

## MODEL KLASIFIKASI DATA REMISI NARAPIDANA MENGUNAKAN ALGORITMA *K-NEAREST NEIGHBORS* DENGAN OPTIMASI ALGORITMA GENETIKA

*Classification Model of Prisoner Remission Data Using K-Nearest Neighbors  
Algorithm with Genetic Algorithm Optimization*

**Titik Khotiah\*<sup>1</sup>, Yuna Sugianela<sup>2</sup>, Dewi Setiowati<sup>3</sup>, Feri Arianto<sup>4</sup>**

<sup>1,2,3</sup> Teknologi Informasi, Teknik, ITB-AD Lamongan, Indonesia  
e-mail: \*<sup>1</sup>titikaye@gmail.com, <sup>2</sup>xxx@xxxx.xxx, <sup>3</sup>xxx@xxx.xxx

### **Abstrak**

Dalam sistem baru pembinaan narapidana, remisi merupakan pengurangan hukuman terhadap narapidana dan anak pidana yang berkelakuan baik dengan tujuan untuk berbuat baik dan segera menjalani kehidupan di masyarakat. Kriteria pemberian remisi perlu diperjelas sehingga dapat menutup peluang remisi menjadi komoditas. Untuk itu perlu dilakukan klasifikasi terhadap pemberian remisi bagi narapidana. Pendekatan melalui computer science diterapkan sebagai model yang dapat membantu Klasifikasi. Metode Klasifikasi yang akan diterapkan dalam penelitian ini yaitu algoritma *K-Nearest Neighbors* (*K-NN*). Salah satu kelemahan dari *K-NN* adalah pemilihan nilai *k*, jika *k* sangat kecil maka akan mengakibatkan hasil klasifikasi terasa kaku. Sebaliknya jika *k* terlalu besar dapat menyebabkan hasil klasifikasi semakin kabur. Model klasifikasi remisi tahanan pertama dengan *K-NN* tanpa optimasi bertujuan untuk menentukan nilai *k* optimal secara manual. Selanjutnya *K-NN* dioptimasi dengan Algoritma Genetika untuk menentukan nilai *k* secara optimal. Hasil dari proses analisa untuk menentukan remisi narapidana dengan Metode *K-NN* dioptimasi menggunakan Algoritma Genetika memiliki tingkat akurasi yang lebih tinggi dengan nilai *k* = 7 yaitu 91.36% dari pada penggunaan Metode *K-NN* yaitu sebesar 87.50 %. Hasil penerapan Algoritma Genetika pada Metode *K-NN* pada klasifikasi remisi narapidana, dengan terdiri dari 1.760 data dan 9 indikator dengan 1 label, di proses Nilai Crossover = 0.5, Fitness = ~ dan Mutation = -1, dari 9 indikator dengan 1 label, dihasilkan indikator X1 = Golongan Registrasi, X3 = Warga Negara, X4 = Agama, X5 = Lama Pidana, X6 = Jenis Kejahatan, X7 = Besaran Denda, X9 = Status JC. Sehingga atribut-atribut tersebut adalah atribut yang paling menentukan dalam ketepatan dan keefektifan klasifikasi penentuan Remisi Narapidana.

**Kata kunci**-Remisi Narapidana, *K-Nearest Neighbors*, Algoritma Genetika

### **Abstract**

In the new system of guiding prisoners, remission is a reduction in punishment of convicted convicts and criminal offspring with the purpose to be the best and immediately living life in society. The criteria for remission should be clarified so that it can close the opportunity for remission to become a commodity. For this reason, it's necessary to classify remissions for prisoners. The approach through computer science is applied as a model that can help classification. The classification method that will be applied in this research is the *K-Nearest Neighbors* (*K-NN*) algorithm. One of the weaknesses of *K-NN* is the selection of the value of *k*, if *k* is very small it will result in the classification results feel stiff. Conversely, if *k* is too large it can cause blurred classification results. The results of the analysis process to determine the remission of prisoners with the *K-NN* Method were optimized using the Genetic Algorithm which has a higher level of accuracy with a value of *k* = 7 that is 91.36% than the use of the *K-NN* Method that is equal to 87.50%. The results of the application of Genetic Algorithms to the *K-NN* Method in the classification of prisoners' remission, consisting of 1,760 data and 9 indicators with 1 label, in the process of Crossover Value = 0.5, Fitness = ~ and Mutation = -1, from 9 indicators with 1 label, generated indicator X1 = Registration Class, X3 = Citizen, X4 = Religion, X5 = Criminal Length, X6 = Type of Crime, X7 = Amount of Fines, X9 = Status of JC. So that these attributes are the most decisive attributes in the accuracy and effectiveness of the classification of determining Prisoners Remission.

