

REKAYASA PERANGKAT LUNAK DETEKSI DINI KECENDERUNGAN GANGGUAN KESEHATAN MASYARAKAT TERTINGGAL DAN PESISIR DENGAN *BAYESIAN NETWORK*

Ilham

Universitas Muhammadiyah Gresik
Jl. Sumatra 101 GKB Gresik, Indonesia
Email: Ilham_m.said@yahoo.com

Abstrak: Penelitian ini dilakukan untuk melihat kecenderungan penyakit yang ditimbulkan akibat pola hidup yang tidak sehat pada masyarakat tertinggal dan pesisir desa-desa sekitar kabupaten Gresik dan Tuban dengan menggunakan algoritma hybrid melalui konstruksi struktur Bayesian Network. Masalah yang akan diselesaikan pada penelitian ini adalah belum ada system yang mampu mendeteksi hubungan antara perilaku hidup tidak sehat dengan penyakit yang di timbulkan. Model struktur ini belum pernah diaplikasikan secara langsung di lapangan untuk mendeteksi suatu kejadian sebab akibat misalnya jika suatu perilaku itu tidak sehat maka akan timbul suatu penyakit. Pengaplikasian model ini perlu dilakukan dengan suatu studi lapangan untuk mengetahui dan membuktikan manfaat yang sebenarnya dari konsep konstruksi struktur hybrid bayesian network. Tujuan dari penelitian ini adalah menghasilkan model perangkat lunak yang mampu mendeteksi dini kecenderungan resiko penyakit masyarakat desa tertinggal dan pesisir yang memiliki pola hidup tidak sehat dalam bentuk konstruksi struktur dan menghasilkan probabilitas nilai dengan kecenderungan penyakit. Hasil perbandingan antara struktur asal dengan struktur hasil uji coba menunjukan tingkat kecocokan untuk data uji lengkap selisih 10% dengan struktur aslinya, dan tingkat kecocokan untuk data uji tidak lengkap sebesar lebih dari 20% tergantung jumlah missing value nya. Validitas bahwa orang merokok akan mengalami kecenderungan punya penyakit Tuberculosis, Bronkitis atau Lung cancer melalui uji system adalah 80% sampai 90%.

Kata kunci: Model; *Bayesian Network*; Pesisir; dan *Algoritma Hybrid*

Abstract: This study was conducted to see the trend of diseases caused by unhealthy lifestyles on disadvantaged communities and coastal villages around Gresik and Tuban using hybrid algorithms through the construction of the structure of Bayesian Network. The problem to be solved in this study is no system that can detect a relationship between unhealthy behavior that caused the disease. Model of this structure has never been applied directly in the field to detect a causal events for example, if a behavior is unhealthy will arise disease. Application of this model needs to be done with a field study to determine and prove the actual benefits of the concept of a hybrid construction of Bayesian network structure. The purpose of this research is to produce a software model capable of early detection of disease risk propensity underdeveloped rural and coastal communities who have unhealthy lifestyles in the form of construction of the structure and generates a probability value with a tendency disease. The comparison between the structure of the origin of the structure of the trial results indicate the level of suitability for complete test data difference of 10% to the original structure, and suitability for the test data is incomplete for more than 20% depending on the amount of his missing value. The validity of that smoke will have the tendency has tuberculosis disease, bronchitis or Lung Cancer through the test system is 80% to 90%.

Keywords: Model, Bayesian Network, coastal, Hybrid algorithm

PENDAHULUAN

Bayesian Network merupakan teknik atau metode yang digunakan untuk membangun model klasifikasi data dalam bentuk graf yang dapat merepresentasikan tingkat ketergantungan dari suatu variabel acak dengan menggunakan struktur graf [1]. Model struktur ini belum pernah diaplikasikan secara

langsung di lapangan untuk mendeteksi suatu kejadian sebab akibat misalnya jika suatu perilaku itu tidak sehat maka akan timbul suatu penyakit. Pengaplikasian model ini perlu dilakukan dengan suatu studi lapangan untuk mengetahui dan membuktikan manfaat yang sebenarnya dari konsep konstruksi struktur hybrid bayesian network.

Konsep *Bayesian Network* mempunyai konstruksi struktur yang cerdas dalam membentuk suatu model yang menggambarkan atau merepresentasikan hubungan sebab akibat (*causalitas*) untuk suatu kondisi atau perilaku. Sehingga sistem ini perlu dilakukan implementasi dan uji coba lapangan secara langsung. *Bayesian Network* memiliki dua algoritma yang dapat bekerja pada basisdata lengkap dan tidak lengkap. Algoritma ini terdiri dari algoritma CB (*Constrain Base*) dan algoritma BC (*Bound and Collapse*).

