

**APLIKASI METODE *IMPROVED CHI-SQUARE*
AUTOMATIC INTERACTION DETECTION DENGAN
KOREKSI BIAS PADA KETIDAKLANJUTAN SISWA
KE SM/SEDERAJAT**

SKRIPSI



FADHIL AL ANSHORY

H051181016

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR**

2022

**APLIKASI METODE *IMPROVED CHI-SQUARE*
AUTOMATIC INTERACTION DETECTION DENGAN
KOREKSI BIAS PADA KETIDAKLANJUTAN SISWA
KE SM/SEDERAJAT**

SKRIPSI

**Diajukan sebagai salah satu syarat memperoleh gelar Sarjana Sains pada
Program Studi Statistika Departemen Statistika Fakultas Matematika dan
Ilmu Pengetahuan Alam Universitas Hasanuddin**

FADHIL AL ANSHORY

H051181016

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN**

MAKASSAR

2022

LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh
bahwa skripsi yang saya buat dengan judul:

***Aplikasi Metode Improved Chi Square Automatic
Interaction Detection dengan Koreksi Bias pada
Ketidaklanjutan Siswa ke SM/Sederajat***

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah
dipublikasikan dalam bentuk apapun

Makassar, 11 November 2022



Fadhil Al Anshory
NIM H051181016

APLIKASI METODE *IMPROVED CHI-SQUARE AUTOMATIC INTERACTION DETECTION* DENGAN KOREKSI BIAS PADA KETIDAKLANJUTAN SISWA KE SM/SEDERAJAT

Disetujui Oleh:

Pembimbing Utama

Pembimbing Pertama



Siswanto, S.Si., M.Si.

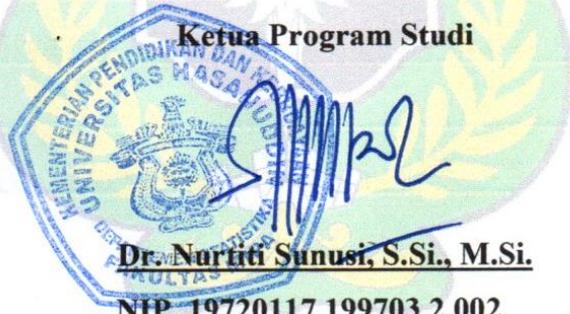
NIP. 19920107 201903 1012



Sri Astuti Thamrin, S.Si., M.Stat., Ph.D

NIP. 19740713 199903 2001

Ketua Program Studi



Dr. Nurtiti Sunusi, S.Si., M.Si.

NIP. 19720117 199703 2 002

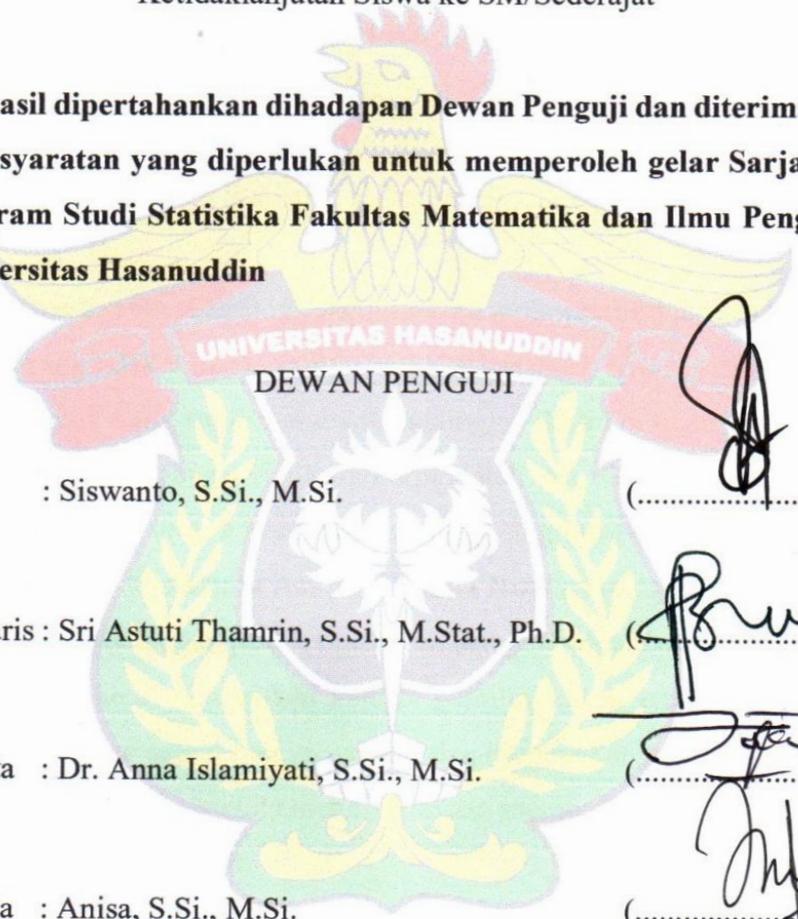
Pada 11 November 2022

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Fadhil Al Anshory
NIM : H051181016
Program Studi : Statistika
Judul Skripsi : Aplikasi Metode *Improved Chi Square Automatic Interaction Detection* dengan Koreksi Bias pada Ketidاكلanjutan Siswa ke SM/Sederajat

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin

- 
1. Ketua : Siswanto, S.Si., M.Si. (.....)
2. Sekretaris : Sri Astuti Thamrin, S.Si., M.Stat., Ph.D. (.....)
3. Anggota : Dr. Anna Islamiyati, S.Si., M.Si. (.....)
4. Anggota : Anisa, S.Si., M.Si. (.....)

Ditetapkan di : Makassar

Tanggal : 11 November 2022

KATA PENGANTAR

Assalamu 'alaikum Warahmatullahi Wabarakatuh

Segala puji hanya milik Allah *Subhanallahu Wa Ta'ala* atas segala limpahan rahmat dan hidayah-Nya yang telah diberikan kepada penulis sampai saat ini. Shalawat dan salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa sallam. Alhamdulillahirobbil'aalamiin*, berkat rahmat dan kemudahan yang diberikan oleh Allah *Subhanallahu Wa Ta'ala*, penulis dapat menyelesaikan skripsi yang berjudul “**Aplikasi Metode Improved Chi Square Automatic Interaction Detection dengan Koreksi Bias pada Ketidaklanjutan Siswa ke SM/Sederajat**” sebagai salah satu syarat memperoleh gelar sarjana pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Dalam penyelesaian skripsi ini, penulis telah melewati perjuangan panjang dan pengorbanan yang tidak sedikit. Namun berkat rahmat dan izin-Nya serta dukungan dari berbagai pihak yang turut membantu sehingga akhirnya tugas akhir ini dapat terselesaikan di waktu yang terbaik menurut Allah swt. Oleh karena itu, penulis menyampaikan ucapan terima kasih yang setinggi-tingginya dan penghargaan yang tak terhingga kepada **Ayahanda Aci** dan **Ibunda Rostina** yang telah membesarkan dan mendidik penulis dengan penuh kesabaran dengan limpahan cinta, kasih sayang dan doa yang tak henti dilangitkan kepada penulis, kakak tercinta penulis yaitu **Gufan Algifari** yang selalu menyemangati dan mendoakan penulis, serta seluruh keluarga besar penulis yang selalu mendoakan dan memberikan bantuan baik dalam bentuk moral ataupun material.

Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan juga penulis ucapkan kepada:

1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin, S.Si., M.Si**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.

3. **Ibu Dr. Nurtiti Sunusi S.Si., M.Si.**, selaku Ketua Departemen Statistika, segenap Dosen Pengajar dan Staf yang telah membekali ilmu dan kemudahan kepada penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.
4. **Ibu Dr. Anna Islamiyati, S.Si., M.Si.** selaku penasehat akademik penulis yang telah ikhlas meluangkan waktu dan pemikirannya untuk memberikan arahan, pengetahuan, motivasi dan bimbingan di tengah kesibukannya.
5. **Bapak Siswanto, S.Si., M.Si.** selaku pembimbing utama penulis yang telah ikhlas meluangkan waktu dan pemikirannya untuk memberikan arahan, pengetahuan, motivasi dan bimbingan di tengah kesibukannya.
6. **Ibu Sri Astuti Thamrin, S.Si., M.Stat., Ph.D.** selaku pembimbing pertama penulis yang telah meluangkan waktunya di tengah kesibukan untuk memberikan arahan, pengetahuan, motivasi, bimbingan untuk penulis.
7. **Ibu Dr. Anna Islamiyati, S.Si., M.Si.** dan **Ibu Anisa, S.Si., M.Si.** selaku tim penguji yang telah memberikan saran dan kritikan yang membangun dalam penyempurnaan penyusunan tugas akhir ini.
8. Teman-teman **INTEGRAL 2018**, terima kasih telah memberi makna dalam cinta, akademik, dan organisasi selama perkuliahan. Pengalaman berharga telah menulis dapatkan dari teman-teman selama berproses bersama.
9. Adik-adik **POISSON 2020**, terima kasih telah menjadi teman bercerita, berbagi, serta menjadi adik bagi penulis. Semoga tanggung jawab dan Amanah yang diberikan selalu dilaksanakan dengan baik.
10. Kakak-kakak **SIMETRIS 2015**, terima kasih atas seluruh pengalaman dan pembelajaran, dan menjadi kakak bagi penulis serta selalu memberikan motivasi, saran, dan nasehat selama berorganisasi.
11. Keluarga besar **Himastat FMIPA Unhas** dan **Himatika FMIPA Unhas**, terima kasih atas seluruh pengalaman, pembelajaran, serta telah menjadi keluarga penulis selama masa perkuliahan.
12. **Keluarga Mahasiswa FMIPA Unhas**, terkhusus **MIPA 2018**, terima kasih untuk cerita, pengalaman, dan ilmu yang sangat berharga selama penulis berorganisasi. Semoga selalu dengan slogan **Takkan Pudar**.

13. Kepada seluruh pihak yang tidak dapat penulis sebutkan satu persatu, terima kasih setinggi-tingginya untuk segala dukungan dan partisipasi yang diberikan kepada penulis semoga bernilai ibadah di sisi Allah *Subhanahu Wa Ta'ala*.

Penulis menyadari bahwa masih banyak kekurangan dalam skripsi ini, untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini memberikan manfaat untuk pembaca.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Makassar, 11 November 2022



Fadhil Al Anshory

**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR
UNTUK KEPENTINGAN AKADEMIK**

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Nama : Fadhil Al Anshory
NIM : H051181016
Program Studi : Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty- Free Right*)** atas tugas akhir saya yang berjudul:

“Aplikasi Metode *Improved Chi Square Automatic Interaction Detection* dengan Koreksi Bias pada Ketidaklanjutan Siswa ke SM/Sederajat”

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar pada tanggal, 11 November 2022

Yang menyatakan



(Fadhil Al Anshory)

ABSTRAK

Badan Pusat Statistika (BPS) pada Survei Sosial Ekonomi Nasional (SUSENAS) dalam Potret Pendidikan Maret 2020 menyebutkan faktor yang memengaruhi siswa tidak melanjutkan pendidikan ke SM/Sederajat yaitu jenis kelamin, pendidikan orang tua, status ekonomi, daerah tempat tinggal, penyandang disabilitas dan kemampuan beradaptasi pada jenjang yang lebih tinggi. *Improved Chi Square Automatic Interaction Detection* (CHAID) dengan koreksi bias merupakan pengembangan dari metode CHAID dengan mengandalkan perhitungan uji *Tschuprow's T* dengan koreksi bias dalam proses pembentukan pohon klasifikasi. Penelitian ini bertujuan untuk mendapatkan klasifikasi faktor-faktor yang memengaruhi siswa tidak melanjutkan pendidikan dari SMP/Sederajat ke SM/Sederajat. Hasil yang diperoleh pada pohon klasifikasi menghasilkan 9 klasifikasi. Berdasarkan hasil pohon klasifikasi, klasifikasi siswa yang tidak melanjutkan pendidikan ke SM/Sederajat yaitu siswa penyandang disabilitas dan tidak memiliki akses teknologi informasi dan komunikasi dengan probabilitas 0.89, siswa yang bekerja bukan penyandang disabilitas dan tidak memiliki akses teknologi informasi dan komunikasi dengan probabilitas 0.73, dan siswa yang tidak bekerja bukan penyandang disabilitas dan tidak memiliki akses teknologi informasi dan komunikasi dengan probabilitas 0.60. Berdasarkan klasifikasi yang didapatkan faktor-faktor yang memengaruhi siswa tidak melanjutkan pendidikan ke SM/Sederajat adalah akses teknologi informasi dan komunikasi, status bekerja, dan penyandang disabilitas. Ketepatan klasifikasi dari hasil klasifikasi keberlanjutan siswa ke SM/Sederajat menggunakan metode *Improved-CHAID* dengan koreksi bias dengan proporsi data *training* 80% dan data *testing* 20% yaitu 72.3033% pada data *training* dan meningkat 73.3300% pada data *testing*.

