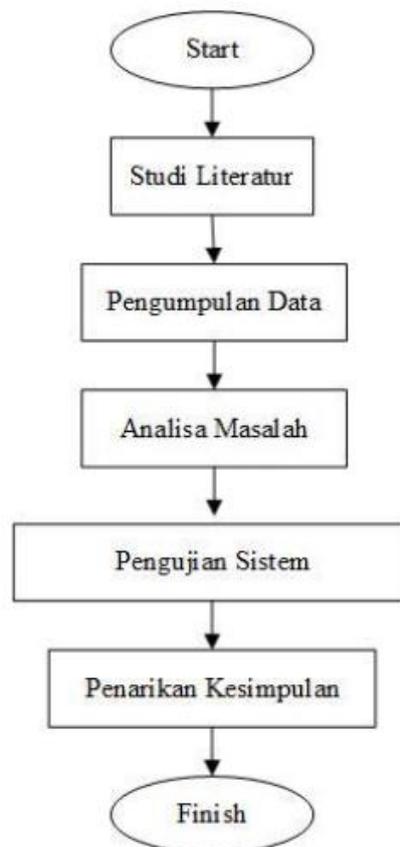
Bantuan Sosial Beras Sejahtera (Bansos Rastra) adalah bantuan pangan dalam bentuk beras (natura) yang diberikan oleh pemerintah untuk disalurkan setiap bulannya kepada Keluarga Penerima Manfaat (KPM) tanpa dikenakan biaya tebus/ harga. Beras yang disalurkan adalah beras berkualitas medium sejumlah 10 Kg. Program Beras Sejahtera (Rastra) merupakan program bantuan sosial, sebagai pendukung program lainnya seperti perbaikan gizi, peningkatan kesehatan, pendidikan dan peningkatan produktivitas Rumah Tangga Sasaran (RTS). Rastra bertujuan untuk mengurangi beban pengeluaran KPM (Keluarga Penerima Manfaat) 3 0yymelalui pemberian bantuan sosial berupa beras berkualitas medium kepada KPM dengan jumlah 10 Kg setiap bulannya (KEMENSOS No.399/HUK/2016).

2.5 Bahan atau Materi Penelitian

Bahan penelitian yang digunakan pada penelitian ini adalah data survey kemiskinan Penduduk Kabupaten Bintan Tahun 2018 yang didapatkan dari Dinas Sosial Kabupaten Bintan.

2.8 Kerangka Pikir Penelitian

Adapun Kerangka Pikir Penelitian yang dilakukan pada implementasi Metode *CART* seperti terlihat pada Gambar 1.

