

# **KLASIFIKASI GANGGUAN MOTORIK KASAR ANAK MENGUNAKAN NAIVE BAYES SERTA OPTIMASI DENGAN PSO DAN ADABOOST**

**Kadek Wibowo**

Program Studi Ilmu Komputer  
STMIK Nusa Mandiri Jakarta  
[kadek.kwo26@gmail.com](mailto:kadek.kwo26@gmail.com)

**Sfenrianto**

Program Studi Ilmu Komputer  
STMIK Nusa Mandiri Jakarta  
[sfenrianto.71@gmail.com](mailto:sfenrianto.71@gmail.com)

**Kaman Nainggolan**

Program Studi Ilmu Komputer  
STMIK Nusa Mandiri Jakarta  
[golan1251@yahoo.co.id](mailto:golan1251@yahoo.co.id)

*ABSTRACT --- Delayed in gross motor movements children is one of the children's growth disorders. Gross motor movements include large muscles such as muscular limbs in babies such as kicking, kicked, grabbed, lifted the neck, and turned. Growth in its ability to be monitored and stimulated so that children can grow and develop optimally. In this study aims to help classify the patients were categorized as patients with normal or impaired at the stage of the initial examination. By comparing algorithms Naive Bayes classifier, Naive Bayes classifier with the optimization of PSO and Naive Bayes classifier based Adaboost in this study, the result of testing the accuracy value Naive Bayes classifier amounted to 88.67%, while the Naive Bayes classifier based Adaboost amounted to 90.00% and the accuracy of the algorithm Naive Bayes classifier with PSO optimization of 98.00%.*

**Keywords: disorders of children, naïve bayes classifier, adaboost, PSO optimization.**

**ABSTRAK ---**Keterlambatan gerak motorik kasar pada anak merupakan salah satu gangguan tumbuh kembang anak. Motorik kasar mencakup gerakan otot-otot besar seperti otot tungkai dan lengan pada bayi berupa gerakan menendang,menjejak , meraih, mengangkat leher, dan menoleh. Pertumbuhan kemampuannya harus terus di pantau dan di stimulasi agar anak dapat tumbuh dan berkembang optimal. Pada penelitian ini bertujuan membantu mengklasifikasi terhadap pasien yang dikategorikan sebagai pasien normal atau yang mengalami gangguan pada tahap pemeriksaan awal. Dengan membandingkan algoritma *Naive Bayes classifier*, *Naive Bayes classifier* dengan optimasi PSO dan *Naive Bayes classifier* berbasis *Adaboost* pada penelitian ini, didapatkan hasil pengujian dengan nilai akurasi *Naive Bayes classifier* sebesar 88.67%, sedangkan *Naive*

*Bayes classifier* berbasis *Adaboost* sebesar 90.00% dan akurasi algoritma *Naive Bayes classifier* dengan optimasi PSO sebesar 98.00%.

**Kata kunci: Naive Bayes classifier, PSO, Adaboost, Keterlambatan gerak motorik kasar anak.**

## **PENDAHULUAN**

Berdasarkan pasal 2 pada Keputusan Menteri Kesehatan Republik Indonesia Nomor 66 Tahun 2014, pemantauan pertumbuhan, perkembangan, dan gangguan tumbuh kembang anak merupakan acuan bagi tenaga kesehatan yang bekerja pada fasilitas pelayanan kesehatan dasar atau primer, kelompok profesi, tenaga pendidik, petugas lapangan keluarga berencana, petugas sosial yang terkait dengan pembinaan tumbuh kembang anak, organisasi profesi dan pemangku kepentingan terkait pertumbuhan, perkembangan dan gangguan tumbuh kembang anak.

Anak dengan gangguan perkembangan mempunyai keterbatasan dalam hal merawat diri, kemampuan bahasa reseptif dan ekspresif, keterbatasan dalam hal belajar, kemampuan untuk hidup mandiri dan sosial. Berdasarkan laporan Departemen kesehatan Republik Indonesia cakupan pelayanan kesehatan anak dalam deteksi tumbuh kembang anak yang mengalami gangguan tumbuh kembang di Indonesia mencapai 45,7%. (Depkes RI, 2010).

Keterlambatan gerak motorik kasar pada anak merupakan salah satu gangguan tumbuh kembang anak. Motorik kasar mencakup gerakan otot-otot besar seperti otot tungkai dan lengan pada bayi berupa gerakan menendang,menjejak , meraih, mengangkat leher, dan menoleh. Pertumbuhan kemampuannya harus terus di pantau dan di stimulasi agar anak dapat tumbuh dan berkembang optimal. Gerak kasar atau motorik kasar adalah aspek yang berhubungan dengan

kemampuan anak melakukan pergerakan dan sikap tubuh yang melibatkan otot-otot besar seperti duduk, berdiri, dan sebagainya (Depkes RI, 2010).

Klinik pela 9 merupakan salah satu klinik pelayanan terpadu untuk menangani gangguan tumbuh kembang anak yang berdiri sejak 6 tahun yang lalu. *Assessment* tumbuh kembang merupakan pemeriksaan awal bagi pasien yang dilakukan oleh tim dokter yang terdiri dari, dokter anak, asisten dokter dan psikolog di klinik pela 9 dalam hal pemeriksaan gangguan tumbuh kembang anak. Dari pemeriksaan awal maka akan diketahui masalah tumbuh kembang anak yang dialami pasien, apakah termasuk gangguan tumbuh kembang motorik kasar atau motorik halus atau bahkan pasien normal tidak mengalami gangguan tumbuh kembang anak.