Kata Kunci: *Improved Chi Square Automatic Interaction Detection*, uji *Tschuprow's T*, pohon klasifikasi, pendidikan, koreksi bias

ABSTRACT

Badan Pusat Statistika (BPS) in Survei Sosial Ekonomi Nasional (SUSENAS) in Potret Pendidikan Maret 2020 stated that the factors that influence students not continuing their education to SM/Equivalent are gender, parental education, economic status, area of residence, persons with disabilities and the ability to adapt at a higher level. Improved Chi Square Automatic Interaction Detection (CHAID) is the development of the CHAID method by relying on Tschuprow's T test calculations in the process of forming a classification tree. This study aims to obtain a classification of the factors that influence students not continuing their education from SMP/Equivalent to SM/Equivalent. The results obtained in the classification tree produce 9 classifications. The results obtained based on the classification tree are Students with disabilities and do not have access to information and communication technology with a probability of 0.89, Students who work are non-disable and do not have access Information and communication technology with a probability of 0.73, and Students who do not work and non-disable, do not have access to information and communication technology with a probability of 0.60. Based on the classification obtained, the factors that influence students not continuing their education to SM/equivalent are access to technology, information, and communication, work status, and persons with disabilities. The classification accuracy of the results of the sustainability classification of students to SM/Equivalent uses the Improved-CHAID method with bias correction with a proportion of 80% training data and 20% testing data, namely 72.3033% on training data and an increase of 73.3300% on testing data.

Keywords: Improved Chi Square Automatic Interaction Detection, Tschuprow's T test, classification tree, education, bias correction

DAFTAR ISI

HALAMAN SAMPUL i

HALAMAN JUDUL..... ii

HALAMAN PERNYATAAN KEOTENTIKAN..... iii

HALAMAN PERSETUJUAN PEMBIMBING iv

HALAMAN PENGESAHAN..... v

KATA PENGANTAR vi

PERSETUJUAN PUBLIKASI KARYA ILMIAH..... ix

ABSTRAK x

ABSTRACT xi

DAFTAR ISI..... xii

DAFTAR TABEL..... xv

DAFTAR GAMBAR xvii

BAB I PENDAHULUAN 1

 1.1 Latar Belakang..... 1

 1.2 Rumusan Masalah 3

 1.3 Batasan Masalah..... 3

 1.4 Tujuan Penelitian..... 3

 1.5 Manfaat Penelitian..... 3

BAB II TINJAUAN PUSTAKA..... 4

 2.1 Metode CHAID 4

 2.2 Algoritma CHAID 5

 2.3 *Improved-CHAID*..... 6

 2.4 Uji *Tschuprow's T* 8

 2.5 Koreksi Bias Uji *Tschuprow's T* 8

 2.6 Pohon Klasifikasi CHAID 9

 2.7 Partisipasi Sekolah..... 10

BAB III METODOLOGI PENELITIAN..... 13

3.1 Jenis dan Sumber Data	13
3.2 Variabel Penelitian	13
3.3 Tahapan Analisis Data.....	15
BAB IV HASIL DAN PEMBAHASAN	16
4.1 <i>Filtering</i> Data	16
4.2 Karakteristik Data.....	16
4.2.1 Jenis Kelamin	16
4.2.2 Pendidikan Orang Tua.....	17
4.2.3 Status Ekonomi.....	18
4.2.4 Daerah Tempat Tinggal.....	18
4.2.5 Angka Melek Huruf.....	19
4.2.6 Penyandang Disabilitas.....	20
4.2.7 Status Bekerja	20
4.2.8 Akses Teknologi Informasi dan Komunikasi.....	21
4.3 Metode <i>Improved-CHAID</i> dengan Koreksi Bias pada Pembentukan Pohon Klasifikasi.....	22
4.3.1 <i>Node 0</i>	22
4.3.2 <i>Node 1</i>	27
4.3.3 <i>Node 2</i>	28
4.3.4 <i>Node 3</i>	30
4.3.5 <i>Node 4</i>	31
4.3.6 <i>Node 5</i>	33
4.3.7 <i>Node 6</i>	33
4.3.8 <i>Node 7</i>	34
4.3.9 <i>Node 8</i>	35
4.3.10 <i>Node 9</i>	36
4.3.11 <i>Node 10</i>	37
4.3.12 <i>Node 11</i>	38
4.3.13 <i>Node 12</i>	39
4.3.14 <i>Node 13</i>	39
4.3.15 <i>Node 14</i>	39

4.3.16 <i>Node</i> 15.....	40
4.3.17 <i>Node</i> 16.....	40
4.4 Pohon Klasifikasi.....	41
4.5 Ketepatan Klasifikasi.....	45
BAB V KESIMPULAN DAN SARAN.....	47
5.1 Kesimpulan.....	47
5.2 Saran.....	47
DAFTAR PUSTAKA	48

DAFTAR TABEL

Tabel 3.1 Deskripsi Variabel.....	13
Tabel 4.1 Tabel Kontingensi Jenis Kelamin terhadap Keberlanjutan Siswa ke SM/Sederajat.....	22
Tabel 4.2 Hasil Uji <i>Chi-Square</i> pada <i>Node 0</i>	23
Tabel 4.3 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 0</i>	26
Tabel 4.4 Hasil Uji <i>Chi-Square</i> pada <i>Node 1</i>	27
Tabel 4.5 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 1</i>	27
Tabel 4.6 Hasil Uji <i>Chi-Square</i> pada <i>Node 2</i>	28
Tabel 4.7 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 2</i>	29
Tabel 4.8 Hasil Uji <i>Chi-Square</i> pada <i>Node 3</i>	30
Tabel 4.9 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 3</i>	30
Tabel 4.10 Hasil Uji <i>Chi-Square</i> pada <i>Node 4</i>	31
Tabel 4.11 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 4</i>	32
Tabel 4.12 Hasil Uji <i>Chi-Square</i> pada <i>Node 6</i>	33
Tabel 4.13 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 6</i>	33
Tabel 4.14 Hasil Uji <i>Chi-Square</i> pada <i>Node 7</i>	34
Tabel 4.15 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 7</i>	35
Tabel 4.16 Hasil Uji <i>Chi-Square</i> pada <i>Node 8</i>	36
Tabel 4.17 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 8</i>	36
Tabel 4.18 Hasil Uji <i>Chi-Square</i> pada <i>Node 9</i>	36
Tabel 4.19 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 9</i>	37
Tabel 4.20 Hasil Uji <i>Chi-Square</i> pada <i>Node 10</i>	37
Tabel 4.21 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 10</i>	38
Tabel 4.22 Hasil Uji <i>Chi-Square</i> pada <i>Node 13</i>	39
Tabel 4.23 Hasil Uji <i>Chi-Square</i> pada <i>Node 14</i>	39
Tabel 4.24 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 14</i>	40
Tabel 4.25 Hasil Uji <i>Chi-Square</i> pada <i>Node 15</i>	40
Tabel 4.26 Hasil Uji <i>Chi-Square</i> pada <i>Node 16</i>	40
Tabel 4.27 Hasil Uji <i>Tschuprow's T</i> pada <i>Node 16</i>	41
Tabel 4.28 Tabel Klasifikasi Metode <i>Improved-CHAID</i> dengan Koreksi Bias...	43

