

BAB II

KAJIAN LITERATUR

2.1 Tinjauan Pustaka

Pada bagian ini akan menjelaskan hal – hal yang berhubungan dengan topik penelitian dan teori pendukungnya. Bagian ini nantinya juga akan membahas mengenai penghilangan derau, pengembalian citra yang kabur dan penghilangan kabut pada citra serta pendekatan lainnya yang berhubungan dengan objek penelitian.

1.1.1 Kanker Serviks

Kanker serviks adalah salah satu penyebab utama kematian akibat kanker di kalangan wanita. Di seluruh dunia, kanker serviks adalah keganasan keempat yang paling sering terjadi pada wanita, dan menghasilkan sekitar 530.000 kasus baru setiap tahun dengan 270.000 kematian. Sekitar 85% kematian di seluruh dunia akibat kanker serviks terjadi di negara-negara berpenghasilan rendah atau berkembang, dan angka kematian 18 kali lebih tinggi di negara-negara berpenghasilan rendah dan menengah dibandingkan dengan negara-negara kaya. Angka kejadian tertinggi terjadi di Amerika Tengah dan Selatan, Karibia, Afrika Sub Sahara, dan Asia Selatan.³ Di Amerika Serikat pada tahun 2016, diperkirakan ada 12.990 kasus dan 4120 kematian akibat kanker serviks, dan usia rata-rata pada saat diagnosis adalah 47 tahun (Small Jr *et al.*, 2017).

Kanker serviks, yang timbul dari serviks, disebabkan oleh perubahan gen yang mengontrol pertumbuhan dan fungsi pembelahan sel. Kanker ini memiliki kemampuan untuk menyebar dari leher rahim ke bagian lain dalam tubuh. Pada tahap awal, tidak ada tanda-tanda yang bisa diamati. Pemeriksaan rutin adalah satu-satunya cara untuk mendeteksi kanker serviks. (Wu, W., Dan Zhou, H. 2017) Gejala muncul pada tahap akhir, termasuk perdarahan vagina dan nyeri panggul. Selain itu, kanker serviks dapat menyebar ke organ lain, seperti perut dan paru-paru. Pada tahap ini, kanker serviks dikatakan stadium lanjut, gejala diantaranya kelelahan, nyeri kaki, patah tulang, penurunan berat badan dan nyeri punggung.

Teknik *magnetic resonance imaging* (MRI) dan teknik *difusi-weighted imaging* (DWI) dapat mendeteksi kanker serviks sampai batas tertentu (exner *et al.*, 2016). Namun, orang-orang di negara berkembang memiliki kesadaran masalah rendah untuk penyaringan rutin. Selain itu, kurangnya keahlian dokter dan peralatan medis yang terbatas membuat kanker serviks menjadi penyebab utama kematian di negara-negara berpenghasilan rendah. Human papillomavirus (HPV) adalah penyebab utama kanker serviks (Wu, W., Dan Zhou, H. 2017). Merokok, penggunaan kontrasepsi, kehamilan ganda dan beberapa faktor lain juga dapat menyebabkan kanker serviks. Misalnya, risiko kanker serviks akan meningkat dua hingga tiga kali jika pasien yang terinfeksi HPV merokok. Sementara itu, ditunjukkan bahwa kejadian kanker serviks untuk wanita yang menggunakan kontrasepsi tiga kali lebih tinggi daripada mereka yang tidak menggunakan kontrasepsi. Selain itu, kejadiannya akan meningkat hingga empat kali lipat jika kontrasepsi digunakan selama lebih dari 10 tahun. Adapun kehamilan ganda, pasien perempuan yang terinfeksi HPV tanpa kehamilan memiliki insiden kanker serviks yang lebih rendah dibandingkan dengan mereka yang memiliki lebih dari satu kehamilan (Kessler *et al.*, 2017).

a. Faktor Risiko Kanker Serviks

Faktor risiko adalah faktor yang mempermudah timbulnya penyakit kanker serviks. Beberapa faktor yang menyebabkan perempuan terpapar HPV (sebagai penyebab dari kanker serviks) adalah sebagai berikut:

1. Usia

Faktor alamiah pencetus kanker serviks adalah wanita usia diatas 40 tahun.

Semakin tua seorang wanita maka makin tinggi risikonya terkena kanker serviks (Kartikawati, 2013). Puncak perkembangan kanker serviks berada pada usia 47 tahun. Sekitar 47% wanita dengan kanker serviks invasif berusia di bawah 35 tahun saat terdiagnosis. Sekitar 10 %, kanker serviks terjadi pada wanita yang lebih tua (> 65 tahun) dan cenderung meninggal karena penyakit karena stadium lanjut mereka saat didiagnosis (Gattoc, *et al*, 2015).

2. Usia pertama kali melakukan hubungan seksual

Sesuai dengan etiologi infeksi, wanita yang memulai hubungan seksual pada usia muda akan meningkatkan risiko terkena kanker serviks karena sel kolumnar serviks lebih peka terhadap metaplasia selama usia dewasa maka wanita yang berhubungan seksual sebelum usia 18 tahun akan berisiko terkena kanker serviks lima kali lipat. Usia pertama kali melakukan hubungan seks merupakan salah satu faktor risiko terpenting karena penelitian para pakar menunjukkan bahwa semakin muda wanita melakukan hubungan seksual maka semakin besar risiko terkena kanker serviks. Wanita yang melakukan hubungan seks pertama sekali pada usia kurang dari 20 tahun mempunyai risiko 3 kali lebih besar daripada wanita yang berhubungan seksual pertama sekali pada usia lebih dari 20 tahun. Umumnya sel-sel mukosa baru matang setelah wanita berusia 20 tahun ke atas (Rasjidi, 2014).

3. Paritas

Berdasarkan penelitian yang dilakukan oleh Hidayat dkk (2013), menyimpulkan bahwa banyaknya anak yang dilahirkan berpengaruh dalam timbulnya penyakit kanker serviks. Paritas merupakan salah satu faktor risiko terjadinya kanker serviks dengan besar risiko 4,55 kali untuk terkena kanker serviks pada wanita dengan paritas >3 dibandingkan wanita dengan paritas 3.

4. Multipartner seks (Berganti-Ganti Pasangan)

Wanita yang memiliki 7 atau lebih kehamilan aterm mungkin memiliki peningkatan risiko kanker serviks (National Cancer Institute, 2012). Berganti ganti pasangan seksual, memungkinkan tertularnya penyakit kelamin, salah satunya Human Papilloma Virus (HPV). Virus ini akan mengubah sel di permukaan mukosa hingga membelah menjadi lebih banyak. Bila hal ini terus menerus terjadi, sel kanker pun akan terus berkembang. Perilaku berganti-ganti pasangan seksual akan meningkatkan penularan penyakit kanker serviks. Risiko terkena kanker serviks meningkat 10 kali lipat pada wanita mempunyai teman seksual 6 orang atau lebih dibandingkan wanita yang mempunyai 1 pasangan seksual (Azis, 2008). Menurut Wahyuni dan Mulyani (2014) berpendapat bahwa partner sex >1 orang akan meningkatkan risiko 6,19 kali lebih besar untuk

mengalami lesi prakanker serviks dibandingkan dengan wanita yang memiliki patner sex 1 orang saja.