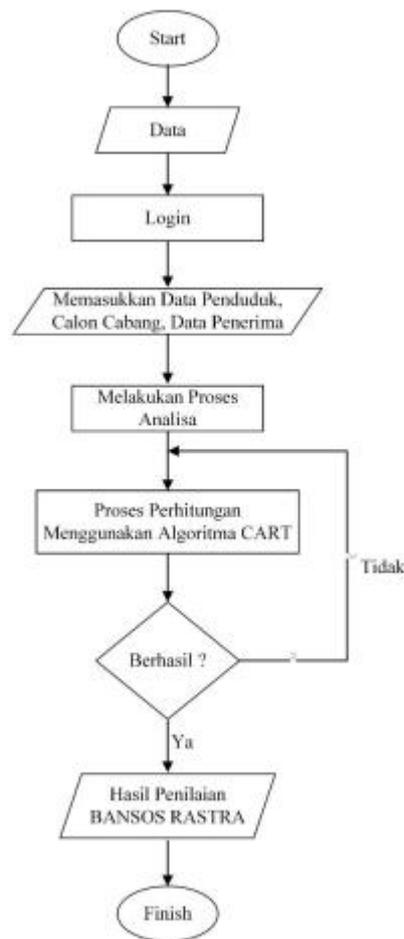


Gambar 1. Kerangka pikir penelitian

Prosedur penelitian yang dilakukan pada penelitian ini dapat digambarkan pada gambar 1. Dijelaskan tentang diagram alir atau serangkaian kegiatan yang dilakukan peneliti secara teratur dan sistematis untuk mencapai tujuan penelitian. Pada gambar 1 dapat dilihat mulai dari dimulai dari melakukan Studi Literatur mencari referensi teori permasalahan yang akan dibahas, lalu pengumpulan data. Setelah itu dilanjutkan dengan proses analisa masalah. Setelah itu dilakukan pengujian sistem jika berhasil maka dapat dilanjutkan dengan kesimpulan dari hasil kesimpulan jika tidak maka kita mengulang pada langkah pengujian kembali hingga proses penelitian berhasil.

2.9 Flowchart Alur Kerja Sistem

Pada gambar 2 Perancangan Alur dengan *Flowchart* dimulai dengan proses *login*, kemudian memasukkan data penduduk, calon cabang dan data penerima, selanjutnya melakukan proses perhitungan dengan algoritma sampai mendapatkan hasil penilaian Bansos Rastra.



Gambar 2. Flowchart Alur Kerja Sistem

III. Hasil dan Pembahasan

3.1 Hasil Analisa Metode CART

Penelitian ini menggunakan metode *Classification and Regression Trees (CART)* untuk melakukan pengklasifikasian penerima Bantuan Sosial Beras Sejahtera (BANSOS RASTRA). Pada Analisis dengan menggunakan metode *CART* langkah pertama adalah membagi data *learning* dan data testing dengan proporsi data *learning* lebih besar dari data testing. Penelitian ini menggunakan data penduduk kuala sempang tahun 2018 yang berjumlah 226 data penduduk yang di dapatkan dari Dinas Sosial Kabupaten Bintan yang akan digunakan pada proses perhitungan menggunakan algoritma *CART*, Karena tidak adanya aturan khusus mengenai pembagian data, dalam penelitian ini digunakan 158 data *learning* dan 68 data testing dengan proporsi 70 % dan 30 %.

3.2 Pembentukan Pohon Maksimal

Pohon maksimal terdiri dari banyak simpul terminal dari penyekatan rekursif secara biner (*binary recursive partitioning*) pada simpul utama (*root node*) maupun pada simpul dalam (*internal node*) dan dibentuk dari data *testing*. Jumlah data mahasiswa lulusan Informatika dengan menggunakan sepuluh data mahasiswa yang digunakan dalam pengklasifikasian ini yang kemudian dibuat model pohon klasifikasi dan data *testing* untuk perhitungan tingkat akurasi klasifikasi. Langkah selanjutnya dalam pembentukan calon cabang pohon maksimal adalah mencari pemilah yang mungkin dari setiap variabel prediktor yang akan menjadi pemilah utama. Variabel *predictor* yang

menjadi pemilah utama adalah variabel yang paling berpengaruh dalam pembentukan pohon klasifikasi dan jalur masuk terpilih sebagai variabel paling berpengaruh dalam pembentukan pohon klasifikasi. Proses pemilihan pohon optimal dimulai dengan membentuk *subtree* yang dipangkas secara terus menerus dari pohon maksimal dan berhenti saat terbentuk pohon yang hanya terdiri dari simpul akar (*root node*). Pemangkasan secara terus menerus ini menghasilkan 12 *subtree*. Pohon optimal yang dipilih adalah subtree yang memiliki *cross validation cost* (CV cost) paling kecil yang berarti pohon tersebut tingkat kesalahan pengklasifikasiannya paling kecil juga. Maka pohon ini dipilih sebagai pohon optimal yang dapat mewakili pohon klasifikasi yang struktur pohonnya sederhana dan memiliki nilai kesalahan pengklasifikasian paling kecil.