Tabel 4.29 Persentase Hasil Klasifikasi Metode <i>Improved-CHAID</i> dengan Koreksi Bias	44
Tabel 4.30 Matriks Konfusi Hasil Klasifikasi pada Data <i>Training</i>	45
Tabel 4.31 Matriks Konfusi Hasil Klasifikasi pada Data <i>Testing</i>	45

DAFTAR GAMBAR

Gambar 2.1 Pohon Klasifikasi CHAID	9
Gambar 2.2 Angka Partisipasi Murni menurut Jenjang Pendidikan	11
Gambar 4.1 Karakteristik Keberlanjutan Siswa ke SM/Sederajat.....	16
Gambar 4.2 Karakteristik Jenis Kelamin terhadap Keberlanjutan Siswa ke SM/Sederajat	17
Gambar 4.3 Karakteristik Pendidikan Orang Tua terhadap Keberlanjutan Siswa ke SM/Sederajat	17
Gambar 4.4 Karakteristik Status Ekonomi Rumah Tangga terhadap Keberlanjutan Siswa ke SM/Sederajat	18
Gambar 4.5 Karakteristik Daerah Tempat Tinggal terhadap Keberlanjutan Siswa ke SM/Sederajat	19
Gambar 4.6 Karakteristik Melek Huruf terhadap Keberlanjutan Siswa ke SM/Sederajat	19
Gambar 4.7 Karakteristik Penyandang Disabilitas terhadap Keberlanjutan Siswa ke SM/Sederajat	20
Gambar 4.8 Karakteristik Status Bekerja terhadap Keberlanjutan Siswa ke SM/Sederajat	21
Gambar 4.9 Karakteristik Akses Teknologi Informasi terhadap Keberlanjutan Siswa ke SM/Sederajat.....	21
Gambar 4.10 Pohon Klasifikasi pada <i>Node 0</i>	26
Gambar 4.11 Pohon Klasifikasi pada <i>Node 1</i>	28
Gambar 4.12 Pohon Klasifikasi pada <i>Node 2</i>	29
Gambar 4.13 Pohon Klasifikasi pada <i>Node 3</i>	31
Gambar 4.14 Pohon Klasifikasi pada <i>Node 4</i>	32
Gambar 4.15 Pohon Klasifikasi pada <i>Node 6</i>	34
Gambar 4.16 Pohon Klasifikasi pada <i>Node 7</i>	35
Gambar 4.17 Pohon Klasifikasi pada <i>Node 10</i>	38
Gambar 4.18 Pohon Klasifikasi Menggunakan Metode <i>Improved-CHAID</i> dengan Koreksi Bias	41

BAB I

PENDAHULUAN

1.1 Latar Belakang

Metode pohon klasifikasi yang pertama kali diperkenalkan oleh Kass pada tahun 1980 adalah metode *Chi-Square Automatic Interaction Detection* (CHAID). Metode CHAID merupakan bagian dari metode *Automatic Interaction Detection* (AID). Metode AID digunakan pada peubah respon yang berskala interval dengan memaksimalkan Jumlah Kuadrat Antar Grup (JKAG) pada masing-masing pemisahan data dengan menggunakan Uji F. Pada AID data secara berturut-turut dibagi dua berdasarkan penduga yang memiliki JKAG terbesar (Wilkinson, 2006 dalam Miftahuddin, 2012).

Metode CHAID digunakan pada peubah respon berskala nominal atau ordinal yang berupa data kategorik. Data kategorik umumnya dianalisis menggunakan Uji *Chi-Square* untuk mengetahui hubungan antar-variabel. CHAID dengan bantuan Uji *Chi-Square* akan membentuk sebuah klasifikasi. Menurut Lehmann dan Eherler (2001), CHAID adalah sebuah metode untuk mengklasifikasikan data kategori yang tujuan dari prosedurnya adalah untuk membagi rangkaian data menjadi *sub*-kelompok berdasarkan pada variabel dependennya. Hasil dari pengklasifikasian dalam CHAID akan ditampilkan dalam sebuah pohon klasifikasi. Kelemahan metode CHAID adalah menghasilkan bias dalam memilih variabel independen yang dapat mengubah interpretasi data pada pohon klasifikasi sehingga perlu adanya pengembangan dari metode CHAID yang disebut dengan *Improved-CHAID* (Damayanti, dkk., 2018).

Improve-CHAID adalah pengembangan dari metode CHAID yang mengandalkan Uji *Tschuprow's T* dalam proses pembentukan pohon klasifikasinya sehingga klasifikasinya lebih baik. CHAID dan *Improved-CHAID* menghasilkan tingkat akurasi yang hampir sama nilainya, tetapi metode *Improved-CHAID* menghasilkan pohon klasifikasi yang lebih sederhana dengan mengurangi efek bias pada interpretasi data (Damayanti, dkk., 2018). Metode *Improved-CHAID* pernah digunakan dalam penelitian yang dilakukan oleh Muhajir (2016) dengan judul Metode *Improved-CHAID (Chi-Squared Automatic Interaction Detection)* pada Analisis Kredit Macet BMT (Baitul Mal wa Tamwil). Penelitian tentang *Improved-*

CHAID juga dilakukan oleh Balasubramaniam dan Gunasundari (2017) dengan judul *Supply Chain Enhancement Using Improved Chaid Algorithm for Classifying the Customer Groups*. Penelitian yang dilakukan oleh Damayanti, dkk. (2018) dengan judul Perbandingan Hasil Pembentukan Pohon Klasifikasi Metode CHAID dan *Improved-CHAID* menghasilkan kesimpulan bahwa metode *Improved-CHAID* lebih baik digunakan dalam metode klasifikasi. Namun, dalam setiap penelitian tentunya masih menghasilkan bias pada perhitungannya, sehingga perlu untuk mengoreksi bias pada perhitungan *Tschuprow's T*.