**Keywords**-Prisoner Remission, *K-Nearest Neighbors*, Genetic Algorithms

## PENDAHULUAN

Keberadaan pembinaan pemasyarakatan di Indonesia sudah berlangsung selama hampir empat dekade yang sebelumnya disebut dengan istilah penjara. Selama ini lapas/rutan identik dengan sebutan tempat pemberian hukuman bagi para pelaku kejahatan sementara itu pelakunya disebut narapidana. Dalam sistem baru pembinaan terhadap narapidana, remisi ditepatkan sebagai sarana dukungan (salah satu) bagi narapidana dalam rangka membina dirinya sendiri. Dikarenakan, remisi tidak hanya berlaku sebagai hukuman atau seperti dalam sistem pemasyarakatan, tidak pula sebagai anugerah seperti dalam sistem kelembagaan pemasyarakatan, akan tetapi dinyatakan sebagai hak dan kewajiban narapidana.

Adapun instansi yang berwenang dalam pemberian remisi adalah Lembaga Pemasyarakatan. Lembaga Pemasyarakatan merupakan salah satu lembaga yang mempunyai wewenang untuk memberi perkiraan remisi dan pengusulan nama-nama narapidana yang selanjutnya akan menerima remisi dan kemudian Pihak Lembaga Pemasyarakatan melaporkan kepada Menteri Hukum dan Perundang-Undangan Republik Indonesia lantas ditetapkan dengan Keputusan Menteri Hukum dan Perundang-Undangan. Akan tetapi dalam pengembangannya, terdapat banyak masalah yang dihadapi Lembaga Pemasyarakatan terkhusus bagi pimpinan atau kepala dan juga petugas atau anggota Lembaga Pemasyarakatan. Masalah tersebut yaitu berkaitan dengan bagaimana proses pengolahan data-data yang terdapat di Lembaga Pemasyarakatan dan menjadikan sebuah keputusan akurat serta cepat apabila keputusan yang akan diperoleh berjumlah lebih banyak. Contohnya adalah sulit ditemukannya alternatif-alternatif putusan dalam rangka penentuan siapa nama-nama narapidana yang berhak untuk memperoleh remisi.

Seiring dengan semakin majunya teknologi khususnya basis data maka timbul teknologi-teknologi untuk aplikasi baru, salah satunya adalah teknologi data mining. *Data mining* dikelompokkan berdasarkan tugasnya, yaitu meliputi Estimasi, Deskripsi, Klasifikasi, Prediksi, Asosiasi dan Pengklusteran<sup>1</sup>. Penelitian tentang Model Klasifikasi Data Remisi Narapidana sampai saat ini masih belum ada. Penelitian sejenis yang pernah dilakukan oleh Eries Surya membahas tentang sistem pendukung keputusan pemberian remisi bagi tahanan<sup>2</sup>. Dalam penelitian tersebut adanya sistem pendukung keputusan digunakan dalam memproses *Profile Matching* dan analisa *gap* yang dirancang khusus berdasar pada data, norma-norma narapidana di suatu Lembaga Pemasyarakatan. Proses *Profile Matching* dibangun untuk penentuan rekomendasi narapidana pada sistem penentuan remisi narapidana dengan berdasarkan pada tiga kategori antara lain kapasitas kegiatan, sikap di Lembaga Pemasyarakatan dan perilaku narapidana.

Model perhitungan yang diterapkan pada penelitian tersebut yaitu model statistik terdiri atas pemetaan *gap* kelayakan. Pemetaan tersebut meliputi aspek kapasitas kegiatan, sikap di suatu Lembaga Pemasyarakatan dan perilaku. Selanjutnya yaitu model pemberian bobot nilai *gap* (selisih), proses hitung dan pengelompokkan *core factor* dan *secondary factor*, model hitung nilai masing-masing aspek, dan jumlah (ranking). Dari beberapa perhitungan yang digunakan mendapatkan hasil bahwa *ranking* narapidana sebagai hasil rekomendasi bagi pembuat keputusan untuk melakukan pemilihan narapidana mana yang cocok memperoleh remisi. Perangkat lunak yang digunakan adalah *Microsoft Access 2002* sebagai penyimpanan data dan *Microsoft Visual Basic 6.0* berperan dalam bahasa pemrograman.