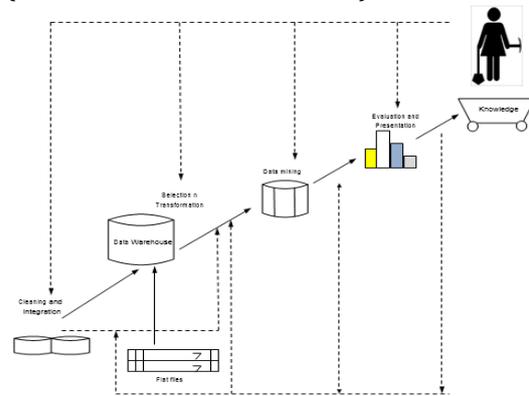
Ada beberapa penelitian yang membahas bagaimana cara mendeteksi dan mengklasifikasikan sebuah data set dengan menggunakan *Naive Bayes*, *Naive Bayes Adaboost* dan *Naive Bayes PSO*. *Naive Bayes* dengan *Adaboost* yang mana dibuat oleh Korada, Kumar dan Deekshitulu (2012) melakukan identifikasi penyakit tanaman jagung menggunakan *Naive Bayes* dengan pembobotan algoritma *Adaboost* menggunakan *Maize Expert System*. Lin dan Jiankun (2011) melakukan penelitian dimana pada penelitian tersebut membandingkan hasil akurasi *Naive Bayes* dengan *Naive Bayes* menggunakan algoritma pembobotan PSO. Penelitian yang dilakukan oleh Karabatak (2015) menjelaskan tentang Algoritma *Naive Bayes* yang digunakan untuk mengklasifikasikan dan mendeteksi penyakit kanker payudara.

Berdasarkan hasil penelitian di atas maka pada penelitian ini akan menggunakan metode *Naive Bayes*, metode ini akan digunakan untuk mengklasifikasikan adanya gangguan gerak motorik kasar anak. Metode *Naive Bayes* dipilih karena proses *learning* dan klasifikasi sangat sederhana dan cepat. Secara umum, metode *Naive Bayes* merupakan metode paling sederhana diantara metode klasifikasi lainnya, *Naive Bayes* dapat menangani data set dengan banyak atribut (Madyatmadja & Aryuni, 2014). Oleh karena itu, pada penelitian ini penulis menggunakan metode *Naive Bayes* serta menggunakan metode optimasinya yaitu *Particle Swarm Optimization* (PSO) dan *Adaboost*.

## BAHAN DAN METODE

### DATA MINING

Data Mining merupakan teknologi baru yang sangat berguna untuk membantu perusahaan-perusahaan menemukan informasi yang sangat penting dari gudang data mereka. Beberapa aplikasi data mining fokus pada prediksi, mereka meramalkan apa yang akan terjadi dalam situasi baru dari data yang menggambarkan apa yang terjadi di masa lalu (Witten, Frank, & Hall, 2011).



Gambar 1. Tahapan data mining

### METODE KLASIFIKASI NAÏVE BAYES

Klasifikasi Bayes juga dikenal dengan *Naive Bayes*, memiliki kemampuan sebanding dengan dengan pohon keputusan dan neural network (Han & Kamber, 2007). Klasifikasi Bayes adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas (Kusrini, 2009). *Naive Bayes* dapat menggunakan penduga kernel kepadatan, yang meningkatkan kinerja jika asumsi normalitas sangat tidak benar, tetapi juga dapat menangani atribut numeric menggunakan diskritisasi diawasi (Witten & Frank, 2011). Berikut adalah bentuk umum dari *teorema bayes*:

$$P(H|X) = \frac{p(X|H)P(H)}{P(X)}$$

Metode algoritma *Naive bayes* merupakan penyederhanaan *metode bayes*. Untuk mempermudah pemahaman, maka *Teorema Bayes* disederhanakan menjadi:

$$P(H|X) = P(X|H) P(X)$$

### 2.5. Particle Swarm Optimization (PSO)

*Particle Swarm Optimization* (PSO) merupakan teknik optimasi *heuristic* global yang diperkenalkan oleh Dokter Kennedy dan Eberhart pada tahun 1995 berdasarkan pada penelitian terhadap perilaku burung dan ikan. Setiap partikel dalam PSO juga dikaitkan dengan kecepatan partikel terbang melalui ruang

pencarian dengan kecepatan yang dinamis disesuaikan untuk perilaku histori mereka. Oleh karena itu, partikel memiliki kecenderungan untuk terbang menuju daerah pencarian yang lebih baik selama proses pencarian. PSO adalah metode pencarian penduduk, yang berasal dari penelitian untuk pergerakan organisme dari kelompok burung atau ikan, seperti algoritma genetika, PSO melakukan pencarian menggunakan populasi (*swarm*) dari individu (partikel) yang diperbaharui dari iterasi untuk iterasi (Fei, Miao, & Liu, 2009).

**ADABOOST**

*Algoritma Adaboost* awalnya diusulkan untuk klasifikasi, namun sekarang Adaboost juga dapat digunakan untuk masalah regresi dan masalah prediksi. Dimana Adaboost merupakan pendekatan yang layak untuk mengurangi kesalahan prediksi. Ide utama dari algoritma ini adalah untuk menilai kinerja prediktor dengan menghitung kesalahan proporsional. Dalam proses pelatihan algoritma AdaBoost, kesalahan dan berat distribusi masing-masing sampel pelatihan diperbarui. Jika kesalahan besar, meningkatkan berat distribusi sampel, jika kesalahan kecil, menurunkan berat distribusi sampel. Kemudian sesuai dengan berat distribusi baru, kita melatih sampel, mengurangi kesalahan mesin berikutnya, dan mendapatkan output prediksi yang lebih baik (Jinna Lu, Honping Hu dan Yanping Bai, 2014).

