

TEKNIK BAYESIAN NETWORK PADA PENGOLAHAN CITRA UNTUK IDENTIFIKASI**¹Erwin, dan ²Rossi Passarella**

^{1,2}Jurusan Sistem Komputer, Universitas Sriwijaya,
e-mail: erwinunsri@gmail.com, passarella.rossi@gmail.com

Abstrak. Teknik identifikasi yang merupakan pengembangan dari teknik biometrik yang berdasarkan pada beberapa ciri alami manusia. Salah satunya menggunakan iris matam manusia. Selaput pelangi mata digunakan untuk mengetahui gambaran kondisi kesehatan manusia. Pada penelitian ini, kondisi usus besar (colon) akan diamati melalui selaput pelangi mata, dimana kondisi colon ini akan tergambar pada zona 2 dan 3 dari selaput pelangi mata, sedangkan kondisi colon akan dikelompokkan menjadi 4 (empat) kondisi, yaitu Colon Normal, Ballooned Sigmoid, Diverticulata, dan Spasm. Kumpulan piksel citra iris mata digunakan untuk mengidentifikasi citra sebagai sistem cerdas dengan menggunakan teknik Bayesian Network. Teknik Bayesian Network berbasis probabilitas yang merepresentasikan suatu himpunan variabel dan conditional interdependencies menggunakan DAG (Directed Acyclic Graph). Teknik ini merupakan pengembangan dari model dasar Naïve Bayes. Data uji merupakan hasil pengambilan data pasien rumah sakit di Palembang diperoleh sebanyak 94 citra. Tingkat akurasi dalam identifikasi untuk Naïve Bayes sebesar 63,83% dan Bayesian Network sebesar 70,21%.

Kata kunci: Identifikasi, Iris mata, Bayesian Network

1. Pendahuluan

Identifikasi iris mata adalah cara mengidentifikasi mata manusia berdasarkan gambaran bentuk pola iris mata (Masek, L. 2003). Otak manusia memiliki kemampuan yang handal dalam melakukan pengenalan iris mata. Namun, mata manusia memiliki keterbatasan untuk dapat mengenali dua buah iris mata dengan pola yang hampir sama, walaupun manusia mampu melakukannya tetapi dibutuhkan waktu lama. Perkembangan penggunaan komputer, diharapkan kemampuan identifikasi iris yang dimiliki oleh manusia dapat diterapkan pada sistem cerdas (Duin, R., & Pekalska, E. 2007).

Pengidentifikasian kondisi colon melalui iris mata bagi sebagian orang awam tidaklah mudah. Diperlukan beberapa pelatihan dan teknik tertentu untuk mengetahui kondisi tubuh manusia, seperti usus besar. Iridologi mampu membantu menganalisis dan menjelaskan gambaran iris mata secara spesifik.

Citra merupakan data yang tidak konsisten atau data yang bias sehingga metode Bayesian ini baik untuk digunakan (Hanson, K. M. 1987). Metode Bayes yang paling sederhana adalah teknik Naïve Bayes yang menggunakan asumsi model fitur yang saling independen merupakan classifier sederhana. Pengembangan dari metode Naïve Bayes adalah teknik Tree Augmented Naïve Bayes (TAN). Kedua metode tersebut menggunakan asumsi bahwa semua variabel akibat dianggap saling bebas bersyarat (*conditionally independent*) karena variabel sebab dan merupakan teknik penyerdehanaan dari kondisi yang sebenarnya yaitu semua variabel bebas (*random variable*) dianggap saling mempengaruhi. Hal ini merupakan dasar dari teknik Bayesian Network.