5. Merokok

Tembakau yang mengandung bahan-bahan karsinogen baik yang dihisap sebagai rokok/sigaret atau dikunyah. Asap rokok menghasilkan *polycyclic aromatic hydrocarbon heterocyclic nitrosamines*. Pada wanita perokok konsentrasi nikotin pada getah serviks 56 kali lebih tinggi dibandingkan di dalam serum. Efek langsung bahan-bahan tersebut pada serviks adalah menurunkan status imun lokal sehingga dapat menjadi kokarsinogen infeksi virus (Kartikawati, 2013). Berdasarkan penelitian yang dilakukan oleh Trimble dkk (2011), pada perokok aktif yaitu mengkonsumsi rokok 10-15 batang perhari menyebabkan resiko neoplasia pada serviks. Wanita yang merokok memiliki risiko 4–13 kali lebih besar untuk mengalami ca serviks daripada wanita yang tidak merokok.

6. Penggunaan Pembersih Vagina (Douching)

Vagina yang sehat justru harus mengandung bakteri *Lactobacillus*, yang merupakan bakteri baik untuk menjaga keasaman vagina agar kuman tak mudah menginfeksi. Kebiasaan menggunakan cairan vagina (douching) akan memberantas bakteri *Lactobacillus* tersebut, sehingga vagina lebih rentan mengalami infeksi. Salah satunya adalah infeksi Human Papilloma Virus (HPV), yang menyebabkan kanker serviks. Penelitian yang dilakukan Neuman (2012) di Utah, Amerika Serikat menyatakan bahwa *douching* Penelitian yang dilakukan Dhorethea (2015), menyatakan bahwa cairan pembersih vagina/ douching yang beredar dipasaran berisi air dan campuran bahan seperti suka, baking soda atau iodium yang biasanya langsung digunakan wanita ke dalam vagina melalui tube. Kebiasaan ini akan mengganggu bakteri sehat (*lactobacillus*) yang sudah ada serta mengganggu keasaman vagina. Douching dapat mengubah lingkungan kimia vagina, membuat leher rahim lebih rentan terhadap perubahan patologis dan kanker serviks. Wanita yang melakukan douching vagina lebih dari sekali seminggu ternyata empat kali lebih rentan diserang kanker serviks (leher rahim). Douching bisa meningkatkan risiko infeksi HPV genital pada vagina. HPV telah diketahui sebagai salah satu pemicu kanker serviks.

7. Pemakaian Kontrasepsi Oral

Wanita yang telah menggunakan kontrasepsi oral (pil KB) selama 5 tahun atau lebih memiliki risiko lebih besar terkena kanker serviks dibandingkan wanita yang tidak pernah menggunakan kontrasepsi oral. Risiko lebih tinggi setelah 10 tahun penggunaan (*National Cancer Institute, 2012*). Menurut Hartmann (2002), Penggunaan kontrasepsi hormonal lebih dari 4 atau 5 tahun dapat meningkatkan risiko terkena kanker serviks 1,5-2,5 kali. Beberapa penelitian menunjukkan bahwa kontrasepsi oral menyebabkan wanita sensitif terhadap HPV yang dapat menyebabkan adanya peradangan pada genitalia sehingga berisiko untuk terjadi kanker serviks. Penggunaan pil KB dapat meningkatkan risiko kanker serviks. Analisis data yang dilakukan oleh *International agency for Research on Cancer (IARC)* pada tahun 2003, menemukan bahwa ada peningkatan risiko kanker serviks dengan penggunaan kontrasepsi oral dihentikan.

8. Riwayat HIV/ AIDS

Faktor risiko lainnya penyebab kanker serviks adalah kondisi immunosupresi atau menurunnya daya tahan tubuh. Daya tahan tubuh berperan penting dalam proses penghancuran sel-sel kanker serta menghambat pertumbuhan dan penyebarannya. Salah satu keadaan immunosupresi bisa ditemui pada penderita AIDS. Virus HIV pada penderita AIDS akan merusak fungsi kekebalan tubuh seseorang, sehingga wanita yang menderita AIDS memiliki risiko tinggi terkena infeksi HPV yang berkembang menjadi kanker serviks. Pada wanita penderita AIDS, perkembangan sel pra-kanker menjadi kanker yang biasanya memerlukan waktu beberapa tahun, dapat terjadi lebih cepat karena immunosupresi. Selain itu, kondisi seperti ini juga bisa ditemui pada wanita yang mengonsumsi obat penurun daya tahan tubuh, seperti wanita penderita penyakit autoimun (daya tahan tubuh yang menyerang organ tubuh sendiri karena menganggap organ tersebut sebagai musuh) atau wanita yang sedang menjalani transplantasi organ tubuh (Krisno, 2011).

9. Riwayat keluarga yang menderita kanker serviks

Riwayat keluarga seperti ibu dan saudara perempuan juga menentukan tingginya potensi terkena kanker serviks. Setidaknya risiko meningkat dua kali lipat di dibandingkan dengan yang tidak memiliki riwayat keluarga. Hal ini terjadi karena dalam riwayat keluarga terdapat sistem imun yang sama, sel yang dibawa oleh faktor keturunan, serta daya tahan tubuh dan faktor terinfeksi yang sama (Pusat info studi Kanker, 2014).

b. Target Hasil Kanker Serviks

Menurut Kartikawati (2013), diagnosis ditegakkan berdasarkan gejala dan hasil pemeriksaan sebagai berikut:

1. IVA (Inspeksi Visual Dengan Asam Asetat)

Metode pemeriksaan dengan mengoles serviks atau leher rahim dengan larutan asam asetat 3-5 % secara inspekulo. Kemudian diamati apakah ada kelainan seperti area berwarna putih. Dibutuhkan waktu 2-3 menit untuk melihat perubahan jaringan epitel. Jika tidak ada perubahan warna, maka dapat dianggap tidak ada infeksi pada serviks. Pemeriksaan ini dapat dilakukan di Puskesmas dengan harga yang relatif murah. Jika terlihat tanda yang mencurigakan, maka metode deteksi lainnya yang lebih lanjut harus dilakukan.

