

## Penerapan Optimasi Portofolio Menggunakan *Strength Pareto Evolutionary Algorithm 2* (SPEA2) untuk Pemilihan Molekul Obat dalam Penghambatan PTP1B pada Penyakit Diabetes Melitus

Yulia Octaviany<sup>1</sup>, Deni Saepudin<sup>2</sup>, Isman Kurniawan<sup>3</sup>

<sup>1,2,3</sup>Fakultas Informatika, Universitas Telkom, Bandung

<sup>1</sup>yuliaoctaviany@students.telkomuniversity.ac.id, <sup>2</sup>denisaepudin@telkomuniversity.ac.id,

<sup>3</sup>ismankrn@telkomuniversity.ac.id

### Abstrak

Pemilihan molekul dalam *drug discovery* memiliki kemiripan dengan masalah dalam portofolio saham. Oleh karena itu, metode portofolio saham dapat diterapkan dalam masalah pemilihan molekul ini. Penelitian ini menggunakan nilai *bio-activity* IC50 untuk menghambat pertumbuhan protein PTP1B pada penyakit diabetes mellitus dengan metode optimasi portofolio. Algoritma SPEA2 digunakan untuk menentukan himpunan pareto-optimal untuk permasalahan optimasi *multi-objective*. Terdapat beberapa parameter yang digunakan seperti ukuran populasi, ukuran arsip, maksimum generasi, probabilitas *crossover* dan probabilitas mutasi. Solusi yang didapat dari algoritma SPEA2 selanjutnya akan dibentuk *Efficient Frontier* yang berbentuk sebuah grafik dua dimensi dimana tiap sumbunya menggambarkan nilai *diversity* dan nilai probabilitas keberhasilan dari portofolio. Data yang digunakan dalam penelitian ini diperoleh dari [www.ebi.ac.uk/chembl](http://www.ebi.ac.uk/chembl) sebanyak 1452 buah data. Pengujian dilakukan sebanyak 3 kali dengan masing – masing pengujian digunakan jumlah molekul yang berbeda, yaitu 5 molekul, 10 molekul dan 20 molekul. Dalam setiap pengujian juga menggunakan jumlah generasi yang berbeda yaitu sebanyak 100, 200 dan 300. Hasil yang didapatkan menunjukkan bahwa semakin besar jumlah generasi tidak mempengaruhi hasil grafik. Karena pada beberapa pengujian, pada jumlah generasi tertentu sudah didapatkan hasil yang cukup optimal karena tidak terjadi perubahan yang signifikan. Jumlah molekul juga mempengaruhi nilai *diversity* dan nilai *expected return*, dimana semakin besar jumlah molekul yang digunakan, nilai *expected return* dan *diversity* yang dihasilkan juga semakin besar. Rentang nilai *diversity* dan nilai *expected return* juga dipengaruhi oleh jumlah molekul.

Kata kunci : portofolio, *drug discovery*, SPEA2, *efficient frontier*

### Abstract

Molecular selection in drug discovery bears resemblance to problems in the stock portfolio. Therefore, the stock portfolio method can be applied in this molecular selection problem. This study used the IC50 bio-activity value to inhibit the growth of the PTP1B protein in diabetes mellitus by using a portfolio optimization method. The SPEA2 algorithm is used to determine the pareto-optimal set for multi-objective optimization problems. There are several parameters used such as population size, archive size, maximum generation, crossover probability and mutatio probability. The solution obtained from the SPEA2 algorithm will then form an Efficient Frontier in the form of a two-dimensional graphic where each axis describes the diversity value and the probability of success of the portfolio. The data used in this study were obtained from [www.ebi.ac.uk/chembl](http://www.ebi.ac.uk/chembl) as many as 1452 pieces of data. The test was carried out 3 times with each test using a different number of molecules, namely 5 molecules, 10 molecules and 20 molecules. Each test also uses a different number of generations, such as 100, 200 and 300. The results obtained indicate that the greater the number of generations does not affect the results of the graph. Because in several tests, a certain number of generations has obtained optimal results because there has been no significant change. The number of molecules also affects the value of diversity and the value of expected return, where the greater the number of molecules used, the greater the value of expected return and resulting diversity. The range of diversity values and expected return values are also influenced by the number of molecules.

