

## Pengaruh Media Terhadap Pengambilan Keputusan Dalam Menjalankan Program Keluarga Berencana Dengan Algoritma *Decision Tree*

Ali Mustopa<sup>1</sup>, Siti Khotimatul Wildah<sup>2</sup>, Ganda Wijaya<sup>3</sup>, Windu Gata<sup>4</sup>, Sarifah Agustiani<sup>5</sup>

<sup>1</sup>STMIK Nusa Mandiri  
e-mail: [14002370@nusamandiri.ac.id](mailto:14002370@nusamandiri.ac.id)

<sup>2</sup>STMIK Nusa Mandiri  
e-mail: [14002399@nusamandiri.ac.id](mailto:14002399@nusamandiri.ac.id)

<sup>3</sup>STMIK Nusa Mandiri  
e-mail: [ganda.gws@nusamandiri.ac.id](mailto:ganda.gws@nusamandiri.ac.id)

<sup>4</sup>STMIK Nusa Mandiri  
e-mail: [windu@nusamandiri.ac.id](mailto:windu@nusamandiri.ac.id)

<sup>5</sup>STMIK Nusa Mandiri  
e-mail: [14002398@nusamandiri.ac.id](mailto:14002398@nusamandiri.ac.id)

**Abstrak** - Indonesia menjadi salah satu negara dengan jumlah penduduk yang beragam sehingga berpotensi mengalami perubahan sosial salah satunya adalah pengaruh media. Media merupakan sebuah konten informasi yang hampir menjadi bagian hidup manusia. Salah satu yang menjadi dampak media yaitu dalam bidang kesehatan salah satunya dalam penentuan program Keluarga Berencana. Keluarga berencana merupakan salah satu program pemerintah Indonesia yang dirancang untuk menekan lajunya pertumbuhan penduduk. Sejak diterapkannya program Keluarga Berencana di Indonesia banyak alat yang digunakan untuk mencegah kehamilan yaitu kontrasepsi. Pemilihan kontrasepsi yang baik tentunya salah satu hal yang penting untuk direncanakan. Dalam menentukan kontrasepsi yang baik tentunya ada pengaruh dari berbagai hal salah satunya adalah media. Pengukuran pengaruh media dalam penentuan program Keluarga Berencana bisa diketahui dengan menerapkan *data mining*. Penelitian dilakukan dengan *data mining* ini menggunakan metodologi standar yang disebut dengan *Cross-Industry Standard Process for Data Mining (CRISP-DM)*. Penggunaan *decision tree* pada penelitian ini dilakukan dengan membandingkan metode yang sama dengan melihat hasil dari tiga model yaitu *Split Validation*, *Cross Validation* dan *Decision Tree Split*. Hasil dari *Split Validation* menghasilkan accuracy sebesar 90,50%, *Cross Validation* menghasilkan accuracy sebesar 91,58% dan *Decision Tree Split* menghasilkan accuracy 89,83%. Hasil terbaik didapatkan dengan menggunakan *cross validation* dimana dengan hasil penelitian terhadap jumlah data sebesar 1473 record diperoleh nilai *accuracy* sebesar 91,58% dan nilai *AUC* yaitu 0,690, dimana hasil tersebut diperoleh dari hasil perhitungan nilai *True Positive (TP)* 1328, *False Negative (FN)* 36, *False Positive (FP)* adalah 88 dan *True Negative (TN)* 21. Paparan media dikatakan bagus atau berpengaruh apabila belum punya anak dan beragama Islam serta pendidikan suaminya SMP dengan taraf hidup rendah tetapi isteri berpendidikan sarjana.

**Kata Kunci:** Keluarga Berencana, Pengaruh Media, *Data Mining*, *Decision Tree*.

**Abstract** - Indonesia has become one of the countries with a diverse population so that it has the potential to experience social change, one of which is the influence of the media. Media is an information content that is almost a part of human life. One of the impacts of the media is in the health sector, one of which is in determining the Family Planning program. Family planning is one of the Indonesian government programs designed to reduce the speed of population growth. Since the implementation of the Family Planning Program in Indonesia many tools have been used to prevent pregnancy, namely contraception. Selection of a good contraction is certainly one important thing to plan. In determining good kontrasepsi certainly there are influences from various things one of which is the media. Measurement of the influence of the media in determining the Family Planning program can be known by applying *data mining*. Research conducted with *data mining* uses a standard methodology called the *Cross-Industry Center Process for Data Mining (CRISP-DM)*. The use of *decision tree* in this study was done by comparing the same method by looking at the results of three models namely *Split Validation*, *Cross Validation* and *Decision Tree Split*. The results of *Split Validation* produce an accuracy of 90.50%, *Cross Validation* produces an accuracy of 91.58% and *Decision Tree Split* produces an accuracy of 89.83%. The best results are obtained by using *cross validation* where with the results

of research on 1473 records the accuracy value is 91.58% and the AUC value is 0.690, where the results are obtained from the calculation of the True Positive (TP) 1328, False Negative (FN) ) 36, False Positive (FP) is 88 and True Negative (TN) 21. Exposure to the media is said to be good or influential if they do not have children and are Muslim and educate their husbands in junior high school with a low standard of living but the wife has a college education.

