PENERAPAN ALGORITMA DECISION TREE C4.5 BERBASIS SELEKSI ATRIBUT CHI SQUARED UNTUK KLASIFIKASI TINGKAT PENGETAHUAN IBU DALAM PEMBERIAN ASI EKSKLUSIF PADA BAYI

Erfan Karyadiputra[1)], Zaenuddin[2)]

[1]Fakultas Teknologi Informasi, Universitas Islam Kalimantan Muhammad Arsyad Al Banjari (Erfan Karyadiputra)
email : erfantsy@gmail.com
[1]Fakultas Teknologi Informasi, Universitas Islam Kalimantan Muhammad Arsyad Al Banjari (Zaenuddin)
email : zaenuddin.uniska@gmail.com

Abstrak

Air susu ibu (ASI) merupakan salah satu anugerah ciptaan Allah Subhanahu Wa Ta'ala yang mengandung nutrisi alamiah terbaik bagi bayi dan mengandung kebutuhan energi serta zat yang dibutuhkan selama enam bulan pertama kehidupan bayi sehingga dapat melindunginya dari serangan berbagai penyakit. Dalam Al-Qur`an, Allah Subhanahu Wa Ta'ala juga telah menegaskan proses penyusuan seorang ibu kepada bayinya selama dua tahun penuh. Masa dua tahun ini merupakan rentang waktu tebaik ditinjau dari segala sisi kesehatan dan kejiwaan bagi bayi seperti yang tercantum dalam Al Qur'an Surah Al-Baqarah [2] ayat 233, Luqman [31] ayat 14, dan Al-Ahqaf [46] ayat 15. Penerapan teknik data mining seleksi atribut chi squared dapat meningkatkan akurasi menjadi 92,86% dan meningkatkan nilai AUC menjadi 0,868 sehingga termasuk kedalam kategori good classification dengan mereduksi sejumlah atribut yang tidak relevan dan didapatkan atribut yang mempengaruhi tingkat pengetahuan ibu dalam pemberian ASI eksklusif pada bayi seperti atribut pendidikan, dukungan petugas kesehatan dan jumlah menyusui anak.

***Keywords** : Algoritma C4.5, Chi Squared, ASI, Data Mining, Klasifikasi, Seleksi Atribut*

1. PENDAHULUAN

Air susu ibu (ASI) merupakan salah satu anugerah ciptaan Allah *Subhanahu Wa Ta'ala* yang berfungsi untuk memenuhi kebutuhan nutrisi pada bayi dan melindunginya dari serangan penyakit (Hasniati, Indah, Asrinawaty, & Kasman, 2015). ASI juga mengandung nutrisi alamiah terbaik bagi bayi karena mengandung kebutuhan energi dan zat yang dibutuhkan selama enam bulan pertama kehidupan bayi (Arizki, Rahmawati, & Mahdiyah, 2010). Dalam Al-Qur`an, Allah *Subhanahu Wa Ta'ala* juga telah menegaskan kelangsungan penyusuan ini selama dua tahun penuh (Asnawati, Bafadholl, & Wahidin, 2019). Masa dua tahun ini merupakan rentang waktu emas ditinjau dari segala sisi kesehatan dan kejiwaan bagi bayi yang terdapat dalam Al Qur'an Surah Al-Baqarah [2] ayat 233, Luqman [31] ayat 14, dan Al-Ahqaf [46] ayat 15. Hal ini juga sejalan dengan pernyataan WHO dan UNICEF) untuk mengatur pola pemberian makan terbaik pada bayi dari sejak lahir sampai usia dua tahun untuk meningkatkan kualitas kesehatan bayi (Agustia, Machmud, & Usman, 2019).

Menurut UNICEF dan WHO tentang laporan anak dunia tahun 2011 yaitu dari 136,7 juta bayi yang lahir di dunia, hanya sebesar 32,6% dari mereka menerima ASI secara eksklusif selama 6 bulan pertama (Hasniati et al., 2015). Hal tersebut sejalan dengan hasil Riset Kesehatan Dasar tahun 2010 di Indonesia yang menunjukkan pemberian ASI eksklusif 6 bulan pertama hanya 15,3%, hal ini karena kesadaran masyarakat dalam mendorong peningkatan pemberian ASI masih relatif rendah (Kementerian Kesehatan Republik, 2019). Di Provinsi Kalimantan Selatan cakupan pemberian ASI eksklusif pada tahun 2011 sekitar 51,18% yang menunjukan mengalami peningkatan dari tahun ke tahun, namun masih jauh dari target nasional yaitu 80% (Wijaya, 2018).

permasalahan pokok terkait masih rendahnya penggunaan ASI eksklusif di Indonesia adalah faktor kurangnya pengetahuan ibu hamil, sosial budaya, keluarga dan masyarakat akan pentingnya ASI ekslusif, serta

instansi kesehatan yang belum sepenuhnya mendukung peningkatan pemberian ASI eksklusif (Kementerian Kesehatan Republik, 2019). Permasalahan ini diperparah dengan semakin gencarnya promosi susu formula dan kurangnya dukungan dari masyarakat, termasuk institusi yang memperkerjakan perempuan belum memberikan tempat bagi ibu menyusui di tempat kerja (Hartatik, 2009).