2. Pemeriksaan pap smear

Pap smear dapat mendeteksi sampai 90% kasus kanker serviks secara akurat dan dengan biaya yang tidak mahal, akibatnya angka kematian akibat kanker leher rahim pun menurun sampai lebih dari 50%. Setiap wanita yang telah aktif secara seksual sebaiknya menjalani *pap smear* secara teratur yaitu 1 kali setiap tahun. Apabila selama 3 kali berturut-turut menunjukkan hasil pemeriksaan yang normal, maka pemeriksaan *pap smear* bisa dilakukan setiap 2 atau 3 tahun sekali. Hasil pemeriksaan *pap smear* adalah sebagai berikut:

- a) Normal.
- b) Displasia ringan (perubahan dini yang belum bersifat ganas).
- c) Displasia berat (perubahan lanjut yang belum bersifat ganas).
- d) Karsinoma in situ (kanker terbatas pada lapisan serviks paling luar).

e) Kanker invasif (kanker telah menyebar ke lapisan serviks yang lebih dalam atau ke organ tubuh lainnya)

3. Thin Prep

Metode ini lebih akurat dibanding pap smear. Jika pap smear hanya mengambil dari sebagian dari sel-sel di serviks atau leher rahim, maka metode ini metode ini memeriksa seluruh bagian serviks atau leher rahim. Hasilnya lebih tepat dan akurat.

a) Pemeriksaan DNA HPV

Pemeriksaan ini dimasukkan pada skrining bersama-sama dengan Pap's smear untuk wanita dengan usia di atas 30 tahun. Penelitian dalam skala besar mendapatkan bahwa Pap's smear negatif disertai DNA HPV yang negatif mengindikasikan tidak akan ada CIN 3 (*Carcinoma Intraepithelial Neoplasia 3*) sebanyak hampir 100%. Kombinasi pemeriksaan ini dianjurkan untuk wanita dengan umur diatas 30 tahun karena prevalensi infeksi HPV menurun sejalan dengan waktu.

b) Kolposkopi

Jika semua tes sebelumnya menunjukkan adanya infeksi atau kejanggalan, prosedur koloskopi akan dilakukan dengan menggunakan alat yang dilengkapi lensa pembesar untuk mengamati bagian yang terinfeksi. Tujuannya untuk menentukan apakah ada lesi atau jaringan yang tidak normal pada serviks atau leher rahim. Jika ada yang tidak normal selanjutnya dilakukan tindakan biopsi.

c) Biopsi

Biopsi merupakan salah satu tindakan medis dimana pengambilan sel atau sebagian jaringan tubuh untuk dianalisa lebih lanjut. Misalnya untuk membedakan benjolan sebagai tumor jinak atau ganas (kanker). Biopsi dilakukan jika pada pemeriksaan panggul tampak suatu pertumbuhan atau luka pada serviks, atau jika hasil pemeriksaan pap smear menunjukkan suatu abnormalitas atau kanker.

d) Tes Schiller

Serviks diolesi dengan larutan yodium, sel yang sehat berubah menjadi coklat, sedangkan sel yang abnormal warnanya menjadi putih atau kuning.

2.1.2. Machine Learning

Menurut Shwartz dan David (2014) *Machine Learning* merupakan sebuah studi tentang algoritma untuk mempelajari sesuatu dalam melakukan beberapa hal tertentu yang dilakukan oleh manusia secara otomatis. Belajar dalam hal ini berkaitan dengan bagaimana menuntaskan berbagai tugas yang ada, atau membuat suatu prediksi kesimpulan baru yang akurat dari berbagai pola yang sudah dipelajari sebelumnya.

Machine Learning merupakan salah satu bidang yang termasuk dalam kecerdasan buatan yang dapat mempengaruhi berbagai aspek yang lain yaitu statistika, matematika dan berbagai aspek teoritikal dari komputer sains. Pada dasarnya *Machine Learning* memiliki tujuan untuk mempelajari sebuah algoritma untuk melakukan sistem belajar secara otomatis dengan kontribusi yang sangat minimal yang dilakukan oleh manusia pada umumnya. (Shwartz dan David, 2014)

Berikut beberapa contoh implementasi dari *Machine Learning* yang pada umumnya diterapkan (Shwartz dan David, 2014):

1. *Face Detection* : Sebuah sistem yang mampu mengenali wajah dari seseorang untuk memberikan representasi identitas dari orang tersebut.
2. *Spam Filtering* : Sebuah sistem yang memiliki kemampuan untuk melakukan penyaringan sebuah pesan apakah termasuk spam atau nonspam.
3. *Customer Segmentation* : Sebuah sistem yang melakukan sebuah prediksi pola perilaku konsumen yang berguna bagi produsen untuk implementasi dan pengaturan biaya dalam hal tertentu misalnya bidang pemasaran.

Machine Learning terbagi menjadi 2 macam konsep pembelajaran, yaitu pertama *Supervised Learning* yang merupakan teknik pembelajaran mesin yang membuat suatu fungsi berdasarkan data latihan yang sudah ada, dalam hal ini dapat dikatakan untuk teknik ini sudah tersedia data latihan secara detil dan teroptimasi dengan baik yang akan dijadikan sebuah model data saat dilakukan

proses ujicoba dengan data tes yang baru dan menghasilkan hasil keluaran yang sesuai diharapkan sebelumnya berdasarkan data latihan yang ada. Kedua adalah *Unsupervised Learning* yang merupakan teknik pembelajaran mesin yang berusaha untuk melakukan representasi pola sebuah input yang berasal dari data latihan dan salah satu yang menjadi perbedaan dengan *Supervised Learning* adalah tidak adanya pengoptimasian dari input data. Dalam *Machine Learning* teknik *Unsupervised Learning* menjadi esensial karena sistem kerja yang diberikan sama dengan cara kerja otak manusia dimana dalam proses pembelajaran tidak ada role model atau informasi dan contoh yang tersedia untuk dijadikan sebagai model dalam melakukan proses uji coba untuk penyelesaian sebuah masalah dengan data yang baru (Shwartz dan David, 2014).

a. **Supervised Learning**

Berikut beberapa contoh algoritma dari konsep pembelajaran *Supervised Learning* menurut Vishwanathan dan Smola (2010):

1. *Logistic Regression* : merupakan suatu teknik statistik yang juga digunakan untuk penyelesaian masalah dengan menghasilkan sebuah prediksi variabel baru berdasarkan beberapa variabel yang sudah ada dan ditentukan sebelumnya dan mencari hubungan keterkaitan antara variabel input data variabel yang sudah ditentukan dengan variabel output atau variabel prediksi yang baru.
2. *K-Nearest Neighbors Algorithm* (KNN) : merupakan sebuah metode yang digunakan dalam penyelesaian masalah optimasi sebuah objek dengan melakukan implementasi *feature space* dimana sebuah objek yang menjadi data *training* dan dijadikan sebuah *model data* diberikan pembobotan nilai dan direpresentasikan ke dalam n-dimensional vektor. Kemudian penyelesaian masalah dilakukan dengan mengukur jarak terdekat objek baru dengan model data yang ada di dalam n-dimensional vektor lalu dilakukan proses pemberian sebuah kategori pada objek baru tersebut.

b. Unsupervised Learning

Berikut beberapa contoh algoritma dari konsep pembelajaran *Unsupervised Learning* menurut Vishwanathan dan Smola (2010):

1. *Clustering* : suatu metode pembentukan pola dasar dari sebuah objek yang digunakan untuk melakukan penyelesaian permasalahan dalam kasus *machine learning* seperti optimasi atau kategorisasi objek ke dalam sebuah kelas / kategori. Terbentuknya pola dasar ini dilakukan dengan menggunakan beberapa *feature* yang ditentukan dan hal ini dilakukan karena data yang dijadikan sebagai *model data* belum terdefinisi kelompok datanya sehingga tidak bisa menentukan optimasi sebuah objek, setelah mendapatkan pola dasar yang terbentuk kemudian dapat dijadikan sebuah *model data* dan dijadikan data *training* dalam kasus kategorisasi objek baru.