**Keywords:** Family Planning, Media Exposure, Data Mining, Decision Tree.

## PENDAHULUAN

Indonesia menjadi salah satu negara dengan jumlah penduduk yang besar dengan berbagai ras, suku dan budaya yang beraneka ragam sehingga berpotensi mengalami perubahan sosial. Pola perilaku masyarakat mengalami pergeseran baik budaya, etika dan norma yang di sebabkan oleh media (Cahyono, 2016). Media merupakan sebuah konten informasi yang hampir tidak mungkin terkena oleh seluruh kalangan masyarakat baik sengaja atau tidak. Hal tersebut menandakan bahwa media telah hadir menjadi bagian hidup manusia (Watie, 2016). Seiring dengan perkembangan zaman, kehadiran media makin beragam dengan merambah ke berbagai bidang salah satunya bidang kesehatan. Salah satu media masuk dalam bidang kesehatan yaitu dengan memberikan dampak seseorang dalam penentuan program Keluarga Berencana.

Keluarga berencana adalah sebuah gerakan dimana dalam tujuannya dapat membentuk keluarga yang sehat dan sejahtera dengan kosep pembatasan kelahiran (Nurdianti, 2014). Program Keluarga Berencana adalah salah satu program pemerintah Indonesia yang dirancang untuk mengendalikan kuantitas lajunya pertumbuhan penduduk di Indonesia. Pertaturan keluarga berencana juga ditetapkan dalam undang-undang nomor 52 tahun 2009 mendukung program Keluarga Berencana sebagai salah satu upaya untuk mewujudkan keluarga sehat dan berkualitas. Pengaturan kehamilan dalam program Keluarga Berencana dilakukan dengan menggunakan alat kontrasepsi (Depkes RI, 2018).

Program Keluarga Berencana sendiri sudah ada sejak abad XIX di Inggris dengan melihat pertumbuhan penduduk terutama buruh yang kala itu jauh dari kehidupan layak karena selain kemiskinan juga memiliki banyak anak. Oleh sebab itu, penggunaan alat kontrasepsi dilakukan untuk mengatur pertumbuhan dan mengakur angka kelahiran secara baik dengan kosep pemahaman secara berkala. (Amalia, 2019). Indonesia sendiri sudah menerapkan Keluarga Berencana (KB) sejak tahun 1957 yang dibentuk oleh para Ikatan Dokter Indonesia (Depkes RI, 2018).

Sejak diterapkannya program Keluarga Berencana di Indonesia ada beberapa alat untuk mencegah kehamilan yang dinamakan sebagai Kontrasepsi (Darmawati, 2017). Cara kerja alat kontrasepsi pada

umunnya berfungsi untuk mencegah agar tidak terjadinya ovulasi, melumpuhkan sperma dan menghalangi bertemunya antara sel telur dan sel sperma. Berdasarkan metodenya alat kontrasepsi terbagi menjadi tiga yaitu KB Alat, Obat dan Operasi adapun jenis-jenisnya terdiri dari Pil KB, Suntik KB, Implant KB, IUD, Kondom, Vasektomi dan Tubektomi.

Pemilihan alat kontrasepsi yang sesuai penting dilakukan pasangan mengingat kehamilan haruslah direncanakan. Perencanaan kehamilan penting dilakukan agar pasangan siap secara mental, fisik, maupun kondisi ekonominya. Sehingga dalam memilih alat kontrasepsi, ada beberapa hal yang perlu kita perhatikan agar alat kontrasepsi yang kita gunakan adalah yang paling efektif dalam menunda kehamilan (Depkes RI, 2018).

Pengaruh media sangat memberikan dampak dalam memilih metode kontrasepsi. Dalam pemilihannya bukanlah suatu hal yang mudah karena efek yang berdampak terhadap tubuh tidak akan diketahui selama belum menggunakannya, sehingga perlu pengetahuan yang luas dan tepat mengenai kekurangan dan kelebihan dari masing-masing metode kontrasepsi yang kemudian disesuaikan dengan kondisi tubuh pengguna. Oleh karena itu, pengaruh daripada paparan media sangat ditentukan mengingat kemajuan media informasi saat ini berjalan begitu pesat.

Dalam menentukan alat kontrasepsi yang efektif, tentunya ada pengaruh dari paparan media yang perlu kita ketahui salah satunya adalah dengan menerapkan *data mining*. *Data mining* merupakan proses pencarian pola atau informasi dengan menggunakan teknik atau metode tertentu dalam suatu data (Yuli, 2017). Metode atau teknik yang digunakan dalam penelitian ini adalah dengan menggunakan Metode klasifikasi yang termasuk kedalam metode *supervised learning* yang digunakan untuk menemukan atribut masukan dan atribut target (Hendrian, 2018). Proses pada *data mining* dengan menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dalam sekumpulan data yang besar dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Sari & Saro, 2018). Pada pemahaman *data mining* lainnya sering digunakan dalam penggabungan dari beberapa disiplin ilmu yang menyatukan Teknik dari pembelajaran mesin, pengenalan pola, statistic, database, dan visualisasi

dalam penanganan suatu permasalahan pengambilan informasi dari database yang besar (Larose & Larose, 2014).