Algoritma CB (*Constrain Base*) adalah algoritma yang dapat bekerja pada basisdata lengkap, sedangkan algoritma BC (*Bound and Collapse*) adalah algoritma yang dapat bekerja pada basis data yang tidak lengkap (*missing value*). Masalah yang akan diselesaikan pada penelitian ini adalah membuktikan bagaimana algoritma *hybrid* ini mampu untuk mengelola data-data perilaku pola hidup tidak sehat masyarakat desa tertinggal dan pesisir agar dapat dideteksi dan diketahui lebih dini akan penyakit ditimbulkannya dengan membentuk konstruksi struktur dan menampilkan nilai-nilai probabilitas kecenderungan.

Tujuan dari penelitian ini adalah menghasilkan model perangkat lunak yang mampu mendeteksi dini kecenderungan resiko penyakit masyarakat desa tertinggal dan pesisir yang memiliki pola hidup tidak sehat dalam bentuk konstruksi struktur dan menghasilkan probabilitas nilai dengan cendrung penyakit yang ditimbulkan. Implementasi perangkat lunak algoritma *hybrid* ini juga dapat mengetahui dan mempelajari fungsi dan cara kerja algoritma secara benar sesuai dengan kebutuhan pengaplikasian sehingga dapat diterapkan dalam uji deteksi dini resiko penyakit atau gangguan terhadap kesehatan masyarakat kedepannya sehingga dapat diantisipasi lebih dini.

PEMBAHASAN

Metode Konstruksi Struktur *Bayesian Network*

Ada dua pendekatan yang digunakan untuk membangun *framework Bayesian Network* [CHE01] yaitu metode *search and scoring* dan metode *dependency analysis*. Masing-masing metoda ini memandang *Bayesian Network* dari sudut yang berbeda. Pada metoda *search and scoring*, *Bayesian Network* dipandang sebagai sebuah struktur yang merepresentasikan JPD dari variabel- variabel, sedangkan pada metoda *dependency analysis*, *Bayesian Network* dipandang sebagai sebuah struktur yang merepresentasikan sekumpulan kebebasan kondisional di antara *node-node*.

Berikut ini diberikan penjelasan lebih rinci mengenai kedua pendekatan tersebut [CHE97[b]:

1. Metode *Search and Scoring (Scoring Based)*. Metode ini melakukan pengembangan struktur dengan mencari sebuah struktur yang paling cocok dengan data. Dalam metode ini, permasalahan digambarkan sebagai pencarian struktur graf (model) yang baik dari data. Model dibangkitkan dari kombinasi *node* (variabel) yang ada. Pencarian tersebut dilakukan dengan *search method*. Setiap model dievaluasi oleh sebuah *scoring function*.
2. Metode *Dependency Analysis (Constraint Based)*. Pada metode ini, permasalahan digambarkan sebagai pencarian kebebasan (ketidakbergantungan) dari data yang kemudian dipergunakan untuk (meng-*infer*) sebuah struktur. Hubungan kebebasan tersebut data diukur dengan menggunakan salah satu dari beberapa jenis *Conditional Independence test* (CI test) dan hubungan tersebut digunakan sebagai batasan untuk membangun *framework Bayesian Network*.

Algoritma Hybrid *Bayesian Network*

Algoritma *Hybrid* merupakan pengembangan dari Algoritma CB dan BC. Algoritma CB dikembangkan oleh Singh & Voltorta adalah sebagai suatu algoritma yang mengkombinasikan metode *analisis dependency* dan metode *search and scoring* merupakan salah satu algoritma yang tidak terlalu bergantung pada CI test dan tidak membutuhkan *node ordering* tetapi akan membangkitkan sendiri *node ordering* [SIN95].