Keywords: *portfolio, drug discovery, SPEA2, efficient frontier*

## 1. Pendahuluan

### Latar Belakang

Dalam pembuatan obat, awalnya para kimiawan memilih molekul hanya berdasarkan kinerja dari masing – masing molekul. Namun, pemilihan molekul yang hanya berdasarkan dari kinerjanya memiliki risiko tersendiri. Solusinya adalah dengan memilih molekul yang berbeda. Namun, pemilihan molekul yang berbeda juga memiliki

risiko yaitu *budget* yang dikeluarkan akan lebih besar karena setiap molekul memiliki harga yang berbeda – beda [1].

Masalah pemilihan molekul ini mirip dengan masalah portofolio keuangan sehingga metode optimasi portofolio dapat diterapkan untuk menangani masalah ini. Agar portofolio yang dihasilkan optimal, maka nilai *expected return* harus dimaksimalkan untuk suatu tingkat risiko tertentu ataupun dengan meminimalkan risiko untuk suatu tingkat *expected return* tertentu.

Dalam tugas akhir ini, masalah pemilihan molekul yang menghambat pertumbuhan protein PTP1B pada penyakit diabetes mellitus diselesaikan menggunakan optimasi portofolio dengan algoritma Strength Pareto Evolutionary Algorithm 2 (SPEA2). Dalam menyelesaikan masalah ini, nilai *diversity* atau keberagaman molekul dan *expected return* yang direpresentasikan dengan probabilitas kesuksesan suatu molekul akan dimaksimalkan. Jumlah molekul dalam portofolio menjadi *constraint* dalam tugas akhir ini.

### Perumusan Masalah

Berdasarkan latar belakang diatas, permasalahan dalam penelitian ini adalah :

- 1) Bagaimana memilih molekul – molekul yang tepat untuk menghambat molekul PTP1B dengan menerapkan metode optimasi portofolio dengan mempertimbangkan *diversity* dan nilai aktivitas molekul dari gabungan molekul tersebut?
- 2) Bagaimana mengimplementasikan algoritma SPEA2 untuk masalah optimasi portofolio pada *drug discovery*?

### Tujuan

Berdasarkan perumusan masalah diatas, tujuan dari penelitian ini adalah :

- 1) Menemukan kumpulan molekul yang tepat untuk menghambat molekul PTP1B penyebab penyakit diabetes mellitus dengan menerapkan metode optimasi portofolio dimana nilai *diversity* dan nilai aktivitas molekul dari gabungan molekul dipertimbangkan.
- 2) Mampu mengimplementasikan algoritma SPEA2 untuk masalah optimasi portofolio pada *drug discovery*.

### Batasan Masalah

Batasan masalah dalam tugas akhir ini adalah :

- 1) Jumlah molekul dalam satu portofolio
- 2) Data yang digunakan berasal dari [www.ebi.ac.uk/chembl](http://www.ebi.ac.uk/chembl)
- 3) Data yang dipilih adalah data dengan nilai molekul activity lebih dari 10000 nM

## 2. Studi Terkait

### Portofolio

Portofolio adalah sekumpulan aset yang dimiliki investor baik real *asset* ataupun finansial *asset* [2] dan pertama kali dikenalkan oleh Hary M. Markowitz (1952). Portofolio biasa digunakan untuk mengurangi adanya resiko. Portofolio dibentuk dari beberapa *asset* yang dipilih agar menjadi portofolio yang optimal. Terdapat dua objek yang harus diperhatikan dalam pembentukan portofolio yaitu *return* dan *risk*. *Return* adalah pendapatan yang diterima dari suatu investasi [3]. *Return* dapat berbentuk keuntungan maupun kerugian. Sedangkan *risk* adalah perbedaan hasil yang diharapkan dengan realisasinya [4]. *Risk* dapat dihitung dengan variansi. Variansi dapat digunakan untuk menghitung *risk* karena kita dapat melihat persebaran data dengan menghitung variansinya. Dalam pembentukan portofolio, terdapat beberapa fakta yang perlu diperhatikan seperti karakter para investor, jangka waktu portofolio, hasil yang diharapkan dan tingkat risiko yang akan dihadapi [5]. Pada dasarnya, penanam modal akan memilih portofolio dengan *return* yang besar dengan berapapun risikonya atau memilih risiko yang kecil dengan berapapun returnnya [4]. Portofolio sendiri terdiri dari beberapa macam, salah satunya adalah portofolio *Drug Discovery*.

### Portofolio Drug Discovery

Portofolio *drug discovery* adalah kumpulan harga molekul yang selanjutnya akan digunakan untuk membuat suatu obat. Dalam portofolio *drug discovery*, terdapat dua objek yang dipertimbangkan, yaitu nilai *expected return* yang direpresentasikan dengan nilai probabilitas keberhasilan suatu molekul dan nilai *diversity* atau keberagaman molekul [1]. Portofolio yang optimal dapat dibentuk dengan memaksimalkan kedua objek tersebut.