Penelitian terkait mengenai program Keluarga Berencana dengan tujuan lain yaitu jurnal dengan judul Perbandingan Metode *Back Propagation Neural Network (BPNN)* dan Metode *Support Vector Machine (SVM)* untuk Klasifikasi Penggunaan Metode Kontrasepsi di Kota Bandung 2018. Metode terbaik merupakan BPNN karena metode tersebut menghasilkan tingkat akurasi yang paling tinggi dibanding yang lain, yakni 75,84 persen (Rochman, 2019). Selain itu, penelitian terkait lain yaitu jurnal dengan judul Sistem Pendukung Keputusan Dalam Menentukan Metode Kontrasepsi Menggunakan Algoritma Naive Bayes. Pada penelitian tersebut hasil prediksi menggunakan algoritma *Naive Bayes* dengan *Cross Validation* menghasilkan tingkat akurasi sebesar 45% (Saputra, Taufik, Ramdhani, Oktavian, & Marsusanti, 2018).

Pada peneliti yang akan dilakukan pemilihan algoritma *data mining* yang diambil menggunakan metode *Decision Tree*. *Decision Tree* digunakan untuk mempelajari dan memprediksi pola sebuah data dan menggambarkan relasi dari variabel atribut dalam bentuk pohon (Wu et al., 2016). *Decision Tree* adalah struktur yang mirip dengan flowchart dimana setiap internal node melakukan pengujian terhadap variabel atribut, tiap cabangnya merupakan hasil dari pengujian tersebut sedangkan node terluar yaitu leaf menjadi label (Han, Kamber, & Jian, 2015). Adapun jenis algoritma *Decision Tree* yang digunakan yaitu Algoritma C4.5.

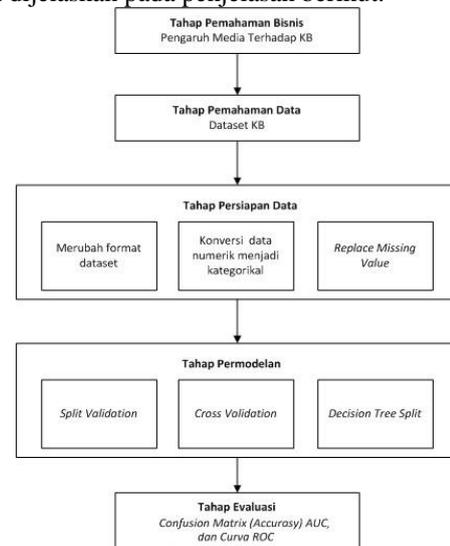
Algoritma C4.5 merupakan algoritma yang dikembangkan dari algoritma ID3 (Larasati & Sutrisno, 2018). Algoritma C4.5 memiliki kelebihan memproses data yang bernilai *nominal, ordinal* maupun kontinyu. Sehingga nilai-nilai setiap atribut yang terdapat pada *dataset* tidak perlu ditransformasikan (Haidar, Sedyono, & Iriani, 2019). Selain itu, dalam menentukan atribut model Algoritma C4.5 menggunakan *gain ratio* dan hasil pohon keputusan C4.5 dipangkas setelah dibentuk (Wajhillah & Yulianti, 2017b). Algoritma C4.5 adalah algoritma *Decision Tree* yang dianggap paling efektif untuk melakukan pengolahan data klasifikasi (Chauhan & Chauhan, 2013). Selain itu, algoritma C4.5 merupakan salah satu pengembangan dari ID3 yang menerapkan konsep pohon keputusan. cara kerjanya secara rekursif mengunjungi tiap simpul keputusan memilih percabangan yang optimal hingga tidak ada hasil percabangan lagi. Setelah itu hasilnya digunakan untuk mengolah data-data yang baru atau disebut test *dataset* (Rahmayuni, 2014).

Penelitian lain yang menggunakan algoritma C4.5 untuk mengetahui keputusan memilih pelanggan potensial untuk menentukan Ya atau Tidaknya dalam memilih keputusan bersedia untuk memilih deposito berjangka. Algoritma dengan model pohon keputusan ini menghasilkan 87% nilai akurasi terhadap keputusan dari *dataset* marketing bank. Konsep algoritma ini membuat klasifikasi berdasarkan model pohon keputusan (Nawawi, Rahayu, Purnama, & Komputer, 2019). Implementasi algoritma C4.5 digunakan untuk memprediksi performa vendor *online*, penelitian ini menghasilkan akurasi sebesar 91,51% untuk menentukan vendor yang sesuai dengan spesifikasi tanpa mengurangi kualitas untuk meneruskan pesanan kepada pelanggan (Rahmawati, Hadianti, Akbar, & Gata, 2018).

Berdasarkan hasil dari pembahasan diatas, maka penelitian ini bertujuan untuk mengetahui pengaruh media terhadap pengambilan keputusan dalam menjalankan program Keluarga Berencana dengan menggunakan algoritma *Decision Tree* dikarenakan pada penelitian sebelumnya penggunaan algoritma dengan model *decision tree* memperoleh nilai akurasi yang baik. Penerapan dengan metode ini diharapkan bisa memprediksi dan mengetahui kinerja dan akurasi yang baik juga dari *dataset* yang digunakan pada penelitian ini.

## METODOLOGI PENELITIAN

Penelitian data mining ini menggunakan metodologi standar yang disebut dengan *Cross-Industry Standard Process for Data Mining (CRISP-DM)*. CRISP-DM merupakan hasil dari beberapa kolaborasi perusahaan, seperti Daimler-Benz, OHRA, NCR Corp, dan SPSS Inc. yang mulai ada sejak tahun 1999 (North, 2012). Adapun penjelasan CRISP-DM akan dijelaskan pada penjelasan berikut.