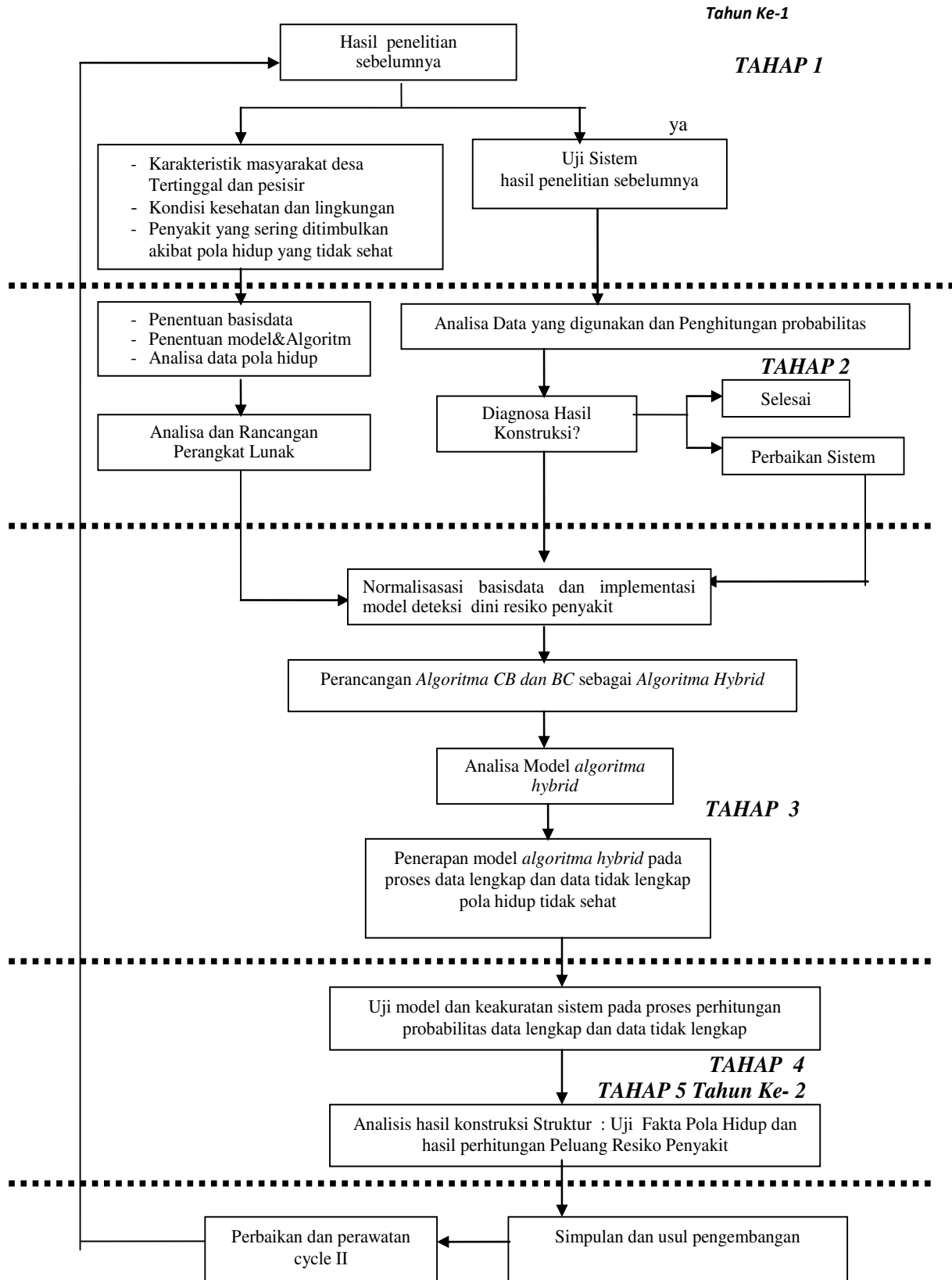
Algoritma CB perlu dikembangkan agar dapat bekerja pada data yang tidak lengkap untuk melakukan membangun struktur *Bayesian Network*. Pengembangan algoritma ini diberi nama dengan Algoritma *Hybrid* yaitu dengan mengkombinasikan Algoritma CB dengan Algoritma BC. Algoritma CB akan membangkitkan *node ordering* kemudian *node ordering* tersebut akan digunakan sebagai *input* oleh Algoritma BC agar membangun *framework Bayesian Network* dari data yang lengkap dan tidak lengkap dapat dilakukan. Di bawah ini adalah susunan lengkap dari Algoritma CB dan BC (algoritma *Hybrid*) yang akan dikembangkan dalam bentuk perangkat lunak sebagai media alat uji data yang akan diteliti

Adapun rancangan proses penelitian ini terlihat pada Gambar 1. Langkah-langkah detail dalam rekayasa algoritma *hybrid* dan implementasi yang dilakukan dalam penelitian ini terlihat pada Gambar 2.

