

**PENERAPAN METODE CART UNTUK MENGLASIFIKASIKAN
TINGKAT PENGANGGURAN TERBUKA
DI PROVINSI SUMATERA BARAT TAHUN 2020**

TUGAS AKHIR

Untuk Memenuhi Sebagian Persyaratan Memperoleh Gelar Ahli Madya



Oleh
SUCCY INDAH SARI
NIM 18037073

**PROGRAM STUDI DIPLOMA III STATISTIKA
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2022**

PERSETUJUAN UJIAN TUGAS AKHIR

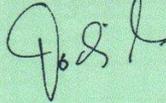
**PENERAPAN METODE CART UNTUK
MENGKLASIFIKASIKAN TINGKAT PENGANGGURAN
TERBUKA DI PROVINSI SUMATERA BARAT TAHUN 2020**

Nama : Succy Indah Sari
NIM/Tahun Masuk : 18037073/2018
Program Studi : DIII Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Padang, 10 Februari 2022

Disetujui oleh:

Pembimbing Akademik



Dodi Vionanda, M.Si. Ph.D
NIP. 19790611 200501 1 002

PERSETUJUAN UJIAN TUGAS AKHIR

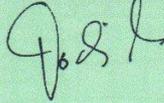
**PENERAPAN METODE CART UNTUK
MENGKLASIFIKASIKAN TINGKAT PENGANGGURAN
TERBUKA DI PROVINSI SUMATERA BARAT TAHUN 2020**

Nama : Sucey Indah Sari
NIM/Tahun Masuk : 18037073/2018
Program Studi : DIII Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Padang, 10 Februari 2022

Disetujui oleh:

Pembimbing Akademik



Dodi Vionanda, M.Si. Ph.D
NIP. 19790611 200501 1 002

SURAT PERNYATAAN TIDAK PLAGIAT

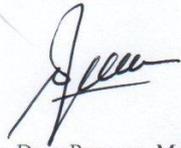
Saya yang bertanda tangan di bawah ini :

Nama : Succy Indah Sari
NIM/TM : 18037073/2018
Program Studi : DIII Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa, Tugas akhir saya dengan judul **“Penerapan Metode CART Untuk Mengklasifikasikan Tingkat Pengangguran Terbuka di Provinsi Sumatera Barat Tahun 2020”** adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam tradisi keilmuan. Apabila suatu saat terbukti saya melakukan plagiat maka saya bersedia diproses dan menerima sanksi akademis maupun hokum sesuai dengan hukum dan ketentuan yang berlaku, baik di institusi UNP maupun di masyarakat dan negara.

Demikian pernyataan ini saya buat dengan penuh kesadaran rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Diketahui oleh,
Ketua Jurusan Statistika,



Dr. Dony Permana, M.Si.
NIP. 19750127 200604 1 001

Saya yang menyatakan,



Succy Indah Sari
NIM. 18037073

Penerapan Metode CART untuk Mengklasifikasikan Tingkat Pengangguran Terbuka di Provinsi Sumatera Barat Tahun 2020

Succy Indah Sari

ABSTRAK

Masalah pengangguran sampai saat ini masih menjadi perhatian utama baik di negara maju maupun negara berkembang. Hal ini disebabkan terjadinya ketidakseimbangan jumlah penduduk angkatan kerja dengan jumlah penduduk yang bekerja. Masalah pengangguran terbuka semakin sulit karena tidak mengenal adanya tunjangan pengangguran. Akibatnya pengangguran cenderung berada dalam kemiskinan karena mereka tidak menerima penghasilan apapun serta dapat meningkatkan kriminalitas yang tinggi. Penelitian ini bertujuan untuk mengetahui model dari klasifikasi tingkat pengangguran terbuka dan faktor-faktor yang berpengaruh terhadap tingkat pengangguran terbuka di Provinsi Sumatera Barat tahun 2020.

Metode CART adalah salah satu metode atau algoritma dari teknik pohon keputusan. CART menghasilkan kumpulan data yang akurat sebagai karakteristik dalam mengklasifikasikan tingkat pengangguran terbuka. Jenis penelitian ini tergolong penelitian terapan. Variabel dalam penelitian ini terdiri dari variabel respons (Y) angkatan kerja dan variabel penjelas (X) yaitu jenis kelamin, pendidikan, usia, status dalam rumah tangga, pengalaman pelatihan kerja, status perkawinan dan tempat tinggal.