Sumber: (Mustopa, Wildah, Wijaya, Gata, & Agustiani, 2020)

Gambar 1. Siklus CRISP-DM

**Business Understanding**

Tahapan *business understanding* adalah sebuah tahapan untuk memahami, apakah tujuan bisnis dapat berjalan dengan baik. Secara umum *business understanding* memberikan gambaran kepada peneliti dalam pemahaman dari suatu proyek bisnis yang akan diteliti (Sutoyo, 2018). Pada penelitian ini mencoba menentukan tujuan dari proyek penelitian dalam perumusan serta mendefinisikan masalah data mining. Sasaran utama dari yang ingin dicapai adalah mengetahui pengaruh media terhadap pengambilan keputusan dalam menjalankan program Keluarga Berencana.

**Data Understanding**

Tahapan ini dilakukan oleh penul dengan cara mengumpulkan, mengidentifikasi, dan memahami data yang kita miliki. Kemudian data tersebut juga harus diverifikasi kebenaran dan realibilitasnya. *Dataset* yang diperoleh dari penelitian ini berasal dari kaggle.com.

**Data Preparation**

Tahapan ini dilakukan dengan persiapan data yang sudah didapatkan. Pada tahapan ini meliputi kegiatan seperti membersihkan data, memformat ulang data dan sebagainya agar bertujuan untuk menyiapkan data agar konsisten sesuai dengan format yang dibutuhkan. Adapun jumlah data yang diteliti sebanyak 1473 atribut. Pada kesempatan penelitian ini, peneliti melakukan penyeleksian data menggunakan *software RapidMiner*.

**Modeling**

Pada tahap *modeling* dalam *data mining* dilakukan proses pemilihan algoritma dengan parameter nilai yang optimal. Hal ini bertujuan untuk representasi komputasi dari hasil pengamatan yang meruakan hasil dari pencarian pola-pola yang terkandung pada data. Pada penelitian ini model algoritma yang dipilih yaitu algoritma *Decision Tree* dengan metode pengujian *split validation*, *cross validation* dan *decision tree split*.

**Evaluation**

Pada tahap *evaluation* dilakukan proses klasifikasi algoritma dengan menguji tingkat akurasi dari algoritma *Decision Tree* dengan melihat hasil *accuracy*, *AUC* dan *Curva ROC*

**HASIL DAN PEMBAHASAN**

**Business Understanding**

*Business Understanding* pada penelitian ini adalah *dataset* yang diperoleh dari hasil survei previlensi Kontrasepsi Nasional Indonesia kepada wanita yang sudah menikah yang didasarkan pada pengambilan keputusan program Keluarga Berencana dengan survei yang diperoleh “Bagus” atau “Tidak Bagus”

pada hasil terhadap pengaruh media. Penggunaan tehknik klasifikasi dengan tingkat prediksi dengan nilai akurasi yang tinggi dapat membantu mengatasi hal tersebut, sehingga hasil dari pengaruh media terhadap pemilihan program Keluarga Berencana dapat lebih mudah, cepat dan akurat. Pada pembahasan ini penulis tertarik untuk membuat klasifikasi terhadap pengguna program Keluarga Berencana yang memiliki pengaruh bagus dan tidak bagus untuk menentukan pengambilan keputusan penggunaan program Keluarga Berencana.

**Data Understanding**

*Data understanding* merupakan *dataset* yg diambil dari *kaggle.com* untuk penelitian ini. Variabel input dari *dataset* yang diambil sebanyak 9 atribut yaitu *Wife's age*, *Wife's education*, *Husband's education*, *Number of children ever born*, *Wife's religion*, *Wife's now working?*, *Husband's occupation*, *Standard-of-living index*, *Contraceptive method used* dan *Media exposure*. Untuk melihat atribut-atribut nya dijelaskan pada tabel 1 berikut.

Tabel 1. Atribut dan Nilai Kategori Data

Nama Atribut	Keterangan
<i>Wife's age</i>	<i>Age of the patient (16–49). Numerical.</i>
<i>Wife's education</i>	<i>(1 =low, 2, 3, 4 = high) . Categorical.</i>
<i>Husband's education</i>	<i>(1 =low, 2, 3, 4 = high) . Categorical.</i>
<i>Number of children ever born</i>	<i>Number (0-16). Numerical.</i>
<i>Wife's religion</i>	<i>(0= Non Islam, 1= Islam) . Binary.</i>
<i>Wife's now working?</i>	<i>(0= Yes, 1= No) . Binary.</i>
<i>Husband's occupation</i>	<i>(1, 2, 3, 4) . Categorical.</i>
<i>Standard-of-living index</i>	<i>(1 =low, 2, 3, 4 = high) . Categorical.</i>
<i>Contraceptive method used</i>	<i>(1=No-use ,2=Long-term,3=Short-term) . Categorical.</i>
<i>Class</i>	<i>Media exposure (0=good, 1= notgood). Class Variable</i>

Sumber: (Mustopa et al., 2020)

**Data Preparation**

Pada tahap ini adalah proses mempersiapkan data sehingga data siap digunakan dalam proses pemodelan. Persiapan data yang dilakukan mulai dari merubah format penulisan data, merubah struktur bahasa pada data tersebut dari bahasa Inggris menjadi bahasa Indonesia agar data lebih mudah dipahami, selanjutnya dilakukan pengkategorian terhadap beberapa atribut diantaranya pengkategorian usia yang awalnya memiliki rentang nilai 16-49 tahun menjadi beberapa kategori yaitu < 20 tahun, 20-35 tahun, dan > 35 tahun, dimana dalam pengkategorian ini tetap mencakup seluruh data yakni 16-49 tahun.