Dengan adanya penelitian tersebut dapat memberi kemudahan dan kecepatan waktu dalam melakukan penentuan atau pemberian keputusan bagi pihak Lembaga Pemasyarakatan berkaitan dengan pemberian remisi bagi narapidana, akan tetapi dalam penelitian tersebut juga tidak lepas dari kelemahan-kelemahan yaitu data dalam sistem tidak mewakili seluruh populasi dan perhitungan sampel dilakukan secara manual. Selain itu, sistem tersebut juga akan bisa berjalan dengan baik jika didukung dengan sumber daya manusia berkompeten, terlatih dan juga ketersediaan perangkat keras secara memadai.

---

<sup>1</sup> Larose, Daniel T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley & Sons. Inc.

<sup>2</sup> Laksana, Eries Surya. 2006. *Sistem Pendukung Keputusan Pemberian Remisi bagi Tahanan*. Skripsi Thesis. STMIK AKAKOM Yogyakarta.

Dari kelemahan tersebut penulis berupaya untuk melakukan Penelitian pada konstelasi Model Klasifikasi Data Remisi Narapidana dengan pendekatan ilmu komputer. Adapun algoritma komputasi yang penulis terapkan dalam penelitian ini adalah algoritma klasifikasi dan optimasi. Beberapa algoritma klasifikasi yang lebih banyak dipergunakan berkaitan dengan proses pengembangan *data mining* yaitu algoritma C4.5, *K-Nearest Neighbors*, *Naive Bayes*, CART. Peneliti sebelumnya menyatakan: “suatu teknik klasifikasi yang dasar dan bersifat sederhana yaitu *K-Nearest Neighbors* (K NN).<sup>3</sup> Kelebihannya yakni tangguh pada data latih yang mempunyai banyak *noise* dan bagus apabila data ujinya besar, prosesnya mudah diterapkan apabila disejajarkan dengan algoritma lain<sup>4</sup>. K NN pun mampu menghasilkan akurasi yang baik pada berbagai bidang *dataset*. Salah satunya berkaitan dengan klasifikasi batik.<sup>5</sup> Dalam penelitian tersebut K NN mampu menghasilkan akurasi yang cukup baik. Beberapa hal yang menjadi pemikiran dalam penggunaan algoritma ini yaitu pemilihan nilai *k*, apabila nilai *k* sangat kecil maka dapat menyebabkan hasil klasifikasi terasa kaku. Sebaliknya jika terlalu besar dapat menyebabkan hasil klasifikasi semakin kabur. Seperti pada penelitian terkait pengklasifikasian penyakit diabetes *mellitus*.<sup>6</sup>

Metode penentuan *k* secara manual telah dilakukan oleh M Adib Al Karomi menghasilkan akurasi 78.15% pada  $k = 7$ .<sup>7</sup> Namun penelitian dengan metode K NN dengan nilai *k* yang kecil menghasilkan akurasi yang lebih besar 81.39%.<sup>8</sup> Tetapi kesederhanaan metode K NN menjadikan nilai plus sehingga dipilah oleh kebanyakan peneliti. Dennes Nur Dwi Irianto dan rekan melakukan penelitian tentang optimasi pada nilai *k* dalam metode K NN. Dengan penggabungan metode K NN dan Algoritma Genetika penelitian ini menghasilkan nilai *k* optimal = 13 dengan akurasi 88% Fitness 0.9.<sup>9</sup>

Berdasarkan kelemahan tersebut dalam penelitian ini penulis akan menerapkan model algoritma K NN untuk menangani permasalahan pada pemilihan nilai *k* dari algoritma tersebut. Metode algoritma K NN dan optimasi Algoritma Genetika dipilih karena merujuk pada penelitian sebelumnya bahwa K NN terbukti memiliki akurasi yang baik pada berbagai *dataset* sedangkan Algoritma Genetika juga baik dalam hal optimasi. Untuk itu penggabungan kedua algoritma tersebut akan diterapkan pada *dataset* baru yaitu data remisi narapidana karena sebelumnya belum ada penelitian terkait dengan klasifikasi remisi narapidana. Dalam hal ini Algoritma Genetika difungsikan untuk mengoptimalkan penentuan nilai *k* secara otomatis dari algoritma K NN. Dari nilai *k* optimal tersebut akan menghasilkan akurasi dari klasifikasi yang merupakan tujuan dari penelitian ini.

## METODE PENELITIAN

Pada penelitian ini, data yang diterapkan adalah data pemberian remisi bagi narapidana. Data remisi tersebut kemudian diolah menggunakan metode *data mining*. Metode *data mining* yang dipilih untuk penelitian ini adalah metode klasifikasi. Metode tersebut akan diterapkan pada data remisi narapidana. Adapun desain eksperimen pada penelitian ini meliputi: pengumpulan data, *preprocessing*,

---

<sup>3</sup> Parvin H, Alizadeh H, Bidgoli B M, *MKNN: Modified K-Nearest Neighbors*. Proceedings of the Word Congress on Engineering and Computer Science 2008 (WCECS 2008). San Francisco. 2008: 831-834.