**Evaluasi Kinerja Algoritma**

Untuk melakukan evaluasi pada metode *Naive Bayes* dengan optimasi menggunakan teknik PSO dan *Adaboost* pengujian model dilakukan beberapa pengujian menggunakan confusion matrix dan kurva *Receiver Operating Characteristic* (ROC).

**1. Cross Validation**

Pendekatan alternatif untuk 'training dan testing' yang sering diadopsi ketika sejumlah contoh kecil (dan yang banyak yang memilih menggunakan terlepas dari ukuran) dikenal sebagai k kali lipat cross-validasi. Jika dataset terdiri kasus N, ini dibagi menjadi bagian-bagian k sama, k biasanya menjadi sejumlah kecil seperti 5 atau 10 (Jika N tidak tepat habis dibagi oleh k, bagian akhir akan memiliki kasus lebih sedikit dari k lain - 1 bagian) Serangkaian berjalan k kini dilakukan. Setiap bagian k pada gilirannya digunakan sebagai ujian menetapkan dan k lainnya - 1 bagian digunakan sebagai training set (Bramer, 2007).

**2. Confusion Matrix**

*Confusion matrix* merupakan dataset hanya memiliki dua kelas, kelas yang satu sebagai positif dan kelas yang lain sebagai negatif (Bramer, 2007). Metode ini menggunakan tabel matrix seperti pada tabel 1.

Tabel 1. Model *Confusion Matrix*

Classification		Predictions		Total
		-1(negative)	+(positif)	
Examples	-1 (negatif)	p	q	p + q
	+1 (positif)	u	v	u + v
	Total	p + u	q + v	m

Sumber : Vercellis, 2009

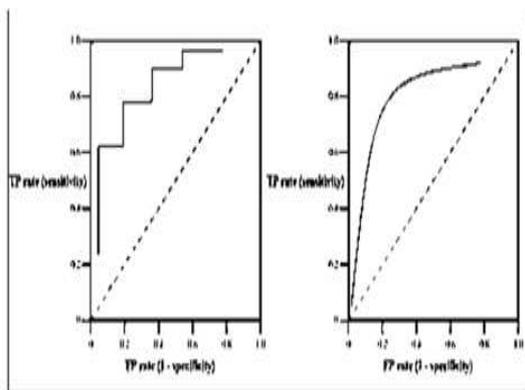
Keterangan :

- p adalah jumlah prediksi yang tepat bahwa *instance* bersifat negatif.
- q adalah jumlah prediksi yang salah bahwa *instance* bersifat positif.
- u adalah jumlah prediksi yang salah bahwa *instance* bersifat negatif.
- v adalah jumlah prediksi yang tepat bahwa *instance* bersifat positif.

**3. Kurva Roc**

Grafik kurva ROC digunakan untuk mengevaluasi akurasi classifier dan untuk membandingkan klasifikasi yang berbeda model (Vercellis, 2009). Sebuah grafik ROC adalah grafik dua dimensi dengan proporsi negatif pada sumbu horisontal dan proporsi positif yang benar di sumbu vertikal (Vercellis, 2009). Kegunaan kurva ROC adalah (Gorunescu, 2011) adalah untuk radar selama Perang Dunia II untuk mendeteksi benda-benda musuh di medan perang, teori deteksi sinyal, dalam psikologi ke rekening untuk deteksi sinyal persepsi, penelitian medis dan dalam mesin pembelajaran dan penelitian data mining, serta masalah klasifikasi. Secara teknis kurva ROC juga dikenal sebagai grafik ROC, dua dimensi grafik dimana tingkat TP diplot pada

sumbu Y dan tingkat FP diplot pada sumbu X. (Gorunescu, 2011).



Sumber : Gorunescu, 2011

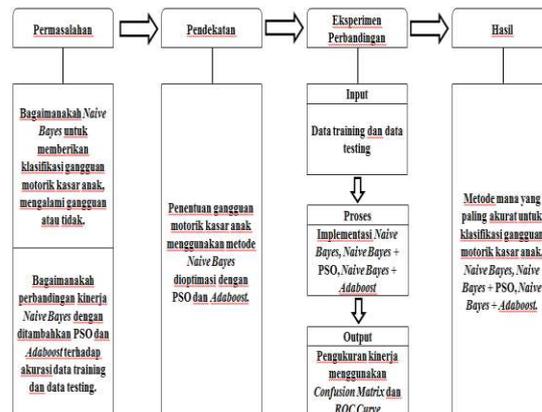
Gambar 2. Model Grafik ROC

Dapat disimpulkan bahwa, satu point pada kurva ROC adalah lebih baik dari pada yang lainnya jika arah garis melintang dari kiri bawah ke kanan atas didalam grafik. Hasil perhitungan dapat divisualisasikan dengan kura ROC atau *Area Under Curve* (AUC). Adapun nilai akurasi dari grafik ROC dimulai dari terendah yaitu 0.50 dan yang tertinggi adalah 1.00 menurut (Gorunescu, 2011) sebagai berikut:

1. Akurasi bernilai 0.90 - 1.00 = *Excellent classification*
2. Akurasi bernilai 0.80 - 0.90 = *Good classification*
3. Akurasi bernilai 0.70 - 0.80 = *Fair classification*
4. Akurasi bernilai 0.60 - 0.70 = *Poor classification*
5. Akurasi bernilai 0.50 - 0.60 = *Failure*