2.1.3. Incomplete data

Incomplete data merupakan ketidaklengkapan data yang berpengaruh pada nilai statistik data. Kasus '*missing data*' ini akan mengakibatkan ketidaklengkapan data (*incomplete data*) dalam suatu model, sehingga akan menghambat analisis statistik yang akan dilakukan. *Missing data* akan menjadi masalah penting pada kasus optimasi dataset. Secara umum metode imputasi dan teknik optimasi adalah dua hal yang berbeda namun poin penting dari optimasi adalah bagaimana mendapatkan data pelatihan yang baik. Karena selain pemilihan metode yang tepat, *akurasi hasil* dari optimasi dipengaruhi oleh karakteristik dan kelengkapan *instance* dari sebuah data (Acuna, 2004). Sehingga dengan melakukan imputasi terhadap *missing data* maka hal tersebut secara langsung dapat mempengaruhi hasil optimasi. Sebaliknya, ketika sebuah kasus *missing data* diabaikan maka dapat dipastikan akan menjadikan sulit memperoleh akurasi yang tinggi untuk hasil optimasi walaupun digunakan algoritma optimasi yang paling handal sekalipun (Zhang, 2011). Beberapa penelitian menunjukkan bahwa penanganan *missing data* dengan menggunakan metode imputasi dapat

meningkatkan akurasi optimasi dibandingkan dengan tanpa imputasi (Farhangfari dkk, 2008) (Laencina dkk, 2009).

a. Pendekatan *Incomplete*

Masalah optimasi untuk data yang tidak lengkap merupakan tantangan bagi pengembangan penelitian data karena teknik yang paling banyak digunakan mengasumsikan data lengkap tetapi tidak menyesuaikan dengan masalah *Incomplete* data. Pada penelitian tersebut Yildirim melakukan perbandingan analisis substitusi nilai rata-rata dengan algoritma *ensemble* untuk mengatasi *incomplete* data serta evaluasi hasil berdasarkan ukuran akurasi dan waktu eksekusi (Yildirim., 2018).

b. Metode *Regression Imputation*

Dalam metode *regression imputation*, model regresi diestimasi untuk memprediksi nilai yang diamati dan dipasang untuk masing-masing variabel dengan nilai-nilai yang hilang (Yuan 2001).

Selanjutnya menduga data hilang dengan menghitung Nilai-nilai dari parameter $\hat{\beta}_i$ dan $\hat{\sigma}_i^2$ terlebih dahulu.

1. Menduga parameter $\hat{\beta}_i$ digunakan metode kuadrat terkecil persamaan, atau

$$\hat{\beta}_i = (x'x)^{-1}(x'y) \tag{2.1}$$

2. Selanjutnya mencari nilai variansi

$$\hat{\sigma}_i^2 = \frac{y'y - \hat{\beta}'_i x'y}{n-p} \tag{2.2}$$

Tahapan penggunaan imputasi pada variabel X_2 .

1. Menghitung nilai variansi Parameter baru β_* dan σ_{*m}^2 di peroleh dari

$$\sigma_{*m}^2 = \frac{\hat{\sigma}_i^2 (n-p)}{g_m} \tag{2.3}$$

g merupakan nilai acak dari distribusi chi-square dan n adalah nilai dari observasi yang tidak hilang.

2. Menghitung koefisien regresi

$$\beta_* = \beta + \sigma_{*m} L_{ij}^T Z \quad (2.4)$$

Di mana σ_{*m} merupakan akar dari σ_{*m}^2 yang diperoleh dari (1), Z merupakan vector variabel acak berukuran p yang berdistribusi normal baku $N(0,1)$ dan L_{ij}^T merupakan matriks segitiga atas dari dekomposisi Cholesky. Menurut Rencher (1934), dekomposisi Cholesky digunakan jika Matriks V adalah matriks simetris yaitu $V = V^T, V = (x'x)^{-1}$.

3. Sehingga nilai data hilang pada variabel X_2 akan diduga dengan menggunakan persamaan

$$Y_{mis} = \beta_{*0} + \beta_{*1} X_1 + sE \quad (2.5)$$

2.1.4. Reduksi

Reduksi merupakan suatu proses di mana data yang diperoleh dari lapangan tersebut dilakukan reduksi, dirangkum dipilih hal-hal yang pokok dan difokuskan pada hal-hal yang penting serta disusun secara sistematis dengan tujuan agar data tersebut menjadi lebih mudah dipahami dan dikendalikan. Data yang sudah direduksi akan memberikan gambaran yang lebih jelas tentang hasil penelitian di lapangan. Dalam proses reduksi ini peneliti tidak asal mengurangi data akan tetapi melakukan seleksi atau memilih data apa yang relevan dan bermakna. Memfokuskan pada pemecahan masalah penemuan pemaknaan atau menjawab pertanyaan penelitian, proses reduksi data berlangsung secara terus menerus selama penelitian berlangsung (Moleong, 2006: 288).

a. Pendekatan Reduksi Atribut

Permasalahan yang sering muncul adalah banyaknya atribut yang ada, berdampak pada lamanya waktu pengambilan keputusan. dalam beberapa tahun terakhir, banyak metode deteksi diusulkan dan diterapkan di lapangan untuk memberikan diagnosis tepat waktu, termasuk pendekatan berbasis data (Wu, W., & Zhou, H., 2017). pada penelitian sejenis Choudhury belum melakukan reduksi

atribut , tetapi menggunakan 3 teknik imbalanced data yaitu SMOTETomek method, under-sampling dan Oversampling (Choudhury, A., 2018).

b. Reduksi *Optimum Index Factor* (OIF)

Optimum Index Faktor (OIF) merupakan salah satu metode statistik untuk memilih kombinasi optimal berdasarkan total standar deviasi dan korelasi koefisien antar band. Adapun rumus OIF yaitu sebagai berikut (Sirait, P., & Arymurthy, 2010):

$$OIF = \frac{\sum_{h=1}^3 S_h}{\sum_{h=1}^3 abs(f,g)} \quad (2.6)$$

Keterangan :

Sh : Simpangan Baku untuk saluran h

abs (f.g) : nilai mutlak untuk sembarang saluran

OIF : Optimum Index Factor

2.1.5. *Genetic Algorithm*

Algoritma ini ditemukan di Universitas Michigan, Amerika Serikat oleh John Holland (1975) melalui sebuah penelitian dan dipopulerkan oleh salah satu muridnya, David Goldberg (1989). Dimana mendefinisikan *Genetic Algorithm* ini sebagai metode algoritma pencarian berdasarkan pada mekanisme seleksi alam dan genetik alam. *Genetic Algorithm* adalah algoritma yang berusaha menerapkan pemahaman mengenai evolusi alamiah pada tugas-tugas pemecahan-masalah (*problem solving*). Pendekatan yang diambil oleh algoritma ini adalah dengan menggabungkan secara acak berbagai pilihan solusi terbaik di dalam suatu kumpulan untuk mendapatkan generasi solusi terbaik berikutnya yaitu pada suatu kondisi yang memaksimalkan kecocokannya atau lazim disebut fitness (Desiani dkk , 2006).