Pengkategorian juga dilakukan pada atribut pendidikan baik untuk suami maupun istri yang mana atribut ini memiliki rentang nilai 1-4 dengan ketentuan 1 adalah nilai terendah dan 4 adalah nilai tertinggi, kemudian di konversi menjadi SD, SMP, SMA dan Sarjana berdasarkan level pendidikan yang ada di Indonesia. Pengkategorian ini juga dilakukan pada beberapa atribut lainnya yaitu pekerjaan, jumlah anak, taraf hidup yang memiliki kategori Masing-masing. Dalam hal ini pengkategorian di ambil dari beberapa jurnal penelitian sebelumnya yang terkait dengan data yang sedang diteliti. Hal ini dimaksudkan agar data tersebut lebih mudah dibaca dan membentuk pola yang lebih mudah dimengerti, karena jumlah data yang terlalu banyak dan beragam membuat data tersebut sulit dipahami. Untuk lebih jelasnya berikut merupakan tabel pengkategorian pada data KB yang dimuat pada Tabel 2 berikut.

Tabel 2. Konversi Data Kedalam Kategorikal

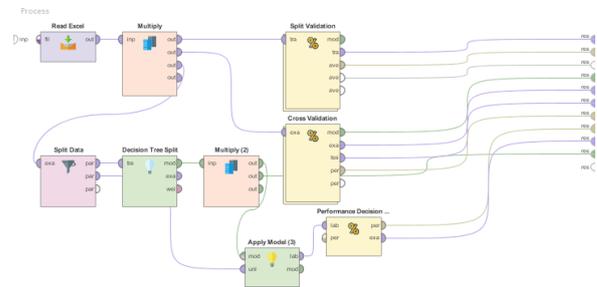
Nama Atribut	Keterangan	Sumber
Usia	<20 Tahun	(Wajhillah & Yulianti, 2017a)
	20-35 Tahun	
	>35 Tahun	
Pendidikan Istri	SD	(Subekti, 2014)
	SMP	
	SMA	
	SARJANA	
Pendidikan Suami	SD	(Subekti, 2014)
	SMP	
	SMA	
	SARJANA	
Jumlah anak	Belum Punya Anak	(Wajhillah & Yulianti, 2017a)
	1 Anak	
	2-4 Anak	
Agama	Non Islam	Kaggle
	Islam	
Status Pekerjaan Istri	Tidak Bekerja	Kaggle
	Bekerja	
Pekerjaan Suami	Petani	https://seafert.
	Wiraswasta	csde.washington
	Pegawai Negeri/Swasta	edu/n/measure/
	Lainnya	doc67.html
Taraf Hidup	Sangat Rendah	Kaggle
	Rendah	
	Sedang	
	Tinggi	
Penggunaan KB	Tidak Menggunakan	Kaggle
	Jangka Pendek	
	Jangka Panjang	
Media	Bagus	Kaggle
	Tidak Bagus	

Sumber: (Mustopa et al., 2020)

### Modeling

Tahapan *modeling* dilakukan setelah data telah siap digunakan, dimana proses persiapan data dilakukan pada tahap sebelumnya. Pemodelan ini dilakukan

dengan menerapkan algoritma *decision tree*. Pada proses pencarian hasil evaluasi dengan mendapatkan akurasi yang baik digunakan dengan pengolahan yang dilakukan dengan tiga model *decision tree* yaitu dengan *split validadadion*, *cross validation* dan *decision tree split*.



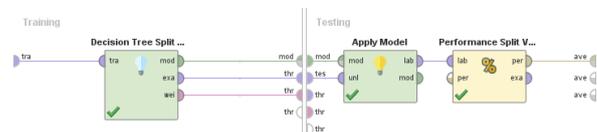
Sumber: (Mustopa et al., 2020)

Gambar 2. Tahapan Pemodelan Data

Pada penerapan tiga model pada gambar 1 memiliki konsep dan penerapan yang berbeda sehingga diharapkan bisa memberikan hasil yang baik dari penerapan ketiga model tersebut. Penjelasan dari ketiga model tersebut sebagai berikut.

### 1. Split Validation

Pada tahapan *split validation dataset* diuji secara eksplisit untuk memprediksi kesesuaian model dengan pengujian hipotesis ketika set diuji. Penerapan penggunaan model *split validation* pada pengujian ini seperti pada gambar berikut.

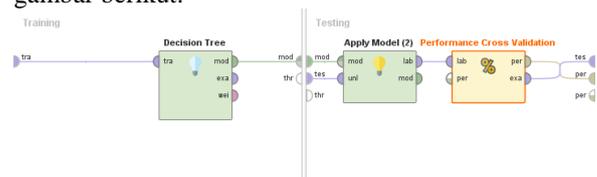


Sumber: (Mustopa et al., 2020)

Gambar 3. Model split validation

### 2. Cross Validation

Pada tahapan *cross validation* memiliki dua subproses yaitu subproses pelatihan (*apply model*) yang digunakan untuk melatih model, model yang terlatih kemudian diterapkan dalam subproses pengujian (*performance cross validation*). Penerapan pada model tersebut digambarkan pada gambar berikut.