Permasalahan yang biasanya terjadi untuk mengidentifikasi kondisi colon melalui iris mata adalah sebagai berikut:

- a. Penentuan posisi atau letak colon dalam peta iridologi jensen yang salah atau tidak sesuai;
- b. Kurang mengertinya pasien untuk menganalisis hasil citra iris mata dengan kondisi colon;
- c. Tingkat akurasi dalam penentuan kondisi colon melalui iris mata;
- d. Diameter pupil mata pasien yang berbeda-beda satu sama lain;
- e. Sulit mendeteksi iris mata pasien yang menderita kolesterol tinggi dan gangguan komplikasi karena terlalu banyak warna putih di iris mata;

Oleh karena itu, dibutuhkan suatu perangkat lunak yang mampu mengolah citra iris mata untuk mengetahui kondisi colon menggunakan ilmu iridologi dengan menggunakan metode *Bayesian Network*, yang memberikan hubungan probabilistik dari penentuan kondisi usus besar melalui iris mata.

1.1 Teknik Identifikasi berdasarkan Iridologi

Mata adalah salah satu indra manusia yang penting. Stimulasi *reseptor* peka cahaya di mata (*fotoreseptor*) menimbulkan indra penglihatan (Corwin, 2009). Struktur mata terdiri dari *sklera, kornea, koroid, iris, pupil*, lensa mata dan *retina*. Dalam sistem biometrik yang digunakan untuk pengidentifikasian dan pendeteksian dalam suatu studi kasus, struktur mata yang paling sering digunakan adalah iris mata.

Konsep identifikasi berbasis iridologi ini cukup bisa dipahami. Iris mata kanan menggambarkan organ tubuh bagian kanan, seperti *pankreas*, kaki kanan, paru-paru kanan, ginjal kanan, kepala bagian kanan, dan sebagainya. Sedangkan pada iris mata kiri, menggambarkan organ tubuh bagian kiri, seperti jantung, kepala bagian kiri dan sebagainya (Jensen, 1980). Konsep kerja menggunakan iridologi ini berbanding terbalik dengan proses penglihatan pada manusia, dimana sebagian serabut mata kanan menyebrang dan memproyeksikan ke sisi kiri otak. Pada saat yang sama, sebagian serabut mata kiri menyebrang dan memproyeksikan ke sisi kanan otak (Corwin, 2009). Pramono *et al* (2006) membangun aplikasi Metode *Backpropagation* untuk pengenalan perubahan organ *pankreas* melalui iris mata. Analisis dari penelitian ini adalah menentukan kondisi gula dalam pengenalan penyakit *Diabetes Mellitus* yang berhubungan dengan organ *pankreas* sebagai penghasil hormon *insulin*.

Rochmad (2009) yang menggunakan teknik *Naïve Bayes* untuk meneliti gangguan *pankreas* dengan asumsi probabilitas *Hypothesis Maximum Appropri Probability (HMAP)*. Pengujian dilakukan pada 98 sampel mata yang terdiri dari 54 citra iris mata yang dikategorikan kondisi *pankreas* normal dan 44 citra mata yang dikategorikan kondisi *pankreas* tidak normal. Hasil perhitungan metode *Bayes* diperoleh keberhasilan senilai 59,15%. Jiang *et al* (2005) mengembangkan teknik *Tree Augmented Naïve Bayes (TAN)*. Kedua metode tersebut menggunakan asumsi bahwa semua variabel akibat dianggap saling *conditionally independent* karena variabel sebab dan merupakan teknik penyerdahaan dari kondisi yang sebenarnya yaitu semua *random variable* dianggap saling mempengaruhi. Hal ini merupakan dasar dari teknik *Bayesian Network* (Gamez *et.al*, 2010, dan Gat-Viks, I *et.al*, 2006).

Penerapan *Bayesian Network* pada *Data Mining* dilakukan Witten (2005) yang menghasilkan teknik-teknik praktis pembelajaran dan efisiensi *Bayesian Network* dilakukan oleh Wong (2004), khusus untuk ekspresi data biologi seperti struktur gen dikembangkan oleh Friedman, N, *et.al* (2000) dan Helman, P *et.al* (2004) untuk data

ekspresi gen. WenChenx, et.al (2008) mengajukan algoritma K2 untuk menemukan struktur node dalam *Bayesian Network* dan Gamez, et.al (2005) menemukan fungsi pembatas dalam pencarian nilai parameter *Bayesian Network*.