Hasil penelitian menggunakan metode CART menghasilkan sebuah diagram pohon yang terdiri dari 8 terminal node. 5 diantaranya memiliki label kelas bekerja dan 3 diantaranya memiliki label kelas tidak bekerja (pengangguran). Berdasarkan hasil terminal node tersebut diperoleh faktor-faktor yang mempengaruhi tingkat pengangguran terbuka di Provinsi Sumatera Barat Tahun 2020 yaitu status perkawinan, tempat tinggal, tingkat pendidikan, jenis kelamin, usia dan status dalam rumah tangga. Selanjutnya diperoleh persentase ketepatan model klasifikasi pada penelitian ini secara tepat sebesar 70,19%.

Kata Kunci : Tingkat Pengangguran Terbuka, Klasifikasi, Metode CART

KATA PENGANTAR



Puji syukur peneliti ucapkan atas kehadiran Allah SWT, atas segala berkah, rahmat, hidayah, serta pertolongan yang diberikan-Nya kepada peneliti, sehingga peneliti bisa menyelesaikan Tugas Akhir ini dengan judul: **“Penerapan Metode CART untuk Mengklasifikasikan Tingkat Pengangguran Terbuka di Provinsi Sumatera Barat Tahun 2020”**. Selanjutnya, shalawat beserta salam untuk nabi Muhammad SAW sebagai suri tauladan bagi seluruh umat.

Peneliti mengucapkan terimakasih kepada semua pihak yang telah memberi saran, serta ide-idenya sehingga tugas akhir ini dapat diselesaikan, yaitu:

1. Bapak Dodi Vionanda, M.Si, Ph.D pembimbing dan penasehat akademik sekaligus Koordinator Program Studi DIII Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang yang telah meluangkan waktunya dalam membimbing dan memberikan masukan-masukan yang bermanfaat bagi peneliti.
2. Bapak Dr. Syafriandi, M.Si dan Ibu Dina Fitria, M.Si dosen penguji.
3. Bapak Dr. Dony Permana, M.Si Ketua Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang.
4. Bapak/Ibu dosen Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang.
5. Teristimewa untuk orang tua dan adik yang telah mencurahkan kasih sayang, do'a, waktu dan dukungan kepada peneliti.
6. Semua sahabat terutama Fitri, Putri, Willy, Riri serta teman dan rekan-rekan Statistika 2018 yang telah banyak membantu dan memberi dukungan kepada

peneliti.

Semoga segala bimbingan, bantuan dan motivasi yang telah diberikan mendapat balasan oleh Allah SWT sebagai amal ibadah.

Dalam penyusunan Tugas Akhir ini peneliti telah berusaha semaksimal mungkin untuk memberikan yang terbaik, namun peneliti menyadari bahwa penulisan Tugas Akhir ini belum sempurna, karena keterbatasan pengetahuan dan kemampuan yang peneliti miliki. Untuk itu kritikan dan saran sangat diperlukan untuk penyempurnaan Tugas Akhir ini. Semoga Tugas Akhir ini dapat bermanfaat bagi peneliti dan pembaca umumnya. Aamiin

Padang, Januari 2022

Succy Indah Sari

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR	ii
DAFTAR ISI	iv
DAFTAR TABEL	vi
DAFTAR GAMBAR	vii
DAFTAR LAMPIRAN	viii
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah	1
B. Batasan Masalah	6
C. Rumusan Masalah	6
D. Tujuan Penelitian	6
E. Manfaat Penelitian	7
BAB II KAJIAN TEORI	8
A. Angkatan Kerja	8
B. Faktor-faktor yang diduga mempengaruhi Tingkat Pengangguran Terbuka (TPT)	9
C. Metode CART	12
D. Rpart	18
E. Undersampling	19
F. Ketepatan Klasifikasi	20
BAB III METODOLOGI PENELITIAN	22
A. Jenis Penelitian	22
B. Jenis dan Sumber Penelitian	22
C. Populasi dan Sampel	22
D. Variabel Penelitian	23
E. Teknik Analisis Data	23
BAB IV HASIL DAN PEMBAHASAN	25
A. Deskripsi Data	25
B. Analisis Data	31
C. Pembahasan	34
BAB V PENUTUP	37
A. Kesimpulan	37
B. Saran	38

DAFTAR PUSTAKA	39
LAMPIRAN.....	42

DAFTAR TABEL

Tabel	Halaman
1. Data Angkatan Kerja di Provinsi Sumatera Barat Tahun 2018-2020.....	2
2. Confusion Matrix	20
3. Variabel Penelitian	23
4. Proporsi Angkatan Kerja.....	32
5. Jumlah Proporsi Angkatan Kerja	33
6. Ketepatan Klasifikasi Pohon	34