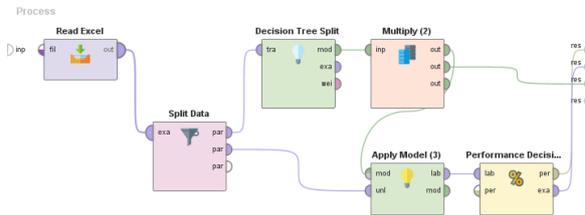
Sumber: (Mustopa et al., 2020)

Gambar 4. Model cross validation

### 3. Decision Tree Split

Pada penerapan model *decision tree split* dimaksudkan membuat keputusan tentang nilai afiliasi ke kelas atau nilai perkiraan nilai target yang

telah di pisahkan. Pada model ini penerapan atribut yang telah diseleksi/split berdasarkan atribut yang berpengaruh. Pada saat split data dilakukan dengan perbandingan pembagian *ratio* terhadap data *training* dengan data *testing*. Pada penerapan model digambarkan pada gambar berikut.



Sumber: (Mustopa et al., 2020)  
 Gambar .5 Model decision tree split

### Evaluation

Hasil pengujian dengan model algoritma yang diuji diperoleh *confusion matrix* nilai *accuracy* dan nilai *AUC* (Area Under Curve) sebagai berikut.

#### 1. Split Validation

Pada model *split validation* menghasilkan model *confusion matrix* yang akan membentuk *matrix* yang terdiri dari *true positive*, *true negative*, *false positive*, dan *false negative*. Di bawah ini merupakan gambaran dari hasil *confusion matrix* yang diperoleh dari pengujian pada data *testing*:

accuracy: 90.50%			
	true Bagus	true Tidak Bagus	class precision
pred. Bagus	389	22	94.65%
pred. Tidak Bagus	20	11	35.48%
class recall	95.11%	33.33%	

Sumber: (Mustopa et al., 2020)  
 Gambar 6. Accuracy split validation

Berdasarkan gambar diatas terdapat rincian jumlah *True Positive (TP)* 382, *False Negative (FN)* 20, *False Positive (FP)* adalah 22 dan *True Negative (TN)* 22. Sehingga menghasilkan nilai *accuracy* sebesar 90,50%. Pada model *split validation* tidak memiliki nilai *AUC* dikarenakan hasil dari *result performance* tidak terdapat hasil *AUC* yang diberikan.

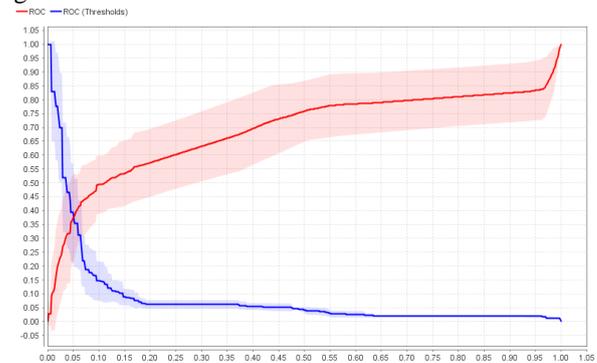
#### 2. Cross Validation

Pengujian pada *model cross validation* menghasilkan model *confusion matrix* yang akan membentuk *matrix* yang terdiri dari *true positive*, *true negative*, *false positive*, dan *false negative*. Di bawah ini merupakan gambaran dari hasil *confusion matrix* yang diperoleh dari pengujian pada data *testing*:

accuracy: 91.58% +/- 2.11% (micro average: 91.58%)			
	true Bagus	true Tidak Bagus	class precision
pred. Bagus	1328	88	93.79%
pred. Tidak Bagus	36	21	36.84%
class recall	97.36%	19.27%	

Sumber: (Mustopa et al., 2020)  
 Gambar 6. Accuracy cross validation

Berdasarkan gambar diatas terdapat rincian jumlah *True Positive (TP)*1328, *False Negative (FN)* 36, *False Positive (FP)* adalah 88 dan *True Negative (TN)* 21. Sehingga menghasilkan nilai *accuracy* sebesar 91,58%. Kemudian pada model *cross validation* menghasilkan nilai *AUC* sebesar 0,690 dimana dalam grafik *ROC* yang dapat diketahui pada gambar berikut.



Sumber: (Mustopa et al., 2020)  
 Gambar 7. AUC cross validation

#### 3. Decision Tree Split

Pengujian pada model *decision tree split* dengan menggunakan *ratio* untuk data *training* dan *testing*. Kemudian pada model *decision tree split* menghasilkan model *confusion matrix* yang akan membentuk *matrix* yang terdiri dari *true positive*, *true negative*, *false positive*, dan *false negative*. Sebelum menentukan *ratio* terbaik, peneliti melakukan percobaan untuk mengetahui *ratio* terbaik pada tabel berikut.