Beberapa penilitian terkait tentang klasifikasi tingkat pengetahuan ibu terhadapa ASI eksklusif diantaranya penelitian yang dilakukan oleh Yusfita Kumala Dewi yang mengklasifikasi tingkat pengetahuan ibu dalam pemberian ASI eksklusif menggunakan metode C4.5 dengan menghasilkan akurasi sebesar 55.24 % dengan label 3 kategori dan 84,52 % dengan label 2 kategori. Penelitian selanjutnya dilakukan oleh Yusuf Indra Wijaya yang melakukan klasifikasi tingkat pengetahuan ibu dalam pemberian ASI eksklusif menggunakan metode C4.5 berbasis PSO dan menghasilkan akurasi sebesar 92.77 % dengan label 2 kategori (Wijaya, 2018). Pada penelitian ini akan dilakukan klasifikasi tingkat pengetahuan ibu dalam pemberian ASI eksklusif menggunakan metode decision tree berbasis seleksi atribut atau seleksi fitur untuk mereduksi fitur-fitur yang tidak relevan dan diketahui atribut yang berpengaruh (Anisah, H, & Pujiastuti, n.d.).

2. METODE PENELITIAN

a. Pengumpulan Data

Tabel 1. Variabel Data ASI

Variabel	Keterangan	Kategori	
Y	Tingkat Pengetahuan	1	Kurang Baik
		2	Baik
X1	Umur	1	< 20 Tahun
		2	20 - 35 Tahun
		3	> 35
X2	Pendidikan	1	Tingkat Dasar
		2	Tingkat Menengah
		3	Tingkat Atas
X3	Pekerjaan	1	Tidak Bekerja
		2	Bekerja
X4	Tempat Bersalin	1	Bukan Fasilitas Kesehatan
		2	Fasilitas Kesehatan
X5	Paritas	1	Primipara
		2	Multipara
		3	Grandemultipara
X6	Menyusui Anak Ke -	1	Kesatu
		2	Kedua
		3	Ketiga
		4	Keempat
		5	Kelima
		6	keenam
X7	Dukungan Petugas Kesehatan	1	Sangat Tidak Mendukung
		2	Tidak Mendukung
		3	Mendukung
		4	Sangat Mendukung
X8	Promosi Susu Formula	1	Sangat Tidak Mendukung
		2	Tidak Mendukung
		3	Mendukung
		4	Sangat Mendukung
X9	Dukungan Orang Tua	1	Tidak Mendukung
		2	Mendukung
X10	Dukungan Suami	1	Tidak Mendukung
		2	Mendukung
X11	Dukungan Teman	1	Tidak Mendukung
		2	Mendukung
X12	Dukungan Tetangga	1	Tidak Mendukung
		2	Mendukung
X13	Budaya	1	Sangat Tidak Mendukung
		2	Tidak Mendukung
		3	Mendukung
		4	Sangat Mendukung

b. Pengolahan Data Awal

Pada tahap ini, data yang telah dikumpulkan melalui pengisian kuesioner kepada responden ibu menyusui yang berada di wilayah Puskesmas Tapin Utara. Kemudian diolah dengan melakukan *pre-processing* dengan mengidentifikasi beberapa atribut-atribut yang tidak relevan sehingga dapat diproses ketahapan selanjutnya dalam data mining.

c. Eksperimen dan Pengujian Metode

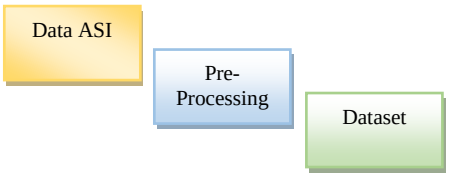
Pada tahap ini dilakukan eksperimen dalam mengimplementasikan pengujian algoritma C4.5 menggunakan rapidminer.

d. Eksperimen dan Pengujian Metode

Tahap selanjutnya melakukan evaluasi dan validasi menggunakan *X-Validation* dan melakukan seleksi atribut *chi squared*.

3. HASIL DAN PEMBAHASAN

Sebelum melakukan eksperimen dan pengujian model, data-data yang dikumpulkan terlebih dahulu diolah agar dapat diproses dalam *data mining*.



Gambar 1. Pre-Processing

Setelah data terkumpul, selanjutnya hitung Jumlah total data, kemudian hitung jumlah total data dengan variabel baik serta tidak baik seperti tabel berikut.