Asumsi bias pada perhitungan *Tschuprow's T* yang diduga oleh Tschuprow pada tahun 1925 dibuktikan oleh Bartlett pada tahun 1937. Dengan melakukan koreksi bias terhadap *Tschuprow's T* akan mengurangi efek bias pada perhitungan *Tschuprow's T* sehingga menghasilkan solusi yang lebih akurat. *Improve-CHAID* dengan koreksi bias akan menghasilkan pohon klasifikasi yang lebih sederhana dibandingkan *Improved-CHAID* tanpa koreksi bias. Hal ini dipengaruhi karena dalam perhitungan *Tschuprow's T* dengan koreksi bias memungkinkan nilai *Tschuprow's T* bernilai negatif, sehingga pada perhitungannya diberikan fungsi kendala agar nilai *Tschuprow's T* bernilai *non-negatif* (Bergsma, 2013).

Pengaplikasian metode *Improve-CHAID* dengan koreksi bias sangat cocok untuk diterapkan pada data yang bersifat kategorik. Survey yang dilakukan oleh Badan Pusat Statistika (BPS) pada Survei Sosial Ekonomi Nasional (SUSENAS) yang berkaitan dengan pendidikan umumnya menghasilkan data yang bersifat kategorik. Beberapa faktor yang memengaruhi siswa tidak melanjutkan pendidikan dari SMP/ sederajat ke SM/ sederajat dalam Potret Pendidikan yang diterbitkan oleh BPS pada SUSENAS Maret 2020 yaitu jenis kelamin, pendidikan orang tua, status ekonomi, daerah tempat tinggal, penyandang disabilitas dan kemampuan beradaptasi pada jenjang yang lebih tinggi. Sehingga, perlu dilakukan klasifikasi berdasarkan faktor-faktor yang memengaruhi.

Berdasarkan penjelasan yang dipaparkan sebelumnya, maka akan dilakukan analisis tentang faktor-faktor yang memengaruhi siswa tidak melanjutkan pendidikan dari SMP/ sederajat ke SM/ sederajat dengan menggunakan metode *Improved-CHAID* dengan koreksi bias. Kemudian, akan dilakukan perhitungan akurasi hasil klasifikasi dari metode *Improved-CHAID* dengan koreksi bias.

1.2 Rumusan Masalah

Berdasarkan latar belakang, maka rumusan masalah dari penelitian ini adalah sebagai berikut:

1. Bagaimana klasifikasi faktor-faktor yang memengaruhi siswa tidak melanjutkan pendidikan dari SMP/ sederajat ke SM/ sederajat dengan menggunakan metode *Improved-CHAID* dengan koreksi bias?
2. Bagaimana akurasi hasil klasifikasi dari metode *Improved-CHAID* dengan koreksi bias?

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah data yang digunakan dengan kriteria inklusi siswa berumur 15 hingga 24 tahun yang berdomisili di Sulawesi Selatan yang memilih untuk lanjut/tidak ke Sekolah Menengah Atas atau Sederajat.

1.4 Tujuan Penelitian

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

1. Mendapatkan klasifikasi faktor-faktor yang memengaruhi siswa tidak melanjutkan pendidikan dari SMP/ sederajat ke SM/ sederajat dengan menggunakan metode *Improved-CHAID* dengan koreksi bias.
2. Mendapatkan akurasi hasil klasifikasi dari *Improved-CHAID* dengan koreksi bias.

1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat bermanfaat terhadap pihak yang berminat pada penelitian tentang faktor-faktor yang memengaruhi siswa tidak melanjutkan pendidikan dari SMP/ sederajat ke SM/ sederajat.

1. Bagi Kementerian Pendidikan dan Budaya, penelitian dapat dijadikan bahan untuk meningkatkan kualitas pendidikan Indonesia.
2. Bagi Masyarakat, penelitian ini mampu menjadi sumber informasi mengenai faktor-faktor yang memengaruhi siswa tidak melanjutkan pendidikan.
3. Bagi Peneliti, penelitian ini mampu menambah wawasan dari peneliti tentang analisis dengan menggunakan metode *Improved-CHAID* dengan koreksi bias.

BAB II

TINJAUAN PUSTAKA

2.1 Metode CHAID

Metode CHAID lebih dikenal dengan metode pohon klasifikasi. Inti dari pendekatan ini adalah membagi data menjadi kelompok-kelompok yang lebih kecil berdasarkan hubungan antara variabel dependen dan variabel independen. Analisis CHAID digunakan bila data yang digunakan adalah data dengan variabel kategorik. Metode CHAID hanya efektif bila diterapkan pada data dengan jumlah observasi yang banyak. Metode ini menghasilkan pohon klasifikasi yang mudah diinterpretasikan karena peneliti dapat melihat secara langsung proses pemisahan dan penggabungan variabel independen berlangsung (Syamsi dan Nugraha, 2019).

Hal yang esensial pada metode CHAID ini adalah statistik *Chi-Square*. Dalam penerapannya terdapat hipotesis yaitu:

H_0 : Variabel independent dan variabel dependen tidak signifikan berhubungan

H_1 : Variabel independent dan variabel dependen signifikan berhubungan

Statistik yang terdapat pada data berdistribusi *Chi-Square* yaitu *Pearson Chi-Square* (χ^2) dan *likelihood ratio Chi-Square* (G^2) (Nugraha, 2014). Statistik uji *Chi-Square* dirumuskan oleh persamaan sebagai berikut (Alamudi, dkk., 1998):

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \left[\frac{(o_{ij} - e_{ij})^2}{e_{ij}} \right] \text{ dengan } e_{ij} = \frac{r_i c_j}{n} \quad (2.1)$$

χ^2 berdistribusi *Chi-Square* dengan derajat bebas: $(r - 1)(c - 1)$. o_{ij} adalah pengamatan baris ke- i dan kolom ke- j . e_{ij} adalah frekuensi harapan baris ke- i dan kolom ke- j . r_i adalah baris ke- i untuk $i = 1, 2, \dots, r$ dan c_j adalah kolom ke- j untuk $j = 1, 2, \dots, c$. Total pengamatan sebanyak n .

Hasil dari uji *Chi-Square* akan berpengaruh pada penyusunan pohon klasifikasi. Hal ini dikarenakan pada metode CHAID akan melakukan pengujian masing-masing variabel independen terhadap variabel dependen. Kemudian, akan disusun berdasarkan tingkat signifikansi dari hasil pengujian *Chi-Square* (Gallagher, 2000).

2.2 Algoritma CHAID

Tahapan dalam metode CHAID, yaitu penggabungan (*merging*), pemisahan (*splitting*) dan penghentian (*stopping*). Diagram pohon ditumbuhkan melalui ketiga tahapan tersebut. Selanjutnya akan dilakukan secara berulang pada *sub* kelompok yang terbentuk (Sulviana, dkk., 2018).