#### KERANGKA PEMIKIRAN

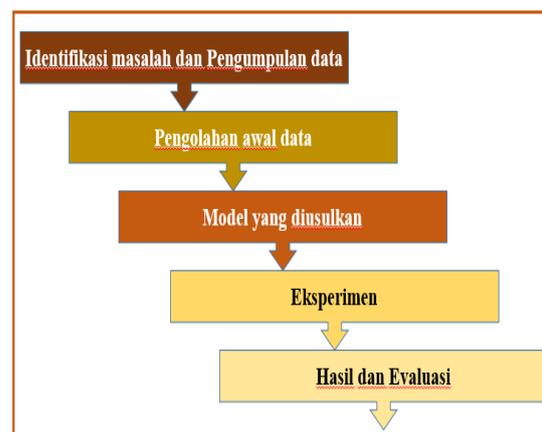
Komponen dari model kerangka pemikiran pada penelitian ini terdiri dari permasalahan, pendekatan, eksperimen perbandingan dan hasil. Kerangka pemikiran tersebut digambarkan sebagai berikut:



Gambar 3. Kerangka Pemikiran

#### Metode Penelitian

Dalam penelitian ini penulis menggunakan data pemeriksaan awal pasien yang melakukan pemeriksaan mengenai gangguan gerak motorik kasar anak pada klinik pela 9 cabang kebayoran. Data pemeriksaan awal gangguan gerak motorik kasar anak ini diolah dengan *RapidMiner Studio* menggunakan metode *Naive Bayes* serta menggunakan metode optimasinya yaitu *PSO* dan *Adaboost* sehingga diperoleh metode mana yang paling akurat dan dapat digunakan sebagai acuan dalam memprediksi adanya gangguan gerak motorik kasar pada anak.



Gambar 4. Tahapan Penelitian

#### JADWAL PENELITIAN

Sistem penjadwalan membantu dalam penyusunan penelitian supaya pelaksanaan lebih teratur dan rapi, selain itu juga dapat membantu pelaksanaan pengerjaan apa yang harus dilakukan. Berikut tabel jadwal penelitian dapat dilihat dibawah ini:

Tabel 2. Jadwal Penelitian

NO	Kegiatan	Bulan / Tahun 2016							
		JAN	FEB	MAR	APR	MEI	JUN	JUL	AGU
1	Tahap Persiapan Penelitian								
	a. Penyusunan dan Pengajuan Judul	■							
	b. Pengajuan Proposal	■	■						
	c. Perijinan Riset	■	■						
2	Tahap Pelaksanaan								
	a. Riset Penelitian		■						
	b. Pengumpulan Data		■						
	c. Analisa Data			■					
	d. Pengolahan Data			■	■				
3	Tahap Penyusunan Laporan								■

**Pengumpulan Data**

Dalam penelitian ini, pengumpulan data adalah data pasien yang memeriksakan adanya gangguan gerak motorik kasar anak atau tidak, pada Klinik Pela 9 cabang Kebayoran. Dimana dari atribut atau variable yang ada dapat diketahui bahwa pasien tersebut mengalami gangguan gerak motorik kasar atau normal tidak mengalami gangguan gerak motorik kasar. Dataset diambil dari populasi data pasien selama 24 bulan, yaitu populasi data dari bulan September 2013 sampai dengan Agustus 2015 untuk hasil penelitian yang baik. Didapat 600 record data pasien pemeriksaan awal gangguan gerak motorik kasar anak, dengan atribut atau variable sebanyak 13 variable dan yang dijadikan implikasi untuk penelitian ini sebanyak 10 variable dengan 1 variabel yaitu kategori sebagai label dan 9 variabel menjadi indikator. Variabel tersebut terdiri dari jenis kelamin, usia, berat badan, tinggi badan, genetik, lingkungan, pendidikan orangtua, stimulasi orangtua, anak ke-n dan kategori sebagai label.

**Pengolahan Awal Data**

ada pengolahan awal data pada penelitian ini perlu dilakukan pengambilan data untuk hasil penelitian yang baik. Jumlah Dataset yang diperoleh, diambil dari populasi data pasien selama 24 bulan, yaitu populasi data dari bulan September 2013 sampai dengan Agustus 2015 sebanyak 600 record data pasien dengan 13 variabel awal. Namun dalam pengolahan data tidak semua atribut dapat digunakan karena harus melalui beberapa tahap pengolahan awal data (preparation data). Untuk mendapatkan data yang berkualitas, beberapa teknik yang dilakukan sebagai berikut (vecellis, 2009):

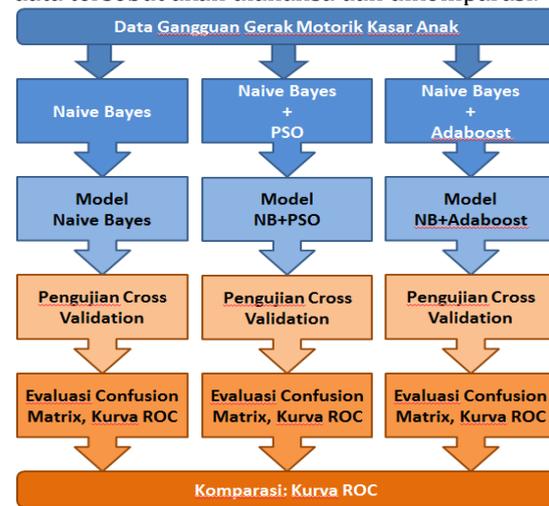
1. Data validation, untuk mengidentifikasi dan menghapus data yang ganjil (outlier/noise), data yang tidak konsisten,

dan data yang tidak lengkap (missing value).