DAFTAR GAMBAR

Gambar	Halaman
1. Proses Undersampling.....	19
2. Deskripsi Status Angkatan	25
3. Kerja Berdasarkan Jenis Kelamin	26
4. Angkatan Kerja Berdasarkan Tingkat Pendidikan.....	27
5. Angkatan Kerja Berdasarkan Usia	28
6. Angkatan Kerja Berdasarkan Status dalam Rumah Tangga	29
7. Angkatan Kerja Berdasarkan Pengalaman Pelatihan Kerja.....	29
8. Angkatan Kerja Berdasarkan Status Perkawinan.....	30
9. Angkatan Kerja Berdasarkan Tempat Tinggal.....	31
10. Perbandingan Jumlah Angkatan Kerja Setelah Undersampling	32
11. Diagram Pohon Klasifikasi	33

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Angkatan Kerja Provinsi Sumatera Barat Tahun 2020	42
2. Karakteristik Responden Menggunakan Syntax Crosstab	43
3. Menampilkan Data dengan R.....	45
4. Output Proporsi Data Angkatan Kerja	46
5. Output rpart dan rpart.plot.....	47
6. Output Undersampling	48
7. Diagram Pohon CART Setelah Undersampling	49
8. Akurasi Prediksi	51

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Masalah pengangguran sampai saat ini masih menjadi perhatian utama baik di negara maju maupun negara berkembang. Hal ini disebabkan terjadinya ketidakseimbangan jumlah penduduk angkatan kerja dengan jumlah penduduk yang bekerja. Menurut Suasti dan Sri (2009:185) angkatan kerja merupakan penduduk usia kerja (15 tahun keatas) yang memiliki pekerjaan, penduduk yang masih mencari pekerjaan ataupun penduduk yang tidak memiliki pekerjaan. Jika terjadi ketidakseimbangan antara jumlah penduduk angkatan kerja dengan jumlah penduduk yang bekerja hal ini akan menyebabkan adanya jumlah penduduk yang masih mencari pekerjaan atau disebut dengan pengangguran.

Tingkat pengangguran cenderung lebih tinggi pada negara berkembang. Salah satu masalah pengangguran yang perlu diselesaikan adalah masalah pengangguran terbuka. Menurut Susanto dan Pratama (2019) dalam penelitiannya pengangguran terbuka mengacu pada tenaga kerja yang berusaha mencari pekerjaan namun mereka tidak mendapatkan pekerjaan sama sekali sehingga mereka benar-benar menganggur.

Masalah pengangguran terbuka semakin sulit karena tidak mengenal adanya tunjangan pengangguran seperti Indonesia. Akibatnya pengangguran cenderung berada dalam kemiskinan karena mereka tidak menerima penghasilan apapun. Menurut Wulandari dan Marwan (2019) selain kemiskinan, pengangguran terbuka juga dapat menyebabkan tingkat kriminalitas yang lebih tinggi dan daya beli masyarakat yang lebih rendah.

Berdasarkan data Badan Pusat Statistik (BPS), di Pulau Sumatera Provinsi Sumatera Barat menduduki posisi ketiga tingkat pengangguran terbuka tertinggi setelah Provinsi Kepulauan Riau dan Provinsi Sumatera Utara. Tabel 1 menyajikan data angkatan kerja di Provinsi Sumatera Barat Tahun 2018-2020.

Tabel 1. Data Angkatan Kerja di Provinsi Sumatera Barat Tahun 2018-2020

Tahun	Jumlah Angkatan Kerja	Jumlah Bekerja	Jumlah Pengangguran	Tingkat Pengangguran
2018	2.629.104	2.480.405	148.699	5,66%
2019	2.684.389	2.540.040	144.349	5,38%
2020	2.772.133	2.581.524	190.609	6,88%

Sumber : *Badan Pusat Statistik*

Tabel 1 menunjukkan perkembangan jumlah penduduk angkatan kerja setiap tahunnya. Terlihat perkembangan jumlah penduduk angkatan kerja setiap tahunnya semakin meningkat. Seiring meningkatnya jumlah penduduk angkatan kerja, maka semakin tinggi ketidakseimbangan jumlah penduduk angkatan kerja dengan jumlah penduduk yang bekerja sehingga menyebabkan semakin tingginya tingkat pengangguran. Tingkat pengangguran tertinggi terjadi pada tahun 2020 dengan tingkat pengangguran sebesar 6,88%.