2.2.1 Penggabungan (*Merging*)

Tahap penggabungan untuk masing-masing variabel independen dalam penggabungan kategori yang tidak signifikan adalah:

1. Untuk setiap variabel independen, buat tabel kontingensi dua arah dengan variabel dependen.
2. Untuk setiap pasangan kategori yang dapat dipilih, hitung statistik *Chi-Square* dari *sub* tabel kontingensi $2 \times J$ yang dibentuk dari pasangan kategori, di mana variabel dependen memiliki kategori sebanyak J .
3. Jika ada kategori yang digabungkan dengan 3 kategori atau lebih, lakukan pemeriksaan apakah suatu kategori variabel bebas harus dipisahkan atau tidak. Jika didapatkan nilai *Chi-Square* signifikan, pisahkan dari kategori yang lain. Jika lebih dari satu kategori perlu dipisah, pembagian dilakukan dengan *Chi-Square* tertinggi. Kemudian kembali ke langkah 2.
4. Perhitungan *p-value* terkoreksi *Bonferroni* berdasarkan tabel yang telah digabung dengan rumus:

$$1 - (1 - \alpha)^M > \alpha \quad (2.2)$$

Dengan M adalah pengali *bonferroni* dan α adalah salah tipe I. Koreksi *Bonferroni* dilakukan ketika berisi kategori variabel independen gabungan. Koreksi *Bonferroni* untuk kategorinya dikurangi menurut jenis variabelnya. Variabel independen berskala nominal dihitung dengan menggunakan rumus sebagai berikut:

Variabel independen nominal

$$M = \sum_{i=1}^{r-1} (-1)^i \frac{(r-1)^c}{i!(r-1)!} \quad (2.3)$$

Dengan M adalah pengali *bonferroni*, c adalah banyaknya kategori asal, r banyaknya kategori baru, dan i adalah kategori baru ke- i .

2.2.2 Pemisahan (*Splitting*)

Tahap pemisahan adalah pembentukan *sub*-kelompok. Pemisah pada *sub*-kelompok dipilih dari variabel independen terbaik. Hal ini diukur dengan memperhatikan tingkat signifikansi uji *Chi-Square* masing-masing variabel independen terhadap variabel dependennya.

2.2.3 Penghentian (*Stopping*)

Tahap Penghentian adalah tahapan untuk menghentikan pembentukan *sub*-kelompok. Hal ini didasarkan pada keseluruhan variabel independen tidak lagi signifikan berpengaruh pada variabel dependen. Jika masih terdapat variabel yang signifikan berpengaruh, maka yang perlu diperhatikan kembali adalah jumlah pengamatan dari *sub*-kelompok yang terbentuk.

2.3 Improved-CHAID

Terdapat dua tahap dalam Algoritma *Improved-CHAID*. Pertama memisahkan dataset menjadi *training* dan *testing* data. Kedua melakukan proses analisis CHAID dengan *Tschuprow's T* (Belaid, 2001).

Pada tahap pertama dataset dibagi menjadi dua yaitu *training* data dan *testing* data. *Training* data yaitu data yang digunakan untuk menemukan aturan pengklasifikasian pada pohon klasifikasi. *Testing* data yaitu data yang diujikan pada pohon klasifikasi yang telah terbentuk.

Pada tahap kedua, terdapat lima langkah analisis CHAID dengan *Tschuprow's T* sebagai berikut (Kass, 1980):

Untuk tiap variabel independen X_1, X_2, \dots, X_k dan variabel dependen Y

1. Untuk setiap variabel independen, buat tabel kontingensi dua arah dengan variabel dependen berdasarkan *Tschuprow's T*.
2. Hitung statistik *Tschuprow's T* untuk setiap pasangan kategori yang dapat dipilih untuk digabung menjadi satu, untuk menguji kebebasannya dalam sebuah sub tabel kontigensi 2 x 2 atau minimal dua kelompok yang persis sama dengan jumlah nilai kelas.
3. Hitung *p-value* terkoreksi *Bonferroni* didasarkan pada tabel yang telah digabung.

4. Pilihlah variabel independen terbaik, yaitu variabel independen dengan nilai *p-value* yang terendah, dan kemudian melakukan pembagian kelompok dengan variabel independen (yaitu gunakan masing-masing kategori variabel independen tersebut yang telah digabung secara optimal, untuk menentukan *sub* pembagian dari kelompok induk menjadi *sub* kelompok yang baru). Jika tidak ada variabel independen dengan nilai *p-value* yang signifikan, maka tidak dilakukan pembagian kelompok yang baru.
5. Kembali ke langkah 1 untuk membahas subkelompok berikutnya. Berhenti ketika semua *sub* kelompok telah dianalisis dan juga berisi sangat sedikit pengamatan.

Terdapat ketentuan dalam proses pengklasifikasian menggunakan *Improved-CHAID* yaitu (Belaid, 2001):

1. Proses pemisahan yaitu dengan memerhatikan nilai *p-value* yang dihitung dari nilai *Chi-Square* kemudian dibandingkan dengan nilai *Tschuprow's T*. Jika nilai *p-value* lebih besar dari nilai *Tschuprow's T*, proses pembentukan pohon klasifikasi tidak dilakukan. Jika kita menurunkan taraf signifikansi dari *p-value*, kita menghasilkan pohon klasifikasi yang lebih pendek.
2. Proses penggabungan kategori yaitu dengan membandingkan pasangan kategori variabel independen yang baru. Kategori yang digabungkan adalah dua kategori yang mirip. Kategori pada variabel independen digabungkan jika perbedaan antar kategori tidak signifikan. Signifikansi statistik ditentukan dengan membandingkan nilai *p-value* yang dihitung dari uji *Chi-Square* dengan taraf signifikansi yang ditentukan oleh peneliti. Jika taraf signifikansinya diturunkan, maka didapatkan pohon klasifikasi yang lebih besar. Jika taraf signifikansi ditingkatkan, maka didapatkan pohon klasifikasi yang lebih kompak.
3. Pada poin-2 dilakukan beberapa kali tes pada proses penggabungan pasangan kategori variabel independen untuk menemukan solusi terbaik di setiap *node*, sehingga peluang untuk mendapatkan solusi terbaik perlu ditingkatkan, maka nilai *p-value* yang sebenarnya dari pengujian harus dikoreksi dengan koreksi Bonferroni.