2. Data integration and transformation, untuk meningkatkan akurasi dan efisiensi algoritma. Data yang digunakan dalam penulisan ini bernilai kategorikal. Data ditransformasikan kedalam bentuk angka untuk memudahkan pengelolaan didalam software rapidminer.

**Model Yang Diusulkan**

Pada tahap pemodelan usulan ini akan membahas metode algoritma yang diuji dengan pemrosesan data training gangguan gerak motorik kasar anak yang didapat dari Klinik Pela 9 cabang Kebayoran. Hasil dari pengujian data tersebut akan dianalisa dan dikomparasi.



Gambar 5. Model Yang Diusulkan

**Metode Naïve Bayes Classifier**

Data yang digunakan pada model Naïve Bayes adalah data pasien selama 24 bulan di Klinik Pela 9 Cabang Kebayoran, yaitu populasi data dari bulan September 2013 sampai dengan Agustus 2015 sebanyak 600 record. Dalam membuat model algoritma Naïve Bayes terlebih dahulu kita mencari nilai probabilitas prior dan posterior hipotesis untuk masing-masing class P(H). Hipotesis yang ada yaitu pasien yang hasil pemeriksaan awal mengalami gangguan dan normal.

**1. Menghitung Probabilitas Prior**

Probabilitas Prior dilakukan dalam bentuk persamaan di bawah ini:

$$\begin{aligned}
 \text{Total data} &= 600 \\
 \text{Data Pasien Gangguan} &= 311 \\
 \text{Data Pasien Normal} &= 289 \\
 P(\text{Gangguan}) &= 311 : 600 = 0,51833 \\
 P(\text{Normal}) &= 289 : 600 = 0,48167
 \end{aligned}$$

Setelah didapatkan nilai *probabilitas* untuk tiap hipotesis dari *class*, maka langkah selanjutnya adalah melakukan penghitungan terhadap kondisi *probabilitas* tertentu (*Probabilitas X*) dengan menggunakan data berdasarkan *probabilitas* tiap hipotesis (*Probabilitas H*) atau yang dinamakan dengan *probabilitas Prior*. Selanjutnya untuk mengetahui hasil perhitungan dari *Probabilitas Prior*, maka dilakukan penghitungan dengan cara merinci jumlah kasus dari tiap-tiap atribut variabel data, adapun hasil perhitungan *probabilitas prior* dengan menggunakan *Algoritma Naive Bayes* dapat dilihat pada tabel dibawah ini:

**Tabel 5. Perhitungan Probabilitas Prior Naive Bayes**

Atribut	Jumlah Kasus (S)	Normal(Si)		Gangguan(Si)		P(X Ci)	
		Normal	Gangguan	Normal	Gangguan	Normal	Gangguan
Total	600	289	311	0.48167	0.51833		
Jenis Kelamin	1	323	177	146	0.61246	0.46945	
	2	277	112	165	0.38754	0.53055	
Usia	1	61	11	50	0.03806	0.16077	
	2	168	70	98	0.24221	0.31511	
	3	184	114	70	0.39446	0.22508	
	4	74	25	49	0.08651	0.15756	
	5	113	69	44	0.23875	0.14148	
Berat Badan(g)	1	128	51	77	0.17647	0.24759	
	2	105	38	67	0.13149	0.21543	
	3	367	200	167	0.69204	0.53698	
	4	0	0	0	-	-	
Tinggi Badan(cm)	1	131	35	96	0.12111	0.30868	
	2	234	121	113	0.41869	0.36334	
	3	124	40	84	0.13841	0.27010	
	4	70	59	11	0.20415	0.03537	
	5	45	34	11	0.11765	0.03537	
Genetik	1	211	59	152	0.20415	0.48875	
	2	389	230	159	0.79585	0.51125	
Lingkungan	1	209	89	120	0.30796	0.38585	
	2	391	200	191	0.69204	0.61415	
Pendidikan Orangtua	1	153	35	118	0.12111	0.37942	
	2	143	79	64	0.27336	0.20579	
	3	147	73	74	0.25260	0.23794	
	4	157	102	55	0.35294	0.17685	
Stimulasi	1	154	42	112	0.14533	0.36013	
	2	205	104	101	0.35986	0.32476	
	3	241	143	98	0.49481	0.31511	
Anak Ke-n	1	148	75	73	0.25952	0.23473	
	2	256	159	97	0.55017	0.31190	
	3	196	55	141	0.19031	0.45338	

Sumber : Hasil Pengolahan Data (2016)

## 2. Menghitung Probabilitas Posterior

Tahapan selanjutnya adalah menggunakan *Probabilitas Prior* untuk menentukan *class* terhadap temuan kasus baru, dengan cara terlebih dahulu menghitung *Probabilitas Posteriornya*, hal tersebut dilakukan apabila ditemukan kasus baru dalam pengolahan data. Berikut tabel *probabilitas posterior* untuk menghitung kasus baru yang ditemukan:

**Tabel 6. Perhitungan Probabilitas Posterior NB**

Data X		P(X Ci)	
Atribut	Nilai	Normal	Gangguan
Jenis Kelamin	2	0.38754	0.53055
Usia	2	0.24221	0.31511
Berat Badan	3	0.69204	0.53698
Tinggi Badan	2	0.41869	0.36334
Genetik	2	0.79585	0.51125
Lingkungan	2	0.69204	0.61415
Pendidikan Orangtua	1	0.12111	0.37942
Stimulasi	1	0.14533	0.36013
Anak Ke-	3	0.19031	0.45338