Kenaikan tingkat pengangguran pada tahun 2020 naik secara signifikan dari tahun sebelumnya, yaitu sebesar 1,5%. Jika dibandingkan pada tahun-tahun sebelumnya kenaikan tingkat pengangguran di Provinsi Sumatera Barat hanya berkisar 0,3%. Menurut Kementerian Keuangan Direktorat Jendral Perbendaharaan Provinsi Sumatera Barat, peningkatan tingkat pengangguran terbuka di Provinsi Sumatera Barat terjadi karena seiring dengan pandemi *Coronavirus Disease 2019* (COVID-19) dan pemberlakuan Pembatasan Sosial Berskala Besar (PSBB).

Penelitian sebelumnya yang terkait dengan tingkat pengangguran terbuka dilakukan oleh Wulandari (2021), metode statistika yang dipakai adalah metode

CHAID (*Chi-Square Automatic Interaction Detection*) dengan variabel independen yang digunakan jumlah kelamin, tingkat pendidikan, usia, status dalam rumah tangga, pengalaman pelatihan kerja, status perkawinan dan tempat tinggal. Penelitian ini disimpulkan bahwa variabel independen yang signifikan mempengaruhi tingkat pengangguran terbuka di Provinsi Sumatera Barat tahun 2020 adalah status perkawinan, tempat tinggal, pendidikan, umur, pelatihan kerja.

Penelitian lain juga dilakukan oleh Yulinda dkk (2020) menggunakan analisis statistika deskriptif tentang tingkat pengangguran terbuka (TPT) perempuan di Kabupaten Subang mengidentifikasi salah satu penyebab pengangguran adalah bonus demografi yang berdampak pada peningkatan usia kerja khususnya bagi perempuan. Pendapat serupa yang mendukung pernyataan di atas adalah Saskara dan Kaluge (2009) dalam penelitian mengenai analisis faktor yang mempengaruhi pengangguran perempuan. Penelitian ini menyatakan secara statistik, kaum perempuan menempati posisi yang kurang menguntungkan dalam berbagai aspek kehidupan dibandingkan dengan laki-laki salah satunya dalam mencari pekerjaan. Berdasarkan kedua pendapat ini secara langsung *gender*/jenis kelamin menjadi salah satu faktor penyebab terjadinya pengangguran.

Bersumber pada penelitian sebelumnya mengenai faktor yang memengaruhi tingkat pengangguran terbuka, maka didapatkan variabel dependen dan variabel independen. Variabel dependen yang diteliti untuk penelitian ini yaitu angkatan kerja yang berkategori bekerja dan tidak bekerja (pengangguran). Variabel independen yang diteliti untuk penelitian ini yaitu jenis kelamin, tingkat pendidikan, usia, status dalam rumah tangga, pengalaman pelatihan kerja, status perkawinan dan tempat tinggal.

Banyaknya faktor yang mempengaruhi tingkat pengangguran terbuka, maka pada penelitian ini akan digunakan metode klasifikasi berstruktur pohon untuk memodelkan tingkat pengangguran terbuka. Metode ini merupakan metode yang direkomendasikan untuk mengkaji hubungan antara variabel dependen dan variabel independen. Data yang dianalisis pada metode ini memiliki kompleksitas yang tinggi yaitu jumlah observasi yang besar atau jumlah peubah yang banyak.

Penelitian yang dilakukan oleh Bekti dkk (2017) tentang perbandingan analisis pada pengaruh lokasi karakteristik konsumen dalam memilih minimarket dengan metode regresi logistik dan CART. Dari hasil penelitian di peroleh bahwa nilai akurasi pengklasifikasian pada model regresi logistik adalah 88% sedangkan dengan menggunakan CART diperoleh nilai akurasi sebesar 90%. Penelitian lainnya yang dilakukan oleh Suwardika (2017) tentang perbandingan ketepatan metode klasifikasi dalam mengelompokkan data hepatitis menggunakan metode *Support Vector Machine (SVM)*, *Classification and Regression Trees (CART)* dan regresi logistik biner. Berdasarkan perbandingan analisis diperoleh ketepatan metode CART memiliki hasil ketepatan klasifikasi yang lebih baik dari pada SVM dan regresi logistik biner.