Nilai dari probabilitas keberhasilan dapat dihitung dengan formulasi berikut :

$$p_i = \frac{a_i}{\sum_{i=1}^{N_{Total}} a_i} \quad (2.1)$$

Nilai  $a_i$  adalah nilai logaritma dari aktivitas molekul ke I, yang didapatkan dari nilai IC50. Nilai  $a_i$  dapat direpresentasikan sebagai berikut :

$$a_i = \frac{1}{e^{IC_i}} \quad (2.2)$$

Penjumlahan dari seluruh probabilitas kesuksesan pada portofolio sama dengan 1. Setelah nilai probabilitas masing – masing molekul didapatkan selanjutnya, nilai tersebut akan digunakan untuk menghitung *expected return* dari masing – masing molekul dengan formula :

$$E(x) = Gp \cdot x - B = \left( \sum_{i=1}^{N_{Total}} Gp_i x_i \right) - B \quad (2.3)$$

Ket :

G = *gain*(keuntungan)

B = *Budget*

$X_i$  = Bit molekul

Nilai *expected return* dari suatu portofolio dapat diformulasikan sebagai berikut :

$$E(x) = \sum_{i=1}^{N_{Total}} E(\bar{x}_i) \quad (2.4)$$

Ket :

$E(\bar{x}_i)$  = nilai *expected return* untuk portofolio x

Objek lain yang dipertimbangkan dalam masalah ini adalah *diversity* atau keberagaman molekul. *Diversity* digunakan untuk mengurangi risiko kegagalan suatu molekul untuk alasan yang sama. Nilai *diversity* untuk dua molekul didapatkan dari persamaan Tanimoto Similarity yang digambarkan sebagai berikut :

$$Sim_T(a, b) = \frac{N_{ab}}{N_a + N_b + N_{ab}} \quad (2.5)$$

Dimana :

$N_{ab}$  = Jumlah bit 1 pada molekul a dan b

$N_a$  = jumlah bit 1 pada molekul a

$N_b$  = Jumlah bit 1 pada molekul b

Setelah nilai  $Sim_T$  didapat, selanjutnya menghitung nilai *diversity* menggunakan rumus :

$$d(x_i, x_j) = 1 - Sim_T \quad (2.6)$$

Untuk menentukan nilai *diversity* dari suatu portofolio, maka dibentuk sebuah matriks korelasi yang berisi hasil dari persamaan berikut :

$$f(d) = e^{-\theta d(s_i, s_j)} \quad (2.7)$$

Nilai  $\theta$  adalah suatu nilai tetap dimana berfungsi untuk memberikan bobot lebih pada jumlah molekul dan mengurangi kesamaan dari molekul – molekul tersebut. Semakin kecil nilai dari  $\theta$  yang ditetapkan, maka semakin besar nilai *diversity* yang dihasilkan.

Selanjutnya, matriks korelasi yang sudah terbentuk dilakukan *invers*. Hasil dari *invers* matriks korelasi selanjutnya dijumlah untuk tiap – tiap elemen yang terdapat didalamnya. Hasil dari penjumlahan tiap elemen tersebut menghasilkan nilai *diversity* untuk satu portofolio.

### Molecular Fingerprint

*Molecular fingerprint* adalah *bit string* yang dari struktur molekul yang digunakan untuk menghitung nilai *similarity* antar molekul. *Molecular fingerprint* didapatkan dari struktur *SMILES* pada *dataset*. Terdapat beberapa jenis *molecular fingerprint*, salah satunya *Topological Fingerprint*. *Topological fingerprint* dimana apabila terdapat struktur dalam molekul maka bit string akan bernilai 1 dan bernilai 0 apabila struktur tidak ada dalam molekul [6].

### Drug Discovery

*Drug discovery* dapat adalah sebuah proses untuk menemukan kandidat molekul untuk menemukan obat baru. Untuk membuat obat baru dibutuhkan waktu yang lama dengan biaya yang sangat besar. Adapun proses untuk *drug discovery* dapat dijelaskan sebagai berikut:

#### 1. Pre – discovery

Pada tahap awal ini, para ahli mempelajari terlebih dahulu penyebab dari terjadinya suatu penyakit.

#### 2. Target Identification

Tahap kedua adalah menentukan target yang akan menjadi sasaran untuk obat barunya.