3.3 Analisa Perbandingan Pohon Keputusan

Setelah pohon dibentuk, selanjutnya akan dilakukan perbandingan dengan data yang merupakan data uji coba, dimana data tersebut dilakukan pengklasifikasian menggunakan algoritma *CART*. Kemudian akan terbentuk kelas sehingga dapat dibandingkan dan dihitung nilai *error*nya.

Tabel 1. Hasil Pohon Klasifikasi

Calon Mutakhir	Jumlah	
	Status Kelayakan dan Prediksi Tidak Benar	Status Kelayakan dan Prediksi Benar
12 calon Cabang Mutakhir	52	7

Tabel 1 menunjukkan bahwa dari data penduduk yang digunakan bahwa terdapat Status kelayakan dan Prediksi Benar sebanyak 52 Orang, dengan menggunakan 12 calon cabang mutakhir maka presentasi akurasi adalah sekitar 76.47 %.

Setelah diuraikan dengan penjelasan diatas, maka untuk menghitung nilai akurasi dilakukan dengan cara: $Total Akurasi = 52/68 \times 100\% = 76,47\%$

IV. Kesimpulan

Kesimpulan yang dapat diambil dari penelitian ini adalah sebagai berikut :

- Algoritma *Classification and Regression Trees (CART)* dapat digunakan dalam pengklasifikasian penerima Bantuan Sosial Beras Sejahtera.
- Data penduduk Kuala Sempang dapat diklasifikasi dengan menggunakan Algoritma *Classification and Regression (CART)* dengan melihat data yang memenuhi nilai kesesuaian terbesar berdasarkan data Bantuan Sosial Beras Sejahtera. Namun dalam penyusunan *candidate split* dan perhitungan keseluruhan calon cabangnya sangat sulit jika data yang diolah dalam jumlah yang besar.
- Tingkat akurasi yang dihasilkan dengan menggunakan metode *Classification and Regression Trees (CART)* pada Studi Kasus ini adalah 76,47 %.

V. Daftar Pustaka

- Bekti, H. B. (2015). Mahir Membuat Website dengan Adobe Dreamweaver CS6, CSS dan JQuery. Yogyakarta: Andi.
- Breiman, L., Friedman, J, H., Olshen, R., & Stone, C., (1993). Classification and Regression Trees. New York: Chapman Hall.

- Fadlan, C., Ningsih, S., Windarto, P, A., 2018, Penerapan Metode Naïve Bayes Dalam Klasifikasi Kelayakan Keluarga Penerima Beras Rastra, JUTIM, Vol 3 No.1.
- Fahmi, M., 2018, Mari Mengenal Bansos Rastra (2018) : Bantuan Sosial Dari Pemerintah Untuk Mengurangi Beban Masyarakat Kurang Mampu,
- Fajar , H, A., 2009, Data Mining, Yogyakarta: Andi, 2009.
- Greenea, M, M., Hughesb, L, Tonda., Hanlonc, A, H., Huangc, L., Sommersc, S, M., Meghanic, H, S., 2019 Predicting cervical cancer screening among sexual minority women using Classification and Regression Tree analysis, Preventive Medicine Reports 13 153-159.
- Haviluddin, Haryono, A. T., & Rahmawati, D. (2016). Aplikasi Program PHP dan MySQL. Mulawarman University.
- J, R, Lewis., 2000, An Introduction to Classification and Regression Tree (CART). Analysis, Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, California: Departement of Emergency Medicine.
- Kadir, A. (2014). Pengenalan Sistem Informasi Edisi Revisi. Yogyakarta: Andi. Kusri and Luthfi, T, E., 2009, Algoritma Data Mining, I ed, Yogyakarta: Theresia
- Ari Prabawati, Ed. Larose., T, Daniel, 2005, Discovering Knowledge in Data: An Introduction to Data Mining, New Jersey: John Willey & Sons. Inc.
- Munir, 2002, Perencanaan Pembangunan Daerah Dalam Perspektif Otonomi. NTB, Bappeda