<sup>4</sup> Han, J., Kamber, M., & Pei, J. 2012. *Data Mining: Concepts and Techniques*.

<sup>5</sup> Wijayanto, Hanang. 2015. *Klasifikasi Batik Menggunakan Metode K-Nearest Neighbors Berdasarkan Gray Level Co-Occurrence Matrices (GlcM)*, no. 5.

<sup>6</sup> Indrayanti., dkk. 2014. *Optimasi Parameter K pada Algoritma K-Nearest Neighbors untuk Klasifikasi Penyakit Diabetes Mellitus*. ISBN:978-602-118-5-1.

<sup>7</sup> Karomi, M Adib Al. 2015. *Optimasi Parameter K Pada Algoritma KNN untuk Klasifikasi Heregistrasi Mahasiswa Program Studi Teknik Informatika STMIK Widya Pratama Jl . Patriot 25 Pekalongan Email : Adib.comp@gmail.com.* IC-TECH X (285): 5

<sup>8</sup> Utami, Lilyani Asri. 2017. *Melalui Komparasi Algoritma Support Vector Machine dan K-Nearest Neighbors Berbasis Particle Swarm Optimization*. 13 (1): 103–12.

<sup>9</sup> Dennes Nur, Dwi Iriantoro, Candra Dewi, and Delvi Fitriani. 2018. *Klasifikasi Pada Penyakit Dental Caries Menggunakan Gabungan K-Nearest Neighbors dan Algoritme Genetika*. 2 (8): 2926–33.

esperiment dan pengujian model, evaluasi dan validasi hasil. Dalam penelitian ini pengumpulan data dilakukan di Lembaga Pemasarakatan Kelas I Semarang. Teknik pengumpulan data menggunakan metode observasi. Data yang diperoleh sebanyak 1760 record, terdiri dari 9 variabel atau atribut. *Preprocessing* yang dilakukan dalam penelitian ini yaitu mengkonversi data. Data awal yang digunakan adalah 1760 data yang bernilai kategorial. Sehingga data tersebut perlu diterjemahkan ke dalam format numerik agar memudahkan dalam pengolahan data secara keseluruhan. Sedangkan desain eksperimen dan pengujian model dalam penelitian ini meliputi: pencarian nilai  $k$ , K-NN dioptimasi dengan GA, K-NN dioptimasi melalui penentuan nilai  $k$  dengan GA. Selanjutnya proses evaluasi penghitungan dengan *Confussion Matrix* untuk memperoleh hasil akurasi.

## HASIL DAN PEMBAHASAN

### 1) Hasil Pengumpulan Data

Data yang digunakan pada penelitian ini adalah data remisi narapidana pada tahun 2018 di Lembaga Pemasarakatan Kelas I Semarang. Adapun *dataset* tersebut terdiri dari 1760 record dan 9 atribut dengan 1 Label. Dari data yang ada, masih ditentukan berberapa data yang tidak memiliki nilai sehingga perlu dilakukan tahapan *preprocessing* pada data tersebut.

### 2) *Preprocessing*

Pada tahap ini, tujuan pengolahan data awal untuk mengetahui apakah dataset yang akan digunakan dalam perhitungan algoritma *K-Nearest Neighbors* dan *Algoritma Genetika* sesuai dengan aturan pada proses komputasi yang ada, sehingga perlu dilakukan tahapan konversi data untuk mengetahui jumlah sebaran data dari masing-masing atribut.

Pada tahap ini data polynominal akan dikonversi menjadi data numerik agar dapat dihitung secara komputasi. Adapun tahapan-tahapan untuk mengkonversi ke data numerik sebagai berikut :

1. Menghitung kemunculan data pada setiap atribut.
2. Menentukan dan mengurutkan *ranking* berdasarkan nilai terbesar hingga terkecil pada frekuensi kemunculan data.