Tabel 2. Dataset

	Jumlah Atribut	Jumlah Data	Baik	Kurang Baik
Total	13	173	129	44

Kemudian menghitung jumlah *entrophy* dari tiap-tiap jumlah dan atribut data.

$$\textbf{\textit{Entrophy}} = \left(-\frac{Jumlah\ Baik}{Jumlah\ Total}\right) x\ Log^2\left(\frac{Jumlah\ Baik}{Jumlah\ Total}\right) + \left(-\frac{Jumlah\ Kurang\ Baik}{Jumlah\ Total}\right) x\ Log^2\left(\frac{Jumlah\ Kurang\ Baik}{Jumlah\ Total}\right) \quad (1)$$

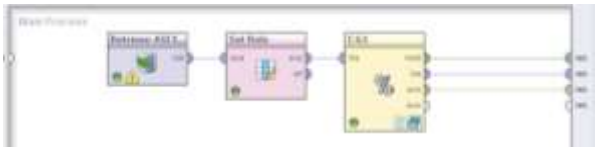
Kemudian hitung *gain* dari setiap *entrophy* tersebut seperti tabel berikut

Tabel 3. Perhitungan Entrophy & Gain

		Jumlah	Baik	Kurang Baik	Entrophy	Gain
Total		173	129	44	0.818077319	
Umur						0.001611816
	<20 Tahun	6	5	1	0.650022422	
	20-35 Tahun	133	98	35	0.831474388	
	>35 Tahun	34	26	8	0.787126586	
Pendidikan						0.016112219
	Tingkat Dasar	40	26	14	0.934068055	
	Tingkat Menengah	100	75	25	0.811278124	
	Tingkat Atas	33	28	5	0.613619002	
Pekerjaan						0.000131561
	Bekerja	45	34	11	0.802353443	
	Tidak Bekerja	128	95	33	0.823427432	
Tempat Bersalin						0.002526944
	Fasilitas Kesehatan	103	79	24	0.783221118	
	Bukan Fasilitas Kesehatan	70	50	20	0.863120569	
Paritas						0.008961528
	Primipara	48	34	14	0.870864469	
	Multipara	122	92	30	0.804717519	
	Grandemultipara	3	3	0	0	
Menyusui Anak						0.016984714
	Kesatu	49	35	14	0.863120569	
	Kedua	78	61	17	0.756404195	
	Ketiga	31	23	8	0.823811633	
	Keempat	12	7	5	0.979868757	
	Kelima	2	2	0	0	
	Keenam	1	1	0	0	
Dukungan Petugas						0.022252206
	Sangat Tidak Mendukung	2	1	1	1	
	Mendukung	56	36	20	0.940285959	
	Sangat Mendukung	115	92	23	0.721928095	
Dukungan Suami						0.000351611
	Tidak Mendukung	5	4	1	0.721928095	
	Mendukung	168	125	43	0.820576827	
Dukungan Orang Tua						0.002147522
	Tidak Mendukung	44	31	13	0.875663392	
	Mendukung	129	98	31	0.795555547	
Dukungan Teman						0.001328817
	Tidak Mendukung	120	88	32	0.836640742	
	Mendukung	53	41	12	0.771709477	
Dukungan Tetangga						0.000001373
	Tidak Mendukung	126	94	32	0.817507288	
	Mendukung	47	35	12	0.819600435	
Promosi Susu						0.009173709
	Sangat Tidak Mendukung	11	10	1	0.439496987	
	Tidak Mendukung	56	42	14	0.811278124	
	Mendukung	93	68	25	0.839762274	
	Sangat Mendukung	13	9	4	0.890491644	

						0.0053454 64
Budaya	Sangat Tidak Mendukung	1	1	0	0	
	Tidak Mendukung	24	18	6	0.8112781 24	
	Mendukung	94	72	22	0.7849920 89	
	Sangat Mendukung	54	38	16	0.8767162 89	

Selanjutnya melakukan implementasi pengujian algoritma C4.5 menggunakan rapidminer.



Gambar 2. Implementasi C4.5

Selanjutnya evaluasi dan validasi melalui pengujian *confusion matrix* dan kurva ROC sehingga diketahui akurasi dan modelnya.

accuracy: 87.89% +/- 8.43% (mikro: 87.88%)

	true Baik	true Kurang Baik	class precision
pred. Baik	107	9	92.24%
pred. Kurang Baik	8	16	66.67%
class recall	93.04%	64.00%	

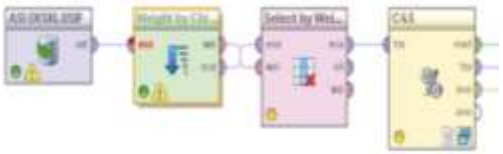
Gambar 3. Hasil Akurasi C.45

Dari hasil pengujian 10-*fold cross-validation*, didapatkan akurasi sebesar 87,89% dengan nilai AUC sebesar 0,679.