Tabel 2. Perbandingan *ratio* data

Ratio	Accuracy	AUC
1:9	87,71%	0,682
2:8	90,15%	0,631
3:7	89,82%	0,627
4:6	88,45%	0,643
5:5	90,35%	0,656
6:4	89,83%	0,654
7:3	89,14%	0,587
8:2	88,81%	0,648
9:1	87,76%	0,663

Sumber: (Mustopa et al., 2020)

Berdasarkan tabel 2 dapat dihasilkan perbandingan *ratio* terbaik jika dilakukan dengan *ratio* 5:5 pada model *decision tree split*. Di bawah ini merupakan gambaran dari hasil *confusion matrix* yang diperoleh dari pengujian pada data *testing*:

accuracy: 90.35%			
	true Bagus	true Tidak Bagus	class precision
pred. Bagus	651	40	94.21%
pred. Tidak Bagus	31	14	31.11%
class recall	95.45%	25.93%	

Sumber: (Mustopa et al., 2020)  
 Gambar 8. Accuracy decision tree split

Berdasarkan gambar diatas terdapat rincian jumlah *True Positive (TP)* 651, *False Negative (FN)* 31, *False Positive (FP)* adalah 40 dan *True Negative (TN)* 14. Sehingga menghasilkan nilai *accuracy* sebesar 90.35%. Kemudian pada model *decision tree split* menghasilkan nilai *AUC* sebesar 0.656 dimana dalam grafik *ROC* yang dapat diketahui pada gambar berikut.



Sumber: (Mustopa et al., 2020)

Gambar 9. *AUC decision tree split*

Berdasarkan hasil dari ketiga model tersebut untuk melihat hasil terbaik digambarkan pada tabel berikut.

Tabel 3. Konversi Data Kedalam Kategorikal

	<i>Split Validation</i>	<i>Cross Validation</i>	<i>Decision Tree Split</i>
<b>Accuracy</b>	90,50%	91,58%	90,35%
<b>AUC</b>	-	0,690	0,656

Sumber: (Mustopa et al., 2020)

Dari tabel diatas untuk mengetahui pengaruh media terhadap pengambilan keputusan program KB, algoritma *decision tree* dengan model *cross validation* memiliki nilai *accuracy* yang tinggi sebesar 91,58% dengan nilai *AUC* yaitu 0,690.

## KESIMPULAN

Berdasarkan hasil penelitian pada dataset program Keluarga Berencana yang diperoleh dari hasil survei prevalensi Kontrasepsi Nasional Indonesia kepada wanita yang sudah menikah yang didasarkan pada pengambilan keputusan program Keluarga Berencana. Paparan media dikatakan bagus atau berpengaruh apabila belum punya anak dan beragama islam serta pendidikan suaminya SMP dengan taraf hidup rendah tetapi isteri berpendidikan sarjana. Perbandingan tiga model yaitu *split validation*, *cross validation* dan *decision tree split* menghasilkan *accuracy* nilai yang didapatkan dari *cross validation* dengan hasil penelitian terhadap jumlah data sebesar 1473 *record* diperoleh nilai *accuracy* sebesar 91,58% dan nilai *AUC* yaitu 0,690, dimana hasil tersebut diperoleh dari hasil perhitungan nilai *True Positive (TP)* 1328, *False Negative (FN)* 36, *False Positive (FP)* 88 dan *True Negative (TN)* 21. Selain itu, dapat disimpulkan

bahwa pengolahan data yang berbeda seperti pengkonversian data numerik menjadi kategorikal, pengkategorian atribut, dan perbedaan dalam pengambilan jumlah sample dapat memberikan akurasi yang berbeda-beda meskipun menggunakan data dan algoritma yang sama.

## REFERENSI

- Amalia, M. (2019). Sejarah Keluarga Berencana Di Dunia Internasional dan Di Indonesia. Retrieved from <http://stikesypib.ac.id/blog/sejarah-keluarga-berencana-di-dunia-internasional-dan-di-indonesia/>
- Cahyono, A. S. (2016). Pengaruh media sosial terhadap perubahan sosial masyarakat di Indonesia. *Jurnal Ilmu Sosial & Ilmu Politik Diterbitkan Oleh Fakultas Ilmu Sosial & Politik, Universitas Tulungagung*.
- Chauhan, H., & Chauhan, A. (2013). Implementation of decision tree algorithm c4.5. *International Journal of Scientific and Research Publications*.
- Darmawati. (2017). Faktor-Faktor Yang Mempengaruhi Wanita Usia Subur Memilih Kontrasepsi Suntik. *Idea Nursing Journal*, 153–158.
- Depkes RI. (2018). Kementerian Kesehatan Republik Indonesia. *Kementerian Kesehatan RI*.
- Haidar, I. R., Sedyono, E., & Iriani, A. (2019). Analisa Prediksi Mahasiswa Drop Out Menggunakan Metode Decision Tree Dengan, *17(2)*, 97–106.
- Han, J., Kamber, M., & Jian, P. (2015). *Data Mining: Concepts and Techniques Third Edition*.
- Hendrian, S. (2018). Algoritma Klasifikasi Data Mining Untuk Memprediksi, *11(3)*, 266–274.
- Larasati, D. A. H. D., & Sutrisno, T. (2018). Tourism Site Recommendation in Jakarta Using Decision Tree Method Based on Web Review. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3268964>
- Larose, D. T., & Larose, C. D. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition. Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition*. <https://doi.org/10.1002/9781118874059>
- Mustopa, A., Wildah, S. K., Wijaya, G., Gata, W., & Agustiani, S. (2020). *Pengaruh Media Terhadap Pengambilan Keputusan Dalam Menjalankan Program Keluarga Berencana Dengan Algoritma Decision Tree*. Jakarta.
- Nawawi, H. M., Rahayu, S., Purnama, J. J., & Komputer, S. I. (2019). Algoritma c4.5 untuk