Berdasarkan penelitian terdahulu, maka variabel dalam penelitian ini adalah angkatan kerja sebagai variabel dependen dan variabel independen adalah tingkat pendidikan, jenis kelamin, tempat tinggal, jenis pekerjaan, usia, status dalam rumah tangga, pengalaman pelatihan, dan status perkawinan. Berdasarkan variabel-variabel yang telah disebutkan dan analisis data yang telah dilakukan oleh penelitian terdahulu, untuk permasalahan dalam penelitian ini analisis yang cocok digunakan adalah metode CART. Menurut penelitian Bekti dkk (2017) dan

Suwardika (2017) metode CART akurat untuk klasifikasi.

Menurut Kayri dan Kayri (2015) CART dikenal sebagai klasifikasi tunggal, adalah proses non-parametrik dan tidak memerlukan asumsi tentang bentuk fungsional data. Pendekatan non-parametrik tidak bergantung pada asumsi tertentu, sehingga memberikan lebih banyak fleksibilitas dalam analisis data. Namun, masih memiliki tingkat akurasi yang tinggi dan mudah diimplementasikan. Menurut Mardika dkk (2016) CART termasuk metode dengan pendekatan regresi nonparametrik yang digunakan untuk menghasilkan pohon keputusan. CART terdiri dari pohon regresi (*regression tree*) dan pohon klasifikasi (*classification tree*). Ketika peneliti menemukan pola numerik, maka menggunakan metode pohon regresi. Sedangkan jika pola kategorik menggunakan pohon klasifikasi. Pada penelitian ini, data berupa pola kategori maka hasil yang diperoleh berupa pohon klasifikasi.

Menurut Rahmi, Hazmira dan Hafifatul (2017) CART terbilang sederhana namun merupakan metode yang kuat. Tujuan CART yaitu menghasilkan kumpulan data yang akurat sebagai karakteristik klasifikasi. CART juga digunakan juga untuk menggambarkan hubungan antara variabel respon (dependen/tak bebas) dengan satu atau lebih variabel prediktor (independen/bebas).

Berdasarkan latar belakang tersebut, dalam penelitian ini akan dilakukan klasifikasi tingkat pengangguran terbuka menggunakan metode CART. Untuk itu penelitian ini akan dilakukan dengan judul **“Penerapan Metode CART untuk Mengklasifikasikan Tingkat Pengangguran Terbuka di Provinsi Sumatera Barat Tahun 2020”**.

B. Batasan Masalah

Batasan ruang lingkup dalam penelitian ini adalah pengelompokan faktor-faktor yang dapat mempengaruhi pengangguran terbuka yaitu usia, tingkat pendidikan, jenis kelamin, daerah tempat tinggal, status dalam rumah tangga, pengalaman pelatihan kerja dan status perkawinan.

C. Rumusan Masalah

Berdasarkan permasalahan diatas, masalah yang dibahas dapat dirumuskan sebagai berikut :

1. Bagaimana model untuk memprediksi klasifikasi Tingkat Pengangguran Terbuka di Provinsi Sumatera Barat Tahun 2020 menggunakan metode CART?
2. Faktor apa saja yang berpengaruh signifikan terhadap klasifikasi Tingkat Pengangguran Terbuka di Provinsi Sumatera Barat Tahun 2020 dengan menggunakan metode CART?

D. Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka diperoleh tujuan penelitian ini adalah sebagai berikut:

1. Mendapatkan model pada data klasifikasi Tingkat Pengangguran Terbuka di Provinsi Sumatera Barat Tahun 2020.
2. Untuk mengetahui faktor apa saja yang berpengaruh signifikan terhadap klasifikasi Tingkat Pengangguran Terbuka di Provinsi Sumatera Barat Tahun 2020 dengan menggunakan metode CART.

E. Manfaat Penelitian

Berdasarkan uraian pada tujuan penelitian, maka manfaat penelitian adalah sebagai berikut :

1. Bagi penulis, dapat menambah pengetahuan tentang kemampuan dalam menerapkan metode statistik yang diperoleh
2. Bagi Pembaca, dapat digunakan sebagai referensi mengenai Penerapan Metode CART untuk Mengklasifikasikan Tingkat Pengangguran Terbuka di Provinsi Sumatera Barat Tahun 2020.
3. Bagi pemerintah, dapat memberikan solusi mengenai bagaimana cara mengurangi tingkat pengangguran di Sumatera Barat dengan melihat acuan pada hasil klasifikasi yang paling mempengaruhi berdasarkan penelitian.