Sumber : Hasil Pengolahan Data (2016)

Selanjutnya setelah mengetahui nilai *probabilitas* dari setiap atribut terhadap *probabilitas* tiap *class* atau yang dirumuskan dalam bentuk persamaan  $P(X|Ci)$ , maka langkah berikutnya adalah melakukan penghitungan terhadap total keseluruhan *probabilitas* tiap *class*. Berikut persamaan untuk menghitung *probabilitas* tiap *class*:

$$P(X|\text{Pasien} = \text{Normal}) = 0.38754 \times 0.24221 \times 0.69204 \times 0.41869 \times 0.79585 \times 0.69204 \times 0.12111 \times 0.14533 \times 0.19031 = 0.00005$$

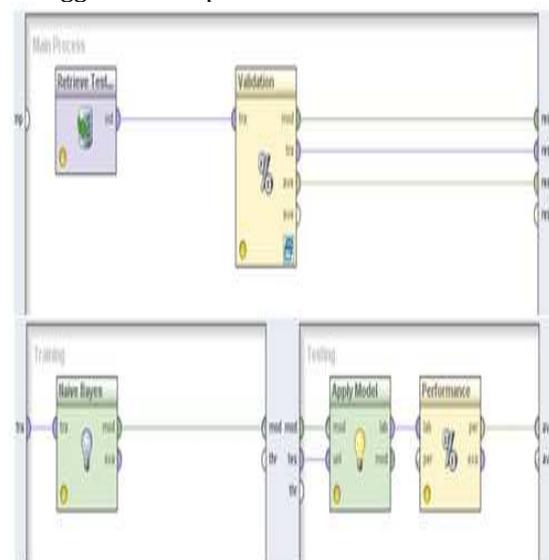
$$P(X|\text{Pasien} = \text{Gangguan}) = 0.53055 \times 0.31511 \times 0.53698 \times 0.36334 \times 0.51125 \times 0.61415 \times 0.37942 \times 0.36013 \times 0.45338 = 0.00063$$

$$P(X|\text{Pasien} = \text{Normal}) P(\text{Normal}) = 0.00005 \times 0.48167 = 0.00002$$

$$P(X|\text{Pasien} = \text{Gangguan}) P(\text{Gangguan}) = 0.00063 \times 0.51833 = 0.00033$$

Dari hasil perhitungan tersebut diketahui nilai  $P(X|\text{Gangguan})$  lebih besar dari pada nilai  $P(X|\text{Normal})$ , sehingga dapat disimpulkan bahwa untuk kasus tersebut akan masuk ke dalam klasifikasi *Gangguan*.

Berikut gambar pengujian modelnya menggunakan *RapidMiner* :



**Gambar 6. Pengujian Model Algoritma Naïve Bayes Classifier**

Dari hasil eksperimen menggunakan algoritma Naïve Bayes dengan menggunakan data sebanyak 600 record didapatkan hasil akurasi sebesar **88.67 %** dengan masih terdapat **11.33%** kesalahan.

**HASIL DAN PEMBAHASAN**

Untuk mendapatkan nilai akurasi yang benar, maka diperlukan proses klasifikasi dan alat ukur yang tepat, karena keakuratan klasifikasi merupakan alat ukur yang bisa menunjukkan bagaimana cara untuk mengklasifikasi sehingga dapat mengidentifikasi data objek dengan benar (Gorunescu, 2011). Proses pengujian terhadap nilai akurasi dapat dilakukan dengan cara melakukan evaluasi tingkat akurasi dari algoritma. Tools yang digunakan untuk pengujian adalah *software rapid miner* serta menggunakan model *Confussion Matrix* dan kurva ROC.

**Hasil Pengujian Metode Naïve Bayes Classifier**

**a) Confussion Matrix**

**Tabel 10. Confussion Matrix Algoritma Naïve Bayes Classifier**

accuracy:88.67% +/- 3.71%(mikro:88.67%)			
	true Normal	true Gangguan	class precision
pred. Normal	261	37	87.58%
pred. Gangguan	31	271	89.74%
class recall	89.38%	87.99%	

Sumber : Hasil Pengolahan Data (2016)

Jumlah *True Positive* (TP) adalah 271 record diklasifikasikan sebagai pasien mengalami GANGGUAN dan *False Negative* (FN) sebanyak 37 record diklasifikasikan sebagai pasien mengalami GANGGUAN tetapi pasien NORMAL. Berikutnya 261 record untuk *True Negative* (TN) diklasifikasikan sebagai pasien NORMAL, dan 31 record *False Positive* (FP) diklasifikasikan sebagai pasien NORMAL tetapi pasien mengalami GANGGUAN.

Tingkat akurasi dengan menggunakan algoritma *Naïve Bayes Classifier* adalah sebesar 88.67% dan dapat dihitung untuk mencari nilai *accuracy*, *sensitivity*, *specificity*, *ppv*, dan *npv* pada persamaan dibawah ini:

$$\text{Sensitivity} = \frac{TP}{TP+FN} = \frac{271}{271+37} = 0.879$$

$$\text{Specificity} = \frac{TN}{TN+FP} = \frac{261}{261+31} = 0.893$$

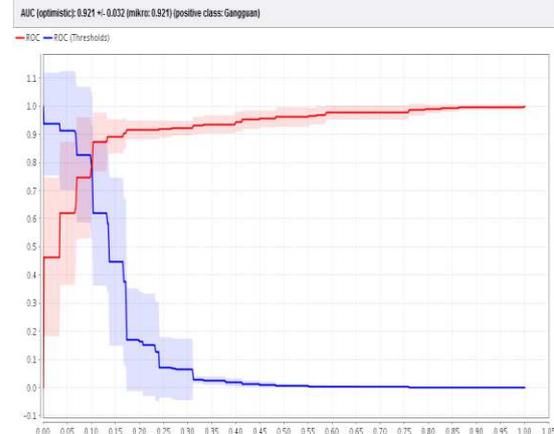
$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{271+261}{271+261+31+37} = 0.886$$