Erwin dkk (2013) telah melakukan identifikasi gangguan colon berdasarkan citra iris mata menggunakan metode Naïve Bayes dengan tingkat akurasi sebesar 62.5%. Hasil dari penelitian ini memiliki error sebesar 37.5% dengan 25 data yang benar dan 15 data yang salah pengidentifikasiannya dari jumlah total sebesar 40 citra training. Selanjutnya, Rossi Passarella dkk (2013) membuat database menggunakan 60 orang subjek, yang terdiri dari 35 orang secara historis memiliki masalah dengan usus besar, sedangkan 25 orang subjek lainnya tidak diketahui latar belakangnya. Hasil penelitian ini menunjukkan bahwa hanya 8% dari 35 orang subjek yang tidak teridentifikasi penyakit usus besarnya.

2. Teorema Bayes

Teorema Bayes adalah suatu pendekatan untuk sebuah ketidakpastian yang diukur dengan probabilitas. Metode Bayes merupakan struktur grafis untuk mewakili hubungan probabilistik antara sejumlah besar variabel dan melakukan inferensi probabilistik dengan variabel. Misalkan, sebuah hubungan probabilistik antara penyakit dan gejala. Metode Bayes memiliki beberapa tipe, diantaranya *Naive Bayes* merupakan *classifier* probabilistik sederhana berdasarkan teorema *Bayes*, menerapkan model probabilitas yang mendasari model fitur yang independen. *Tree Augmented Naive Bayes* (TAN) merupakan pohon perpanjangan dari *Naive Bayes*, dimana node kelas langsung menunjuk ke semua node atribut dan atribut node hanya dapat memiliki satu orang tua dari atribut node lain (selain node kelas) (Jiang et al, 2005). Teknik *Bayesian Network*, tipe ini merupakan tipe umum Bayesian yang berbasis probabilitas yang merepresentasikan suatu himpunan variabel dan *conditional interdependencies* melalui suatu DAG (*Directed Acyclic Graph*).

2.1 Representasi Teorema Bayes

Misalkan X merupakan himpunan atribut data dan Y merupakan variabel kelas. Jika variabel kelas memiliki hubungan dengan atribut, maka dapat dinyatakan X dan Y sebagai variabel acak dan hubungan peluang menggunakan $P(Y|X)$. Menaksir nilai peluang *posterior* secara akurat untuk setiap kombinasi label kelas yang mungkin dan nilai atribut adalah masalah sulit. Teorema *Bayes* bermanfaat karena menyediakan hubungan peluang *posterior* dari peluang *prior* $P(Y)$, peluang kelas bersyarat $P(X|Y)$ dan $P(X)$:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \quad \dots\dots (1)$$

Ketika membandingkan peluang *posterior* untuk nilai Y berbeda, istilah *dominator*, $P(X)$, selalu tetap, sehingga dapat diabaikan. Peluang *prior* $P(Y)$ dapat dengan mudah diestimasi dari *training set* dengan menghitung pecahan *training record* yang dimiliki tiap kelas.

2.2 Teknik Naïve Bayes

Naive bayes mengestimasi peluang kelas bersyarat dengan mengasumsikan bahwa atribut adalah independen secara bersyarat yang diberikan dengan label kelas y . Asumsi independen bersyarat dapat dinyatakan dalam bentuk berikut :

$$P(X|Y = y) = \prod_{i=1}^d P(X_i|Y = y) \quad \dots\dots\dots (2)$$

dengan tiap himpunan atribut $X = \{X_1, X_2, \dots, X_d\}$ terdiri dari d atribut.

Asumsi independen bersyarat, termasuk menghitung peluang bersyarat untuk setiap kombinasi X , hanya memerlukan mengestimasi peluang bersyarat untuk tiap X_i yang diberikan Y . pendekatan selanjutnya lebih praktis karena tidak mensyaratkan *training set* sangat besar untuk memperoleh estimasi peluang yang baik.

