

BAB V

PEMBAHASAN DAN UJI COBA HASIL PENELITIAN

5.1. Penyempurnaan

Berikut adalah penyempurnaan yang sudah dilakukan pada aplikasi untuk memaksimalkan kinerja sistem.

5.1.1. Kamus Sinonim *Offline*

Pada saat *generate* kalimat, kata kunci yang telah diberikan akan dicari sinonimnya. Pencarian sinonim tersebut dilakukan dengan cara mengakses sebuah halaman *website* yang menyediakan sinonim kata dalam Bahasa Inggris. Permasalahan yang muncul adalah ketika tidak terdapat koneksi internet, maka dari itu dibuat sebuah mekanisme dimana pada saat selesai mencari sinonim, hasilnya akan disimpan kedalam *file* teks sehingga untuk setiap kata kunci hanya dilakukan pencarian ke internet sebanyak satu kali, selebihnya akan mengakses melalui *file* teks yang sudah dibuat. Hal ini membuat proses pencarian sinonim lebih cepat terhadap kata kunci yang sudah pernah diberikan sebelumnya karena tidak perlu mengakses internet, hanya perlu membaca *file* teks yang telah disimpan. Selain itu hal ini juga memungkinkan untuk mengakses sinonim kata kunci yang sudah pernah diberikan pada kondisi tidak terdapat koneksi internet.

5.1.2. Penyempurnaan Pada Proses *Semantic*

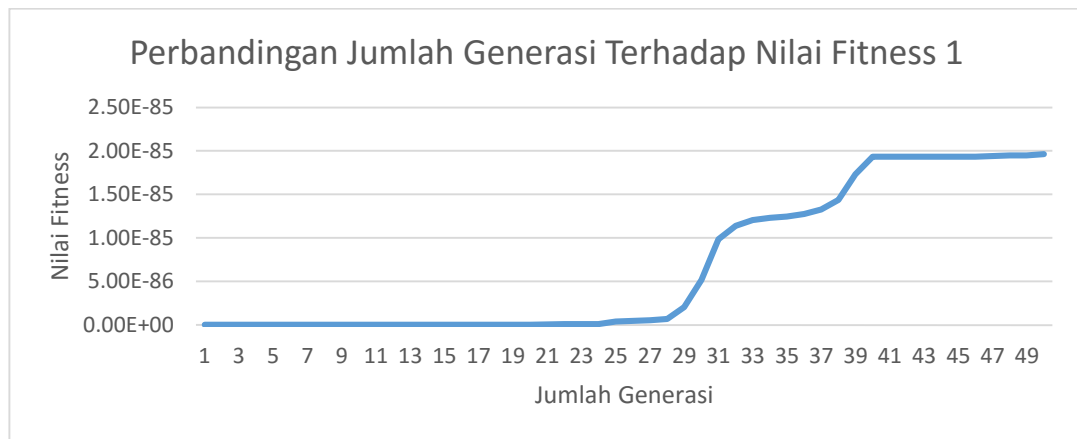
Salah satu yang menjadi penentu kualitas kalimat yang dihasilkan pada sistem adalah nilai *semantic*. Nilai *semantic* ini didapatkan melalui sebuah *tools* (WS4J) dalam bahasa Java yang dijalankan melalui sistem. Waktu proses untuk mendapat nilai *semantic* ini cukup lama yaitu untuk mendapatkan nilai *semantic* antar 2 buah kata memerlukan waktu hampir 3 detik. Hal tersebut menjadi masalah karena untuk mendapat nilai *semantic* dari sebuah kalimat, setiap kata dalam kalimat tersebut dipasangkan dan dinilai satu persatu, contoh untuk sebuah kalimat yang terdiri dari 21 kata, jumlah kemungkinan pasangan katanya adalah 210 pasang. Maka untuk menjalankan mendapat nilai *semantic* untuk kalimat tersebut memerlukan 210×3 detik, atau sekitar 10,5 menit.

Maka dari itu untuk menangani masalah ini dibuat sebuah mekanisme yang sama dengan penyempurnaan saat mencari sinonim, yaitu menyimpan nilai *semantic* yang telah didapat ke dalam sebuah *file* teks. Hal tersebut menghasilkan waktu proses semakin berkurang seiring dengan semakin sering sistem dijalankan. Hal ini terbukti dari pada saat membuat individu awal yang berjumlah 200 memerlukan waktu sekitar 9 jam, setelah

sistem dijalankan beberapa kali waktu tersebut menjadi semakin berkurang hingga saat ini untuk membuat 200 individu awal diperlukan waktu hanya sekitar 2 menit saja.