#### 3. Target Validation

Setelah menentukan target, para ahli dapat menunjukkan bahwa target yang ditunjuk memang benar – benar terlibat dalam penyakit dan dapat bereaksi dengan obat.

#### 4. Drug Discovery

Pada tahap keempat ini, para ahli mencari *lead compound* yang dapat bereaksi dengan target.

#### 5. Early Safety Test

Setelah para ahli menemukan *lead compound* yang tepat, selanjutnya *lead compound* tersebut diuji untuk menilai keamanan dari senyawa tersebut.

#### 6. Lead Optimization

Senyawa yang masih bertahan sesaat setelah diuji, selanjutnya dioptimalisasi.

#### 7. Preclinical Testing

Senyawa yang sudah dioptimalisasi selanjutnya diuji secara *in vitro* dan *in vivo*.

### IC50

Menurut IUPAC (2006) IC50 adalah sebuah senyawa yang dapat menghambat system yang diberikan sebesar 50% [7]. Dengan kata lain untuk aktivitas PTP1B nilai IC50 berarti sebuah konsentrasi yang dapat menurunkan konsentrasi PTP1B sebesar 50% dari konsentrasi PTP1B awal. Semakin kecil nilai IC50 maka semakin baik [8].

### Strength Pareto Evolutionary Algorithm 2 (SPEA2)

Algoritma SPEA2 adalah sebuah algoritma pengembangan dari algoritma SPEA. Pada algoritma ini, semua fungsi objektif akan dimaksimalkan, tidak ada satu fungsi objektif pun yang diabaikan sehingga semua fungsi objektif sama pentingnya. Algoritma SPEA2 dapat direpresentasikan sebagai berikut [9] :

#### a) Inisialisasi (Inputan Awal)

Inputan awal dalam algoritma ini adalah ukuran populasi ( $P_t$ ), ukuran arsip ( $A_t$ ), maksimum generasi (Max\_gen), Probabilitas *Crossover* ( $P_c$ ) dan Probabilitas mutasi ( $P_m$ ).

### b) Perhitungan Nilai Fitness

Hitung nilai  $S(i)$  atau jumlah individu yang mendominasi individu lain untuk semua individu populasi ( $P_t$ ) dan arsip ( $A_t$ ).  $S(i)$  dapat direpresentasikan sebagai berikut :

$$S(i) = |\{j | j \in P_t + A_t \wedge i \succ j\}| \quad (2.8)$$

Dimana:

$|\cdot|$  = jumlah populasi anggota himpunan

$+$  = multiset union

$\succ$  = pareto dominance

Selanjutnya menghitung nilai *raw fitness* atau biasa disebut  $R(i)$  untuk setiap individu dengan menjumlahkan semua nilai  $S(i)$  antar individu :

$$R(i) = \sum_{j \in P_t + A_t, j \succ i} S(j) \quad (2.9)$$

Dimana :

$+$  = multiset union

$\succ$  = pareto dominance

Selanjutnya menghitung nilai *density* atau  $D(i)$  untuk setiap individu dengan menggunakan *Euclidian distance*. Nilai  $D(i)$  dapat dihitung menggunakan rumus :

$$D(i) = \frac{1}{d_i^k + 2} \quad (2.10)$$

Nilai  $d_i^k$  dapat dihitung menggunakan *Euclidean distance*. Angka 2 digunakan untuk memastikan bahwa nilainya lebih besar dari 0. Setelah nilai  $R(i)$  dan  $D(i)$  didapatkan, maka selanjutnya menghitung nilai *fitness total*  $F(i)$  dengan rumus sebagai berikut :

$$F(i) = R(i) + D(i) \quad (2.11)$$

### c) Environmental Selection

Selanjutnya nilai *fitness* yang sudah didapatkan, diurutkan secara *descanding*. Setelah diurutkan, lalu dipilih nilai *fitness* terbaik sesuai dengan ukuran populasi dan arsip yang sudah ditentukan sebelumnya. Populasi yang telah didapatkan, selanjutnya akan digunakan untuk melakukan *crossover* dan mutasi. Sedangkan arsip yang sudah didapatkan, akan digunakan untuk generasi baru.