Generasi ini akan merepresentasikan perbaikan-perbaikan pada populasi awalnya. Dengan melakukan proses ini secara berulang, algoritma ini diharapkan dapat mensimulasikan proses evolusioner. Pada akhirnya, akan didapatkan solusi-solusi yang paling tepat bagi permasalahan yang dihadapi. Untuk menggunakan *Genetic Algorithm*, solusi permasalahan direpresentasikan sebagai khromosom. Tiga aspek yang penting untuk penggunaan *Genetic Algorithm* (Desiani dkk , 2006):

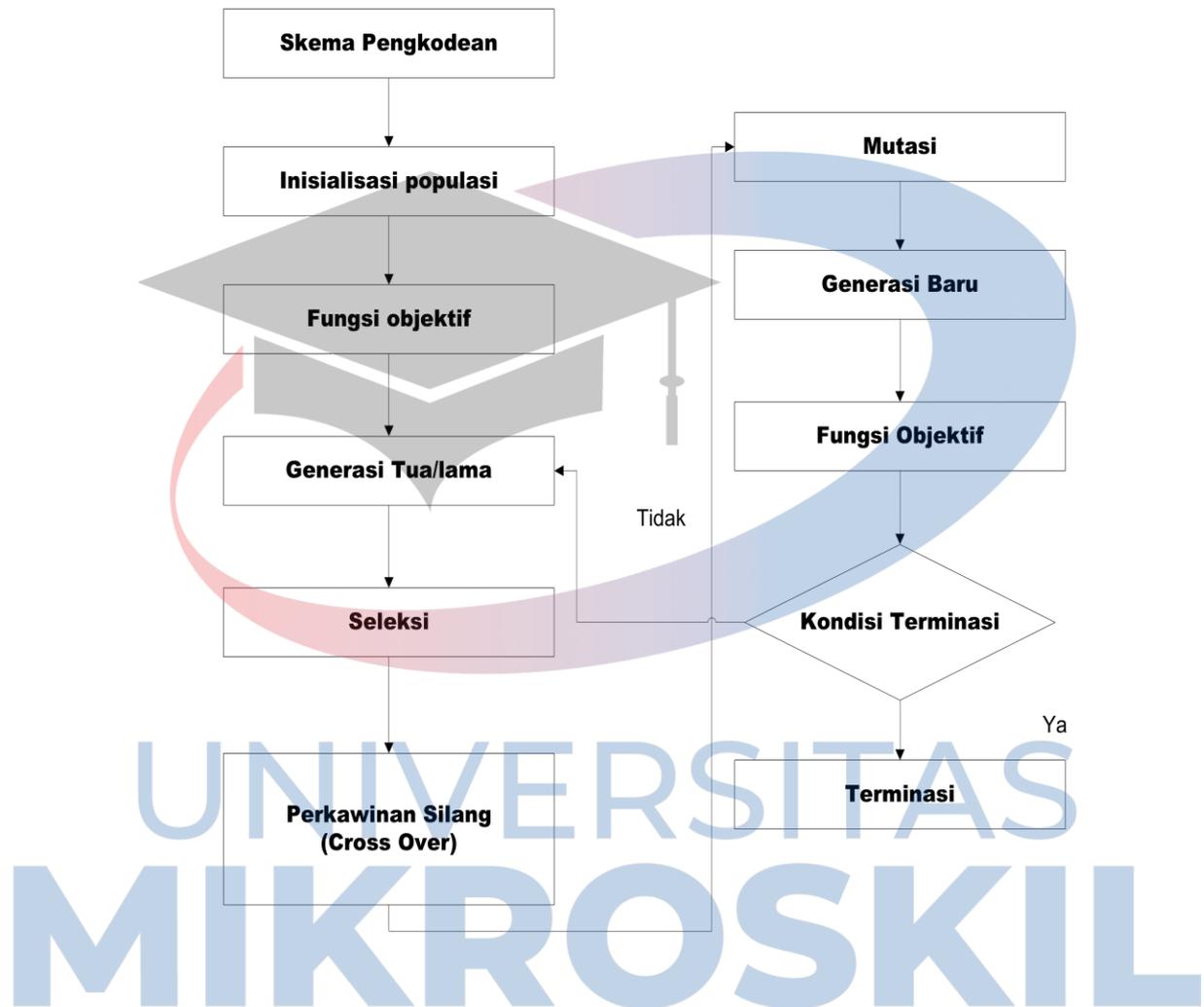
1. Defenisi fungsi fitness
2. Defenisi dan implementasi representasi genetik
3. Defenisi dan implementasi operasi genetik

Jika ketiga aspek di atas telah didefinisikan, *Genetic Algorithm* akan bekerja dengan baik. Tentu saja, *Genetic Algorithm* bukanlah solusi terbaik untuk memecahkan segala masalah. Sebagai contoh, metode tradisional telah diatur untuk mencari penyelesaian dari fungsi analitis convex yang “berperilaku baik” yang variabelnya sedikit. Pada kasus-kasus ini, metode berbasis kalkulus lebih unggul dari *Genetic Algorithm* karena metode ini dengan cepat menemukan solusi minimum ketika *Genetic Algorithm* masih menganalisa bobot dari populasi awal. Untuk problem-problem ini pengguna harus mengakui fakta dari pengalaman ini dan memakai metode tradisional yang lebih cepat tersebut (Desiani dkk , 2006).

Akan tetapi, banyak persoalan realistik yang berada di luar golongan ini. Selain itu, untuk persoalan yang tidak terlalu rumit, banyak cara yang lebih cepat dari *Genetic Algorithm*. Jumlah besar dari populasi solusi, yang merupakan keunggulan dari *Genetic Algorithm*, juga harus mengakui kekurangannya dalam kecepatan pada sekumpulan komputer yang dipasang secara seri-fitness function dari tiap solusi harus dievaluasi. Namun, bila tersedia komputer-komputer yang paralel, tiap prosesor dapat mengevaluasi fungsi yang terpisah pada saat yang bersamaan. Karena itulah, *Genetic Algorithm* sangat cocok untuk perhitungan yang paralel (Desiani dkk , 2006).

Genetic Algorithm memberikan suatu pilihan bagi penentuan nilai parameter dengan meniru cara reproduksi genetik, pembentukan kromosom baru

serta seleksi alami seperti yang terjadi pada makhluk hidup. *Genetic Algorithm* secara umum dapat diilustrasikan dalam diagram alir berikut ini (Desiani dkk , 2006):



Gambar 2.1 Diagram Alir *Genetic Algorithm*
(sumber : Desiani dkk , 2006.)