$$\text{PPV} = \frac{TP}{TP+FP} = \frac{271}{271+31} = 0.897$$

$$\text{Precision} = \frac{TN}{TN+FN} = \frac{261}{261+37} = 0.875$$

**b) Evaluasi ROC Curve**

Kurva ROC mengekspresikan confusion matrix dari tabel 10. Garis horizontal adalah false positives dan garis vertikal *true positives*. Menghasilkan nilai AUC sebesar 0.921 dengan nilai akurasi klasifikasi baik (*good classification*).



**Gambar 9. Nilai AUC dalam grafik ROC Algoritma Naïve Bayes Classifier**

**Hasil Pengujian Metode Naïve Bayes Classifier + Adaboost**

**a) Confussion Matrix**

**Tabel 11. Confussion Matrix Algoritma Naïve Bayes Classifier + Adaboost**

accuracy:90.00% +/- 3.65%(mikro:90.00%)			
	true Normal	true Gangguan	class precision
pred. Normal	260	28	90.28%
pred. Gangguan	32	280	89.74%
class recall	89.04%	90.91%	

Sumber : Hasil Pengolahan Data (2016)

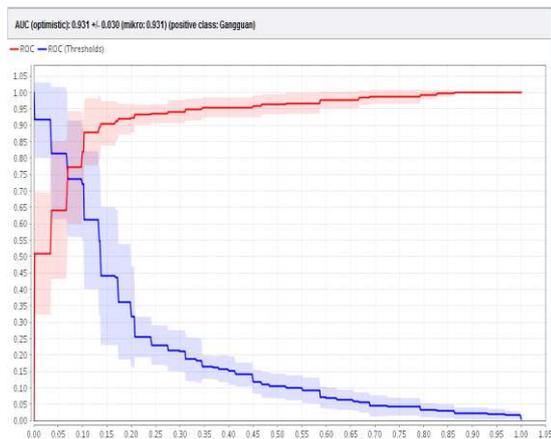
Jumlah *True Positive* (TP) adalah 280 record diklasifikasikan sebagai pasien mengalami GANGGUAN dan *False Negative*

(FN) sebanyak 28 record diklasifikasikan sebagai pasien GANGGUAN tetapi pasien NORMAL. Berikutnya 260 record untuk True Negative (TN) diklasifikasikan sebagai pasien NORMAL, dan 32 record False Positive (FP) diklasifikasikan sebagai pasien NORMAL tetapi pasien mengalami GANGGUAN.

Tingkat akurasi dengan menggunakan algoritma *Naive Bayes Classifier* berbasis *Adaboost* adalah sebesar 90.00%.

**b) Evaluasi ROC Curve**

Kurva ROC mengekspresikan confusion matrix dari tabel 10. Garis horizontal adalah false positives dan garis vertikal *true positives*. Menghasilkan nilai AUC sebesar 0.931 dengan nilai akurasi klasifikasi baik (*good classification*).



**Gambar 10. Nilai AUC dalam grafik ROC Algoritma Naive Bayes Classifier+Adaboost**

**Hasil Pengujian Metode Naive Bayes Classifier + PSO**

**a) Confusion Matrix**

**Tabel 12. Confusion Matrix Algoritma Naive Bayes Classifier + PSO**

accuracy: 98.00% +/- 1.25% (mikro: 98.00%)			
	true Normal	true Gangguan	class precision
pred. Normal	283	4	98.61%
pred. Gangguan	8	305	97.44%
class recall	97.25%	98.71%	

Sumber : Hasil Pengolahan Data (2016)

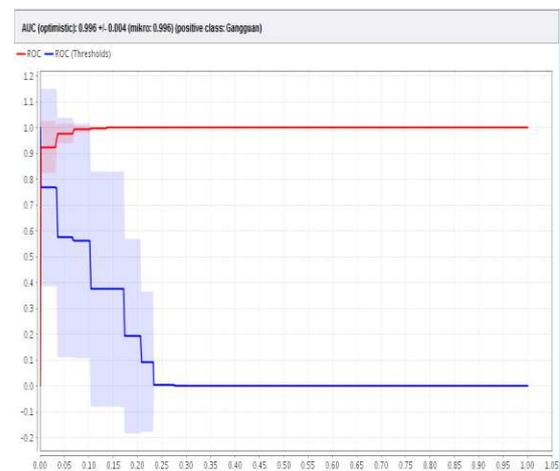
Jumlah True Positive (TP) adalah 305 record diklasifikasikan sebagai pasien

mengalami GANGGUAN dan False Negative (FN) sebanyak 4 record diklasifikasikan sebagai pasien GANGGUAN tetapi pasien NORMAL. Berikutnya 283 record untuk True Negative (TN) diklasifikasikan sebagai pasien NORMAL, dan 8 record False Positive (FP) diklasifikasikan sebagai pasien NORMAL tetapi pasien mengalami GANGGUAN.

Tingkat akurasi dengan menggunakan algoritma *Naive Bayes Classifier* berbasis *Adaboost* adalah sebesar 98.00% .