Gambar 4. Model Pohon Keputusan

Selanjutnya melakukan seleksi atribut menggunakan *chi squared* untuk meningkatkan akurasi dan mengetahui atribut yang berpengaruh.



Gambar 5. Seleksi Atribut Chi Squared

Dari hasil percobaan seleksi atribut *chi squared* dengan mengganti nilai *weight relation top k* maka didapatkan seleksi atribut dengan nilai *top k* = 3 memiliki akurasi terbaik sebesar 92,90% dengan nilai AUC sebesar 0,868.

accuracy: 92.90% +/- 5.54% (mikro: 92.88%)

	true Baik	true Kurang Baik	class precision
pred. Baik	114	9	92.68%
pred. Kurang Baik	1	16	94.12%
class recall	99.13%	64.00%	

Gambar 6. Hasil Akurasi Chi Squared

Dari hasil seleksi atribut didapatkan atribut yang berpengaruh dalam klasifikasi tingkat pengetahuan ibu dalam pemberian ASI eksklusif yaitu atribut pendidikan, dukungan petugas kesehatan dan jumlah menyusui anak seperti yang tergambarkan pada pohon keputusan berikut



Gambar 7. Model Pohon Keputusan

4. KESIMPULAN

Dari hasil penelitian yang dilakukan dari tahap awal pengujian hingga akhir pengujian, maka dapat disimpulkan bahwa penerapan teknik seleksi atribut *chi squared* dapat meningkatkan akurasi dan kehandalan klasifikasi algoritma C4.5 dengan mereduksi sejumlah atribut yang tidak relevan sehingga didapatkan atribut yang mempengaruhi tingkat pengetahuan ibu dalam pemberian ASI eksklusif pada bayi.

5. REFERENSI

Agustia, N., Machmud, R., & Usman, E. (2019). FAKTOR YANG BERHUBUNGAN DENGAN PEMBERIAN ASI EKSKLUSIF PADA IBU BEKERJA DI KABUPATEN OGAN KOMERING ULU, *8*(3), 573–582.

Anisah, S., H, A. S., & Pujiastuti, A. (n.d.). KLASIFIKASI TEKS

MENGGUNAKAN CHI SQUARE FEATURE SELECTION UNTUK MENENTUKAN KOMIK BERDASARKAN PERIODE, MATERI DAN FISIKDENGAN ALGORITMA NAIVE BAYES, 59–66.

Arizki, W., Rahmawati, D., & Mahdiyah, D. (2010). HUBUNGAN PENGETAHUAN IBU TENTANG ASI EKSKLUSIF TERHADAP PEMBERIAN ASI PADA BAYI 0-6 BULAN DI PUSKESMAS S.PARMAN BANJARMASIN, 39–48.

Asnawati, Bafadholl, I., & Wahidin, A. (2019). PEMBERIAN ASI PADA ANAK DALAM PERSPEKTIF AL-QUR ` AN, (1), 85–98. https://doi.org/10.30868/at.v4i01.429

Hartatik, T. (2009). HUBUNGAN PENGETAHUAN DAN SIKAP IBU DENGAN PEMBERIAN ASI EKSKLUSIF DI KELURAHAN GUNUNGPATI KECAMATAN GUNUNGPATI KOTA SEMARANG TAHUN 2009.

Hasniati, Y., Indah, M. F., Asrinawaty, & Kasman. (2015). DETERMINAN PEMBERIAN ASI EKSKLUSIF DI KABUPATEN BARITO KUALA KALIMANTAN SELATAN, 39–43.

Karyadiputra, E., & Hijriana, N. (2019). PENERAPAN ALGORITMA DECISION TREE C4. 5 UNTUK KLASIFIKASI PENENTUAN DAFTAR PRIORITAS PENGEMBANGAN JEMBATAN. *Technologia: Jurnal Ilmiah*, *10*(1), 43-46.

Karyadiputra, E., & Hijriana, N. (2018). ANALISIS PENERAPAN ALGORITMA NAIVE BAYES UNTUK KLASIFIKASI PRIORITAS PENGEMBANGAN JALAN DI PROVINSI KALIMANTAN SELATAN. *Technologia: Jurnal Ilmiah*, *9*(2), 105-108.

Kementerian Kesehatan Republik. (2019). BANYAK SEKALI MANFAAT ASI BAGI BAYI DAN IBU, 1–2.

Wijaya, Y. I. (2018). KLASIFIKASI TINGKAT PENGETAHUAN IBU DALAM PEMBERIAN ASI EKSKLUSIF MENGGUNAKAN ALGORITMA DECISION TREE C4.5 BERBASIS PSO, *9*(2), 120–124.