Menurut Golberg (1989) mengemukakan bahwa *Genetic Algorithm* mempunyai karakteristik-karakteristik yang perlu diketahui sehingga dapat terbedakan dari prosedur pencarian atau optimasi yang lain, yaitu:

1. *Genetic Algorithm* dengan pengkodean dari himpunan solusi permasalahan berdasarkan parameter yang telah ditetapkan dan bukan parameter itu sendiri.
2. *Genetic Algorithm* pencarian pada sebuah solusi dari sejumlah individu-individu yang merupakan solusi permasalahan bukan hanya dari sebuah individu.
3. *Genetic Algorithm* informasi fungsi objektif (fitness), sebagai cara untuk mengevaluasi individu yang mempunyai solusi terbaik, bukan turunan dari suatu fungsi.
4. *Genetic Algorithm* menggunakan aturan-aturan transisi peluang, bukan aturan-aturan deterministik.

Variabel dan parameter yang digunakan pada *Genetic Algorithm* adalah:

1. Fungsi fitness (fungsi tujuan) yang dimiliki oleh masing-masing individu untuk menentukan tingkat kesesuaian individu tersebut dengan kriteria yang ingin dicapai.
2. Populasi jumlah individu yang dilibatkan pada setiap generasi.
3. Probabilitas terjadinya persilangan (crossover) pada suatu generasi.
4. Probabilitas terjadinya mutasi pada setiap individu.
5. Jumlah generasi yang akan dibentuk yang menentukan lama penerapan *Genetic Algorithm*.

Secara umum struktur dari suatu *Genetic Algorithm* dapat mendefinisikan dengan langkah-langkah sebagai berikut:

1. Membangkitkan populasi awal Populasi awal ini dibangkitkan secara random sehingga didapatkan solusi awal. Populasi itu sendiri terdiri atas sejumlah kromosom yang merepresentasikan solusi yang diinginkan.
2. Membentuk generasi baru Untuk membentuk generasi baru, digunakan operator reproduksi/ seleksi, crossover dan mutasi. Proses ini dilakukan berulang-ulang sehingga didapatkan jumlah kromosom yang cukup untuk membentuk generasi baru dimana generasi baru ini merupakan representasi dari solusi baru. Generasi baru ini dikenal dengan istilah anak (offspring).

3. Evaluasi solusi Pada tiap generasi, kromosom akan melalui proses evaluasi dengan menggunakan alat ukur yang dinamakan fitness. Nilai fitness suatu kromosom menggambarkan kualitas kromosom dalam populasi tersebut. Proses ini akan mengevaluasi setiap populasi dengan menghitung nilai fitness setiap kromosom dan mengevaluasinya sampai terpenuhi kriteria berhenti. Bila kriteria berhenti belum terpenuhi maka akan dibentuk lagi generasi baru dengan mengulangi langkah 2. Beberapa kriteria berhenti sering digunakan antara lain: berhenti pada generasi tertentu, berhenti setelah dalam beberapa generasi berturut-turut didapatkan nilai fitness tertinggi tidak berubah, berhenti dalam n generasi tidak didapatkan nilai fitness yang lebih tinggi. Fungsi fitness tersebut sebagai berikut (Golberg, 1989):

$$Fitness = \frac{1}{1 + P_{inalti}} \quad (2.7)$$

Dimana:

$$P_{inalti} = \sum BP \sum NP$$

Dari persamaan diatas nilai fitness ditentukan oleh nilai penalty. Penalty tersebut menunjukkan jumlah pelanggaran kendala pada suatu kromosom. Semakin tinggi nilai fitness akan semakin besar kemungkinan kromosom tersebut terpilih ke generasi berikutnya. Jadi nilai penalty berbanding terbalik dengan nilai fitness, semakin kecil nilai penalty (jumlah pelanggaran) semakin besar nilai fitnessnya. Jadi fungsi fitness (Golberg , 1989):

$$Fitness = \frac{1}{1 + \sum BP \sum NP} \quad (2.8)$$

Keterangan:

Bp : Bobot Pelanggaran

Np : Indikator Pelanggaran

a. Pengkodean

Pengkodean adalah suatu teknik untuk menyatakan populasi awal sebagai calon solusi suatu masalah ke dalam suatu kromosom sebagai suatu kunci pokok persoalan ketika menggunakan *Genetic Algorithm*. Berdasarkan jenis symbol yang digunakan sebagai nilai suatu gen, metode pengkodean dapat dioptimalkan sebagai berikut (Gen, M., dan R. Cheng, 2000):

1. Pengkodean biner merupakan cara pengkodean yang paling umum digunakan karena adalah yang pertama kali digunakan dalam *Genetic Algorithm* oleh Holland. Keuntungan pengkodean ini adalah sederhana untuk diciptakan dan mudah dimanipulasi. Pengkodean biner memberikan banyak kemungkinan untuk kromosom walaupun dengan jumlah nilai-nilai yang mungkin terjadi pada suatu gen yang sedikit (0 dan 1). Di pihak lain, pengkodean biner sering tidak sesuai untuk banyak masalah dan kadang pengoreksian harus dilakukan setelah operasi crossover dan mutasi.
2. Pengkodean bilangan riil adalah suatu pengkodean bilangan dalam bentuk riil. Masalah optimisasi fungsi dan optimisasi kendala lebih tepat jika diselesaikan dengan pengkodean bilangan riil karena struktur topologi ruang genotif untuk pengkodean bilangan riil identik dengan ruang fenotifnya, sehingga mudah membentuk operator genetik yang efektif dengan cara memakai teknik yang dapat digunakan yang berasal dari metode konvensional.
3. Pengkodean bilangan bulat adalah metode yang mengkodekan bilangan dalam bentuk bilangan bulat. Pengkodean ini baik digunakan untuk masalah optimisasi kombinatorial.
4. Pengkodean struktur data adalah model pengkodean yang menggunakan struktur data. Pengkodean ini digunakan untuk masalah kehidupan yang lebih kompleks seperti perencanaan jalur robot, dan masalah pewarnaan Graph.

b. Operator Genetik

Genetic Algorithm merupakan proses pencarian yang heuristik dan acak sehingga penekanan pemilihan operator yang digunakan sangat menentukan keberhasilan *Genetic Algorithm* dalam menemukan solusi optimum suatu masalah yang diberikan. Hal yang harus diperhatikan adalah menghindari terjadinya konvergensi premature, yaitu mencapai solusi optimum yang belum waktunya, dalam arti bahwa solusi yang diperoleh adalah hasil optimum lokal (Kusumadewi, 2003).

