

**KLASIFIKASI DENGAN *ITERATIVE DICHOTOMISER 3* PADA  
DAERAH RAWAN PENYAKIT DEMAM BERDARAH *DENGUE*  
DI JAWA TENGAH**



oleh

AYUNING CAHYA NILAMSARI

M0116008

**SKRIPSI**

ditulis dan diajukan untuk memenuhi sebagian persyaratan  
memeroleh gelar Sarjana Matematika

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS SEBELAS MARET SURAKARTA**

2020

*commit to user*

KLASIFIKASI DENGAN *ITERATIVE DICHOTOMISER 3* PADA DAERAH  
RAWAN PENYAKIT DEMAM BERDARAH *DENGUE* DI JAWA TENGAH

SKRIPSI

AYUNING CAHYA NILAMSARI

NIM. M0116008

dibimbing oleh

Pembimbing I

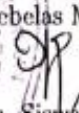
Pembimbing II

Dr. Dewi Retno Sari S., S.Si., M.Kom.  
NIP. 19700720 199702 2 001

Prof. Tri Atmojo K. M.Sc., Ph.D  
NIP. 19630826 198803 1 002

dipertahankan di hadapan Dewan Penguji  
dan dinyatakan telah memenuhi syarat  
pada hari Selasa tanggal 21 Juli 2020

Jabatan	Nama dan NIP	Tanda Tangan	Tanggal
Ketua	Titin Sri Martini, S.Si., M.Kom. NIP. 19750120 200812 2 001		07/09/2020
Sekretaris	Bowo Winarno, S.Si., M.Kom. NIP. 19810430 200812 1 001		07/09/2020
Penguji I	Dr. Dewi Retno Sari S., S.Si., M.Kom. NIP. 19700720 199702 2 001		07/09/2020
Penguji II	Prof. Tri Atmojo K. M.Sc., Ph.D NIP. 19630826 198803 1 002		07/09/2020

Disahkan  
di Surakarta pada tanggal .. 08 SEP 2020  
Kepala Program Studi Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Sebelas Maret Surakarta  
  
Dr. Drs. Siswanto, M.Si.  
NIP. 19670813 199203 1 002

## PERNYATAAN

Dengan ini saya menyatakan bahwa skripsi saya yang berjudul Klasifikasi dengan *Iterative Dichotomiser 3* pada Daerah Rawan Penyakit Demam Berdarah *Dengue* di Jawa Tengah belum pernah diajukan untuk memperoleh gelar kesarjaan pada suatu perguruan tinggi, dan sepanjang pengetahuan saya juga belum pernah ditulis atau dipublikasikan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar rujukan.



Surakarta, 21 Juli 2020

Ayuning Cahya Nilamsari

## RINGKASAN

Demam berdarah *dengue* (DBD) adalah penyakit yang disebabkan oleh virus *dengue* dan DBD tidak menular melalui kontak manusia dengan manusia tetapi hanya dapat ditularkan melalui nyamuk yaitu nyamuk *Aedes aegypti*. DBD banyak ditemukan di daerah tropis dan subtropis salah satunya Asia Tenggara. Beberapa faktor yang menyebabkan persebaran penyakit DBD adalah kepadatan penduduk, curah hujan dan ketinggian suatu wilayah. Tingginya kepadatan penduduk memengaruhi persebaran penyakit DBD karena semakin tinggi tingkat kepadatan penduduk maka persebaran penyakit akan semakin cepat. Kemudian tingginya curah hujan yaitu semakin tinggi curah hujan maka semakin banyak habitat nyamuk dan berikutnya adalah ketinggian suatu wilayah yang memengaruhi daya hidup nyamuk.

Daerah rawan penyakit DBD dapat diketahui melalui pengolahan *dataset* dengan metode yang ada pada data *mining*. Dari faktor-faktor tersebut dapat dilakukan suatu klasifikasi dengan pohon keputusan berdasarkan tingkat kerawannya. Algoritme dalam klasifikasi yang dapat digunakan adalah algoritme ID3. Pemilihan algoritme ini dikarenakan algoritme ID3 dapat digunakan dalam memproses data *mining* jika data atribut yang terdapat pada *dataset* berupa data kategorikal. Algoritme ID3 memerlukan nilai *entropy* dan *information gain* dengan *information gain* digunakan sebagai metode pemilihan atributnya. ID3 membangun pohon keputusan secara *top-down* (dari atas ke bawah), dimulai dengan atribut mana yang pertama kali harus dicek dan diletakkan sebagai simpul akar. Simpul akar dipilih dari nilai *information gain* terbesar. Atribut yang digunakan dalam penelitian ini adalah faktor-faktor yang menyebabkan persebaran penyakit DBD, yaitu kepadatan penduduk, curah hujan dan ketinggian wilayah. Data-data yang digunakan merupakan data dari setiap kabupaten/kota di Jawa Tengah tahun 2018, yaitu data kepadatan penduduk diperoleh dari Badan Pusat Statistik (BPS), data curah hujan diperoleh dari data Badan Meteorologi Klimatologi dan Geofisika (BMKG) dan data ketinggian wilayah tiap kabupaten/kota diperoleh dari data Badan Pusat Statistik (BPS).

Berdasarkan hasil klasifikasi menggunakan algoritme ID3 pada daerah rawan DBD, diperoleh 6 aturan klasifikasi. Terdapat 6 kabupaten/kota dengan tingkat kerawanan yaitu rawan dan 27 kabupaten/kota dengan tingkat kerawanan yaitu tidak rawan. Dilakukan pengujian akurasi dengan *confusion matrix* dan tingkat keakurasian yang diperoleh sebesar 81.8%.

## SUMMARY

Dengue Hemorrhagic Fever (DHF) is a disease caused by the dengue virus and DHF can't be transmitted through human contact. It only can be transmitted through mosquitoes such as *Aedes Aegypti*. DHF is found in many tropical and subtropical regions, one of them being Southeast Asia. Some factors that cause the spread of DHF are population density, rainfall and height of an area. High population density affects the spread of dengue disease because if the level of the population density is higher then the rate the disease will spread increases. Due to the higher rainfall, more mosquito habitats are created and also the height of an area can affect the mosquito population due to the height of the area affects the viability of mosquitoes.

DHF prone areas can be identified by processing with methods of data mining. From these factors can be determined by a classification within a decision tree based on the level of vulnerability. The classification algorithm that can be used is the ID3 algorithm. The reason to choose this algorithm is because the ID3 algorithm can be used in processing data mining, if the attribute data which is contained in the dataset is categorical data. ID3 algorithm requires the value of entropy and information gain. Information gain to be used as a method of selecting its attributes. ID3 constructs a decision tree by employing a top-down and starting with which attribute must first be checked and placed as a root node. The root node is chosen from the largest information gain value. The attributes that are used in this research and which are the factors that cause the spread of dengue fever are population density, rainfall and height of an area. The data that's used is data from each district or city in Central Java in 2018, namely population density data obtained from Badan Pusat Statistik (BPS), rainfall data obtained from Badan Meteorologi dan Geofisika (BMKG) and the height of an area for each regency or city is obtained from Badan Pusat Statistik (BPS).

Based on the results of the classification using the ID3 algorithm in DHF prone areas, 6 classification rules are obtained. There are 6 districts or cities with a vulnerability level that is vulnerable and 27 districts or cities with a vulnerability level that is not prone. Accuracy testing was performed with confusion matrix and the level of accuracy obtained was 81.8%.

## MOTO

Semua hal terjadi karena suatu alasan dan rencana Allah akan selalu indah.



*commit to user*



## PERSEMBAHAN

Karya ini kupersembahkan untuk  
bapak (Sudarmadi), ibu (Ismu Widiastuti) dan adik saya (Yumna Amalia  
Putri).



*commit to user*

## PRAKATA

Bismillahirrahmanirrahim, Segala puji dan syukur penulis panjatkan kepada Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini. Dengan selesainya skripsi ini penulis mengucapkan terimakasih kepada

1. Dr. Dewi Retno Sari S., M.Kom. sebagai Pembimbing I atas bimbingan mengenai materi, arahan, motivasi, dan saran.
2. Prof. Drs. Tri Atmojo Kusmayadi M.Sc., Ph.D. sebagai Pembimbing II atas bimbingan, arahan dan saran.
3. Tim riset grup *Mathematical Softcomputing* 2019/2020 khususnya Fajhria Budi Ramadhanti, Panji Bimo Nugroho Setio dan Afif Zakiy Abdullah atas diskusi materi dan saran.

Semoga skripsi ini bermanfaat.

Surakarta, 21 Juli 2020

Ayuning Cahya Nilamsari

*commit to user*



## DAFTAR ISI

HALAMAN JUDUL . . . . .	i
PENGESAHAN . . . . .	ii
PERNYATAAN . . . . .	iii
RINGKASAN . . . . .	iv
SUMMARY . . . . .	v
MOTO . . . . .	vi
PERSEMBAHAN . . . . .	vii
PRAKATA . . . . .	viii
DAFTAR ISI . . . . .	ix
DAFTAR TABEL . . . . .	x
DAFTAR GAMBAR . . . . .	xi
 <b>I PENDAHULUAN</b>	 <b>1</b>
1.1 Latar Belakang Masalah . . . . .	1
1.2 Perumusan Masalah . . . . .	3
1.3 Tujuan Penelitian . . . . .	3
1.4 Manfaat Penelitian . . . . .	3
 <b>II LANDASAN TEORI</b>	 <b>4</b>
2.1 Tinjauan Pustaka . . . . .	4
2.2 Teori Penunjang . . . . .	4
2.2.1 Pohon Keputusan . . . . .	5
2.2.2 <i>Iterative Dichotomiser 3</i> <i>mit.to user</i> . . . . .	6
2.2.3 <i>Confusion Matrix</i> . . . . .	7

2.2.4 Demam Berdarah <i>Dengue</i> (DBD) . . . . .	7
2.3 Kerangka Pemikiran . . . . .	8
<b>III METODOLOGI PENELITIAN</b>	<b>9</b>
3.1 Data Penelitian . . . . .	9
3.2 Langkah Penelitian . . . . .	10
<b>IV HASIL DAN PEMBAHASAN</b>	<b>11</b>
4.1 Deskripsi Atribut Data . . . . .	11
4.2 Pembentukan Pohon Keputusan . . . . .	12
4.2.1 Menentukan Simpul Akar . . . . .	12
4.2.2 Menentukan Simpul Cabang . . . . .	13
4.3 Pengujian Akurasi Pohon Keputusan ID3 . . . . .	15
4.4 Interpretasi Pohon Keputusan ID3 . . . . .	16
<b>V PENUTUP</b>	<b>18</b>
5.1 Kesimpulan . . . . .	18
5.2 Saran . . . . .	18
<b>DAFTAR RUJUKAN</b>	<b>19</b>
<b>LAMPIRAN</b>	<b>20</b>
Lampiran 1 . . . . .	21
Lampiran 2 . . . . .	23
Lampiran 3 . . . . .	26
Lampiran 4 . . . . .	27

## DAFTAR TABEL

2.1	Tabel <i>Confusion Matrix</i> . . . . .	7
3.1	Atribut penelitian . . . . .	9
4.1	Rincian nilai atribut . . . . .	11
4.2	Hasil Perhitungan <i>Entropy</i> dan <i>Information Gain</i> . . . . .	12
4.3	Hasil Perhitungan Simpul Cabang untuk 0-500 Mdpl . . . . .	13
4.4	Hasil Perhitungan <i>Information Gain</i> . . . . .	14
4.5	Hasil Perhitungan Simpul Cabang untuk 500-1000 Mdpl . . . . .	14
4.6	Hasil Perhitungan <i>Information Gain</i> . . . . .	15
4.7	Hasil Perhitungan <i>Confusion Matrix</i> . . . . .	16
5.1	Data Status Kerawanan Tiap Kabupaten/kota . . . . .	21
5.2	Hasil Perhitungan Simpul Cabang untuk 0-500 Mdpl dengan AN .	23
5.3	Hasil Perhitungan Simpul Cabang untuk 0-500 Mdpl dengan N . .	23
5.4	Hasil Perhitungan Simpul Cabang untuk 0-500 Mdpl dengan BN .	24
5.5	Hasil Perhitungan Simpul Cabang untuk 500-1000 Mdpl dengan N	24
5.6	Hasil Perhitungan Simpul Cabang untuk 500-1000 Mdpl dengan BN	25
5.7	Status Kerawanan Tiap Kabupaten/kota di Jawa Tengah . . . . .	27

*commit to user*

## DAFTAR GAMBAR

2.1	Struktur pohon keputusan . . . . .	5
3.1	<i>Flowchart</i> langkah penelitian . . . . .	10
5.1	Pohon Keputusan ID3 . . . . .	26

*commit to user*