Untuk selanjutnya atribut yang memiliki tipe data *polynominal* di atas dilakukan *ascending* menjadi *descending*. Kemudian data yang telah *didescending* dinormalisasi dengan menggunakan persamaan pada rumus berikut ini :

$$\begin{aligned}
 X^1 &= \frac{x - \text{min value}}{\text{max value} - \text{min value}} \\
 &= \frac{45447 - 3572}{56174 - 3572} \\
 &= 0,7960725
 \end{aligned}$$

Keterangan :

Data awal	: 45447
<i>Min Value</i>	: 3572
<i>Max Value</i>	: 56174
Hasil Normalisasi	: 0,7960725

### 3) Analisa dengan Algoritma *K-Nearest Neighbors*

Adapun cara penerapan dari perhitungan K NN adalah sebagai berikut :

- a) Menentukan nilai  $k=1$
- b) Menentukan data *training* dan *testing*
- c) Hitung *distance* data *testing* ke setiap data *training*

Dalam penelitian ini Nilai  $k$  yang digunakan yakni  $k = 1$  sampai dengan  $k = 13$ . Analisa data dalam penelitian ini menggunakan *Tools* untuk mengetahui data miniing yaitu *Rapidminer*. Dalam *Rapidminer* data akan diolah dengan melakukan pemilihan data yang akan diuji dan pengujian dalam Algoritma K NN.

Dari skema tersebut dimana hasil perhitungan K NN yang kemudian di validasi dengan

*Confusion Matrix*. Berdasarkan *Confusion Matrix* ditentukan nilai akurasi hasil perhitungan K NN dengan K=12 dan divalidasi menggunakan *10-Fold Cross Validation* menghasilkan nilai sebesar 87, 78 %

#### 4) Optimasi Nilai *k* Menggunakan Algoritma Genetika

Pada penelitian ini pembentukan *Chrommosome* dilakukan untuk menentukan *gen-gen* yang dijadikan sebagai uji coba, pembentukan *Chrommosome* diambil dari *dataset* Remisi Narapidana telah dinormalisasi. Pemberian nilai inisialisasi secara random dengan range 0- pada populasi sebanyak sembilan *Chrommosome* dan sembilan *gen*. Evaluasi *Chrommosome* yaitu dengan membangkitkan nilai fungsi\_objektif [F/O] pada tiap-tiap *Chrommosome* yang telah dibangkitkan:  $F/O [i] = k1\_1 - k1\_2 - k1\_3 - k1\_4 - k1\_5. k1\_6 - k1\_7 - k1\_8 - k1\_9$ . Proses seleksi yang dilakukan yaitu *chrommosome* yang mempunyai fungsi\_objektif kecil akan mempunyai mempunyai nilai probabilitas yang tinggi atau kemungkinan besar akan terpilih. Adapun persamaan yang digunakan untuk fungsi *Fitness* :  $Fitness[i] = (1/(1+fungsi\_objektif))$ .

Hasil yang diperoleh dari pencarian nilai probabilitas yaitu yang mempunyai nilai *Fitness* yang paling besar adalah P [6] = 0,230 sehingga *Chrommosome* ini terpilih pada generasi selanjutnya. Setelah nilai *Cumulative Probabilitasnya* diketahui, selanjutnya melakukan seleksi dengan menggunakan *Roulete Wheel*. Langkah awal sebelum proses seleksi adalah membangkitkan bilangan acak R dalam range 0-1. Jika  $R[k] < C[1]$  maka *Chrommosome* 1 dipilih sebagai induk, selain itu pilih *Chrommosome* ke- *k* sebagai induk dengan syarat  $C[k-] < R < C[k]$ . Setelah itu *Roulete Wheel* diputar sebanyak jumlah populasi yang ada pada penelitian ini yaitu 9, sehingga harus mengulang sebanyak 9 kali putaran, sebelumnya membangkitkan bilangan acak R untuk setiap kali putaran dipilih satu *Chrommosome* untuk populasi baru.

Pada bilangan acak pertama R[1] lebih kecil dari C[1] maka pilih *Chrommosome* [1] sebagai *Chrommosome* baru pada populasi selanjutnya, cara tersebut di atas diulangi hingga didapatkan hasil populasi baru. Proses yang dilakukan setelah proses seleksi adalah *Crossover* (kawin silang). Salah satu metode yang dapat digunakan adalah *one-cut point*, yaitu menentukan *allele* secara acak pada satu posisi dalam *Chrommosome* induk kemudian saling menukar *gen*. *Chrommosome* yang menjadi induk dipilih secara acak dan jumlah *Chrommosome* yang mengalami *Crossover* dipengaruhi oleh parameter *Crossover\_Probability* (*cp*). Adapun *Crossover\_Rate* adalah 25% maka diharapkan dalam satu generasi ada 50%, 4 *Chrommosome* dari satu generasi mengalami proses *Crossover*. Langkah awal yang dilakukan adalah membangkitkan bilangan R secara *random* sebanyak jumlah populasi yaitu 9. Kemudian akan dipilih *chrommosome* ke *k* sebagai induk jika  $R[k] < cp$ . Dari bilangan acak R di atas R[2], R[5] dan R[6] lebih kecil dari *cp*[25%] maka diperoleh *chrommosome* yang dipilih adalah *chrommosome* [2], *chrommosome* [4], *chrommosome* [5], dan *chrommosome* [10].