### d) Mating Selection

*Mating selection* adalah sebuah proses yang digunakan untuk mencari *parent* yang terbaik berdasarkan nilai *fitness* yang didapatkan. Dalam penelitian ini, metode yang digunakan untuk *mating selection* adalah *random selection*, yaitu *parent* yang dihasilkan dipilih secara *random*

### e) Crossover dan Mutasi

*Crossover* adalah perkawinan silang antara dua individu (*parent*) yang didapatkan dari hasil proses *mating selection* dan menghasilkan dua individu baru (*offspring*). Individu yang akan dipilih berdasarkan nilai Probabilitas *Crossover* ( $P_c$ ) yang sudah ditentukan. Mutasi dilakukan dengan *swap mutation*, yaitu dimana jumlah kromosom yang akan mengalami mutasi dalam satu populasi akan ditentukan melalui parameter *mutation rate* ( $P_m$ ), lalu gen yang telah terpilih secara acak akan ditukar dengan sesudahnya.

### PTP1B (Protein Tirosin Fosfatase)

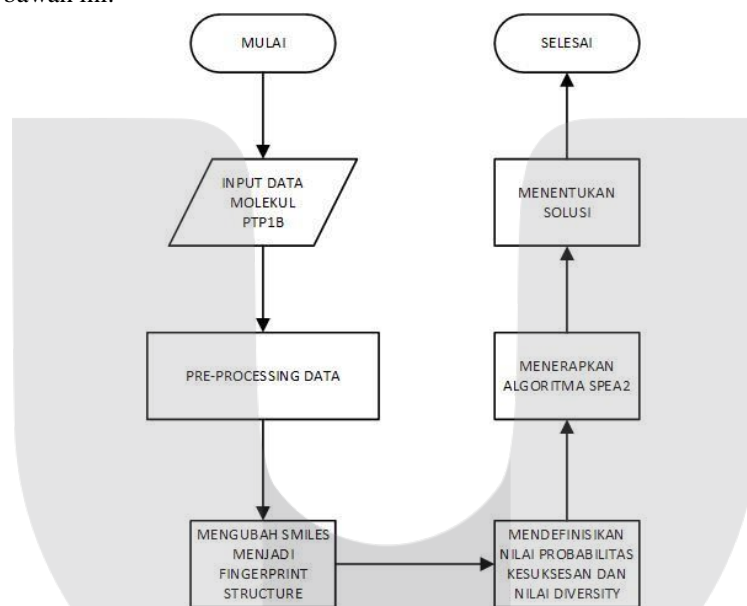
PTP1B adalah sebuah protein yang menjadi target potensial untuk pengobatan diabetes mellitus tipe 2 dan berperan penting dalam pathogenesis diabetes mellitus tipe 2. PTP1B adalah PTP seluler yang terlibat dalam regulasi negatif insulin serta sinyal leptin. PTP1B inhibitor berpotensi meningkatkan resistensi insulin dan menormalkan glukosa plasma dan level insulin tanpa terjadi *hypoglycemia*. Terdapat beberapa inhibitor PTP1B seperti benzofuran, asam aminobenzoik dan vanadium kompleks [11].

### Efficient Frontier

*Efficient frontier* pertama kali dikenalkan oleh Harry Markowitz pada tahun 1995. *Efficient frontier* adalah sekumpulan portofolio efisien diantara semua portofolio yang ada, dimana nilai *expected return* maksimal untuk tingkat resiko tertentu ataupun nilai resiko minimal untuk tingkat *expected return* tertentu [12]. Portofolio yang tidak berada pada garis *efficient frontier* adalah portofolio yang tidak efisien karena pada nilai resiko tertentu terdapat nilai *expected return* yang lebih maksimal dari sebuah portofolio lain ataupun pada nilai *expected return* tertentu terdapat nilai resiko yang lebih minimum dari sebuah portofolio [13]. *Efficient frontier* berbentuk grafik yang berisi nilai *expected return* dan resiko dari portofolio yang ada.

### 3. Sistem yang Dibangun

Sistem yang dibangun dalam tugas pengujian ini digambarkan dalam bentuk diagram alir pada gambar di bawah ini.



Gambar 1 Perancangan Sistem

#### Input Data Molekul

Data yang digunakan dalam tugas akhir ini adalah data molekul yang diambil dari [www.ebi.ac.uk/chembl/](http://www.ebi.ac.uk/chembl/). data yang digunakan adalah molekul PTP1B dan nilai IC50 dari masing – masing molekul

#### Pre-Processing Data

Data yang sudah didapatkan selanjutnya dilakukan *pre-processing*, yaitu memilih jenis data yang digunakan, menghapus data yang kosong dan menghapus data yang memiliki nama molekul sama

#### Mengubah SMILES menjadi *Fingerprint Structure*

Mengubah struktur SMILES pada *dataset* menjadi bit string *Molecular Fingerprint* agar dapat dilakukan perhitungan *similarity* antar molekul.