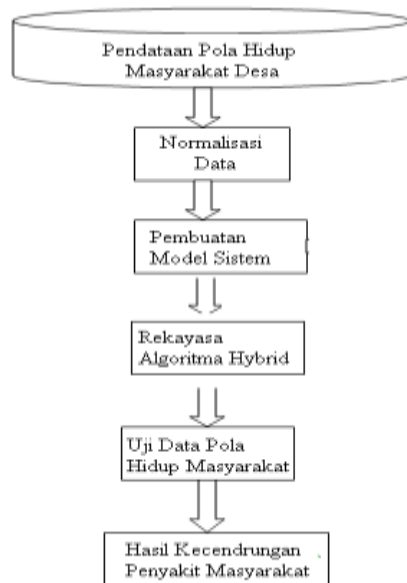
Penjelasan langkah algoritma *Hybrid* pada Gambar 2 adalah di mulai dari mengumpulkan data

melalui survey dan observasi pola hidup masyarakat desa tertinggal di kabupaten Gresik dan Tuban. Kemudian melakukan normalisasi terhadap data yang

di kumpulkan, penerapan model, rekayasa algoritma hybrid(CB dan BC), Uji data (implementasi system dilapangan) dan hasil resiko penyakit.

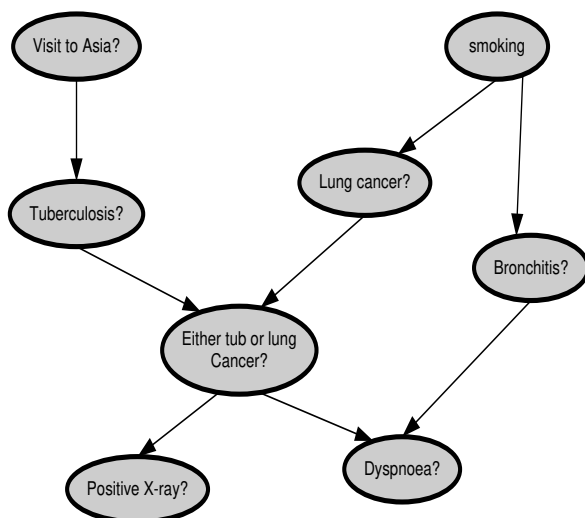


Gambar 1. Rancangan Proses Sistem



Gambar 2. Langkah Implementasi Sistem

Bayesian Network untuk kasus *Visit to Asia* yang akan dibuktikan tingkat kecocokan atau kesamaan dengan konstruksi struktur pada Gambar 3.



Gambar 3. Bayesian Network Kasus VisittoAsia

Hasil Uji Data pola Hidup Tidak Sehat

Setelah menyelesaikan tahap implementasi, maka selanjutnya tahapan yang akan dilakukan adalah pengujian terhadap perangkat lunak yang telah dihasilkan. Struktur yang dihasilkan untuk kasus uji pertama ini akan dibandingkan dengan struktur asal *Bayesian Network* Tabel 1.

Hasil perbandingan antara struktur asal dengan struktur hasil uji coba menunjukkan tingkat kecocokan untuk data uji lengkap selisih 10% dengan struktur aslinya, dan tingkat kecocokan untuk data uji tidak lengkap sebesar lebih dari 20% tergantung jumlah

missing value nya. Validitas bahwa orang merokok akan mengalami kecendrungan punya penyakit Tuberculosis, Bronkitis atau Lung cancer melalui uji system adalah 80% sampai 90%.

Tabel 1. Data Hasil Uji Kecocokkan Struktur

Data	Tingkat Kecocokkan Hasil Konstruksi Struktur			
	Konstruksi Struktur	Konstruksi Struktur Hasil	Persentase (%) Kecocokkan Konstruksi Struktur Asal	Persentase (%) Kecocokkan Konstruksi Struktur Hasil
Data	Visit-Asia -> Tuberculosis	Visit-Asia -> Tuberculosis	10	10
Smoking	Smoking -> Lung Cancer	Smoking -> Lung Cancer	10	10
Missing	Smoking -> Bronchitis	Tuberculosis -> TubOrLung	10	10
Value 50	TubOrLung -> X-Ray	Lung Cancer -> TubOrLung	10	10
	Bronchitis -> Dyspnoea	Smoking -> Bronchitis	10	10
	Tuberculosis -> TubOrLung	TubOrLung -> X-Ray	10	10
	Lung Cancer -> TubOrLung	Bronchitis -> Dyspnoea	10	10
	TubOrLung -> Dyspnoea		10	-
			80	70
Data	Visit-Asia -> Tuberculosis	Visit-Asia -> Tuberculosis	10	10
Smoking	Smoking -> Lung Cancer	Smoking -> Lung Cancer	10	10
Missing	Smoking -> Bronchitis	Lung Cancer -> TubOrLung	10	10
Value 50	TubOrLung -> X-Ray	X-Ray -> TubOrLung	10	-
	Bronchitis -> Dyspnoea	Tuberculosis -> TubOrLung	10	10
	Tuberculosis -> TubOrLung	Smoking -> Bronchitis	10	10
	Lung Cancer -> TubOrLung	Bronchitis -> Dyspnoea	10	10
	TubOrLung -> Dyspnoea		10	-
			80	60

KESIMPULAN

Dari hasil uji coba salah satu kasus pola hidup tidak sehat terdapat perubahan pada struktur yang dihasilkan. Perubahan struktur ini dapat terjadi akibat pengaruh data yang digunakan. Data yang digunakan dalam pengujian ini jumlahnya sangat terbatas dan tidak ada jaminan bahwa data tersebut sudah mencukupi sebagai data uji untuk menghasilkan struktur yang baik. Pada satu sisi hasil uji itu menunjukkan perubahan struktur tidak ditemukan walaupun ada data yang hilang sebanyak 50%, hal ini disebabkan karena sampel data yang dimiliki oleh data tersebut berjumlah lebih besar dari dua kasus yang lainnya. Sehingga meskipun data yang hilang adalah 50% namun sudah cukup membuktikan bahwa kecendrungan hasil struktur pola hidup dengan resiko penyakitnya mendekati kecocokan dengan nilai 80% dengan struktur aslinya.

DAFTAR PUSTAKA

- [1] Sitohang, B., & Saptawati, P. (2006). Improvement of CB & BC Algorithm (CB* Algorithm) for Learning Structure of Bayesian Networks as Classifier in Data Mining. Sekolah Teknik Elektro dan Informatika, *ITB.J.ICT*, Vol.1, No.1, 2007, 29-41
- [2] Sandhyaduhita, P.I. (2005). *Algoritma CB: Algoritma yang Dibangun dengan Dua Pendekatan untuk Konstruksi Struktur Bayesian Network dalam Data Mining*. Program Studi Teknik Informatika, STEI, ITB.

- [3] Selvia Lorena Br Ginting (2008). Studi algoritma CB Dalam Data Mining untuk Konstruksi Struktur Bayesian Network dari Basis Data Incomplete, Konferensi Nasional Sistem dan Informatika 2008; Bali, November 15, 2008, KNS,108-038
- [4] Cheng, J., Bell, D., & Liu, W. (1997). *An Algorithm for Bayesian Belief Networks Construction from Data*. Proceeding of Ai & STAT '97 (pp.83-90). Ft. Lauderdale, Florida.
- [5] Cheng, J., Bell, D., & Liu, W. (1998). *Learning Bayesian Networks from Data: An Efficient Approach Based on Information Theory*. Faculty of Informatics, University of Ulster, U.K.
- [6] Sebastiani, P., & Ramona, M. (1997). *Bayesian Inference with Missing Data Using Bound and Collapse*. Report KMi-TR-58, Knowledge Media Institute, The Open University.
- [7] Singh, M., & Valtorta, M. (1995). *Construction of Bayesian Network Structures from Data: A Brief Survey and an Efficient Algorithm*. Dept. of Computer Science, University of South Carolina, Columbia, USA.
- [8] Simanjuntak, H. (2006). *Pengembangan Algoritma CB untuk Konstruksi Struktur Bayesian Network dari Data Tidak Lengkap*. Program Studi Teknik Informatika, STEI, ITB.
- [9] Maharani, H. (2005). *Konstruksi Struktur Bayesian Network dalam Data Mining untuk Basis Data Incomplete dengan Metode Bound and Collapse*. Program Studi Teknik Informatika, STEI, ITB.
- [10] Neapolitan, R.E. (2004). *Learning Bayesian Networks*. USA: Pearson Pentice Hall.
- [11] Cheng, Jie, et al. (2001). *Learning Bayesian Network from data: An Information-Theory Based Approach*. Department of Computing Science, University of Alberta., Faculty of Informatics, University of Ulster, Toronto, Canada.