Selanjutnya menentukan posisi titik potong (*one-cut-point*). Pada penentuan posisi titik potong *Crossover* dengan cara membangkitkan bilangan acak dengan *range* 1-[*n-Chrommosome*] *one-cut-point* yang dibangkitkan secara acak adalah 1–4. Langkah selanjutnya setelah *Crossover* adalah mutasi. Proses mutasi yang dilakukan yaitu dengan mengganti salah satu *gen* yang terpilih dengan nilai baru yang diperoleh secara acak. Penentuan jumlah *Chrommosome* yang mengalami mutasi dalam satu populasi ditentukan oleh parameter *Mutation\_Rate* dengan langkah pertama yaitu menghitung panjang total\_gen yang ada pada satu populasi.

$$\begin{aligned} \text{Total\_gen} &= (\text{jumlah gen dalam kromosom}) * \text{jumlah populasi} \\ &= 9 * 1.760 \\ &= 15.840 \end{aligned}$$

#### Keterangan:

Jumlah *gen* dalam *Chrommosome* : 9

Jumlah populasi : 1.760

Adapun cara untuk menentukan posisi titik *gen* yang mengalami mutasi yaitu dengan membangkitkan bilangan integer secara acak antara 1 sampai total\_gen, yaitu 1 sampai 15.840. Jika bilangan acak yang dibangkitkan lebih kecil dari pada variabel *Mutation\_Rate* (*pm*) , maka pilih posisi tersebut sebagai sub-*Chrommosome* yang mengalami mutasi. Kita tentukan *pm* semisal 20% diharapkan ada 20% dari total\_gen yang mengalami populasi.

#### Informasi Artikel:

Submitted: Mei2022, Accepted: Mei2022, Published: Mei2022

ISSN: 2685-4902(media online), Website: <http://jurnal.umus.ac.id/index/php/intech>

mengalami mutasi pada *gen* 25, 234, 460, 470 dan 1112, sehingga posisi yang mengalami mutasi adalah *Chrommosome* 1 pada *gen* 1 dan 9, *Chrommosome* 4 pada *gen* 5 dan 2 dan *Chrommosome* 2 pada *gen* 7. Kemudian pada titik *gen* yang mengalami mutasi tersebut diganti dengan nilai *random* yang telah ditentukan. Sebelumnya bangkitkan dulu nilai *random* dengan rentang 0-1. Hasil mutasi yaitu titik mutasi terjadi pada *chrommosome* [1] pada *gen* 2 dan 5, *chrommosome* [4] pada *gen* 2 dan 5 dan *chrommosome* [5] pada *gen* 2 dan 5, *chrommosome* [6] pada *gen* 1 dan 5, *chrommosome* [8] pada *gen* 4, dan 5, dan pada *chrommosome* [9] pada *gen* 3.

Apabila proses mutasi telah diselesaikan maka berarti satu iterasi dalam algoritma genetika yang disebut dengan satu generasi telah selesai, untuk mencari nilai yang maksimal atau minimum, *chrommosome* - *chrommosome* akan mengalami proses yang sama seperti generasi sebelumnya. Proses ini terus berulang hingga sejumlah generasi memperoleh *chrommosome* yang terbaik. Setelah mendapatkan *chrommosome* terbaik maka nilai tersebut digunakan untuk nilai bobot optimal pada atribut input pada *KNN* dan *Algoritma Genetika* yang selanjutnya digunakan pada tahapan pengujian. Adapun nilai bobot yang didapatkan dengan menggunakan *tools Rapidminer* adalah seperti yang ditampilkan pada tabel 4.11 di bawah ini :

**Tabel 1.**  
**Nilai Bobot *Tools Rapidminer***

inst#	Atribut	Weight
1	X1	1.0
3	X3	1.0
4	X4	1.0
5	X5	1.0
6	X6	1.0
7	X7	1.0
9	X9	1.0

Dalam penelitian ini Nilai *k* yang digunakan yakni  $k = 1$  sampai dengan  $k = 13$ . Analisa data dalam penelitian ini menggunakan *Tools* untuk mengetahui *data mining* yaitu *Rapidminer*. Dalam rapid miner data akan diolah dengan melakukan pemilihan data yang akan diuji dan pengujian dalam Algoritma K NN dengan Algoritma Genetika. Dari skema tersebut dimana hasil perhitungan K NN yang kemudian di validasi dengan *Confusion Matrix* menghasilkan tabel *Confusion Matrix* sebagai berikut, dari perhitungan untuk  $K = 1$  sampai dengan  $K = 13$ .