### Mendefinisikan Nilai Probabilitas Kesuksesan dan Nilai *Diversity*

Menghitung nilai probabilitas kesuksesan menggunakan persamaan (2.1) yang nantinya akan digunakan untuk menghitung nilai *expected return*. Lalu, nilai *diversity* akan dihitung menggunakan persamaan (2.4).

### Menerapkan Algoritma SPEA2

Setelah nilai probabilitas kesuksesan dan nilai *diversity* dihitung, selanjutnya mencari solusi dengan menggunakan algoritma SPEA2.

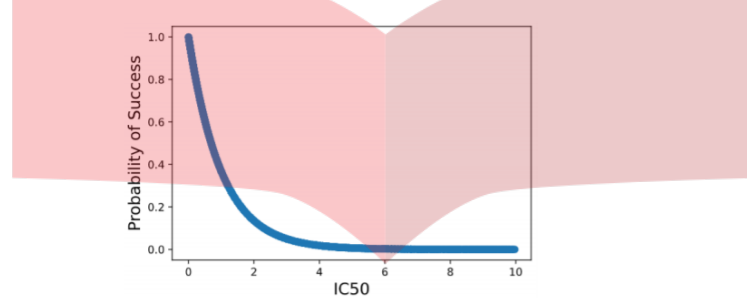
### Menentukan Solusi

Solusi yang sudah dihitung menggunakan algoritma SPEA2, selanjutnya dibentuk grafik *efficient frontier* agar dapat dipilih solusi yang optimal.

## 4. Evaluasi

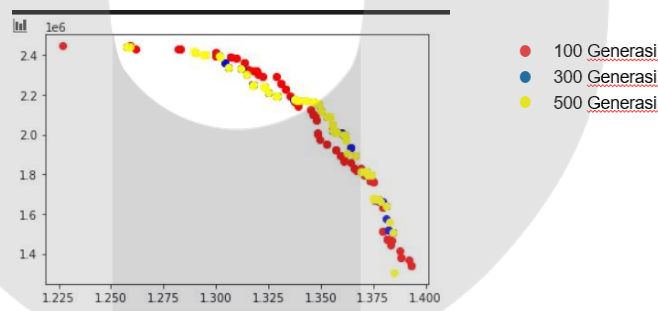
### 4.1. Hasil Pengujian

Berikut adalah grafik hubungan antara data aktivitas molekul dengan probabilitas keberhasilan suatu molekul :

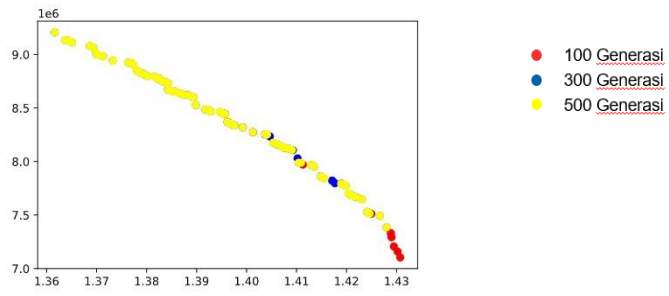


**Gambar 2 Hubungan antara data aktivitas molekul dengan probabilitas keberhasilan**

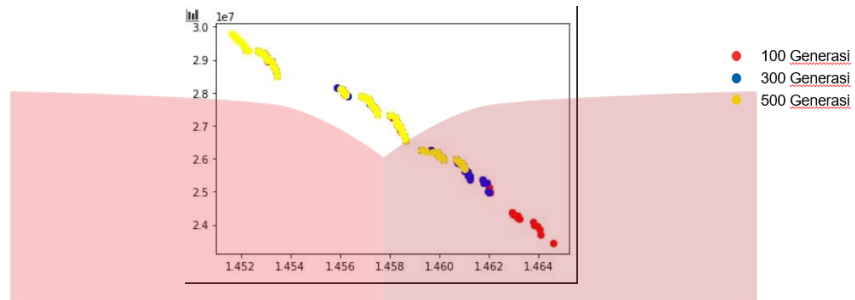
Berdasarkan gambar grafik diatas, nilai IC50 memiliki hubungan dengan probabilitas keberhasilan suatu molekul dimana semakin besar nilai IC50 maka nilai keberhasilan suatu molekul semakin kecil. Selanjutnya pengujian dilakukan dengan menggunakan 100, 300 dan 500 generasi. Pengujian dilakukan sebanyak tiga kali dengan masing – masing pengujian menggunakan jumlah molekul yang berbeda untuk satu portofolio yaitu 5, 10 dan 20 molekul. Parameter yang digunakan dalam penelitian ini adalah nilai populasi ( $P_t$ ) sebesar 100, nilai ukuran arsip ( $A_t$ ) sebesar 100, Probabilitas *Crossover* ( $P_c$ ) sebesar 0,8 dan probabilitas mutasi ( $P_m$ ) sebesar 0,1. Hasil yang didapatkan dari *running* program adalah grafik *efficient frontier* dengan masing – masing grafik merepresentasikan jumlah molekul yang sama dengan nilai maksimum generasi yang berbeda.