2.4 Uji *Tschuprow's T*

Tschuprow's T merupakan uji yang digunakan untuk melihat hubungan antara dua variabel nominal dengan koefisien estimasi sederhana (Bergsma, 2013). Koefisien estimasi sederhana yang dimaksud adalah uji *Chi-Square* (χ^2). Nilai χ^2 tidak dibatasi dan berada pada interval $[0, +\infty)$, memungkinkan untuk terjadi bias pada interpretasi data. Solusi untuk mengatasi kerugian tersebut adalah dengan menormalkan nilai χ^2 yang berada pada interval $[0,1]$ sehingga diperoleh persamaan berikut (Balasubramaniam dan Gunasundari, 2017):

$$T_s = \frac{\hat{\phi}^2}{\sqrt{(r-1)(c-1)}} \quad (2.4)$$

Dengan r adalah jumlah baris, c adalah jumlah kolom, dan $\hat{\phi}^2$ adalah *mean square contingency*. *Tschuprow's T* dalam metode *Improved-CHAID* bertujuan untuk menentukan variabel independen yang akan digunakan dalam pembentukan pohon klasifikasi. *Tschuprow's T* dan Tabel Kontingensi digunakan untuk mempermudah perhitungan dan menghasilkan klasifikasi yang lebih baik. Selain itu, tabel kontingensi memudahkan untuk melihat dan membaca data dalam bentuk kategori.

2.5 Koreksi Bias Uji *Tschuprow's T*

Bias pada $\hat{\phi}^2$ diduga oleh *Tschuprow* (1925) dan dibuktikan oleh *Bartlett* (1937) sehingga:

$$E[\hat{\phi}^2] = \frac{1}{n-1}(r-1)(c-1) \quad (2.5)$$

Dengan $E[\hat{\phi}^2]$ adalah ekspektasi bias terhadap nilai $\hat{\phi}^2$, n adalah jumlah sampel, r adalah jumlah baris, dan c adalah jumlah kolom. Diasumsikan bahwa dugaan benar, *Tschuprow* (1925) merumuskan koreksi bias untuk $\hat{\phi}^2$ sebagai berikut:

$$\tilde{\phi}^2 = \hat{\phi}^2 - E[\hat{\phi}^2] \quad (2.6)$$

Dengan $\tilde{\phi}^2$ adalah nilai $\hat{\phi}^2$ dengan koreksi bias yang diasumsikan lebih mendekati nilai sebenarnya. Karena asumsi 2.10 maka nilai $\tilde{\phi}^2$ kemungkinan bernilai negatif atau nilai menjadi menjadi tidak bias apabila $\tilde{\phi}^2 = 0$ dan nilai negatif tidak memungkinkan pada rentang $[0,1]$, sehingga akan digunakan estimator non-negatif.

$$\tilde{\phi}_+^2 = \text{Max}(0, \tilde{\phi}^2) \quad (2.7)$$

Perlu diperhatikan bahwa nilai $\tilde{\phi}_+^2$ selalu bernilai positif dan tidak selalu bernilai nol, sehingga $\tilde{\phi}_+^2$ akan memiliki bias. Untuk mendapatkan estimasi koreksi bias *Tschuprow's T* maka nilai $\tilde{\phi}^2$ ditukar dengan $\tilde{\phi}_+^2$. Perhitungan ini akan menghasilkan nilai RMSE yang tidak nol jika terdapat ketergantungan pada populasi dependen. Maka, nilai r dan c dikoreksi menggunakan nilai sebagai berikut:

$$\tilde{r} = r - \frac{1}{n-1}(r-1)^2 \tag{2.8}$$

$$\tilde{c} = c - \frac{1}{c-1}(c-1)^2 \tag{2.9}$$

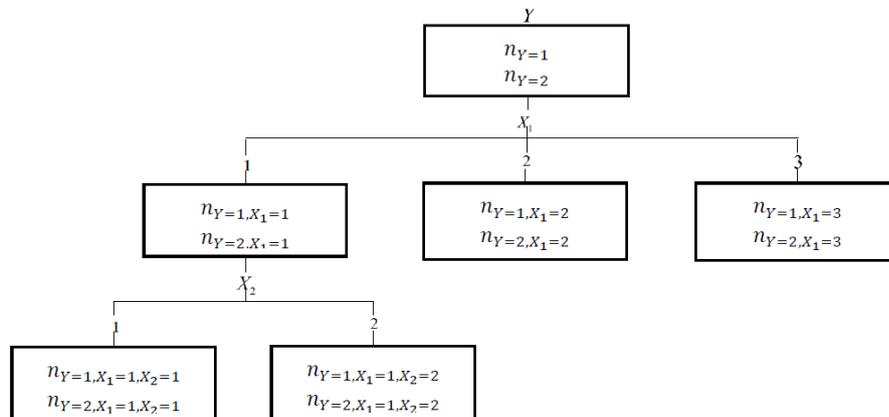
Dengan \tilde{r} adalah jumlah baris dengan koreksi bias, \tilde{c} adalah jumlah kolom dengan koreksi bias, r adalah jumlah baris, dan c adalah jumlah kolom. Sehingga, rumus yang digunakan untuk mengoreksi bias pada *Tschuprow's T* sebagai berikut:

$$\tilde{T}_s = \frac{\tilde{\phi}_+^2}{\sqrt{(\tilde{r}-1)(\tilde{c}-1)}} \tag{2.10}$$

Dengan \tilde{r} adalah jumlah baris dengan koreksi bias, \tilde{c} adalah jumlah kolom dengan koreksi bias, dan $\tilde{\phi}_+^2$ adalah nilai $\hat{\phi}^2$ dengan koreksi bias.

2.6 Pohon Klasifikasi CHAID

Pohon klasifikasi CHAID akan menampilkan sebuah diagram pohon CHAID yang menggambarkan pembentukan klasifikasi dari hasil analisis CHAID. Diagram pohon CHAID dimaksudkan mampu mengeksplorasi data dan menemukan keterkaitan antara variabel independen dan variabel dependen. Dalam pembentukan diagram pohon CHAID akan muncul beberapa sub-kelompok yang disusun mulai dari kelompok utama seperti gambar 2.1 (Eherler dan Lehmann, 2001):



Gambar 2.1 Pohon Klasifikasi CHAID

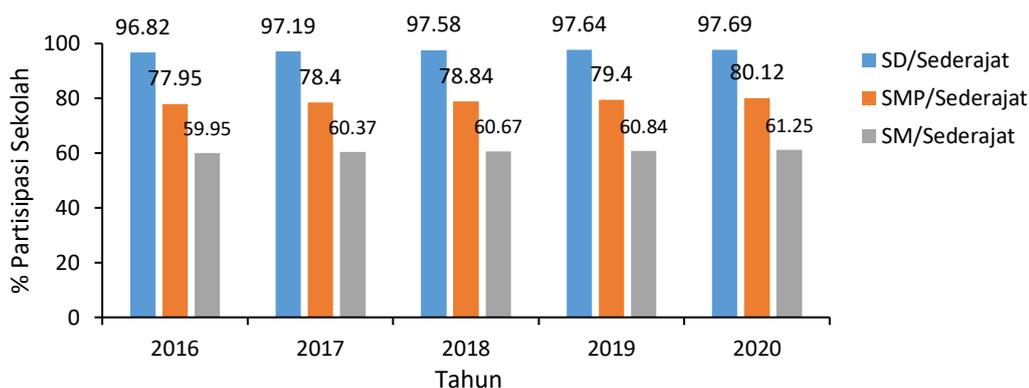
Dengan $n_{Y=1, X_n=i}$ dan $n_{Y=2, X_n=i}$ merupakan jumlah sampel pada *node* yang dibentuk dengan pemisah variabel independen X_n berdasarkan kategori variabel dependen, dan i adalah kategori variabel independen X_n .