**Tabel 2.**  
**Nilai *Confusion Matrix* Metode K NN dengan Algoritma Genetika**

<i>kNN</i>	<i>Confusion Matrix</i>
$k=1$	86.65% +/- 2.75%
$K=2$	85.51% +/- 1.88%
$k=3$	86.82% +/- 1.22%
$k=4$	88.18% +/- 2.02%
$k=5$	89.32% +/- 1.89%
$k=6$	89.43% +/- 1.26%
$K=7$	89.72% +/- 1.15%
$k=8$	89.49% +/- 2.24%
$k=9$	89.32% +/- 3.25%
$k=10$	89.49% +/- 1.82%
$k=11$	88.86% +/- 1.97%
$K=12$	88.75% +/- 2.14%
$k=13$	88.69% +/- 2.12%

Berdasarkan *Confusion Matrix* pada tabel dapat di tentukan nilai akurasi hasil

perhitungan K NN berbasis Algoritma Genetika dengan  $k = 7$  dan divalidasi menggunakan *10-Fold Cross Validation* adalah sebagai berikut :

	<i>True Dapat</i>	<i>True Tidak Dapat</i>	<i>Class Precision</i>
Pred. dapat	328	75	81.39%
Pred. tidak dapat	106	1251	92.19%
<i>Class recall</i>	75.58%	94.34%	

Hasil akurasi K NN berbasis Algoritma Genetika yang dicapai menggunakan *10-Fold Cross Validation* adalah 89, 72 %.

#### 5) Analisa Klasifikasi K NN dengan Algoritma Genetika

Setelah menentukan Nilai *Crossover*, *Fitness* dan *Mutation* dan menghitung nilai  $k$  terbaik, yaitu  $k = 7$  dengan Akurasi terbesar 89.72% +/- 1.15%. Untuk proses selanjutnya klasifikasi akan dilakukan dengan metode K NN dan Algoritma Genetika. Dalam proses tersebut menghasilkan bobot atribut (*weight*) berbeda dari yang lain, dilambangkan dengan 0 dan 1 untuk bobot yang berbeda, untuk bobot yang berbeda sesuai tujuan dalam penelitian ini maka atribut tersebut mendapatkan perhatian khusus agar klasifikasi remisi narapidana dapat tepat sasaran dan efektif. dapat dilihat pada tabel 4.13 Data bobot atribut hasil klasifikasi KNN dan algoritma genetika berikut :

**Tabel3.**  
**Data Bobot Atribut Hasil Klasifikasi K NN dan Algoritma Genetika**

inst#	Atribut	Weight
1	Golongan Registrasi (X1)	1.0
2	Usia (X2)	0.0
3	Warga Negara (X3)	1.0
4	Agama (X4)	1.0
5	Lama Pidana (X5)	1.0
6	Jenis Kejahatan (X6)	1.0
7	Besaran Denda (X7)	1.0
8	Tgl BA 8 (X8)	0.0
9	Status (X9)	1.0

Berdasarkan hasil pada tabel diatas yaitu dengan Nilai *Crossover*, *Fitness* dan *Mutation* dan menghitung nilai K terbaik, yaitu K=7 dengan Akurasi terbesar 89.72% +/- 1.15%, klasifikasi data remisi narapidana dengan Metode K NN dan Algoritma Genetika yang terdiri dari 1.760 data dan 9 indikator dengan 1 label, bobot (*weight*) atribut yang berbeda yang memunculkan nilai 1 dihasilkan pada X1 = Golongan Registrasi, X3 = Warga Negara, X4 = Agama, X5 = Lama Pidana, X6 = Jenis Kejahatan, X7 = Besaran Denda, X9 = Status JC. Sehingga atribut-atribut tersebut adalah atribut yang paling menentukan dalam ketepatan dan ke efektifan klasifikasi penentuan remisi narapidana.

#### 6) Analisa Hasil Komparasi

Hasil dari pengujian metode yang dilakukan adalah untuk klasifikasi data remisi narapidana dengan *K-Neareast Neighbors* dan *K-Neareast Neighbors* berbasis *Algoritma Genetika* dalam untuk menentukan nilai komparasi *accuracy*, Maka dapat rangkum hasilnya pada tabel 4.15.