**Gambar 3 Grafik *Efficient frontier* menggunakan 5 molekul dengan nilai maksimum generasi 100, 300, dan 500**



Gambar 4 Grafik *Efficient frontier* menggunakan 10 molekul dengan nilai maksimum generasi 100, 300 dan 500



Gambar 5 Grafik *Efficient frontier* menggunakan 20 molekul dengan nilai maksimum generasi 100, 300 dan 500





Berikut adalah salah satu populasi yang dihasilkan dari pengujian menggunakan 5 molekul.



Gambar 6 Populasi yang dihasilkan dari percobaan menggunakan 5 molekul

Setelah dilakukan pengujian menggunakan 5 molekul dengan 100 generasi, didapatkan 100 portofolio dengan dengan masing – masing portofolio berisi molekul yang berbeda – beda. Dalam

satu portofolio juga memiliki nilai *expected return* dan *diversity* yang berbeda. Berikut adalah 5 portofolio dengan nilai *expected return* tertinggi.

	Mol	X	Y
0	[CHEMBL403468, CHEMBL269166, CHEMBL564439, CHE...	2.430847e+06	1.311172
1	[CHEMBL378614, CHEMBL269166, CHEMBL564439, CHE...	2.433790e+06	1.310290
2	[CHEMBL1086227, CHEMBL378614, CHEMBL269166, CH...	2.466171e+06	1.281734
3	[CHEMBL564439, CHEMBL378614, CHEMBL1083271, CH...	2.378748e+06	1.316815
4	[CHEMBL1086227, CHEMBL269166, CHEMBL408382, CH...	2.283135e+06	1.329968

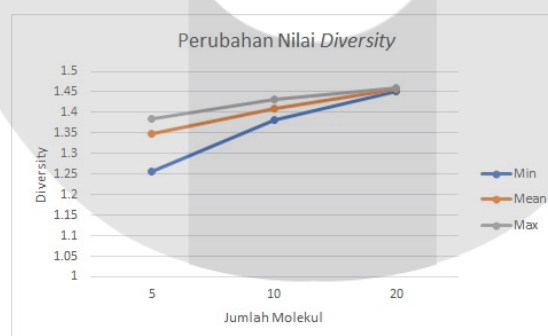
**Gambar 7 Molekul dengan nilai *expected return* tertinggi pada pengujian menggunakan 5 molekul dengan 100 generasi**

#### 4.2. Analisis Hasil Pengujian

Berdasarkan hasil pengujian diatas, setiap nilai maksimum generasi memiliki hasil grafik yang berbeda. Pada pengujian menggunakan 5 molekul, jumlah generasi 500 sudah cukup optimal karena tidak ada perubahan yang signifikan dari 300 generasi sampai 500 generasi. Pada pengujian 10 dan 20 molekul, 300 generasi sudah cukup optimal karena dari 100 generasi sampai 300 generasi tidak ada perubahan yang signifikan. Jumlah molekul yang digunakan untuk satu portofolio pun mempengaruhi hasil grafik. Semakin banyak jumlah molekul yang digunakan, nilai *diversity* dan *expected return* yang dihasilkan juga semakin tinggi. Semakin tinggi jumlah molekul yang digunakan juga mempengaruhi rentang untuk nilai *expected return* dan nilai *diversity*. Rentang nilai *expected return* akan membesar jika jumlah molekul yang digunakan juga semakin besar. Namun, pada rentang nilai *diversity* akan semakin menuju nilai asimtotik tertentu. Hal ini dapat digambarkan dengan grafik berikut :