**b) Evaluasi ROC Curve**

Kurva ROC mengekspresikan confusion matrix dari tabel 10. Garis horizontal adalah false positives dan garis vertikal *true positives*. Menghasilkan nilai AUC sebesar 0.996 dengan nilai akurasi klasifikasi baik (*good classification*).



**Gambar 11. Nilai AUC dalam grafik ROC Algoritma Naive Bayes Classifier+PSO**

Teknik validasi yang dilakukan pada penelitian ini menggunakan *tenfold cross validation* dengan membagi data secara acak ke dalam 10 bagian dimana dari setiap bagian akan dilakukan proses klasifikasi. Dari hasil pengujian di atas, hasil pengujian metode Naive Bayes dengan optimasinya menggunakan PSO dan metode Naive Bayes dengan optimasinya menggunakan *Adaboost* ditampilkan pada tabel di bawah ini:

**Tabel 13. Hasil Perbandingan Metode Naive Bayes, Naive Bayes + PSO dan Naive Bayes + Adaboost**

Record Data	Accuracy			AUC		
	NB	NB+PSO	NB+Adaboost	NB	NB+PSO	NB+Adaboost
600	88.67%	98.00%	90.00%	0.921	0.996	0.931

Sumber: Hasil Pengolahan Data (2016)

Hasil pengujian algoritma Naïve Bayes classifier, Naive Bayes classifier dengan optimasi PSO dan Naive Bayes classifier berbasis Adaboost. Didapat bahwa pengujian Naive Bayes classifier dengan optimasi PSO lebih baik dari pada Naive Bayes classifier berbasis Adaboost. Dengan record data pasien sebanyak 600, pengujian menggunakan Naive Bayes classifier berbasis Adaboost didapat akurasi sebesar 90.00% dengan masih terdapat kesalahan data sebesar 10.00%. Sedangkan akurasi pengujian Naive Bayes classifier dengan optimasi PSO meningkat 8.00% yaitu sebesar 98.00% dengan tingkat kesalahan data rendah, dengan menggunakan data yang sama.

### KESIMPULAN

Dari hasil pengujian yang telah dilakukan dan kesimpulan yang sudah dipaparkan maka saran untuk penelitian ini antara lain :

1. Penelitian ini diharapkan dapat digunakan oleh klinik tumbuh kembang anak atau rumah sakit yang menyediakan pemeriksaan gangguan tumbuh kembang anak untuk menentukan pasien mengalami gangguan gerak motorik kasar anak atau pasien tidak mengalami gangguan dengan kata lain normal.
2. Jumlah data yang digunakan pada penelitian ini sebanyak 600 data dengan jumlah atribut 9, untuk pengukuran yang lebih akurat lagi disarankan untuk menambah jumlah data yang lebih besar.

Dalam pengembangan penelitian kedepannya disarankan menggunakan algoritma pembobotan selain PSO dan Adaboost seperti Bee Colony, Bat algoritma dan yang lainnya untuk mencari akurasi tertinggi.

### REFERENSI

Ariani, Yosoprawoto, M (2012). *Children Age and Mother Literacy as the Risk Factors for Children Development Disorder*. Jurnal Kedokteran Brawijaya, Vol. 27 No.2.

Bramer, M. (2007). *Principle of Data Mining*. Berlin: Springer.

Departemen Kesehatan Republik Indonesia, 2010. *Pedoman Nasional Tumbuh Kembang Anak*. Jakarta: Gramedia.

Fei, S. W., Miao, Y. B., & Liu, C. L. (2009). *Chinese Grain Production Forecasting Method Based On Particle Swarm Optimization-Based Support Vector Machine*. Recent Patents On Engineering 2009 , 3, 8-12.

Gorunescu, F. (2011). *Data Mining Concepts, Models and Techniques*. Intelligent System Reference Library, Vol 12, ISBN 978-3-642-19721-5.

Han, J., & Kamber, M. (2007). *Data Mining: Concepts and Techniques, Second Edition*. Soft Computing. San Francisco: Morgan Kaufmann Publishers.

Karabatak, Murat. (2015). *A New Classifier for Breast Cancer Detection Based on Naïve Bayesian*. Journal of Software Engineering, Measurement 72, pp 32-36.

Korada, N. K., Kumar, N. S. P., & Deekshitulu, Y. V. N. H. (2012). *Implementation of Naive Bayesian and Ada-Boost Algorithm Using Maize Expert System*. IJIST Vol.2, No.3, May 2012. DOI: 10.5121/ijist.2012.2305.

Kusrini & Ema Taufiq L. (2009). *Algoritma Data Mining*. Yogyakarta: Andi.

Lin, Jie., & Jiankun, Yu. (2011). *Wighted Naive Bayes Classification Algorithm Based on Particle Swarm Optimization*. IEEE 978-1-61284-486-2/11/2011.

Lu, Jinna., Hu, Hongping., & Yanping, Bai. (2014). *Generalized Radial Basic Function Neural Network Based on an Improved Dynamic Particle Swarm Optimization and Adaboost Algorithm*. Neurocomputing, Vol 152, pp. 305-315.

Madyatmadja, E. D & Mediana Aryuni. (2014). *Comparative Study of Data Mining Model for Credit Card Application Scoring Bank*. Journal of Theoretical and Applied Information Technology, Vol. 59 No 2, pp. 269-274.

Soetjningsih. (1998). *Tumbuh Kembang Anak*. Jakarta: EGC.

Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. United Kingdom: John Wiley & Sons. Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. United Kingdom: John Wiley & Sons.

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining Practical Machine Learning Tools And Techniques*. Burlington, Usa: Morgan Kaufmann Publishers.

Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. United Kingdom: John Wiley & Sons.