Untuk mengklasifikasi test *record*, *Naive Bayes* menghitung peluang posterior untuk tiap kelas Y :

$$P(Y|X) = \frac{P(Y) \prod_{i=1}^d P(X_i|Y)}{P(X)} \quad \dots\dots\dots (3)$$

$P(X)$ adalah tetap untuk seluruh Y , cukup untuk memilih kelas yang memaksimalkan istilah numerator, $P(Y) \prod_{i=1}^d P(X_i|Y)$.

2.3 Teknik Bayesian Network

Asumsi independen bersyarat digunakan pada *Naive Bayes* mungkin terlalu rapuh, khususnya untuk masalah identifikasi dengan atribut yang dihubungkan dengan sesuatu. Bagian ini mengembangkan pendekatan lebih fleksibel untuk memodelkan peluang kelas bersyarat $P(X|Y)$.

2.3.1 Representasi Model

Bayesian Network menyediakan representasi grafis dari hubungan peluang bersama dengan set variabel acak. Ada dua unsur kunci *Bayesian network* :

1. *Directed acyclic graph* (dag) mengencode hubungan dependen antar set variabel.
2. Tabel peluang mengasosiasikan tiap node ke node orangtua selanjutnya.

Properti *Bayesian Network* dinyatakan sebagai berikut:

Properti 1 (Independensi Bersyarat). *Node* pada *Bayesian Network* independen secara bersyarat dengan *non descendant*-nya, jika orangtuanya diketahui.

Disamping kondisi independen bersyarat yang dikenakan dengan topologi *network*, tiap *node* juga diasosiasikan dengan tabel peluang.

1. Jika *node* X tidak memiliki orangtua, maka tabel hanya berisi peluang *prior* $P(X)$.
2. jika *node* X hanya memiliki satu orangtua, Y , maka tabel berisi peluang bersyarat $P(X|Y)$.
3. jika *node* X memiliki banyak orangtua $\{Y_1, Y_2, \dots, Y_k\}$, maka tabel berisi peluang bersyarat $P(X|Y_1, Y_2, \dots, Y_k)$.

2.3.2 Pembuatan Model

Pembuatan model di dalam *Bayesian Network* melibatkan tiga langkah berikut.

1. Membuat struktur *network*.
2. Mengestimasi nilai peluang di dalam tabel yang dihubungkan dengan tiap *node*.
3. Topologi *network* dapat diperoleh dengan mengencode *knowledge* (pengetahuan) subyektif dari *expert domain*.

3. Studi Kasus: Identifikasi Gangguan Colon dengan Naive Bayes dan Bayesian Network

Dalam penelitian ini, teknik Naive Bayes dan Bayesian Network yang digunakan untuk menghitung probabilitas dari pixel citra terhadap colon. Berikut perhitungan peluang dari masing-masing kategori :

$$P(\text{colon}) = \frac{|\text{citra}|}{|\text{data}|} \dots\dots\dots (4)$$

dengan :

$P(\text{colon})$ = peluang setiap citra pada sekumpulan citra;

$|\text{citra}|$ = frekuensi citra training pada tiap kategori;

$|\text{data}|$ = jumlah citra training yang ada;

$$P(\text{pixel}_i | \text{colon}_j) = \frac{n_k + 1}{n + |\text{nilai}|} \dots\dots\dots (5)$$

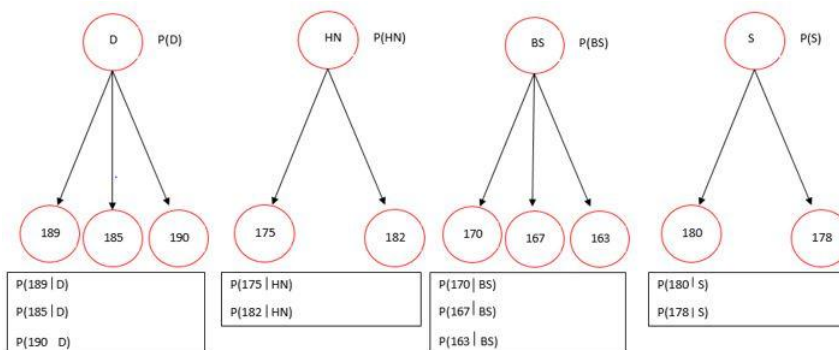
dengan :

$P(\text{pixel}_i | \text{colon}_j)$ = peluang kategori colon j terhadap $\text{pixel } i$

n_k = nilai kemunculan pixel_i pada tiap kategori colon j

n = jumlah seluruh pixel pada kategori colon j

$|\text{nilai}|$ = banyak pixel yang digunakan di data training.