Diagram pohon analisis CHAID mengikuti aturan “dari atas ke bawah” (*top-down stopping rule*). Pohon klasifikasi CHAID pada Gambar 2.1 menunjukkan pengklasifikasian dengan membentuk beberapa sub-kelompok. Kedalaman (*depth*) pohon klasifikasi dihitung berdasarkan tingkatan yang terbentuk pada *node-node sub-kelompok* teratas hingga yang terakhir. Pada kedalaman pertama, terbentuk 3 *node* yang diberi tanda *node* 1, 2, dan 3 dan yang menjadi pemisah adalah variabel independen X_1 . Pada kedalaman kedua, terbentuk 2 *node* yang diberi tanda *node* 4, dan 5 dan yang menjadi pemisah adalah variabel independen X_2 untuk *sub-kelompok* yang terbentuk pada *node* 1. Pada masing masing *node* juga ditunjukkan jumlah pengamatan berdasarkan kategori pada variabel dependen. (Kunto dan Hasana, 2006).

Ketepatan hasil klasifikasi dapat dihitung dari nilai akurasi yang didefinisikan sebagai persentase dari observasi yang benar dalam pengklasifikasian terhadap jumlah total observasi (Syamsi dan Nugraha, 2019)

2.7 Partisipasi Sekolah

Manusia merupakan modal utama pembangunan nasional. Untuk itu, pemerintah Indonesia berkomitmen dalam agenda pembangunan RPJMN IV tahun 2020-2024 untuk meningkatkan sumber daya manusia yang berkualitas dan berdaya saing, salah satunya yaitu melalui peningkatan pemerataan layanan pendidikan berkualitas. Bentuk peningkatan pemerataan layanan pendidikan berkualitas dapat dilihat dari persentase partisipasi sekolah penduduk. Berdasarkan partisipasi sekolah, penduduk dikelompokkan dalam tiga kategori, yaitu: tidak/belum pernah sekolah, masih bersekolah, dan tidak bersekolah lagi (BPS, 2020).



Gambar 2.2 Angka Partisipasi Murni Menurut Jenjang Pendidikan

Indikator lain yang dihasilkan dalam publikasi ini adalah Angka Partisipasi Murni (APM). APM mengukur ketepatan usia penduduk dalam berpartisipasi untuk mengenyam suatu jenjang pendidikan tertentu. Secara umum, nilai APM akan selalu lebih rendah dari APK karena APK memperhitungkan jumlah penduduk di luar usia sekolah pada jenjang pendidikan yang bersangkutan, sedangkan APM hanya sebatas usia pada jenjang yang bersesuaian (BPS, 2020).

Faktor sosial, ekonomi, sosial dan budaya adalah faktor-faktor yang terkait untuk memengaruhi anak putus sekolah. Banyaknya penduduk miskin yang ada akan semakin memperbesar kesempatan anak untuk putus sekolah di daerah tersebut (Shahidul dan Karim, 2015).

Persentase anak tidak sekolah lebih tinggi pada kelompok penduduk yang tinggal di perdesaan, berjenis kelamin laki-laki, atau merupakan penyandang disabilitas. Hal ini sejalan dengan studi yang dilakukan oleh UNICEF *Global Initiative on Out of School Children: Indonesia Case Study 2015* mengenai Anak Tidak Sekolah (ATS) dimana faktor yang menyebabkan anak tidak sekolah di antaranya adalah daerah tempat tinggal yang terpencil/tertinggal, kemiskinan, serta belum optimalnya sarana pendidikan untuk penyandang disabilitas. Sedangkan kelompok penduduk yang memiliki AMH lebih rendah adalah penduduk yang tinggal di daerah perdesaan, penduduk berjenis kelamin perempuan, penyandang disabilitas, atau penduduk yang tinggal di rumah tangga dengan status ekonomi terbawah (BPS, 2020).

Pendidikan dalam konstitusi negara Indonesia disebutkan pada Undang-Undang Dasar Republik Indonesia 1945 Pasal 31. Ayat 1 mengatakan bahwa Setiap warga negara berhak mendapat pendidikan, sehingga jelas bahwa tidak ada satu

orangpun yang boleh menghalangi seseorang untuk berpendidikan. Dilanjutkan, ayat 2 setiap warga negara wajib mengikuti pendidikan dasar dan pemerintah wajib membiayai. Oleh karena itu, pendidikan seharusnya menjadi barang wajib untuk seluruh warga negara Indonesia dan pemerintah wajib membiayainya.

Dalam RPJMN 2015 – 2019 ditemukan tujuh dari sepuluh yang berkaitan dengan pendidikan yang sejalan dengan target SDGs (Komnas HAM, 2017). Upaya-upaya untuk mewujudkan pembangunan pendidikan tersebut dirumuskan dalam Rencana Pembangunan Pendidikan Nasional Jangka Panjang 2005 – 2025 yang selanjutnya secara bertahap mengalami penyesuaian sesuai kondisi terkini (Temu, dkk., 2019).

Salah satu program yang dilaksanakan oleh Pemerintah Indonesia adalah Program Indonesia Pintar melalui pelaksanaan Wajib Belajar 12 Tahun pada RPJMN 2015-2019 yang memiliki sasaran untuk meningkatkan angka keberlanjutan pendidikan yang ditandai dengan menurunnya angka putus sekolah. Pada tahun 2019, angka putus sekolah semakin membesar pada jenjang pendidikan yang lebih tinggi (BPS, 2020).

Dalam Potret Pendidikan yang diterbitkan oleh BPS pada Maret 2020 adalah siswa yang tidak melanjutkan pendidikan dari SMP/ sederajat ke SM/ sederajat menempati persentase tertinggi, yakni 18,93 persen. Anak tidak bersekolah merupakan salah satu indikator yang dilaporkan Indonesia pada *Voluntary National Reviews* 2019. Penghitungan indikator anak tidak sekolah ini sudah menyesuaikan dengan standar global (BPS, 2020).