**Gambar 8 Grafik pengaruh jumlah molekul terhadap nilai *expected return***



**Gambar 9 Grafik pengaruh jumlah molekul terhadap nilai *diversity***

## 5. Kesimpulan

Dari pembahasan dan hasil pengujian yang telah dilakukan maka dapat ditarik beberapa kesimpulan. Algoritma SPEA2 dapat digunakan untuk mencari individu terbaik dengan memaksimalkan nilai *Expected Return* untuk tingkat *diversity* tertentu. Besar nilai maksimum generasi mempengaruhi grafik yang dihasilkan. Semakin besar generasi belum tentu lebih baik, karena seperti pada percobaan 10 dan 20 molekul, 300 generasi sudah cukup optimal sebab tidak terjadi perubahan yang signifikan dari 100 generasi sampai 300 generasi. Banyak jumlah molekul juga mempengaruhi hasil grafik. Semakin banyak molekul yang digunakan untuk satu portofolio, nilai *expected return* dan nilai *diversity* yang dihasilkan pun semakin besar. Rentang nilai *expected return* dan *diversity* pun dipengaruhi oleh banyak jumlah molekul. Semakin banyak jumlah molekul yang digunakan, rentang nilai *expected return* pun akan semakin tinggi. Namun, pada nilai *diversity*, rentang yang dihasilkan semakin menuju nilai asimtotik tertentu.

Untuk penelitian selanjutnya, penulis menyarankan perlu dilakukan pengujian juga dapat dari satu kali agar grafik yang dihasilkan semakin terlihat konvergen. Parameter yang digunakan juga sebaiknya diubah untuk mengetahui nilai parameter yang optimal.

## Daftar Pustaka

- [1] I. Yevseyeva, E. B. Lenselink, A. d. Vries, A. P. IJzerman, A. H. Deutz and M. T. Emmerich, "Application of Portofolio Optimization to Drug Discovery," *Infomation Science*, 2018.
- [2] A. Halim, Analisis Investasi di Aset Keuangan, Jakarta: Mitra Wacana Media, 2015.
- [3] D. A. Lestari, "Sharpe Square Ratio (SSR) untuk Ukuran Performansi Portofolio," *Telkom University*, vol. 4, 2017.
- [4] E. V. Nurdhian and S. M. A. Dr. Norita, "Analisis Pembentukan Portofolio Optimal Menggunakan Capital Asset Pricing Model Serta Penilaian Kinerja Portofolio Berdasarkan Metode Sharpe Ratio, Treynor Ratio dan Jensen," *Telkom University*, vol. 4, 2017.
- [5] L. Chandra and Y. D. Hapsari, "Analisis Pembentukan Portofolio Optimal dengan Menggunakan Model Markowitz untuk Saham LQ 45 Periode 2008 - 2012," *Universitas Katolik Indonesia Atma Jaya Jakarta*, vol. 1, 2013.
- [6] A. Cereto-Massague, M. J. Ojeda, C. Valls, M. Mulero, S. Garcia-Vallve and G. Pujadas, "Molecular Fingerprint Similarity Search in Virtual Screening," *Methods*, pp. 58-63, 2015.
- [7] A. E. Febrinda, M. Astawan, T. Wresdiyati and N. D. Yuliana, "Kapasitas Antioksidan dan Inhibitor Alfa Glukosidase Ekstrak Umbi Bawang Dayak," *Institut Pertanian Bogor*, vol. 24, 2013.
- [8] A. Widiasanti, D. Rohdiana and N. Ekatama, "Aktivitas Antioksidan Ekstrak Teh (*Camellia Sinensis*) dengan Metode DPPH (2,2 Difenil -1- Pikrilhidrazil)," <http://ejournal.upi.edu/index.php>, 2016.
- [9] E. Zitzler, M. Laumanns and L. Thiele, "SPEA2 : Improving the Strength Pareto Evolutionary Algorithm," *ETH Zurich Research Collection*, 2001.
- [10] S. N. Fadhila, D. Saepudin and R. F. Umbara, "Penerapan Algoritma Genetika Multi-objective SPEA-II pada Optimasi Portofolio Saham," *e-Proceeding of Engineering*, vol. 2, p. 7929, 2015.
- [11] J. Sun, C. Qu, Y. Wang, H. Huang, M. Zhang, H. Li, Y. Zhang, Y. Wang and W. Zou, "PTP1B, A Potential Target of Type 2 Diabetes Melitus," *Molecular Biology*, 2016.
- [12] T. Nurislamiaty, D. Saepudin and A. A. Rohmawati, "Optimasi Portofolio Risiko Terkecil Efficent Frontier Mean Semivariance dengan Metode Interior Point," *Telkom University*, vol. 5, p. 8285, 2018.

- [13] A. Pratiwi, D. Saepudin and R. f. Umbara, "Optimasi Portofolio Mean-semivariance dengan Algoritma Genetika amultiobjective Evolutuinary NSGA II," *Telkom University*, vol. 5, p. 8269, 2018.