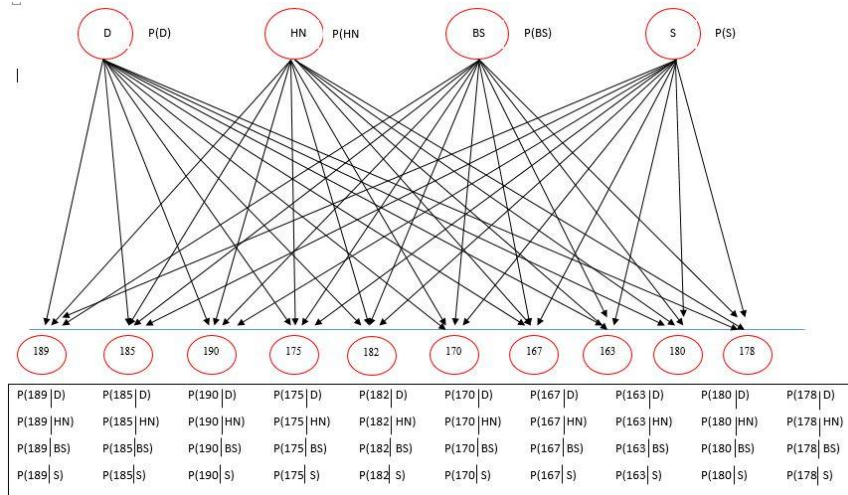
Gambar 1. Graf untuk Identifikasi Colon disease menggunakan Citra Iris Mata dengan Naive Bayes

Gambar 1 memperlihatkan graf Naive Bayes untuk memodelkan pasien *colon disease* atau masalah gangguan usus besar. Tiap variabel dalam diagram diasumsikan bernilai kontinu. *Node* orangtua untuk *colon Diverticulata* (D), Hampir Normal (HN), *Balloned Sigmoid* (BS) dan *Spasm* (S). *Node* anak untuk *colon disease* bersesuaian dengan nilai *pixel* 189, 185, 190, 175, 182, 170, 167, 163, 180, dan 178. Pada teknik Naive Bayes, hubungan antara node anak dengan orangtua menunjukkan pengaruh atau diyatakan dalam peluang bersyarat, misalnya node Diverticulata dipengaruhi oleh citra dengan nilai pixel 189, 185 dan 190 sedangkan nilai pixel lainnya diabaikan karena nilai peluangnya sangat kecil(kurang dari 0,0005).

Sedangkan, pada Gambar 2 memperlihatkan *directed acyclic graph* (dag) Bayesian Network. Misalnya node Diverticulata dipengaruhi oleh citra dengan nilai

pixel semua nilai pixel 189, 185, 190, 175, 182, 170, 167, 163, 180, dan 178 begitu pula untuk node yang lainnya. Perhitungan nilai peluang selengkapnya disajikan dalam tabel 1 dan 2 dengan menggunakan persamaan 4 dan 5.

Persamaan 4 dan 5 digunakan dalam proses training untuk menentukan nilai probabilitas dari citra training. Untuk proses identifikasi dengan citra uji, dengan teknik Naïve Bayes, yaitu dengan mengalikan nilai probabilitas semua kategori dengan probabilitas tiap pixel yang diambil dari masing-masing citra. Dari hasil perkalian tiap kategori, ditentukan nilai probabilitas terbesar. Nilai probabilitas terbesar adalah kategori citra uji.