Operator genetik yang digunakan setelah proses evaluasi tahap pertama membentuk populasi baru dari generasi sekarang. Operator-operator tersebut adalah operator seleksi, crossover dan mutasi (Desiani dkk, 2006).

c. Seleksi

Seleksi bertujuan memberikan kesempatan reproduksi yang lebih besar bagi anggota populasi yang paling fit. Langkah pertama dalam seleksi ini adalah pencarian nilai fitness. Masing-masing individu dalam suatu wadah seleksi akan menerima probabilitas reproduksi yang tergantung pada nilai objektif dirinya sendiri terhadap nilai objektif dari semua individu dalam wadah seleksi tersebut. Nilai fitness inilah yang nantinya akan digunakan pada tahap seleksi berikutnya (Kusumadewi, 2003).

Kemampuan *Genetic Algorithm* untuk memproduksi kromosom yang lebih baik secara progresif tergantung pada penekanan selektif (selective pressure) yang diterapkan ke populasi. Penekanan selektif dapat diterapkan dalam dua cara. Cara pertama adalah membuat lebih banyak kromosom anak yang dipelihara dalam populasi dan memilih hanya kromosom-kromosom terbaik bagi generasi berikut. Walaupun orang tua dipilih secara acak, metode ini akan terus menghasilkan kromosom yang lebih baik berhubungan dengan penekanan selektif yang diterapkan pada individu anak tersebut (Kusumadewi, 2003).

Cara lain menerapkan penekanan selektif adalah memilih orang tua yang lebih baik ketika membuat keturunan baru. Dengan metode ini, hanya kromosom sebanyak yang dipelihara dalam populasi yang perlu dibuat bagi generasi

berikutnya. Walaupun penekanan selektif tidak diterapkan ke level keturunan, metode ini akan terus menghasilkan kromosom yang lebih baik, karena adanya penekanan selektif yang diterapkan ke orangtua (Kusumadewi, 2003).

Ada beberapa metode untuk memilih kromosom yang sering digunakan antara lain adalah seleksi roda rolet (roulette wheel selection), seleksi ranking (rank selection) dan seleksi turnamen (tournament selection). Dalam penelitian ini, metode yang digunakan adalah seleksi roda rolet (roulette wheel selection). Pada seleksi ini, orang tua dipilih berdasarkan fitness mereka. Lebih baik kualitas suatu kromosom, lebih besar peluangnya untuk terpilih. Probabilitas suatu individu terpilih untuk crossover sebanding dengan fitness-nya. Cara penyeleksian ini merupakan peniruan dari permainan roda rolet (Kusumadewi, 2003).

d. Crossover

Crossover (perkawinan silang) bertujuan menambah keanekaragaman string dalam populasi dengan penyilangan antar-string yang diperoleh dari sebelumnya. Beberapa jenis crossover tersebut adalah (Kusumadewi, 2003):

a. Crossover 1-titik

Pada *crossover* dilakukan dengan memisahkan suatu string menjadi dua bagian dan selanjutnya salah satu bagian dipertukarkan dengan salah satu bagian dari string yang lain yang telah dipisahkan dengan cara yang sama. Proses yang demikian dinamakan operator *crossover* satu titik seperti diperlihatkan pada gambar berikut:

Tabel 2.1 Contoh Crossover 1-titik

Kromosom orang tua1	11001011
Kromosom orang tua2	11011111
Keturunan	11001111

b. Crossover 2-titik

Proses crossover ini dilakukan dengan memilih dua titik crossover. Kromosom keturunan kemudian dibentuk dengan barisan bit dari awal kromosom sampai titik crossover pertama disalin dari orangtua pertama,

bagian dari titik crossover pertama dan kedua disalin dari orangtua kedua, kemudian selebihnya disalin dari orangtua pertama lagi.

Tabel 2.2 Contoh Crossover 2-titik

Kromosom orang tua1	11001011
Kromosom orang tua2	11011111
Keturunan	11011111

c. Crossover seragam

Crossover seragam menghasilkan kromosom keturunan dengan menyalin bit-bit secara acak dari kedua orangtuanya.

Tabel 2.3 Contoh Crossover seragam

Kromosom orang tua1	11001011
Kromosom orang tua2	11011111
Keturunan	11011111

d. Mutasi

Mutasi merupakan proses mengubah nilai dari satu atau beberapa gen dalam suatu kromosom. Operasi *crossover* yang dilakukan pada kromosom dengan tujuan untuk memperoleh kromosom-kromosom baru sebagai kandidat solusi pada generasi mendatang dengan *fitness* yang lebih baik, dan lama-kelamaan menuju solusi optimum yang diinginkan. Akan tetapi, untuk mencapai hal ini, penekanan selektif juga memegang peranan yang penting. Jika dalam proses pemilihan kromosom-kromosom cenderung pada kromosom yang memiliki *fitness* yang tinggi saja, konvergensi *premature*, yaitu mencapai solusi yang optimal lokal sangat mudah terjadi (Murniati, 2009: 24).

Untuk menghindari konvergensi *premature* tersebut dan tetap menjaga perbedaan (*diversity*) kromosom-kromosom dalam populasi, selain melakukan penekanan selektif yang lebih efisien, operator mutasi juga dapat digunakan (Murniati, 2009: 24).

Proses mutasi dalam system biologi berlangsung dengan mengubah isi *allele* gen pada suatu *locus* dengan *allele* yang lain. Proses mutasi ini bersifat acak sehingga tidak selalu menjamin bahwa setelah proses mutasi akan diperoleh kromosom dengan *fitness* yang lebih baik. Operator mutasi merupakan operasi yang menyangkut satu kromosom tertentu. Beberapa cara operasi mutasi diterapkan dalam *Genetic Algorithm* menurut jenis pengkodean terhadap phenotype, antara lain (Desiani dkk , 2006):

a. Mutasi dalam Pengkodean Biner

Mutasi pada pengkodean biner merupakan operasi yang sangat sederhana. Proses yang dilakukan adalah menginversi nilai bit pada posisi tertentu yang terpilih secara acak (atau menggunakan skema tertentu) pada kromosom, yang disebut inverse bit.

Tabel 2.4 Contoh Mutasi pada pengkodean biner

Kromosom Sebelum Mutasi	1 0 0 1 0 1 1 1
Kromosom Setelah Mutasi	1 0 0 1 0 0 1 1

b. Mutasi dalam Pengkodean Permutasi

Proses mutasi yang dilakukan dalam pengkodean biner dengan mengubah langsung bit-bit pada kromosom tidak dapat dilakukan pada pengkodean permutasi karena konsistensi urutan permutasi harus diperhatikan. Salah satu cara yang dapat dilakukan adalah dengan memilih dua posisi (*locus*) dari kromosom dan kemudian nilainya saling dipertukarkan.

Tabel 2.5 Contoh Mutasi pada pengkodean permutasi

Kromosom Sebelum Mutasi	1 2 3 4 6 5 8 7 9
Kromosom Setelah Mutasi	1 2 7 4 6 5 8 3 9

c. Mutasi dalam Pengkodean

Nilai Mutasi pada pengkodean nilai hampir sama dengan yang dilakukan pada pengkodean biner, tetapi yang dilakukan bukan menginversi nilai bit. Penerapannya bergantung pada jenis nilai yang digunakan. Sebagai contoh untuk nilai riil, proses mutasi dapat dilakukan seperti yang dilakukan pada

pengkodean permutasi, dengan saling mempertukarkan nilai dua gen pada kromosom.