**Tabel 4.**  
**Tabel Hasil Perbandingan Performance Algoritma**

	<i>K-Neareast Neighbors</i>	<i>K-Neareast Neighbors</i> berbasis <i>Algoritma Genetika</i>	<b>Kemajuan</b>
<b>Accuracy</b>	87.78%	89.72%	<b>1.94 %</b>

Hasil dari proses analisa di atas untuk menentukan klasifikasi data remisi narapidana

dengan Metode K NN dan Algoritma Genetika yang terdiri dari 1.760 data dan 9 indikator dengan 1 label, didapatkan hasil analisa Algoritma K NN, dengan  $k$  yang terbaik adalah  $k = 12$ , yaitu dengan Akurasi terbesar 87.78% +/- 2.89%, kemudian K NN dan Algoritma Genetika dengan Nilai *Crossover* = 0.5, *Fitness* = ~ dan *Mutation* = -1 dan menghitung nilai  $k$  terbaik, yaitu  $k = 7$  dengan Akurasi terbesar 89.72% +/- 1.15%, untuk bobot (weight) atribut yang berbeda yang memunculkan nilai 1 dihasilkan pada atribut X1 = Golongan Registrasi, X3 = Warga Negara, X4 = Agama, X5 = Lama Pidana, X6 = Jenis Kejahatan, X7 = Besaran Denda, X9 = Status JC. Sehingga atribut-atribut tersebut adalah atribut yang paling menentukan dalam ketepatan dan keefektifan klasifikasi penentuan Remisi Narapidana.

## KESIMPULAN

Dari hasil Analisa dan komputasi beberapa metode dan mengacu pada tujuan penelitian, dapat disimpulkan bahwa Metode K NN menggunakan metode Algoritma Genetika yang digunakan untuk menghitung dan menentukan K dan meningkatkan akurasi pendataan remisi narapidana menghasilkan beberapa kesimpulan sebagai berikut:

- 1) Hasil dari proses analisa untuk menentukan remisi narapidana yang terdiri dari 1.760 data dan 9 indikator dengan 1 label Metode K NN yaitu sebesar 87.78 %
- 2) Hasil penerapan Algoritma Genetika pada Metode K NN dengan Algoritma Genetika pada klasifikasi remisi narapidana, dengan terdiri dari 1.760 data dan 9 indikator dengan 1 label, memiliki tingkat akurasi yang lebih tinggi yaitu 89.72 %, dari 9 indikator dengan 1 label, dihasilkan indikator X1 = Golongan Registrasi, X3 = Warga Negara, X4 = Agama, X5 = Lama Pidana, X6 = Jenis Kejahatan, X7 = Besaran Denda, X9 = Status JC. Sehingga atribut-atribut tersebut adalah atribut yang paling menentukan dalam ketepatan dan keefektifan klasifikasi penentuan Remisi Narapidana.

## DAFTAR PUSTAKA

- [1] Larose, Daniel T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*, John Willey & Sons. Inc
- [2] Parvin H, Alizadeh H, Bidgoli B M, *MKNN: Modified K-Nearest Neighbors*. Proceedings of the World Congress on Engineering and Computer Science 2008 (WCECS 2008). San Francisco. 2008: 831-834.
- [3] Ngai E W T, Hu Y, Wong Y H, Chen Y, Sun X. *The Application of Data Mining Techniques in Financial Fraud Detection: A Classification Framework and An Academic Review of Literature*. Decision Support Systems. 2011; 50(3): 559-569.
- [4] Laksana, Eries Surya. 2006. *Sistem Pendukung Keputusan Pemberian Remisi bagi Tahanan*. Skripsi Thesis. STMIK AKAKOM Yogyakarta.
- [5] Han, J., Kamber, M., & Pei, J. 2012. *Data Mining: Concepts and Techniques*.
- [6] Wijayanto, Hanang. 2015. *Klasifikasi Batik Menggunakan Metode K-Nearest Neighbors Berdasarkan Gray Level Co-Occurrence Matrices (GlcM)*, no. 5.
- [7] Indrayanti., dkk. 2014. *Optimasi Parameter K pada Algoritma K-Nearest Neighbors untuk Klasifikasi Penyakit Diabetes Mellitus*. ISBN:978-602-118-5-1.
- [8] Karomi, M Adib Al. 2015. *Optimasi Parameter K Pada Algoritma KNN untuk Klasifikasi Heregistrasi Mahasiswa Program Studi Teknik Informatika STMIK Widya Pratama Jl . Patriot 25 Pekalongan Email : Adib.comp@gmail.com.* IC-TECH X (285): 5
- [9] Utami, Lilyani Asri. 2017. *Melalui Komparasi Algoritma Support Vector Machine dan K-Nearest Neighbors Berbasis Particle Swarm Optimization*. 13 (1): 103–12.
- [10] Dennes Nur, Dwi Iriantoro, Candra Dewi, and Delvi Fitriani. 2018. *Klasifikasi Pada Penyakit Dental Caries Menggunakan Gabungan K-Nearest Neighbors dan Algoritme Genetika*. 2 (8): 2926–33.