Gambar 2. Struktur Directed Acyclic Graph (Dag) Bayesian Network untuk Mengidentifikasi Gangguan Usus Besar dengan Citra Iris Mata pada Pasien

Perangkat lunak yang dibangun membutuhkan pengetahuan awal untuk mengidentifikasi suatu himpunan data atau citra berdasarkan kategori kelas. Dari hasil pengumpulan data klinis, didapatkan himpunan citra uji. Citra uji telah mendapatkan pengetahuan awal mengenai kondisi *colon* dari pakar di salah satu rumah sakit di Palembang. Terdapat 94 citra uji yang telah diidentifikasi berdasarkan empat jenis kondisi *colon*, yaitu Hampir Normal sebanyak 29 citra, *Balloned Sigmoid* sebanyak 24 citra, *Diverticulata* sebanyak 21 citra dan *Spasm* sebanyak 20 citra.

Nilai peluang untuk setiap kategori *colon* menggunakan persamaan 4 dan disajikan di tabel 1.

Tabel 1
Peluang Tiap Kategori P(C)

No	Kategori Colon	P (Colon)
1	Diverticulata (D)	0.223404
2	Hampir Normal (HN)	0.308511
3	Balloned Sigmoid (BS)	0.255319
4	Spasm (S)	0.212766

Selain itu, dari 94 data citra iris mata tersebut menghasilkan nilai peluang terhadap nilai dari kategori P (Colon | Nilai) dengan menggunakan persamaan 5 seperti terlihat pada tabel 2. Terdapat 10 *pixel* dari semua kategori *colon* yang digunakan sebagai *pixel* acuan sesuai dengan peta iridologi untuk melakukan perhitungan *Bayesian Network* pada proses identifikasi, yaitu nilai *pixel* 189, 185, 190, 175, 182, 170, 167, 163, 180, dan 178.

Setelah mendapatkan pengetahuan mengenai *pixel* acuan dan frekuensi masing-masing dari perangkat lunak, maka perhitungan *Bayesian Network* bisa dilakukan. tabel 2. menyajikan hasil perhitungan untuk mengetahui nilai peluang bersyarat $P(\text{Nilai}|\text{Colon})$ setiap *pixel*.

Tabel 2
Nilai Peluang $P(\text{Nilai Pixel} | \text{Colon})$

Colon	189	185	190	175	182	170	167	163	180	178
D	0.368059	0.324524	0.278675	0.000523	0.000523	0.000523	0.000523	0.000523	0.000523	0.000523
HN	0.000527	0.000527	0.000527	0.426936	0.295529	0.000527	0.000527	0.000527	0.000527	0.000527
BS	0.000190	0.000190	0.000190	0.000190	0.000190	0.29352	0.363689	0.341449	0.000190	0.000190
S	0.000458	0.000458	0.000458	0.000458	0.000458	0.000467	0.000458	0.000458	0.418224	0.306542

Hasil identifikasi untuk setiap katagori dengan menggunakan teknik Naïve Bayes dan Bayesian Network disajikan dalam tabel 3 berikut.

Tabel 3
Hasil Identifikasi Citra Uji per Katagori dengan Teknik Naïve Bayes dan Bayesian Network

Kondisi	Jumlah Citra Uji	Identifikasi Naïve Bayes				Identifikasi Bayesian Network			
		Benar	Salah	Akurasi	Kesalahan	Benar	Salah	Akurasi	Kesalahan
Hampir Normal	29	19	10	63,83%	36,17%	21	8	70,21%	29,79%
Diverticulata	21	14	7			16	5		
Balloned Sigmoid	24	14	10			15	9		
Spasm	20	13	7			14	6		
	94	60	34			66	28		

Dengan menggunakan hasil pada tabel 3, diperoleh tingkat akurasi dan *kesalahan* dalam metode Bayesian Network untuk mengidentifikasi kondisi colonberbasis Iridologi disajikan. Taksiran dengan Bayesian Network memiliki kesalahan sebesar 29,79% dengan 66 citra yang benar dan 28 citra yang salah dari jumlah total sebesar 94 citra uji. Ada penurunan tingkat kesalahan (*error*) pada proses identifikasi antara penggunaan metode Bayesian Network bila dibandingkan dengan metode Naïve Bayes.