- memprediksi pengambilan keputusan memilih deposito berjangka. *Jurnal Techno Nuasa Mandiri*, 16(1), 65–72.
- North, M. (2012). *Data Mining for the Masses. Computer*.
- Nurdianti, S. R. (2014). Analisis Faktor-Faktor Hambatan Komunikasi Dalam Sosialisasi Program Keluarga Berencana Pada Masyarakat Kebon Agung-Samarinda. *Ilmu Komunikasi*.
- Rahmawati, E., Hadianti, S., Akbar, M. F., & Gata, W. (2018). Penerapan Algoritma Cuntuk Memprediksi Performa Vendor Online, 224–231.
- Rahmayuni, I. (2014). Perbandingan performansi algoritma c4.5 dan cart dalam klasifikasi data nilai mahasiswa prodi teknik komputer politeknik negeri padang. *Teknoif*. <https://doi.org/10.1016/j.jnc.2008.09.001>
- Rochman, A. G. (2019). Perbandingan Metode Backpropagation Neural Network (BPNN) dan Metode Support Vector Machine (SVM) untuk Klasifikasi Penggunaan Metode Kontrasepsi di Kota Bandung 2018. *Prosiding Seminar Nasional Statistika/ Departemen Statistika FMIPA Universitas Padjadjaran*, 8(1), 4.
- Saputra, R. A., Taufik, A. R., Ramdhani, L. S., Oktapian, R., & Marsusanti, E. (2018). Sistem Pendukung Keputusan Dalam Menentukan Metode Kontrasepsi Menggunakan Algoritma Naive Bayes. *Seminar Nasional Inovasi Dan Tren*, 106–111.
- Sari, F., & Saro, D. (2018). Implementasi Algoritma C4.5 Dalam Menentukan Lokasi Prioritas Penyuluhan Program Keluarga berencana di kecamatan dumai timur. *Jurnal Penelitian Pos Dan Informatika*. <https://doi.org/10.17933/jppi.2018.080105>
- Subekti, P. (2014). Model regresi logistik multinomial untuk menentukan pilihan sekolah lanjutan tingkat atas pada siswa SMP. *Cauchy*, 3(2), 91. <https://doi.org/10.18860/ca.v3i2.2577>
- Sutoyo, I. (2018). IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI DATA PESERTA DIDIK. *Jurnal Pilar Nusa Mandiri*. <https://doi.org/10.33480/pilar.v14i2.926>
- Wajhillah, R., & Yulianti, I. (2017a). Penerapan Algoritma C4.5 Untuk Prediksi Penggunaan Jenis Kontrasepsi Berbasis Web. *Klik - Kumpulan Jurnal Ilmu Komputer*, 4(2), 160. <https://doi.org/10.20527/klik.v4i2.98>
- Wajhillah, R., & Yulianti, I. (2017b). Penerapan Algoritma C4.5 Untuk Prediksi Penggunaan Jenis Kontrasepsi Berbasis Web. *Klik - Kumpulan Jurnal Ilmu Komputer*. <https://doi.org/10.20527/klik.v4i2.98>
- Watie, E. D. S. (2016). Komunikasi dan Media Sosial (Communications and Social Media). *Jurnal The Messenger*. <https://doi.org/10.26623/themessenger.v3i2.270>
- Wu, X., Wang, Q., Liu, Y., Li, Y., Kurt, A., Gulbagci, B., ... Hall, M. A. (2016). Data Mining Theories, Algorithms, and Examples. *Journal of Chemical Information and Modeling*. <https://doi.org/10.1017/CBO9781107415324.004>
- Yuli, M. (2017). Data Mining: Klasifikasi Menggunakan Algoritma C4.5. *Jurnal Edik Informatika*, 2(2), 213–219.

## PROFIL PENULIS

### Ali Mustopa

Tahun 2019 lulus dari Program Strata Satu (S1) Program Studi Sistem Informasi Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri Nusa Mandiri dan Sekarang sedang melanjutkan pendidikan Program Magister (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri.

### Siti Khotimatul Wildah

Tahun 2019 lulus dari Program Strata Satu (S1) Program Studi Sistem Informasi Universitas Bina Sarana Informatika dan Sekarang sedang melanjutkan pendidikan Program Magister (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri.

### Ganda Wijaya

Tahun 2013 lulus dari Program Strata Satu (S1) Program Studi Sistem Informasi Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri Nusa Mandiri dan tahun 2015 lulus dari Magister (S2) Program Studi Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri.

### Windu Gata

Tahun 1999 lulus dari Program Strata Satu (S1) Program Studi Teknik Informatika Universitas Budi Luhur, Tahun 2008 lulus dari Program Magister (S2) Program Studi Ilmu komputer Universitas Budi Luhur dan Tahun 2015 lulus dari program Doktor (S3) Program Studi Manajemen Pendidikan Universitas Negeri Jakarta.

### Sarifah Agustiani

Tahun 2019 lulus dari Program Strata Satu (S1) Program Studi Sistem Informasi Universitas Bina Sarana Informatika dan Sekarang sedang melanjutkan pendidikan Program Magister (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri.