



**APLIKASI *CLIQUE* PADA  
NONPARAMETRIC MAXIMUM LIKELIHOOD ESTIMATOR  
UNTUK DATA TERSENSOR BIVARIAT**

(Studi Kasus : Pasien Kanker Paru-Paru)

**SKRIPSI**

Oleh

**Eka Ayu Ningsih**

**NIM 151810101054**

**JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS JEMBER  
2019**



**APLIKASI *CLIQUE* PADA  
NONPARAMETRIC MAXIMUM LIKELIHOOD ESTIMATOR  
UNTUK DATA TERSENSOR BIVARIAT**

(Studi Kasus : Pasien Kanker Paru-Paru)

**SKRIPSI**

diajukan guna melengkapi tugas akhir dan memenuhi salah satu syarat  
untuk menyelesaikan studi pada Program Studi Matematika (S-1)  
dan mencapai gelar Sarjana Sains

Oleh

**Eka Ayu Ningsih**

**NIM 151810101054**

**JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS JEMBER  
2019**

## PERSEMBAHAN

Puji syukur kehadiran Tuhan Yang Maha Esa, skripsi ini saya persembahkan untuk:

1. Kedua orang tua saya, Bapak Nurhamim dan Ibu Kastiyem yang telah mendoakan dan mendukung serta memberikan kasih sayang;
2. Keluarga besar yang selalu memberi doa dan dukungan;
3. Seluruh dosen dan guru sejak sekolah dasar sampai perguruan tinggi, yang telah memberikan ilmu dan membimbing dengan penuh kesabaran;
4. Almamater Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember, SMA Negeri Yosowilangun, SMP Negeri 2 Parigi Moutong dan SD Negeri 2 Parigi Moutong;
5. Teman-teman angkatan 2015 (Sigmafifteen) yang selalu membantu, memberi dukungan dan kenangan;

**MOTO**

Jangan patah semangat dengan apapun yang terjadi, jika kita  
menyerah, habislah sudah.  
(Top Ittipat)<sup>\*)</sup>



---

<sup>\*)</sup>Sugmakanan, Songyos (Producer). 2011. *The Billionaire*. GMM Tai Hub. Thailand, 124 mins.

**PERNYATAAN**

Saya yang bertanda tangan di bawah ini :

Nama : Eka Ayu Ningsih

NIM : 151810101054

menyatakan dengan sesungguhnya bahwa skripsi yang berjudul “Aplikasi *Clique* pada *Nonparametric Maximum Likelihood Estimator* untuk data tersensor bivariat” adalah benar-benar hasil karya sendiri, kecuali kutipan yang sudah saya sebutkan sumbernya, belum pernah diajukan pada institusi manapun, dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa ada tekanan dan paksaan dari pihak manapun serta bersedia mendapat sanksi akademik jika ternyata di kemudian hari pernyataan ini tidak benar.

Jember, Januari 2019  
Yang menyatakan,

Eka Ayu Ningsih  
NIM. 151810101054

**SKRIPSI**

**APLIKASI *CLIQUE* PADA  
NONPARAMETRIC MAXIMUM LIKELIHOOD ESTIMATOR  
UNTUK DATA TERSENSOR BIVARIAT**

Oleh

Eka Ayu Ningsih

NIM 151810101054

Pembimbing

Dosen Pembimbing Utama : Dr. Mohamat Fatekurohman, S.Si., M.Si

Dosen Pembimbing Anggota : Dian Anggraeni, S.Si., M.Si

**PENGESAHAN**

Skripsi berjudul “Aplikasi *Clique* pada *Nonparametric Maximum Likelihood Estimator* untuk Data Tersensor Bivariat” telah diuji dan disahkan pada:

hari, tanggal :

tempat : Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas  
Jember

Tim Penguji,

Ketua,

Anggota I,

Dr. Mohamat Fatekurohman, S.Si., M.Si

Dian Anggraeni, S.Si., M.Si

NIP. 196906061998031001

NIP. 198202162006042002

Anggota II,

Anggota III,

Prof. Drs. I Made Tirta, M.Sc., Ph.D.

Dr. Firdaus Ubaidillah, S.Si., M.Si.

NIP. 195912201985031002

NIP. 197006061998031003

Mengesahkan

Dekan,

Drs. Sujito, Ph.D.

NIP. 196102041987111001



## RINGKASAN

**Aplikasi *Clique* pada *Nonparametric Maximum Likelihood Estimator* untuk Data Tersensor Bivariat**; Eka Ayu Ningsih; 151810101054; 2019; 56 halaman; Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam.

Penyakit kanker paru disebabkan karena munculnya tumor ganas di organ paru maupun di organ tubuh lainnya. Pada umumnya penyakit kanker paru mulai dapat diketahui setelah munculnya gejala yang dialami pasien. Waktu awal pasien terjangkit kanker paru hingga dinyatakan sembuh disebut sebagai masa inkubasi. Lamanya masa inkubasi pasien dalam menjalani rawat inap dapat dianalisa menggunakan analisis survival.

Data kanker paru yang digunakan pada penelitian ini menggunakan tipe tersensor interval bivariat yang direpresentasikan dalam bentuk himpunan data persegi yang selanjutnya disederhanakan dalam bentuk *clique*. Data tersensor dapat direpresentasikan sebagai graf irisan. Adanya data tersensor menyebabkan jumlah data yang diperoleh menjadi relatif sedikit sehingga distribusi masa inkubasi tidak mengikuti distribusi apapun. Akibatnya distribusi masa inkubasi akan menggunakan metode nonparametrik. Perhitungan NPMLE (*Nonparametric Maximum Likelihood Estimator*) pada data tersensor interval bivariat terdiri dari dua tahap. Tahap pertama yakni penetapan daerah irisan dan tahap kedua yakni mencari *maximum likelihood*.

Terdapat beberapa metode perhitungan NPMLE yang telah ada sebelumnya. Metode tersebut yakni: (1) Metode dengan algoritma *Expectation Maximization* (EM), (2) Metode Newton Raphson, (3) Metode Pohon Reduksi. Adapun Penentuan *maximal clique* sebelumnya telah digunakan untuk mengelompokkan (*clustering*) graf melalui *complement matrix*. Algoritma Genetika sebelumnya juga pernah digunakan untuk menyelesaikan sistem persamaan linier.

Pada penelitian ini dilakukan penentuan *maximal cliques* dengan menggunakan algoritma Bron-Kerbosch dan memaksimumkan fungsi *likelihood* dengan algoritma Genetika sehingga diperoleh estimator NPMLE. Penyelesaian perhitungan algoritma Bron-Kerbosch dan algoritma Genetika dilakukan dengan



menggunakan program MATLAB R2015b. Hasil dari *maximal cliques* yang diperoleh yakni;  $M_1 = \{R_1, R_2, R_5, R_7, R_8, R_{13}, R_{16}, R_{23}, R_{25}\}$ ,  $M_2 = \{R_1, R_2, R_5, R_7, R_8, R_{16}, R_{24}, R_{25}\}$ ,  $M_3 = \{R_1, R_7, R_8, R_{13}, R_{22}, R_{23}\}$ ,  $M_4 = \{R_2, R_5, R_8, R_{16}, R_{17}, R_{24}, R_{25}\}$ ,  $M_5 = \{R_3, R_5, R_{17}, R_{25}\}$ ,  $M_6 = \{R_3, R_{19}\}$ ,  $M_7 = \{R_4, R_6, R_{14}, R_{18}\}$ ,  $M_8 = \{R_4, R_{15}, R_{21}\}$ ,  $M_9 = \{R_6, R_9, R_{12}, R_{14}, R_{18}\}$ ,  $M_{10} = \{R_{10}, R_{14}\}$ ,  $M_{11} = \{R_{11}, R_{19}\}$  dan  $M_{12} = \{R_{20}, R_{21}\}$ .  $R_i$  menyatakan daerah persegi, dimana satu daerah persegi mewakili satu individu. Berdasarkan hasil tersebut, tidak terdapat daerah persegi yang saling asing. Paling sedikit terdapat dua daerah persegi yang saling beririsan dan paling banyak terdapat 9 daerah persegi yang saling beririsan. Daerah persegi yang saling beririsan merepresentasikan interval masa inkubasi yang sama. Adapun nilai estimator yang dihasilkan yaitu;  $p_1 = 0,0625$ ,  $p_2 = 0,0626$ ,  $p_3 = 0,1249$ ,  $p_4 = 0,1251$ ,  $p_5 = 0,0626$ ,  $p_6 = 0,0625$ ,  $p_7 = 0,0781$ ,  $p_8 = 0,1249$ ,  $p_9 = 0,0625$ ,  $p_{10} = 0,0469$ ,  $p_{11} = 0,0624$ ,  $p_{12} = 0,1250$  dan nilai fungsi *likelihood*  $L = 1,1636 \times 10^{-20}$ . Berdasarkan hasil tersebut, diperoleh nilai-nilai estimator yang beragam dan tak nol, sehingga dapat dilakukan perhitungan masa inkubasi yang sama untuk semua daerah irisan maksimal. Nilai dari setiap estimator akan dikalikan dengan 168 (7 hari x 24 jam) sehingga outpunya berupa masa inkubasi yang sama dalam satuan jam.

## PRAKATA

Puji syukur kepada Allah SWT atas segala rahmat dan karuniaNya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Aplikasi *Clique* pada *Nonparametric Maximum Likelihood Estimator* untuk Data Tersensor Bivariat”. Skripsi ini disusun untuk memenuhi salah satu syarat untuk menyelesaikan pendidikan strata satu (S1) pada Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Penyusunan dan penyelesaian skripsi ini tidak lepas dari bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan ucapan terima kasih kepada:

1. Dr. Mohamat Fatekurohman, S.Si., M.Si dan Dian Anggaraeni, S.Si., M.Si, selaku Dosen Pembimbing yang telah memberikan arahan dan dorongan serta telah meluangkan waktu, pikiran, tenaga dan perhatian selama penulisan skripsi ini;
2. Prof. Drs. I Made Tirta, M.Sc., Ph.D dan Dr. Firdaus Ubaidillah, S.Si., M.Si, selaku Dosen Penguji yang telah memberikan kritik dan saran sebagai penyempurnaan skripsi ini;
3. Seluruh dosen Fakultas MIPA Universitas Jember yang telah memberi banyak ilmu pengetahuan selama masa perkuliahan; staff dan karyawan atas segala bantuan yang telah diberikan selama menjadi mahasiswa Fakultas MIPA Universitas Jember;
4. Kedua orang tua tercinta yang telah mendoakan, membantu baik secara moril maupun materiil, serta telah membesarkan dan mengajarkan makna kehidupan serta seluruh keluarga besar yang selalu menjadi penyemangat;
5. Rekan seperjuangan Rere, Retno, Febi dan Tutik yang telah membantu dalam hal pengurusan berkas data penelitian;
6. Sahabatku Ellenda dan Amalia serta Sigmafifteen'15 yang selalu memberikan dukungan dan semangat;
7. Semua pihak yang membantu kelancaran skripsi ini.

DAFTAR ISI

	Halaman
HALAMAN JUDUL .....	i
HALAMAN PERSEMBAHAN .....	ii
HALAMAN MOTTO .....	iii
HALAMAN PERNYATAAN .....	iv
HALAMAN PEMBIMBINGAN .....	v
HALAMAN PENGESAHAN .....	vi
RINGKASAN.....	vii
PRAKATA .....	ix
DAFTAR ISI .....	x
DAFTAR GAMBAR.....	xii
DAFTAR TABEL .....	xiii
<b>BAB 1. PENDAHULUAN.....</b>	<b>1</b>
<b>1.1 Latar Belakang .....</b>	<b>1</b>
<b>1.2 Rumusan Masalah .....</b>	<b>3</b>
<b>1.3 Tujuan .....</b>	<b>3</b>
<b>1.4 Manfaat .....</b>	<b>4</b>
<b>BAB 2. TINJAUAN PUSTAKA .....</b>	<b>5</b>
<b>2.1 Analisis <i>Survival</i>.....</b>	<b>5</b>
<b>2.2 Data Tersensor.....</b>	<b>7</b>
<b>2.3 Teori Graf .....</b>	<b>8</b>
<b>2.4 Representasi Graf pada Data Tersensor Bivariat .....</b>	<b>11</b>
<b>2.5 Maximum Likelihood Estimator (MLE) .....</b>	<b>14</b>
<b>2.6 MLE pada Data Tersensor Bivariat .....</b>	<b>14</b>
<b>2.7 Algoritma Bron-Kerbosch.....</b>	<b>15</b>
<b>2.8 Algoritma Genetika .....</b>	<b>16</b>
<b>2.9 Penyakit Kanker Paru.....</b>	<b>17</b>
<b>BAB 3. METODE PENELITIAN.....</b>	<b>19</b>
<b>3.1 Data Penelitian.....</b>	<b>19</b>

3.2 Langkah-Langkah Penelitian.....	19
<b>BAB 4. HASIL DAN PEMBAHASAN.....</b>	<b>22</b>
4.1 Data Bivariat (Data Pasien Kanker Paru).....	22
4.2 Graf Irisan .....	23
4.3 Matriks Adjasensi.....	24
4.4 <i>Maximal Clique</i> .....	25
4.5 Fungsi <i>Likelihood</i> .....	26
<b>BAB 5. PENUTUP .....</b>	<b>30</b>
5.1 Kesimpulan.....	30
5.2 Saran.....	30
<b>DAFTAR PUSTAKA.....</b>	<b>31</b>
<b>LAMPIRAN .....</b>	<b>33</b>

**DAFTAR GAMBAR**

	Halaman
2.1 Grafik fungsi survival.....	5
2.2 Representasi graf $G$ dan subgraf dari graf $G$ .....	10
2.3 Graf lengkap dengan 6 titik .....	10
2.4 Graf irisan .....	10
2.5 Representasi graf interval .....	10
2.6 Matriks adjasensi dan matriks insidensi dari suatu graf.....	11
2.7 Himpunan data yang tersensor interval bivariat .....	13
2.8 Graf irisan .....	13
2.9 Graf clique .....	13
3.1 Skema Penelitian.....	21
4.1 Himpunan persegi panjang data pengamatan .....	22
4.2 Graf irisan dari himpunan persegi panjang pengamatan .....	23
4.3 Matriks adjasensi representasi graf irisan .....	24
4.4 Graf clique dari 25 pasien kanker paru .....	26
4.5 Matriks clique dari 25 pasien kanker paru .....	26

**DAFTAR TABEL**

	Halaman
2.1 Matriks <i>intersections</i> .....	13





## BAB 1. PENDAHULUAN

### 1.1 Latar Belakang

Penyakit kanker paru-paru merupakan jenis penyakit yang mematikan. Gejala awal dari penyakit kanker paru-paru yakni batuk dalam jangka waktu lama disertai darah, sesak napas hingga penurunan berat badan. Pasien yang didiagnosa menderita penyakit tersebut biasanya akan menjalani serangkaian proses pengobatan. Pengamatan yang bisa dilakukan terhadap pasien yakni mengetahui waktu awal pasien didiagnosa, lama pasien dapat bertahan hidup, lamanya masa inkubasi pasien dan waktu pasien meninggal dunia. Masa inkubasi tercatat dari waktu awal pasien terdiagnosa penyakit kanker paru hingga dinyatakan sembuh. Lamanya masa inkubasi pasien dalam menjalani rawat inap dapat dianalisa menggunakan analisis data survival.

Analisis survival merupakan analisis data yang berkaitan dengan waktu mulai dari *time origin* yang telah ditentukan sampai terjadinya kejadian tertentu atau *end-point*. *Time origin* sering berkaitan dengan awal perekrutan suatu individu dalam suatu pengamatan. Bila *end-point* merupakan kematian seorang pasien maka data yang dihasilkan berupa waktu ketahanan hidup (Collett, 2003). Data akan menjadi tersensor jika diketahui waktu tahan hidupnya meskipun tidak secara pasti (Kleinbaum & Klein, 2005). Data tersensor dapat direpresentasikan kedalam bentuk graf yang digambarkan pada koordinat 2 dimensi.

Suatu graf  $G=(V,E)$  terdiri dari sebuah himpunan titik  $V$  dan himpunan sisi  $E$  sedemikian hingga setiap titik ujung dari sisi  $E$  akan termuat dalam  $V$  (Rosen, 2000). Pada kasus bivariat, visualisasi data akan lebih mudah jika direpresentasikan ke dalam bentuk daerah persegi panjang. Irisan yang terjadi pada sekumpulan daerah persegi dapat dinyatakan dalam bentuk graf irisan. Pada graf irisan, simpulnya menyatakan daerah persegi panjang individu. Sisi akan menghubungkan antar simpul apabila terdapat irisan antar daerah persegi. Keterhubungan antar simpul graf irisan dinyatakan dalam bentuk matriks adjasensi. Matriks adjasensi digunakan untuk mempermudah dalam pencarian *maximal intersections* yang merupakan representasi nyata dari *maximal clique*.



Secara teori graf, *clique* adalah subgraf lengkap dimana masing-masing titiknya dihubungkan oleh sisi (Skiena, 2008). Dalam kaitannya dengan data tersensor, suatu *clique* merupakan himpunan indeks yang berkorespondensi dengan individu yang memiliki daerah yang saling beririsan (Gentleman & Vandal, 2002). Pada graf *clique*, simpulnya merepresentasikan *maximal intersections*, sehingga graf *clique* merupakan penyederhanaan dari graf irisan. Representasi graf *clique* dinyatakan ke dalam matriks *clique*. Matriks *clique* berperan penting dalam perhitungan *Nonparametric Maximum Likelihood Estimator* (NPMLE).

Salah satu alternatif untuk mengestimasi parameter dalam model matematika yaitu *maximum likelihood estimation* (Kleinbaum & Klein, 2005). Terdapat 2 jenis metode untuk mengestimasi parameter yaitu parametrik dan nonparametrik. Metode parametrik membutuhkan beberapa asumsi yang harus dipenuhi sedangkan metode nonparametrik tidak membutuhkan asumsi-asumsi tertentu. Adanya data tersensor mengakibatkan data yang diperoleh relatif sedikit sehingga tidak dapat ditentukan secara pasti model distribusi data tersebut. Estimasi nonparametrik akan lebih cocok jika dilakukan pada data yang tersensor.

Kholidati (2017) melakukan perhitungan NPMLE menggunakan algoritma EM (*Expectation Maximization*). Penelitian dilakukan pada data univariat pasien hepatitis A dan hepatitis B di Rumah Sakit Kaliwates. Hasil yang diperoleh yakni hanya data hepatitis B yang bisa diperoleh fungsi *likelihood*-nya. Pada tahun yang sama Acita membandingkan perhitungan NPMLE menggunakan metode Newton Raphson dan metode pohon reduksi melalui representasi dari matriks insidensi. Penelitian dilakukan pada data tersensor interval bivariat pada pasien HIV/AIDS di RSUD Balung. Hasil yang diperoleh yakni perhitungan NPMLE dengan menggunakan pohon reduksi lebih baik dibandingkan metode Newton Raphson karena pohon reduksi memiliki nilai fungsi *likelihood* yang lebih besar.

Pada tahun yang sama pula, Antoro melakukan penelitian tentang penerapan algoritma Bron-Kerbosch dalam pengelompokan (*clustering*) graf menggunakan *complement matrix*. Penelitian tersebut menggunakan data graf yang merepresentasikan rute penerbangan domestik dari salah satu maskapai di

Indonesia. Hasil yang diperoleh yaitu terdapat 41 *maximal cliques* dan 14 *cluster*. Setahun sebelumnya, Azhari menerapkan algoritma Genetika pada penyelesaian sistem persamaan linier. Masalah yang diteliti yaitu enam buah sistem persamaan linier yang diubah ke dalam bentuk matriks. Algoritma Genetika tersebut menghasilkan solusi pendekatan pada beberapa sistem persamaan linier yang memiliki karakteristik matriks tertentu dengan nilai *fitness* yang mendekati 1.

Penelitian ini akan menghitung NPMLE dari data bivariat dengan mengaplikasikan graf *clique*. Matriks adjasensi digunakan untuk membangkitkan matriks *clique* yang berperan sebagai matriks indikator pada perhitungan NPMLE. Pembentukan matriks *clique* menggunakan bantuan algoritma Bron-Kerbosch melalui bantuan program MATLAB R2015b. Selain itu, program tersebut telah menyediakan script untuk memaksimumkan suatu fungsi dengan menggunakan algoritma Genetika, sehingga pada penelitian ini perhitungan NPMLE menggunakan algoritma Genetika.

## 1.2 Rumusan Masalah

Adapun rumusan masalah dalam penelitian ini adalah :

- a. Bagaimana memperoleh *maximal cliques* menggunakan algoritma Bron-Kerbosch berdasarkan data pasien kanker paru-paru ?
- b. Bagaimana menghitung NPMLE menggunakan algoritma Genetika berdasarkan data pasien kanker paru-paru ?

## 1.3 Tujuan

Adapun tujuan dalam penulisan penelitian ini adalah :

- a. Diperolehnya *maximal cliques* menggunakan algoritma Bron-Kerbosch berdasarkan data pasien kanker paru-paru.
- b. Menjelaskan perhitungan NPMLE menggunakan algoritma Genetika berdasarkan data pasien kanker paru-paru.

#### **1.4 Manfaat**

Manfaat yang diperoleh dari penelitian ini bagi pembaca akan menambah pengetahuan baru tentang keberadaan algoritma yang dapat digunakan untuk menyelesaikan permasalahan analisis survival melalui pendekatan graf. Sedangkan manfaat bagi peneliti selanjutnya, penelitian ini dapat dijadikan sebagai referensi dan acuan untuk melakukan penelitian selanjutnya.

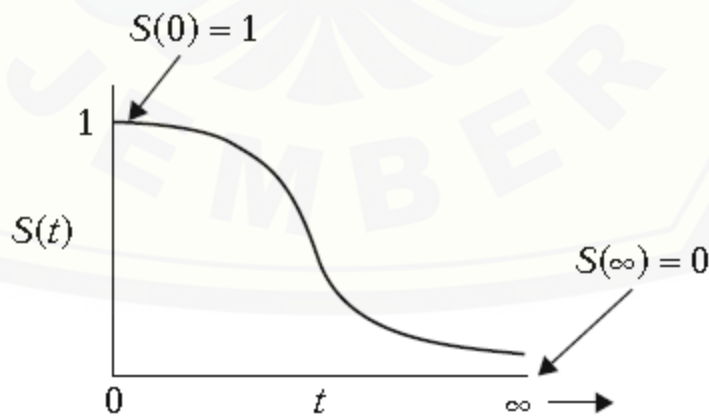


## BAB.2 TINJAUAN PUSTAKA

### 2.1 Analisis Survival

Analisis survival adalah ungkapan yang digunakan untuk menjelaskan analisis data yang berkaitan dengan waktu mulai dari *time origin* yang telah ditentukan sampai terjadinya kejadian tertentu atau *end-point*. *Time origin* sering berhubungan dengan awal perekrutan suatu individu dalam suatu pengamatan yang bersifat percobaan. Bila *end-point* merupakan kematian seorang pasien maka data yang dihasilkan berupa waktu survival (Collett, 2003). Dalam analisis survival, variabel yang menjadi perhatian yakni waktu sebagai waktu survival. Hal itu akan memberikan informasi waktu bahwa seseorang telah bertahan selama beberapa periode pengamatan. Kejadian yang diamati biasanya merupakan kejadian buruk yang dialami oleh individu, misalnya kematian (Kleinbaum & Klein, 2005).

Fungsi survival yang dinotasikan dengan  $S(t)$  memberikan peluang bahwa seseorang bertahan lebih lama daripada beberapa waktu  $t$  yang telah ditentukan.  $S(t)$  menghasilkan peluang bahwa variabel acak  $T$  melampaui waktu  $t$  yang telah ditetapkan sebelumnya. Secara teori, selang interval  $t$  berawal dari 0 sampai tak hingga, fungsi survival dapat direpresentasikan melalui grafik berupa kurva *smooth*. Seperti yang diilustrasikan oleh Gambar 2.1, dimana  $t$  berada di sumbu- $x$ .



Gambar 2.1 Grafik fungsi survival

Semua fungsi survival memiliki karakteristik sebagai berikut :

- Fungsi tersebut tidak naik tetapi tidak konstan; yakni bergerak ke bawah seiring peningkatan variabel  $t$ .
- Pada saat  $t=0$ ,  $S(0)=I$ ; yakni pada saat awal pengamatan, karena belum ada individu yang mengalami kejadian sehingga peluang waktu bertahan hidup pada saat 0 nilainya 1.
- Pada saat waktunya  $t=\infty$ ,  $S(\infty)=0$  ; yakni jika periode pengamatan meningkat tanpa batas, hasilnya tidak ada individu yang akan bertahan hidup, sehingga kurva survivalnya menuju angka 0 (Kleinbaum & Klein, 2005).

Misalkan  $T \geq 0$  menyatakan suatu variabel acak non negatif dan menunjukkan waktu survival setiap subyek. Menurut Lee & Wang (2003), distribusi dari  $T$  dapat dinyatakan dalam beberapa cara yaitu :

#### a. Fungsi Kepadatan Peluang

Fungsi kepadatan peluang atau PDF (*Probability Density Function*) dari fungsi survival adalah peluang suatu individu yang meninggal atau mengalami kejadian sesaat dalam interval waktu  $t$  sampai  $t+\Delta t$ . Fungsi kepadatan peluang  $f(t)$  dinyatakan sebagai berikut ;

$$f(t) = \lim_{\Delta t \rightarrow 0} \left[ \frac{P(t < T < (t+\Delta t))}{\Delta t} \right] = \lim_{\Delta t \rightarrow 0} \left[ \frac{F(t+\Delta t) - F(t)}{\Delta t} \right] \quad (2.1)$$

dengan  $F(t)$  merupakan fungsi distribusi kumulatif kontinu dari  $T$ , didefinisikan sebagai peluang suatu individu mengalami kejadian kurang dari atau sama dengan waktu  $t$  yaitu,

$$F(t) = P(T \leq t) = \int_0^t f(x) dx \quad (2.2)$$

Berdasarkan Persamaan (2.2) diperoleh:

$$f(t) = \frac{d(F(t))}{dt} = F'(t) \quad (2.3)$$

#### b. Fungsi Survival

Fungsi survival  $S(t)$  didefinisikan sebagai peluang suatu individu dapat bertahan hidup dengan waktu survival sampai dengan waktu  $t$  ( $t \geq 0$ ) yaitu :

$$S(t) = P(T \geq t) \quad (2.4)$$



Sesuai dengan definisi fungsi distribusi kumulatif  $F(t)$  dari  $T$ , fungsi survival dapat dinyatakan dengan,

$$\begin{aligned} S(t) &= 1 - P(T \leq t) \\ &= 1 - F(t) \end{aligned}$$

Fungsi survival juga dapat dinyatakan dalam bentuk fungsi kepadatan peluang yaitu,

$$S(t) = P(T \geq t) = \int_t^{\infty} f(x) dx$$

Diperoleh hubungan antara fungsi kepadatan peluang, fungsi distribusi kumulatif dari  $T$  dan fungsi survival yaitu ;

$$\begin{aligned} F(t) &= 1 - S(t) \\ &= \frac{d(F(t))}{dt} = \frac{d(1 - S(t))}{dt} \\ &= F'(t) = -S'(t) \\ &= f(t) = -S'(t) \end{aligned}$$

diperoleh,

$$f(t) = F'(t) = -S'(t) \quad (2.5)$$

## 2.2 Data Tersensor

Penyensoran dapat terjadi ketika kita memiliki beberapa informasi tentang waktu ketahanan individu, tetapi tidak diketahui secara pasti waktu ketahanannya. Alasan secara umum mengapa penyensoran dapat terjadi karena : individu tidak mengalami kejadian sebelum penelitian berakhir, individu hilang selama periode penelitian dan individu mengundurkan diri dari penelitian karena kematian (Klein & Kleinbaum, 2005). Terdapat 3 jenis penyensoran yaitu :

1. Sensor kanan (*right censored*) : terjadi ketika masa pengamatan telah berakhir namun individu belum mengalami kejadian yang menjadi perhatian, sedangkan waktu awal individu dapat teramati secara penuh. Misalkan penelitian dilakukan terhadap pasien leukimia. Kejadian yang menjadi perhatian yaitu kematian. Sebulan kemudian pasien berhenti dari pengobatan sebelum masa studi berakhir. Hal ini mengakibatkan pasien hilang dari pengamatan sehingga waktu survivalnya tersensor kanan (Collett, 2003).

2. Sensor kiri (*left censored*) : terjadi ketika waktu survival individu yang sebenarnya kurang dari waktu pengamatan yang telah ditentukan. Misalkan pengamatan pada pasien penderita kanker. Tiga bulan setelah menjalani operasi, pasien dinyatakan sembuh dari penyakit kanker. Selama waktu tersebut, beberapa pasien mungkin dinyatakan kembali menderita penyakit kanker (kambuh). Sehingga waktu kejadian yang sebenarnya kurang dari 3 bulan dan waktu survival pasien ini dikatakan tersensor kiri (Collett, 2003).
3. Sensor interval (*interval censored*) : terjadi ketika kejadian yang menjadi perhatian telah terjadi diantara selang waktu  $a$  dan  $b$ . Misalnya penelitian dilakukan pada pasien berusia 50 tahun yang telah mengidap *diabetic retinopathy*. Berdasarkan rekaman medis pasien mengindikasikan bahwa pasien tersebut pada usia 45 tahun, belum menderita *retinopathy*. Sehingga waktu diagnosa penyakit berada diantara 45 dan 50 tahun (Lee & Wang, 2003).

### 2.3 Teori Graf

Suatu graf  $G$  memuat kedua himpunan  $V$  dan  $E$  yang berhingga, dimana  $V$  merupakan himpunan titik atau simpul sedangkan  $E$  merupakan himpunan sisi (Chartrand & Zhang, 2012). Adapun menurut Voloshin (2009), suatu graf adalah himpunan titik dan garis yang menghubungkan beberapa pasangan titik tertentu. Pada teori graf, titik disebut sebagai simpul sedangkan garis disebut sebagai sisi.

Sebuah graf dengan himpunan sisi yang kosong disebut sebagai *null graph*. Diberikan suatu graf  $G=(V,E)$ , subgraf dari  $G$  adalah graf  $G'=(V',E')$  sedemikian hingga  $V' \subseteq V$  dan  $E' \subseteq E$  seperti yang diilustrasikan pada Gambar 2.2. Suatu graf yang setiap pasangan titik yang berbeda dihubungkan dengan sisi disebut sebagai graf lengkap seperti pada Gambar 2.3 (Rahman, 2017). Suatu *clique* dalam sebuah graf  $G$  adalah subgraf lengkap dari graf  $G$  (Chartrand & Zhang, 2012).

Suatu *clique* dari graf dikatakan *maximal* jika *clique* tersebut tidak termuat pada *clique* lainnya (Bapat, 2010). Penyajian *clique* dapat dinyatakan ke dalam bentuk matriks *clique*. Matriks *clique* yang dinotasikan  $C=[c_{ij}]$  merupakan matriks yang baris dan kolomnya merepresentasikan titik dari graf yang

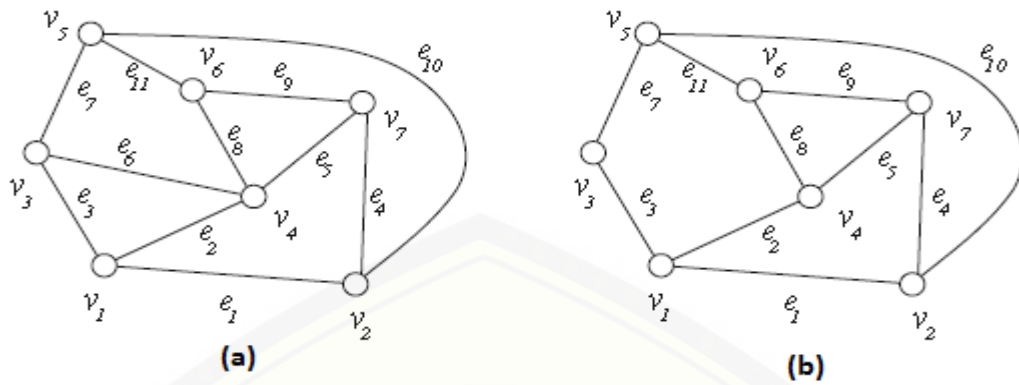


bersesuaian dan jumlah *clique*. Entri dari *clique matrix* akan bernilai 1 jika *clique* ke- $i$  terhubung dengan titik ke- $j$  dan 0 untuk lainnya (Baruah & Baruah, 2012).

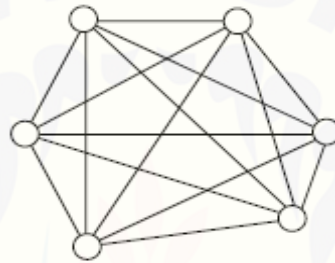
Misalkan  $G$  merupakan graf dimana  $V(G) = \{v_1, v_2, \dots, v_n\}$  dan  $E(G) = \{e_1, e_2, \dots, e_m\}$ . Matriks adjasensi dari  $G$  adalah matriks  $A$  yang berordo  $n \times n$  dengan entrinya bernilai 1 apabila  $v_i v_j \in E(G)$  dan 0 untuk lainnya. Adapun matriks insidensi dari  $G$  adalah matriks  $B$  yang berordo  $n \times m$  dengan entrinya bernilai 1 apabila  $v_i$  terhubung dengan  $e_j$  dan 0 untuk lainnya. Matriks adjasensi dan matriks insidensi dari suatu graf disajikan pada Gambar 2.6 (Chartrand & Zhang, 2012).

Suatu graf  $G = (V, E)$  dikatakan sebagai graf irisan dari suatu keluarga himpunan berhingga  $F$  yang tak kosong, jika terdapat korespondensi satu-satu antara  $V$  dan  $F$  sehingga dua himpunan anggota  $F$  memiliki irisan yang tak kosong, yakni jika dan hanya jika korespondensi titik dalam  $V$  saling terhubung. Untuk masing-masing titik  $v_i$  dari  $G$ , membentuk suatu himpunan  $S_i$  yang memuat sisi yang terhubung dengan  $v_i$ , setiap dua himpunan memiliki irisan yang tak kosong jika dan hanya jika titik yang bersesuaian dihubungkan melalui suatu sisi. Contoh graf irisan; misalkan  $S_+ = \{1, 2, 3, \dots\}$ ,  $S_- = \{-1, -2, -3, \dots\}$ ,  $S_p = \{2, 3, 5, 7, 11, \dots\}$ ,  $S_f = \{1, 1, 2, 3, 5, 8, \dots\}$ ,  $S_e = \{0, \pm 2, \pm 4, \dots\}$ ,  $S_o = \{\pm 1, \pm 3, \dots\}$ , masing-masing merupakan himpunan bilangan bulat positif, bilangan bulat negatif, bilangan prima, bilangan *fibonacci*, bilangan bulat genap dan bilangan bulat ganjil. Representasi graf irisan dinyatakan pada Gambar 2.4 (Pal, 2013).

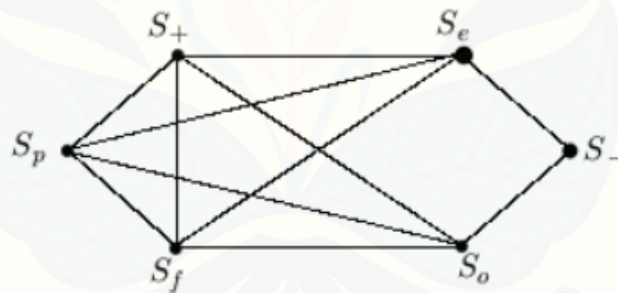
Berdasarkan konsep graf irisan, kemudian didefinisikan suatu graf yakni graf *clique*. Graf *clique*  $C(G)$  dari suatu graf  $G$  adalah graf irisan dari keluarga himpunan semua *clique* yang ada dalam  $G$ . Terdapat pula graf interval yang tergolong ke dalam graf irisan. Suatu graf  $G = (V, E)$  dikatakan sebagai graf interval jika himpunan titik  $V$  dapat dipetakan (korespondensi) satu-satu pada sebuah himpunan interval yang tersusun pada garis bilangan real sehingga dua titik dalam  $G$  dihubungkan jika dan hanya jika terdapat irisan yang tak kosong pada interval yang bersesuaian. Contoh dari graf interval disajikan pada Gambar 2.5 (Pal, 2013).



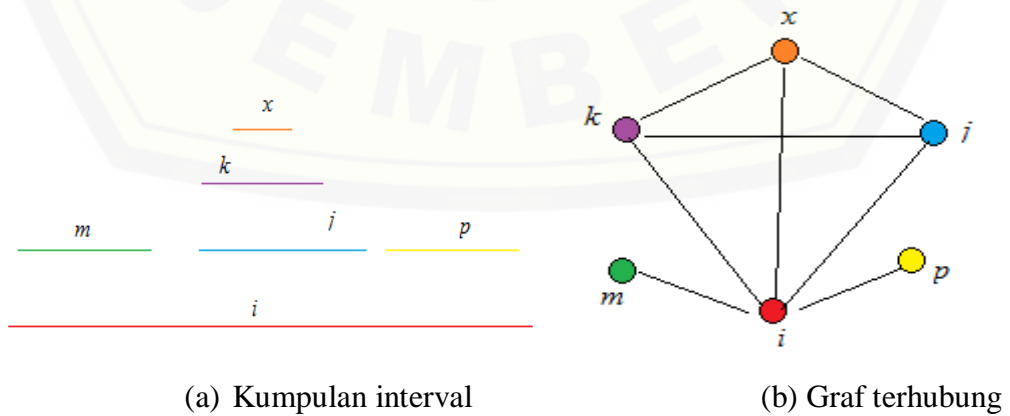
Gambar 2.2 Representasi graf  $G$  dan subgraf dari graf  $G$



Gambar 2.3 Graf lengkap dengan 6 titik



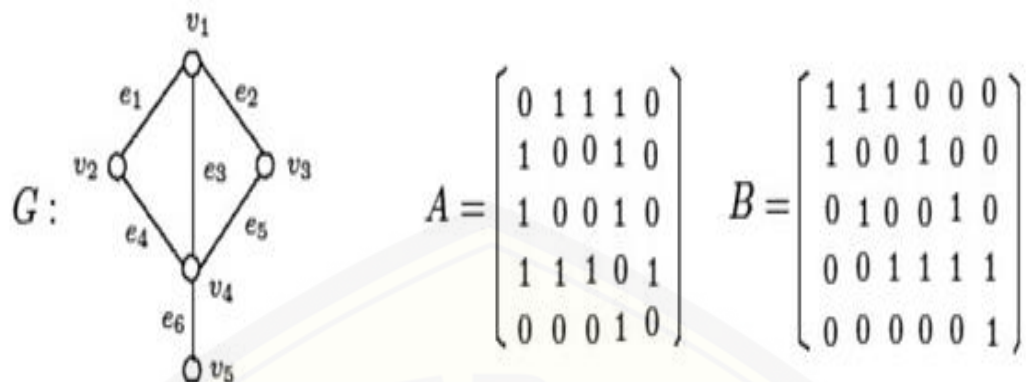
Gambar 2.4 Graf irisan



(a) Kumpulan interval

(b) Graf terhubung

Gambar 2.5 Representasi graf interval



Gambar 2.6 Matriks adjasensi dan matriks insidensi

#### 2.4 Representasi Graf pada Data Tersensor Bivariat

Menurut Gentleman & Vandal (2002), suatu graf pada data tersensor dapat dibangkitkan dengan cara membentuk sebuah titik pada masing-masing persegi pengamatan kemudian titik-titik yang perseginya saling beririsan akan dihubungkan dengan sisi. Waktu kejadian bivariat diasumsikan berdistribusi bebas menurut distribusi  $F(t_1, t_2)$ .  $F(t_1, t_2)$  adalah fungsi distribusi waktu kejadian data atau fungsi survival yang berkorespondensi yang menjadi perhatian. Menurut Lee (1983) dalam Maathuis (2003),  $R = \{R_1, \dots, R_n\}$  menotasikan suatu himpunan persegi panjang. Masing-masing  $R_i$  direpresentasikan sebagai  $(x_{1i}, x_{2i}, y_{1i}, y_{2i})$ . Titik  $(x_{1i}, y_{1i})$  adalah sudut kiri bawah dan  $(x_{2i}, y_{2i})$  adalah sudut kanan atas. Koordinat  $(x_{1i}, x_{2i})$  disebut sebagai interval  $x$  dan  $(y_{1i}, y_{2i})$  disebut sebagai interval  $y$  dari  $R_i$ . Data yang telah berbentuk persegi panjang pengamatan dapat dibentuk *clique*-nya.

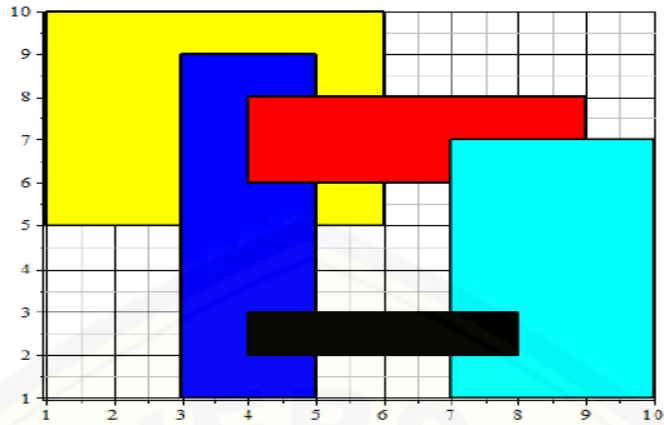
Sebuah *clique* adalah *subset*  $C$  dari  $V$  sehingga setiap anggota  $C$  dihubungkan melalui sebuah sisi pada masing-masing anggota  $C$ . Representasi nyata dari *clique* yakni daerah irisan anggotanya. Sebuah graf irisan (*intersections*) dibangun dari  $R$  dengan aturan berikut. Masing – masing pengamatan  $R_j$  berkorespondensi dengan satu titik  $j$ . Dua buah titik  $j$  dan  $k$  dihubungkan dengan sebuah sisi jika terjadi irisan yang bersesuaian antara  $R_j$  dan  $R_k$  dengan  $j \neq k$ . Ketika  $R_j$  merupakan daerah yang tersusun pada garis lurus, dalam hal ini graf irisan disebut sebagai

graf interval (Vandal & Gentleman, 2001). Representasi dari graf irisan dapat dinyatakan ke dalam bentuk matriks adjasensi.

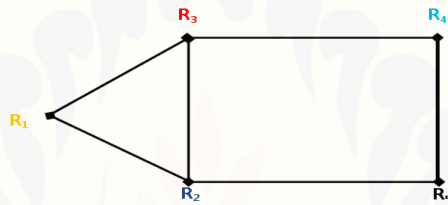
Pada setiap himpunan data yang berhingga terdapat  $m$  *maximal cliques*.  $H_i$  menotasikan representasi nyata dari daerah yang mendefinisikan anggota  $M_j$  yaitu,  $H_i \stackrel{\text{def}}{=} \bigcap_{j \in M_i} R_j$ .  $H$  merupakan himpunan daerah persegi dari graf irisan yang berkoresponden dengan *maximal cliques*, dinotasikan dengan  $H_i = \{H_1, H_2, \dots, H_m\}$ , sedemikian hingga terdapat korespondensi satu-satu antara  $M_i \leftrightarrow H_i$ ,  $i=1, 2, 3, \dots, m$  (Gentleman & Vandal, 2001). *Maximal intersections* yang diperoleh dari himpunan data persegi panjang  $R_1, R_2, \dots, R_n$  disebut sebagai *inner rectangle*. Representasi nyata dari *maximal clique* adalah *maximal intersections* dan sebarang *maximal intersections* merupakan representasi dari *maximal clique*. Setiap pengamatan memuat paling sedikit satu *maximal clique*. Struktur dari graf irisan yang diperoleh dari *maximum clique* dapat diringkas ke dalam bentuk *clique matrix*.

Menurut Fatekurohman, *et al* (2014) misalkan diberikan suatu data dalam bentuk persegi panjang seperti pada Gambar 2.7. Pada kasus ini persegi panjang tersebut direpresentasikan sebagai  $(x_{ij}, y_{ij}, x_{uj}, y_{uj})$ , dimana koordinat  $(x_{ij}, y_{ij})$  adalah sudut kiri atas dan  $(x_{uj}, y_{uj})$  adalah sudut kanan bawah. Masing-masing persegi panjang tersebut yakni  $R_1=(1,10,6,5)$ ,  $R_2 = (3,9,5,1)$ ,  $R_3=(4,8,9,6)$ ,  $R_4=(7,7,10,1)$  dan  $R_5=(4,3,8,2)$ . Terdapat 5 pengamatan yang direpresentasikan dengan persegi panjang  $R_1$ ,  $R_2$ ,  $R_3$ ,  $R_4$  dan  $R_5$ . Seperti yang terlihat masing-masing pada Gambar 2.7 dan 2.8, ada 4 *cliques* yang seluruhnya merupakan *maximal* yaitu  $M_1=\{R_1, R_2, R_3\}$ ,  $M_2=\{R_2, R_5\}$ ,  $M_3=\{R_3, R_4\}$  dan  $M_4=\{R_4, R_5\}$ . Graf *clique* dari *maximal clique* persegi panjang tersebut ditampilkan pada Gambar 2.9. Seluruhnya merupakan *maximal* yaitu  $M_1=\{R_1, R_2, R_3\}$ ,  $M_2=\{R_2, R_5\}$ ,  $M_3=\{R_3, R_4\}$  dan  $M_4=\{R_4, R_5\}$ .

Sebuah graf dari Gambar 2.7 dapat dibangun dengan membuat sebuah titik pada masing-masing persegi panjang dan titik-titik yang memiliki interseksi yang tak kosong dihubungkan dengan sisi seperti pada Gambar 2.8. Himpunan ketetanggaan dari 4 titik lain, masing-masing memuat 2 titik dan titik-titik ini tidak dihubungkan dengan sisi (Gentleman & Vandal, 2002). Matriks *intersections* dari data tersebut ditampilkan pada Tabel 2.1.



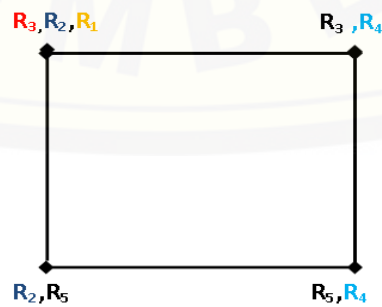
Gambar 2.7 Himpunan data yang tersensor interval bivariat



Gambar 2.8 Graf irisan

Tabel 2.1 Matriks *intersections*

	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$
$H_1$	1	1	1	0	0
$H_2$	0	1	0	0	1
$H_3$	0	0	1	1	0
$H_4$	0	0	0	1	1



Gambar 2.9 Graf *clique*



### 2.5 Maximum Likelihood Estimator (MLE)

Misalkan  $X_1, X_2, \dots, X_n$  sampel acak dengan fungsi peluang  $f(x_i, \theta)$  dengan  $i = 1, 2, \dots, n$ . Apabila  $L$  yaitu fungsi peluang bersama dari  $X_1, X_2, \dots, X_n$  dipandang sebagai fungsi dari  $\theta$  dan  $x_1, x_2, \dots, x_n$  sebagai bilangan tertentu maka  $L(\theta) = \prod_{i=1}^n f(x_i, \theta)$  disebut sebagai fungsi *likelihood*.

Misalkan  $X_1, X_2, \dots, X_n$  sampel acak dengan fungsi peluang  $f(x_i, \theta)$  dan fungsi *likelihood*  $L(\theta)$ . Setiap nilai  $w = h(x_1, x_2, \dots, x_n)$  yang memaksimumkan  $L(\theta)$  yakni  $L(w) \geq L(\theta)$  dinamakan *Maximum Likelihood Estimator* (MLE). Terkadang akan lebih mudah untuk tidak memaksimumkan  $L(\theta)$  melainkan  $\ln L(\theta)$ . Hal ini dapat dilakukan karena  $L(\theta)$  dan  $\ln L(\theta)$  mencapai maksimum pada  $\theta$  yang sama (Bain & Engelhardt, 1992).

### 2.6 MLE pada Data Tersensor Bivariat

Menurut Bogaerts dan Lessaffre (2004), perhitungan NPMLM meliputi 2 tahapan. Tahapan tersebut yaitu menentukan daerah persegi panjang dan menentukan *maximum likelihood*. Misalkan  $R = \{R_1, R_2, \dots, R_n\}$  dinotasikan sebagai himpunan persegi panjang pengamatan, maka fungsi *likelihood* dan *log likelihood* dapat dituliskan dalam bentuk :

$$L(F) = \prod_{i=1}^n P_F(R_i), \text{ dengan } P_F(R_i) = \sum_{j=1}^m \alpha_{ij} p_j \quad (2.6)$$

$$\ln(F) = \sum_{i=1}^n \log(P_F(R_i)) \quad (2.7)$$

Pada pengamatan data persegi panjang akan didapatkan *inner rectangle* yang saling bebas maka untuk setiap himpunan data terbatas terdapat sejumlah *inner rectangle* yang saling bebas dimana MLE menentukan massa probabilitas.

Asumsikan  $F$  kelas dari semua fungsi yang berdistribusi bivariat, MLE merupakan penyelesaian masalah optimisasi dalam dimensi tak hingga:

$$\hat{F}_n = \arg \max_{F \in \mathcal{F}} l_n(F) = \arg \max_{F \in \mathcal{F}} \sum_{i=1}^n \log(P_F(R_i)) \quad (2.8)$$

**Lemma 2.1** (Gentleman & Vandall, 2002)

*Sebarang CDF yang meningkat di luar himpunan  $H$  tidak dapat menjadi MLE untuk  $F$ .*

Berdasarkan lemma 2.1, memperbolehkan untuk memparameterisasi  $F$  dengan memisalkan  $p_i = P_F(H_i)$  dan misalkan  $\alpha_{ij}$  massa probabilitas yang menentukan *inner rectangle*  $R_j$ . Masalah optimisasi dapat dinyatakan dalam bentuk vektor  $\mathbf{p} = \{p_i\}_{i=1}^m$ , dengan kendala  $\sum p_i = 1$  dan  $p_i \geq 0$  untuk semua  $i$ . Fungsi *likelihood* dapat dengan mudah dituliskan kembali dalam bentuk  $\mathbf{p}$  yaitu

$$\begin{aligned} P_F(R_i) &= \sum_{i=1}^m \alpha_{ij} p_i \\ L(F) &= L(\mathbf{p}) = \prod_{j=1}^n [\sum_{i=1}^m \alpha_{ij} p_i] \text{ atau} \\ L(\mathbf{p}|R_1, R_2, \dots, R_n) &= \prod_{j=1}^n (A' p_j) \quad , \end{aligned} \quad (2.9)$$

dengan  $A$  merupakan *clique matrix*.  $L(\mathbf{p}|R_1, R_2, \dots, R_n)$  menyatakan bahwa dimensi  $m$  dari  $\mathbf{p}$  bergantung pada data  $R_1, R_2, \dots, R_n$ .

## 2.7 Algoritma Bron-Kerbosch

Algoritma Bron-Kerbosch pertama kali diperkenalkan oleh Coen Bron dan Joep Kerbosch pada tahun 1973 (Antoro *et al*, 2017). Menurut Conte (2013), berikut ini prosedur dari algoritma Bron-Kerbosch :

1. Sediakan himpunan  $R = \{\}$ ,  $P = \{v_1, v_2, \dots, v_n\}$ ,  $X = \{\}$ .  $R$  adalah hasil yang akan dicatat sebagai *maximal clique*.  $P$  adalah himpunan kandidat yang mungkin menjadi *maximal clique* dan bertetangga dengan  $v$ .  $X$  adalah himpunan titik yang dihapus.
2. Ambil sebuah titik  $v$  dari  $P$ . Tambahkan  $v$  ke  $R$  dan hapus titik yang tak bertetangga dengan  $v$  pada himpunan  $P$  dan  $X$ .
3. Ambil sebarang titik yang lain dari  $P$  yang baru dan ulangi langkah (2) hingga elemen pada  $P$  menjadi kosong. Ketika elemen di  $P$  dan  $X$  telah kosong maka catat elemen di  $R$  sebagai *maximal clique* yang baru. Bila tidak demikian maka  $R$  memuat subset *clique* yang sebelumnya.
4. Selanjutnya melangkah mundur ke titik yang diambil sebelumnya dan atur kembali  $P$ ,  $R$  dan  $X$  seperti langkah sebelumnya.
5. Hapus titik yang ada di  $P$  dan tambahkan titik itu ke  $X$ . Jika elemen di  $P$  telah habis maka melangkah mundur ke langkah (1).



## 2.8 Algoritma Genetika

Konsep dari algoritma Genetika pertama kali diperkenalkan oleh John Holland pada tahun 1975. Berikut ini beberapa definisi dalam algoritma genetika. Gen adalah sebuah nilai yang menyatakan satuan dasar yang membentuk suatu arti tertentu. Kromosom adalah gabungan gen-gen yang membentuk nilai tertentu. Individu menyatakan satu nilai atau keadaan yang menyatakan salah satu solusi yang mungkin dari permasalahan yang diangkat. Populasi merupakan sekumpulan individu yang akan diproses bersama dalam satu siklus proses evolusi. Generasi menyatakan satu-satuan siklus proses evolusi. Nilai Fitness menyatakan seberapa baik nilai dari suatu individu atau solusi yang didapatkan (Basuki, 2003).

Algoritma Genetika merupakan simulasi dari proses evolusi dan operasi genetika. Adapun siklus algoritma Genetika sebagai berikut:

1. Pembentukan kromosom merupakan tahap untuk memberikan kode pada gen dari suatu kromosom. Kromosom merepresentasikan suatu penyelesaian terhadap permasalahan optimasi yang dikerjakan. Satu gen mewakili satu variabel. Nilai dari setiap variabel ditentukan sesuai dengan batasan kendala. Suatu kromosom disebut juga dengan individu.
2. Pembentukan populasi awal merupakan proses membangkitkan sejumlah individu secara acak atau melalui prosedur tertentu. Populasi merupakan himpunan penyelesaian acak awal yang merupakan solusi yang mungkin dari permasalahan optimasi yang dikerjakan. Jumlah individu di dalam suatu populasi disebut dengan *popsize*.
3. Evaluasi *fitness* merupakan proses untuk mengevaluasi setiap populasi dengan menghitung nilai *fitness* setiap individu. Nilai *fitness* didapat dengan mensubstitusikan nilai setiap variabel pada kromosom ke fungsi tujuan.
4. Kriteria berhenti merupakan ketentuan dalam algoritma Genetika untuk memberikan batasan dalam proses algoritma Genetika sampai mendapatkan solusi yang optimum. Kriteria berhenti yang digunakan biasanya berdasarkan pada jumlah generasi dan percobaan.

5. Seleksi individu merupakan tahap penyeleksian individu di dalam suatu populasi. Individu yang memiliki nilai probabilitas yang tinggi memiliki kemungkinan yang lebih besar untuk terpilih pada proses selanjutnya.
6. *Crossover* merupakan proses perpindahan silang gen dalam kromosom antara dua kromosom induk untuk menghasilkan kromosom anak yang baru. Individu yang terpilih menjadi kromosom induk ditentukan dengan probabilitas *crossover* ( $P_c$ ).
7. Mutasi merupakan proses mengubah nilai dari satu atau beberapa gen dalam suatu kromosom. Gen yang mengalami mutasi ditentukan oleh probabilitas mutasi ( $P_m$ ). Nilai dari gen yang digantikan tetap memperhatikan batasan kendala yang dihadapi (Aziz *et al.*, 2016).

## 2.9 Penyakit Kanker Paru

Kanker adalah kelainan yang terjadi pada sel yang menjadi dasar pertahanan tubuh. Pada kondisi normal, sel akan menggandakan diri dan mati secara bertahap. Terkadang sel tidak tumbuh, berkembang dan mati seperti biasanya. Hal ini mungkin disebabkan karena darah atau cairan limfa yang ada di dalam tubuh menjadi tak normal atau membentuk suatu gumpalan yang disebut dengan tumor. Tumor digolongkan menjadi tumor ganas dan tumor jinak. Pada tumor jinak, sel tersebut hanya berkembang pada satu area dan tidak menyebar ke bagian tubuh lainnya. Adapun tumor ganas merupakan pembentuk sel kanker yang memiliki kemampuan untuk menyebar ke seluruh tubuh melalui aliran darah atau cairan limfa.

Paru-paru merupakan organ utama pada sistem pernapasan dan termasuk ke dalam bagian sistem pengeluaran (*respiratory*). Kanker paru merupakan tumor ganas yang tumbuh secara tak terkendali pada salah satu atau kedua organ paru. Kanker paru yang disebabkan karena sel kanker yang hidup di organ paru disebut sebagai kanker paru primer. Kanker paru yang disebabkan karena penyebaran sel kanker dari organ lain disebut sebagai kanker paru sekunder.

Merokok merupakan penyebab 9 dari 10 kasus kanker paru primer. Perokok pasif juga rentan terjangkit kanker paru. Selain itu penyebab kanker paru lainnya

yaitu paparan bahan kimia *asbestos*, riwayat keluarga, serta gaya hidup individu. Gejala umum yang dialami penderita kanker paru yaitu batuk dalam jangka waktu lama yang disertai dahak ataupun darah, sesak napas, nyeri dada, suara serak dan penurunan berat badan (Mothoneos, 2016).



## BAB 3. METODE PENELITIAN

### 3.1 Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data pasien penderita kanker paru-paru sejumlah 25 pasien yang diperoleh dari Rumah Sakit DKT Jember tahun 2018 bulan Januari sampai bulan Maret. Data tersebut merupakan data bivariat sekunder yang tersensor kiri dan interval. Kejadian yang menjadi perhatian yakni terdiagnosa kanker paru. Pada kasus ini variabel yang akan digunakan yaitu :

a. Variabel Bebas ( $X$ )

Variabel bebas penelitian ini adalah interval waktu dugaan awal mula munculnya karsinoma (tumor). Titik awal interval dicatat dari perhitungan mundur  $n$  hari sebelum pasien melakukan perawatan di rumah sakit. Bilangan  $n$  ini dibangkitkan dengan aturan berikut. Pasien yang datang sebelum tanggal 7, maka nilai  $n < 7$  dan pasien yang datang sesudah tanggal 7 maka nilai  $n = 7$ . Apabila nilai  $n < 7$  mengakibatkan data penelitian menjadi tersensor kiri. Hal ini karena interval waktu ( $x$ ) kurang dari interval waktu yang ditentukan oleh peneliti.

b. Variabel Terikat ( $Y$ )

Variabel terikat penelitian ini adalah interval waktu terjangkit penyakit kanker paru-paru. Interval ini dicatat dari tanggal awal pasien memasuki rumah sakit hingga meninggalkan rumah sakit. Hal ini mengakibatkan data menjadi tersensor interval, karena waktu survival individu hanya diketahui pada interval waktu tertentu, yakni pada saat individu melakukan perawatan di rumah sakit.

### 3.2 Langkah-Langkah Penelitian

Skema untuk penelitian ini disajikan pada Gambar 3.1. Penjelasan Gambar 3.1 sebagai berikut :

a. Data bivariat kanker paru-paru

Langkah pertama yang dilakukan adalah mencari data bivariat kanker paru-paru di Rumah Sakit DKT, Jember. Selanjutnya, menentukan variabel-variabel yang akan digunakan dari data bivariat kasus kanker paru-paru.

b. Persegi panjang pengamatan (*Rectangles Observation*)

Data yang telah ditentukan kemudian diubah ke dalam bentuk persegi panjang pengamatan yang direpresentasikan sebagai  $(x_{1i}, x_{2i}, y_{1i}, y_{2i})$ . Titik  $(x_{1i}, y_{1i})$  adalah sudut kiri bawah dan  $(x_{2i}, y_{2i})$  adalah sudut kanan atas.

c. Merepresentasikan persegi pengamatan ke dalam bentuk graf

Data yang berbentuk persegi pengamatan diubah ke dalam bentuk graf irisan.

d. Membentuk matriks adjasensi

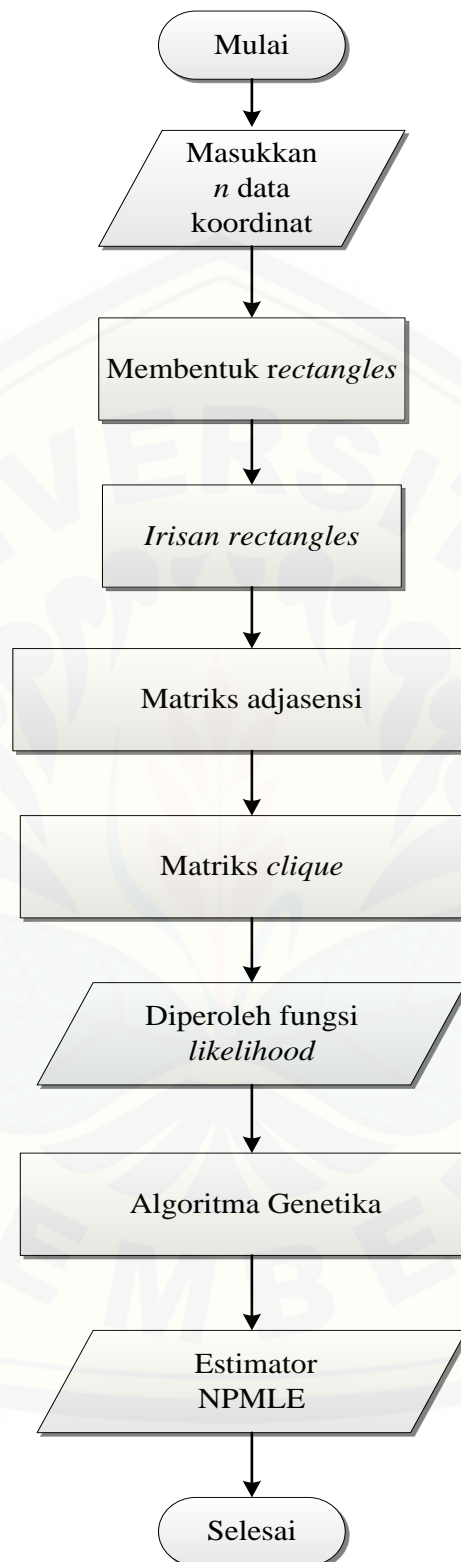
Graf irisan selanjutnya direpresentasikan ke dalam bentuk matriks adjasensi.

e. *Maximal clique*

Dilakukan penentuan jumlah *maximal cliques* berdasarkan matriks adjasensi untuk mendapatkan matriks *clique*. Tahap ini menggunakan bantuan algoritma Bron-Kerbosch.

f. Fungsi *likelihood*

Matriks *clique* direpresentasikan ke dalam bentuk graf *clique*. Elemen  $\alpha_{ij}$  dari matriks *clique* akan disubstitusikan ke dalam fungsi *likelihood*. Selanjutnya fungsi *likelihood* dimaksimumkan dengan bantuan algoritma Genetika.



Gambar 3.1 Skema penelitian



## BAB 5. PENUTUP

### 5.1 Kesimpulan

Berdasarkan hasil dan pembahasan diperoleh kesimpulan sebagai berikut.

- a. *Maximal cliques* yang diperoleh dari graf irisan hasil persegi panjang dengan menggunakan algoritma Bron-Kerbosch yaitu  $M_1 = \{R_1, R_2, R_5, R_7, R_8, R_{13}, R_{16}, R_{23}, R_{25}\}$ ,  $M_2 = \{R_1, R_2, R_5, R_7, R_8, R_{16}, R_{24}, R_{25}\}$ ,  $M_3 = \{R_1, R_7, R_8, R_{13}, R_{22}, R_{23}\}$ ,  $M_4 = \{R_2, R_5, R_8, R_{16}, R_{17}, R_{24}, R_{25}\}$ ,  $M_5 = \{R_3, R_5, R_{17}, R_{25}\}$ ,  $M_6 = \{R_3, R_{19}\}$ ,  $M_7 = \{R_4, R_6, R_{14}, R_{18}\}$ ,  $M_8 = \{R_4, R_{15}, R_{21}\}$ ,  $M_9 = \{R_6, R_9, R_{12}, R_{14}, R_{18}\}$ ,  $M_{10} = \{R_{10}, R_{14}\}$ ,  $M_{11} = \{R_{11}, R_{19}\}$  dan  $M_{12} = \{R_{20}, R_{21}\}$ . Pada hasil tersebut, tidak terdapat daerah persegi yang saling asing. Paling sedikit terdapat dua daerah persegi yang saling beririsan dan paling banyak terdapat 9 daerah persegi yang saling beririsan. Daerah persegi yang saling beririsan, merepresentasikan interval masa inkubasi yang sama.
- b. Perhitungan NPMLE dengan menggunakan algoritma genetika menghasilkan nilai  $p_1 = 0,0625$ ,  $p_2 = 0,0626$ ,  $p_3 = 0,1249$ ,  $p_4 = 0,1251$ ,  $p_5 = 0,0626$ ,  $p_6 = 0,0625$ ,  $p_7 = 0,0781$ ,  $p_8 = 0,1249$ ,  $p_9 = 0,0625$ ,  $p_{10} = 0,0469$ ,  $p_{11} = 0,0624$ ,  $p_{12} = 0,1250$  dan nilai fungsi *likelihood*  $L = 1,1636 \times 10^{-20}$ . Hasil estimator yang diperoleh relatif beragam dan tak nol, sehingga dapat dilakukan perhitungan masa inkubasi yang sama untuk semua daerah irisan maksimal. Nilai dari setiap estimator akan dikalikan dengan 168 (7 hari x 24 jam) sehingga outpunya berupa masa inkubasi yang sama untuk beberapa pasien dalam satuan jam.

### 5.2 Saran

Pada penelitian ini menggunakan algoritma Bron-Kerbosch dan algoritma Genetika dengan menggunakan data bivariat tersensor kiri dan interval. Pada penelitian selanjutnya ;

1. Sebaiknya untuk penentuan *maximal clique* dan Perhitungan NPMLE menggunakan algoritma yang lain.
2. Sebaiknya menggunakan kasus lain yang dapat diperluas pada dimensi 3.



**DAFTAR PUSTAKA**

- Acita, S., M. Fatekurohman, dan D. Anggraeni. 2017. Matriks Insidensi pada NPMLE untuk Data Tersensor Interval Bivariat. *Prosiding Seminar Nasional Matematika dan Pembelajarannya*. 519-524.
- Antoro, S.C., K. A. Sugeng, dan B. D. Handari. 2017. Application of Bron-Kerbosch Algorithm in Graph Clustering Using Complement Matrix. *AIP Conference Proceedings*, American Institut of Physics. 030141.
- Azhari, Z. 2016. Penerapan Algoritma Genetika Pada Penyelesaian Sistem Persamaan Linier. *Skripsi*. Jurusan Matematika FMIPA UNEJ. Jember.
- Aziz, A., B. Prihandono, dan Ilhamsyah. 2016. Algoritma Genetika pada Pemrograman Linear dan Nonlinear. *Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster)*. 5 (3) : 265-274.
- Bain, L. J. dan M. Engelhardt. 1992. *Introduction to Probability and Mathematical Statistics*. Belmont: Duxburry Press.
- Bapat, R. B. 2010. *Graphs and Matrices*. India: Hindustan Book Agency.
- Baruah, A. K. dan N. Baruah. 2012. Clique Matrix of a Graph in Traffic Control Problems. *International Journal of Computer Application*. 53(6): 975-8887.
- Basuki, A. 2003. *Algoritma Genetika: Suatu Alternatif Penyelesaian Permasalahan Searching, Optimasi dan Machine Learning*. Surabaya: Politeknik Elektronika Negeri Surabaya PENS-ITS.
- Bogaerts, K. dan E. Lesaffre. 2004. A New, Fast Algorithm to Find the Region of Possible Suport for Bivariat Interval-Censored Data. *Journal of Computational and Graphical Statistics*. 13(2): 330-340
- Chartrand, G dan P. Zhang. 2012. *A First Course in Graph Theory*. New York: Dover Publications, INC.
- Collett, D. 2003. *Modelling Survival Data in Medical Research*. Second Edition. New York: Chapman & Hall/Crc.
- Conte, A. 2013. *Review of the Bron-Kerbosch Algorithm and Variations*. University of Glasgow: School of Computing Science.
- Fatekurohman, M., Subanar, dan Danardono. 2014. Algorithm of Intersection Graph for Bivariate Censored Data. *Global Journal of Pure and Applied Mathematics*. 11(1): 0973-1768.

- Gentleman, R. dan A. C. Vandal. 2001. Computational Algorithm for Censored Data Problems Using Intersection Graphs. *J. Comp. Graphical Statist.*10 (3): 403-421.
- Gentleman, R dan A. C. Vandal. 2002. Graph-Theoretical Aspects of Bivariate Censored Data. *J. Comp. Graphical Statist.*30 (4): 557-571.
- Kholidati, J. R., M. Fatekurohman, dan A. F. Hadi. 2017. Fungsi *Likelihood* untuk Metode Nonparametrik Maksimum *Likelihood* Estimator pada Data Tersensor Interval. *Prosiding Seminar Nasional Matematika dan Pembelajarannya.* 510-518.
- Kleinbaum, D.G. dan M. Klein. 2005. *Survival Analysis: A Self-Learning Text.* Edisi kedua. United State of America : Springer Science+Business Media, Inc.
- Lee, D. T. 1983. Maximum Clique Problem of Rectangle Graphs. *Advance in Computing Research.* 91-107.
- Lee, E. T dan J. W. Wang. 2003. *Statistical Methods for Survival Data Analysis Third Edition.* New Jersey: John Wiley & Sons, Inc.
- Maathuis, M. H. 2003. *Nonparametric Maximum Estimation for Bivariate Censored Data.* Master's thesis, Delf University of Technology, Delf.
- Mothoneos, J. 2016. *Understanding Lung Cancer A Guide for People with Cancer, Their Families and Friends.* Cancer Council Australia: SOS Print + Media Group.
- Pal, M. 2013. Intersection Graphs: An Introduction. *Annals of Pure and Applied Mathematics.* 4(1): 43-91.
- Rahman, M. S. 2017. *Undergraduated Topics in Computer Science: Basic Graph Theory.* Switzerland: Springer International Publishing AG
- Rosen, K. H. 2000. *Handbook of Discrete and Combinatorial Mathematics.* United States of America: CRC Press LLC.
- Skiena, S. S. 2008. *The Algorithm Design Manual Second Edition.* New York: Springer Science+Business Media.
- Voloshin, V. I. 2009. *Introduction to Graph Theory.* New York: Nova Science Publisher, Inc.
- West, D. G. 2002. *Introduction to Graph Theory Second Edition.* India: Rashtriya Printers.

LAMPIRAN

Lampiran A. Data Pasien Kanker Paru bulan Januari-Maret tahun 2018

No.	NOMOR	KECAMATAN	UMUR (TAHUN)	JK	RUANG PERAWATAN	DIAGNOSA AKHIR	KODE ICD-10	TANGGAL MRS	TANGGAL KRS	LAMA DIRAWAT	HP	KEADAAN PASIEN PULANG
1	70381	PATRANG	25	P	ANGGREK	KP	A16.2	15/01/2018	22/01/2018	7	8	1
2	70980	SUKOREJO	26	P	ANGGREK	KP	A16.2	19/01/2018	23/01/2018	4	5	1
3	55212	PATRANG	65	P	ANGGREK	KP	A16.2	25/01/2018	30/01/2018	5	6	1
4	70582	KALISATI	62	L	MELATI	KP	A16.2	04/01/2018	09/01/2018	5	6	1
5	71079	SEMBAGU	69	P	MELATI	KP	A16.2	20/01/2018	26/01/2018	6	7	1
6	69406	WULUHAN	55	L	TERATAI	KP	A16.2	07/01/2018	11/01/2018	4	5	1
7	70868	KALISAT	26	L	TERATAI	KP	A16.2	15/01/2018	23/01/2018	8	9	1
8	71001	ALIWATE	22	L	TERATAI	KP	A16.2	17/01/2018	24/01/2018	7	8	1
9	70742	RAMBIPUJ	39	L	TERATAI	KP	A16.2	09/01/2018	11/01/2018	2	3	1
10	69406	WULUHAN	55	L	TERATAI	KP	A16.2	12/01/2018	15/01/2018	3	4	1
11	71416	JIMBULSAR	58	L	TERATAI	KP	A16.2	30/01/2018	31/01/2018	1	2	1
12	71742	PANJI	49	P	ANGGREK	KP	A16.2	09/02/2018	10/02/2018	1	2	3
13	67982	PATRANG	65	P	ANGGREK	KP	A16.2	17/02/2018	21/02/2018	4	5	1
14	71606	UMBERSA	76	L	MELATI	KP	A16.2	08/02/2018	13/02/2018	5	6	1
15	71606	UMBERSA	76	L	TERATAI	KP	A16.2	05/02/2018	07/02/2018	5	3	1
16	71606	KEBONSAR	76	L	TERATAI	KP	A16.2	20/02/2018	23/02/2018	5	4	1
17	71913	SEMBAGU	63	L	TERATAI	KP	A16.2	22/02/2018	26/02/2018	5	5	1
18	72548	UMUKMA	59	P	ANGGREK	KP	A16.2	08/03/2018	11/03/2018	5	4	1
19	73192	WULUHAN	75	P	ANGGREK	KP	A16.2	29/03/2018	31/03/2018	5	3	1
20	72347	TAPEN	62	L	TERATAI	KP	A16.2	01/03/2018	04/03/2018	5	4	1
21	72307	KALIWATE	51	L	TERATAI	KP	A16.2	01/03/2018	06/03/2018	5	6	1
22	71656	KALIWATE	26	L	TERATAI	KP	A16.2	15/03/2018	19/03/2018	5	5	2
23	72813	TANGGUL	43	L	TERATAI	KP	A16.2	16/03/2018	21/03/2018	5	6	1
24	70103	SILO	50	L	TERATAI	KP	A16.2	21/03/2018	23/03/2018	5	3	1
25	68817	WULUHAN	37	L	TERATAI	KP	A16.2	20/03/2018	26/03/2018	5	7	1

## Lampiran B. Data interval bivariat pasien kanker paru

NO.	Nomor	Interval X	Interval Y	Tanggal Masuk RS	Tanggal Keluar RS
1.	070381	8-15	15-22	15/1/2018	22/1/2018
2.	070980	12-19	19-23	19/1/2018	23/1/2018
3.	055212	18-25	25-30	25/1/2018	30/1/2018
4.	070582	1-4	4-9	04/01/2018	09/01/2018
5.	071079	13-20	20-26	20/01/2018	26/01/2018
6.	069406	1-7	7-11	7/1/2018	11/1/2018
7.	070868	8-15	15-23	15/1/2018	23/1/2018
8.	071001	10-17	17-24	17/1/2018	24/1/2018
9.	070742	2-9	9-11	09/01/2018	11/01/2018
10.	069406	5-12	12-15	12/01/2018	15/01/2018
11.	071416	23-30	30-31	30/01/2018	31/01/2018
12.	071742	2-9	9-10	09/02/2018	10/02/2018
13.	067982	10-17	17-21	17/02/2018	21/02/2018
14.	071606	1-8	8-13	08/02/2018	13/02/2018
15.	071606	1-5	5-7	05/02/2018	07/02/2018
16.	071606	13-20	20-23	20/02/2018	23/02/2018
17.	071913	15-22	22-26	22/02/2018	26/02/2018
18.	072548	1-8	8-11	08/03/2018	11/03/2018
19.	073192	22-29	29-31	29/03/2018	31/03/2018
20.	072347	1-4	1-4	01/03/2018	04/03/2018
21.	072307	1-6	1-6	01/03/2018	06/03/2018
22.	071656	8-15	15-19	15/03/2018	19/03/2018
23.	072813	9-16	16-21	16/03/2018	21/03/2018

NO.	Nomor	Interval X	Interval Y	Tanggal Masuk RS	Tanggal Keluar RS
24.	070103	14-21	21-23	21/03/2018	23/03/2018
25.	068817	13-20	20-26	20/03/2018	26/03/2018





Lampiran C. Menyelesaikan fungsi *likelihood* menggunakan algoritma Genetika

```

syms p1 p2 p3 p4 p5 p6 p7 p8 p9 p10 p11 p12
>> B=[1      1      1      0      0      0      0      0      0      0      0      0      0;
1      1      0      1      0      0      0      0      0      0      0      0      0;
0      0      0      0      1      1      0      0      0      0      0      0      0;
0      0      0      0      0      0      1      1      0      0      0      0      0;
1      1      0      1      1      0      0      0      0      0      0      0      0;
0      0      0      0      0      0      1      0      1      0      0      0      0;
1      1      1      0      0      0      0      0      0      0      0      0      0;
1      1      1      1      0      0      0      0      0      0      0      0      0;
0      0      0      0      0      0      0      0      1      0      0      0      0;
0      0      0      0      0      0      0      0      0      1      0      0      0;
0      0      0      0      0      0      0      0      0      0      1      0      0;
0      0      0      0      0      0      0      0      1      0      0      0      0;
1      0      1      0      0      0      0      0      0      0      0      0      0;
0      0      0      0      0      0      1      0      1      1      0      0      0;
0      0      0      0      0      0      0      1      0      0      0      0      0;
1      1      0      1      0      0      0      0      0      0      0      0      0;
0      0      0      1      1      0      0      0      0      0      0      0      0;
0      0      0      0      0      0      1      0      1      0      0      0      0;
0      0      0      0      0      0      1      0      0      0      0      1      0;
0      0      0      0      0      0      0      0      0      0      0      0      1;
0      0      0      0      0      0      0      1      0      0      0      0      1;
0      0      1      0      0      0      0      0      0      0      0      0      0;
1      0      1      0      0      0      0      0      0      0      0      0      0;
0      1      0      1      0      0      0      0      0      0      0      0      0;
1      1      0      1      1      0      0      0      0      0      0      0      0];
>> P=[p1; p2; p3; p4; p5; p6; p7; p8; p9; p10; p11; p12];
>> B*P
ans =
 p1 + p2 + p3
  p1 + p2 + p4
  p5 + p6
  p7 + p8
 p1 + p2 + p4 + p5
  p7 + p9
  p1 + p2 + p3
 p1 + p2 + p3 + p4
  p9
  p10
  p11
  p9
  p1 + p3
 p7 + p9 + p10
  p8
 p1 + p2 + p4
  p4 + p5
  p7 + p9
  p6 + p11
  p12
  p8 + p12

```



$p^3$   
 $p^1 + p^3$   
 $p^2 + p^4$   
 $p^1 + p^2 + p^4 + p^5$

```
function [c,ceq]=paru_constraint(x)
```

```
c=[];
```

```
ceq=[x(1)+x(2)+x(3)+x(4)+x(5)+x(6)+x(7)+x(8)+x(9)+x(10)+x(11)+x(12)-1];
```

```
function y=paru_fitness(x)
```

```

y=-(x(3)*x(8)*(x(9)^2)*x(10)*x(11)*x(12)*((x(1) + x(3))^2)*(x(2) + x(4))*(x(4) + x(5))*(x(5) + x(6))*(x(7) +
x(8))*((x(7) + x(9))^2)*(x(6) + x(11))*(x(8) + x(12))*((x(1) + x(2) + x(3))^2)*((x(1) + x(2) + x(4))^2)*(x(7) +
x(9) + x(10))*(x(1) + x(2) + x(3) + x(4))*((x(1) + x(2) + x(4) + x(5))^2));

```

```
%running maksimisasi likelihood
```

```
ObjectiveFunction=@paru_fitness;
```

```
nvars=12;
```

```
LB=[0 0 0 0 0 0 0 0 0 0 0 0];
```

```
UB=[1 1 1 1 1 1 1 1 1 1 1 1];
```

```
ConstraintFunction=@paru_constraint;
```

```
rng default
```

```
>> [x,fval]=ga(ObjectiveFunction,nvars,[],[],[],[],LB,UB,ConstraintFunction)
```

## Lampiran D. Membangun persegi panjang &amp; irisannya

```

%membentuk koordinat Persegi Panjang pengamatan
[FileName,PathName]=uigetfile({'*txt','Pilih file .txt'});
if FileName~=0
    data=importdata([PathName,FileName]);
    [m n]=size(data);
    for i=1:m
        ro(i)={'R' num2str(i)};
    end
    set(handles.uitable1,'data',data,'rowname',ro,...
        'columnname',{'x1','x2','y1','y2'},'columnwidth',{55});
    set(handles.edit1,'string',[PathName FileName]);
    set(handles.figure1,'userdata',ro);
    set(handles.uitable1,'columneditable',true)
end
%menentukan irisan persegi panjang
clc
data=get(handles.uitable1,'data');
ir=iris(data);

dlmwrite('irisan.txt', ir, 'delimiter', '\t', ...
    'precision', 6)
[m n]=size(data);
a=data(:,1:2);
b=data(:,3:4);
[v w]=size(data);
for i=1:v
    text(data(i,1),data(i,4),['\color{black}R' num2str(i)],...
        'fontsize',11,'fontweight','normal','fontname','cambria',...
        'HorizontalAlignment','left','VerticalAlignment','baseline')
end
data(m+1,:)= [min(min(a))-1 max(max(a))+1 min(min(b))-1 max(max(b))+1];
[m n]=size(data);
a=data(:,1:2);
b=data(:,3:4);
x=min(min(a)):max(max(a));
y=min(min(b)):max(max(b));
set(handles.figure1,'CurrentAxes',handles.axes1);
set(handles.axes1,'nextplot','replace');
for i=1:length(x)
    plot([x(i) x(i)],[y(1) y(length(y))],'k:');
    set(handles.axes1,'nextplot','add');
end
for i=1:length(y)
    plot([x(1) x(length(x))],[y(i) y(i)],'k:');
end
for i=1:m-1
    wr=rand(1,3);
    plot([data(i,1) data(i,1)],[data(i,3) data(i,4)],'color',wr,'linewidth',2.5);
    plot([data(i,1) data(i,2)],[data(i,3) data(i,3)],'color',wr,'linewidth',2.5);
    plot([data(i,1) data(i,2)],[data(i,4) data(i,4)],'color',wr,'linewidth',2.5);
    plot([data(i,2) data(i,2)],[data(i,3) data(i,4)],'color',wr,'linewidth',2.5);
    text(data(i,1),data(i,4),['\color{blue}R' num2str(i)],...
        'fontsize',10,'fontweight','normal','fontname','cambria',...

```

```

    'HorizontalAlignment','left','VerticalAlignment','baseline')
end
A=mat_adjacency(ir);
B=BK_MaxClique(A);
B=B';
[m1 n1]=size(B);
for i=1:m1
    for j=1:n1
        if B(i,j)~=0
            C(i,j)=j;
        end
    end
end
for i=1:m1
    a=hapus_nol(C(i,:));
    iri=data(a(1,:));
    for j=1:n1
        if B(i,j)~=0
            iri=iris_wrn(iri,data(j,:));
        end
    end
    ir1(i,:)=iri;
end
[mw nw]=size(ir1);
for i=1:mw
    wr='k';
    plot([ir1(i,1) ir1(i,1)],[ir1(i,3) ir1(i,4)],'color',wr,'linewidth',5);
    plot([ir1(i,1) ir1(i,2)],[ir1(i,3) ir1(i,3)],'color',wr,'linewidth',5);
    plot([ir1(i,1) ir1(i,2)],[ir1(i,4) ir1(i,4)],'color',wr,'linewidth',5);
    plot([ir1(i,2) ir1(i,2)],[ir1(i,3) ir1(i,4)],'color',wr,'linewidth',5);
    a=(ir1(i,1)+ir1(i,2))/2;
    b=(ir1(i,3)+ir1(i,4))/2;
    plot(a,b,'s','markersize',25,'markerfacecolor',rand(1,3),'markeredgecolor','k');
    text(a,b,['\color{white}H' num2str(i)],...
        'fontsize',10,'fontweight','bold','fontname','cambria',...
        'HorizontalAlignment','center')
end
set(handles.axes1,'Color',[1 1 1],'XColor',[0 0 0],'YColor',[0 0 0]);
set(handles.axes1,'XLim',[min(x) max(x)],[min(y) max(y)],...
    'ytick','y','xgrid','off','ygrid','off');
xlabel('Interval Munculnya Tumor','fontname','cambria','fontsize',11);
ylabel('Interval Terjangkit Kanker Paru','fontname','cambria','fontsize',11);

```

## Lampiran E. Membentuk graf irisan &amp; matriks adjasensi

```

%Membentuk matriks adjasensi
function out=mat_adjacency(A)
[m n]=size(A);
out=eye(m,n);
for i=1:m
    for j=1:n
        if A(i,j)~=0
            out(i,A(i,j))=1;
        end
    end
end
%Membangun graf irisan
clc
data=dlmread('irisan.txt');
A=mat_adjacency(data);
[m n]=size(A);
[x y]=poligon(m);
hold off
for i=1:m
    for j=1:n
        if A(i,j)~=0
            plot([x(i) x(j)],[y(i) y(j)],'-');
            hold on
        end
    end
end
for i=1:m
    plot(x(i),y(i),'o','markerfacecolor','y','markeredgecolor','k');
    text(x(i),y(i)+0.03,['\color{black}R' num2str(i)],'fontsize',11,'fontweight',...
        'normal','fontname','cambria','HorizontalAlignment','center','VerticalAlignment','baseline')
end
axis equal
set(hObject,'Color',[1 1 1],'xlim',[min(x)-0.3 max(x)+0.3],'xtick',[],'ytick',[]);

```

Lampiran F. Menentukan *maximal clique*

```

function [ M ] = BK_MaxIS( int_matrix )
no_vertices = size(int_matrix,2);
% output for the maximal independent sets
M = [];
P = [];
S = [];
T = 1:no_vertices;
findIS( P, S, T);
% Recursive function to branch on
function findIS(P, S, T)

% Check if we are done
if check_end(S,T)

% Loop through the edges
for i=1:size(T,2)
% pick x as candidate
x = T(i);
P = [P x];
S_new = remove_nbrs(S, x,0);
T_new = remove_nbrs(T, x, 1);

% if S(j) and T(j) are empty, we got max'l IS
if (isempty(S_new) && isempty(T_new))
new_IS = convert_set_to_bin(P, no_vertices);
M = [M new_IS'];

% go to step 5
elseif ( ~isempty(S_new) && isempty(T_new))
% go to step 5

else
% recursive call
findIS(P, S_new, T_new);
end

% remove x from P
x_ind = find(P == x);
P(x_ind) = [];
% add x to S
S = [S x];

end % for i=1:size(T,2)
end % if check_end(S,T)

end % findIS
% Removes the neighbors of x from inp, and returns it
% Sinp: is the set where the neighborhood will be removed
% x :index of node whose neighborhood will be removed
% removeX : if 1, x is also removed, from Sinp
function [ Sout ] = remove_nbrs( Sinp, x, removeX)

```

```

if nargin ~= 3
    error('Need all three arguments');
end
int_line = int_matrix(x, :); % proper line in the interference matrix
nbrs = find(int_line == 1); % locate where int_line = 1

% if we also need to remove the node itself
if (removeX == 1)
    nbrs = [nbrs x];
end

mems = ismember(Sinp, nbrs); % find locations where there is a match
no_ind = find(mems ~= 1); % find indices of locations where there is not match
Sout = Sinp(no_ind); % return those indices
end % remove_nbrs
% Converts an array of decimals to 0-1 array
% [2 4] => [ 0 1 0 1 ]
% The size of the output array is exactly len
function [Sout] = convert_set_to_bin(Sin, len)

    Sout = zeros(1,len);
    Sout(Sin) = 1;

end %convert_set_to_bin
% check whether we can end the execution
function ret = check_end(S, T)
    ret = 1;
    for k=1:size(S,2)

        %TODO Fix this for not using g to get the neighborhood
        E_k = g_neighbors(int_matrix,S(k)); %neighborhood of k

        if ( isempty(intersect(E_k, T)) )
            ret = 0;
            break; %there exists such a node, so no need to continue
        end % if ( isempty(intersect(E_k, T)) )

    end % for k=1:size(S,2)
end %ret
% returns the neighbors of n in the matrix given by M
function nbrs = g_neighbors(M,n)
    m_row = M(n,:);
    nbrs = find(m_row == 1);
end
end %BKmaximal

function [ M ] = BK_MaxClique( int_matrix )
no_nodes = size(int_matrix,1);
% Find the complement of the graph
int_new = abs(int_matrix - 1);
int_new(logical(eye(no_nodes))) = zeros(1, no_nodes);
M = BK_MaxIS(int_new);
End

```



Lampiran G. Membangun *clique matrix*

```
%Membentuk clique matrix
data=dlmread('irisan.txt');
A=mat_adjacency(data);
[m n]=size(A);
B=BK_MaxClique(A);
for i=1:m
    r(i)={'R' num2str(i)};
end
[b k]=size(B);
for i=1:k
    kol(i)={'H' num2str(i)};
end
set(hObject,'data',B,'rowname',r,'columnname',kol,'columnwidth',{30})
```