4. Kesimpulan dan Saran

Berdasarkan hasil penelitian di atas disimpulkan bahwa proses identifikasi citra iris mata untuk mengetahui kondisi usus besar (*colon*) dengan menggunakan teknik Bayesian Network menghasilkan keakuratan sebesar 70,21% dan lebih akurat bila dibandingkan dengan Naïve Bayes sebesar 63,83%.

Pada tahap berikutnya, model DAG yang ada dapat dikembangkan dengan menambah node variabel pengaruh yang bersifat diagnosis klinis seperti pola makan, menu diet, suhu tu.buh dan lain-lain serta variabel citra lain non iridologi seperti hasil foto rotgen

Daftar Pustaka

- Corwin, Elizabeth J. (2009). *Buku Saku Patofisiologi*. Buku Kedokteran EGC, Jakarta
- Duin, R., & Pekalska, E. (2007). The science of pattern recognition. Achievements and perspectives. *Challenges for Computational Intelligence*, 221-259.

- Erwin, Muhammad Fachrurrozi, Rossi Passarella dan Annisa Darmawahyuni. (2013). *Identifikasi Gangguan Usus Besar Berdasarkan Citra Iris Mata Menggunakan Metode Naïve Bayes*. Seminar Nasional Matematika, Sain dan Teknologi Tahun 2013, Jakarta
- Friedman N, Linial M, Nachman I, Pe'er D .(2000). *Using bayesian network to analyze expression data*. *Comput Biol* 7:601–620
- Gamez, J.A, J.L. Mateo, J.M. Puerto. (2010). *Learning Bayesian Networks by Hill Climbing: Efficient Methods Based on Progressive Restriction of The Neighborhood*, *Data Mining Knowledge Disc*,22:106:148
- Gat-Viks, I, A. Tanay, D. Rajjman and R. Shamir. (2006). *A Probabilistic Methodology for Integrating Knowledge and Experiments on Biological Network*, Vol. 13. No. 2, Pp 115-181
- Hanson, K. M. (1987). Bayesian and related methods in image reconstruction from incomplete data. *Image Recovery: Theory and Application*, 79-125
- Helman, P, R. Veroff, S.R. Atlas and C. Willman. (2004). *A Bayesian Network Classification Methodology for Gene Expression Data*, *Journal of Computational Biology*, Vol 11 No. 4, Pp 581-615
- Jensen, B.(1980), *Iridology Simplificated*, Bernard Jensen Enterproses CA 92025, California
- Jiang, Liangxiao et al. (2005). *Learning Tree Augmented Naive Bayes for Ranking*. University of Geosciences Wuhan, China and University of New Brunswick, Canada.
- Masek, L. (2003). Recognition of human iris patterns for biometric identification. *M. Thesis, The University of Western Australia*, 3.
- Pramono, M. (2006). Aplikasi Metode Backpropagation untuk Pengenalan Perubahan Abnormal Organ Pankreas melalui Iris Mata, SNATI 2006. Yogyakarta
- Rochmad, M.(2009). *Identifikasi Kerusakan Pankreas Melalui Iridology Menggunakan Metode Bayes Untuk Pengenalan Diabetes Mellitus*. Makalah Seminar Nasional Informatika 2009 (semnasIF 2009). Yogyakarta
- Rossi Passarella, Erwin, M. Fachrurrozi dan Sutarno. (2013). *Development of Iridology System Database for Colon Disorders Identification using Image Processing*. *Indian Journal of Bioinformatics and Biotechnology (IJBB)*, Vol 2(6):100-103
- WenChen X, Anantha G, Lin X.(2008). *Improving Bayesian network structure learning with mutual information-based node ordering in the k2 algorithm*. *IEEE Trans Knowl Data Eng* 20(5):628–640
- Witten IH, Frank E.(2005). *Data mining: practical machine learning tools and techniques*, 2nd edn. MorganKaufmann, San Francisco
- Wong ML, Leung KS.(2004). *An efficient data mining method for learning Bayesian networks using anevolutionary algorithm-based hybrid approach*. *IEEE Trans Evol Comput* 8(4):378–404