- d. Mutasi dalam Pengkodean Pohon Mutasi dalam pengkodean pohon dapat dilakukan antara lain dengan cara mengubah operator (+, -, *, /) atau nilai yang terkandung pada suatu verteks pohon yang dipilih. Atau, dapat juga dilakukan dengan memilih dua verteks dari pohon dan saling mempertukarkan operator atau nilainya.

Tidak setiap gen selalu dimutasi tetapi mutasi dikontrol dengan probabilitas tertentu yang disebut dengan mutation rate (probabilitas mutasi) dengan notasi P_m . Jenis operator mutasi antara lain (Desiani dkk, 2006):

a. Mutasi Terarah

Mutasi terarah tergantung dari informasi gen. Informasi gen tersebut berupa nilai pelanggaran gen (violation gen). Ini berarti bahwa setiap gen mempunyai peluang yang berbeda untuk terjadi mutasi. Gen yang mempunyai nilai pelanggaran yang lebih besar maka gen tersebut mempunyai peluang untuk terjadi mutasi. Mutasi ini menghubungkan nilai pelanggaran relatif (nilai pelanggaran suatu gen dibagi dengan nilai pelanggaran total suatu kromosom) dengan probabilitas terjadinya mutasi dari suatu gen pada kromosom. Hubungan tersebut dinyatakan secara matematis sebagai berikut:

$$nr(i) = \frac{n(i)}{1+n_{total}} \quad (2.9)$$

$$pm(i) = (1 + Nr(i))^2 P_m \quad (2.10)$$

Keterangan persamaan:

$nr(i)$: nilai pelanggaran relative gen ke-i

n_{total} : nilai pelanggaran total kromosom

$pm(i)$: probabilitas mutasi gen ke-i

pm : probabilitas mutasi

b. Mutasi Biasa

Mutasi biasa tidak tergantung dari informasi gen. Setiap gen mempunyai peluang yang sama untuk terjadi mutasi.

e. Parameter Genetik

Pengoperasian *Genetic Algorithm* dibutuhkan 4 parameter (Juniawati, 2003) yaitu:

1. Probabilitas Persilangan (*Crossover Probability*) Menunjukkan kemungkinan *crossover* terjadi antara 2 kromosom. Jika tidak terjadi *crossover* maka keturunannya akan sama persis dengan kromosom orangtua, tetapi tidak berarti generasi yang baru akan sama persis dengan generasi yang lama. Jika probabilitas *crossover* 100% maka semua keturunannya dihasilkan dari *crossover*. *Crossover* dilakukan dengan harapan bahwa kromosom yang baru akan lebih baik.
2. Probabilitas Mutasi (*Mutation Probability*) Menunjukkan kemungkinan mutasi terjadi pada gen-gen yang menyusun sebuah kromosom. Jika tidak terjadi mutasi maka keturunan yang dihasilkan setelah *crossover* tidak berubah. Jika terjadi mutasi bagian kromosom akan berubah. Jika probabilitas 100%, semua kromosom dimutasi. Jika probabilitasnya 0%, tidak ada yang mengalami mutasi.
3. Jumlah Individu Menunjukkan jumlah kromosom yang terdapat dalam populasi (dalam satu generasi). Jika hanya sedikit kromosom dalam populasi maka *Genetic Algorithm* akan mempunyai sedikit variasi kemungkinan untuk melakukan *crossover* antara orangtua karena hanya sebagian kecil dari *search space* yang dipakai. Sebaliknya jika terlalu banyak maka *Genetic Algorithm* akan berjalan lambat.

4. Jumlah Populasi Menentukan jumlah populasi atau banyaknya generasi yang dihasilkan, digunakan sebagai batas akhir proses seleksi, persilangan dan mutasi

2.2 Kerangka Pikir/Pola Pemecahan Masalah

Perkembangan dunia Teknologi yang mengarah ke bidang medis pada masa sekarang ini semakin diperhitungkan, salah satunya diagnosa penyakit berbasis sistem pakar. Dan teknologi semacam ini semakin banyak digunakan, seperti masalah kesehatan biasanya memprediksi beberapa penyakit parah pada tahap awal, misalnya, kanker serviks. Kanker serviks adalah salah satu penyakit yang semakin mengancam kesehatan wanita (Deng *et al.*, 2019) kanker serviks menjadi penyebab kematian yang signifikan di negara-negara berpenghasilan rendah, lebih dari setengah juta kasus per tahun, dan menewaskan lebih dari seperempat juta pada periode yang sama (Fernandes *et al.*, 2017).

Permasalahan yang sering muncul adalah kurangnya pengetahuan akan deteksi dini terkait jenis penyakit untuk penderita awal, sehingga menyebabkan penderita yang sudah stadium lanjut baru mengetahui jenis penyakit akibat keterlambatan pada proses pencegahan dini. Banyak Algoritma yang digunakan untuk menganalisis koleksi data faktor risiko kanker serviks, namun beberapa diantaranya mengalami masalah pada proses “*missing data*” karena adanya *incomplete* pada dataset sejenis. Pada penelitian ini *incomplete* data diselesaikan dengan metode *Regression Imputation* yaitu proses regresi linier, dimana data yang tidak lengkap akan dibandingkan dengan data terdekat yang sejenis. Namun terlalu banyak nya atribut juga akan mempengaruhi lamanya waktu pengambilan keputusan, oleh sebab itu dari 32 atribut yang di gunakan penulis bermaksud mereduksi atribut dengan metode komputasi *Optimum Index Factor* (OIF). *Optimum index factor* (OIF) merupakan perhitungan nilai statistik dengan

mengkombinasikan atribut sehingga didapatkan hasil yang optimal dari beberapa kombinasi yang ada.

Sebagian besar penelitian juga berusaha untuk mendapatkan metode yang tepat untuk meningkatkan nilai akurasi, salah satunya mengusulkan metode optimasi untuk beberapa dataset yang tersedia. Optimasi data medis dengan keakuratan hasil -melalui pendekatan optimasi *Genetic Algorithm* telah dilakukan pada penelitian Gorzałczany dan Rudzinski (2017), pada penelitian sejenis Wissanu Thungrut and Naruemon (2019) juga melakukan hal serupa yang mengoptimasikan dataset diabetes dengan metode *Genetic Algorithm*, hasil yang di dapat membuktikan bahwa metode optimasi *Genetic Algorithm* lebih unggul untuk pemecahan masalah sejenis, maka dari itu pada penelitian ini, untuk menambah nilai akurasi dan presisi hasil, penulis menggunakan metode optimasi *Genetic Algorithm*, yang mana diharapkan dapat memberikan hasil yang relatif lebih optimal dibandingkan penelitian sebelumnya (Choudhury, A. et al, 2018)

UNIVERSITAS MIKROSKIL