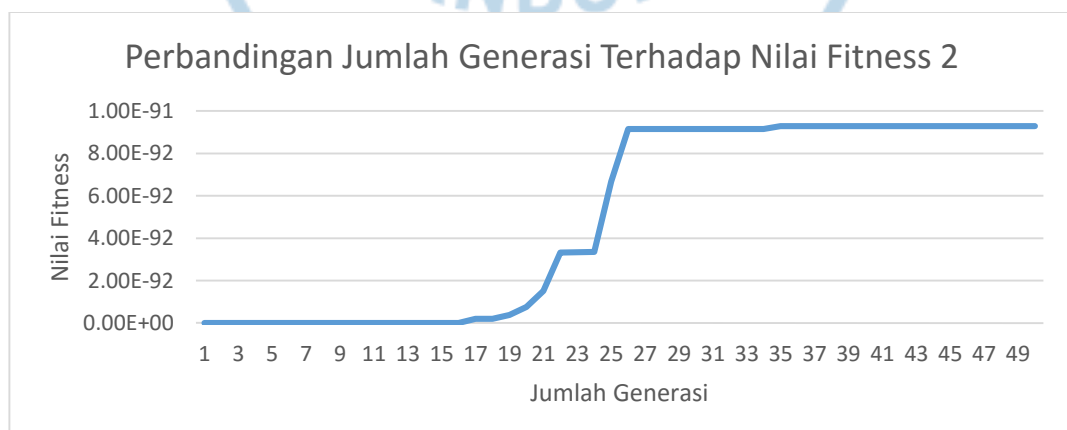
5.2. Pengujian Terhadap Jumlah Generasi

Salah satu faktor yang menentukan dari konfigurasi proses genetika adalah jumlah generasi. Jumlah generasi ini mempengaruhi beberapa hal dalam proses genetika salah satu contohnya adalah nilai fitness setiap individu. Untuk melihat hubungan antara jumlah generasi dengan nilai probabilitas dilakukan percobaan dengan hasil sebagai berikut.



Gambar 5.1 Grafik Perbandingan Nilai Fitness dengan Generasi 1

Gambar 5.1 dan Gambar 5.2 merupakan grafik perbandingan nilai *fitness* dengan jumlah generasi yang menggunakan 163 data *tweet* dan dengan jumlah individu awal 200. Perbedaan terhadap kedua gambar tersebut terletak pada probabilitas mutasi dan kawin silangnya. Gambar 5.1 memiliki probabilitas mutasi 75% dan probabilitas kawin silang 50%, sedangkan Gambar 5.2 memiliki probabilitas mutasi 25% dan probabilitas kawin silang 50%.



Gambar 5.2 Grafik Perbandingan Nilai Fitness dengan Generasi 2

Berdasarkan Gambar 5.1 dan Gambar 5.2, dapat dilihat bahwa parameter yang digunakan pada Gambar 5.1 lebih baik karena dapat menghasilkan nilai *fitness* yang lebih tinggi dibandingkan dengan Gambar 5.2. Grafik dari kedua gambar tersebut hampir serupa, keduanya kenaikan nilai *fitness* yang tidak signifikan pada awal generasi (generasi 1 sampai 28 untuk Gambar 5.1, dan generasi 1 sampai 19 untuk Gambar 5.2), setelah itu mengalami kenaikan nilai *fitness* yang signifikan (generasi 29 sampai 40 untuk Gambar 5.1, dan generasi 20 sampai 27 untuk Gambar 5.2), sampai akhirnya nilai *fitness* kembali tidak mengalami kenaikan yang signifikan.

Selain itu berdasarkan Gambar 5.1 dan Gambar 5.2 dapat dilihat bahwa semakin besar jumlah generasi yang dilakukan maka nilai *fitness* juga semakin baik sampai generasi ke 40 untuk Gambar 5.1 dan generasi 25 untuk Gambar 5.2, tetapi berdasarkan hasil generasi pada Gambar 5.3 dapat dilihat semakin besar jumlah generasi semakin serupa kalimat yang dihasilkan. Hal tersebut terjadi, karena pada setiap proses genetika individu dikawin silangkan hanya individu yang hampir mirip, sehingga menghasilkan individu yang semakin mirip juga.

```

Top 5 Generasi Ke 5
the goal successful and happy of humanity to us wake up the basis each other
just money the same is the same to us the way successful and happy the future
a matter for positive is our special quality to be happy the way a more just money
be miserable more peaceful you succeed to us the way successful and happy the future
the future the way your hands to us our own interest successful and happy the same

Top 5 Generasi Ke 10
the basis to be happy the goal the world can achieve to us for determination
the way to be happy the goal the world the same can achieve more peaceful
each other to be happy of us a sense in nature to us the way
the basis wake up the goal to be happy the same common sense more peaceful
the way to be happy the goal the world the same to us of humanity

Top 5 Generasi Ke 15
common sense the goal the basis more peaceful happier more for determination of humanity
common sense the goal the basis a matter for determination each of us a happy life
the future be happy of religious practice the world a matter use it well wake up
common sense for positive the basis to be happy be miserable a more others happy
common sense the goal the basis to be happy of humanity avoid suffering in nature

```

Gambar 5.3 Hasil Generasi ke 5, 10, dan 15

5.3. Perbandingan Penilaian Menggunakan *Semantic Relatedness*

Gambar 5.4 merupakan kalimat yang dihasilkan oleh sistem sebelum nilai *semantic* ditambahkan kedalam kriteria penilaian kalimat (penilaian *parser*). Jika dibandingkan dengan Gambar 5.3, kalimat yang dihasilkan oleh sistem sesudah nilai *semantic* dimasukan kedalam kriteria penilaian lebih baik, karena jika dibandingkan kalimat antara penilaian *semantic* dan penilaian *parser* pada generasi yang sama, kalimat hasil penilaian *semantic* lebih beragam (tidak terlalu seragam) dibandingkan dengan kalimat hasil penilaian *parser*.

Top-5 Generasi ke 5
do that the basis make others happy be happy of us supporting each other makes us happy
makes us happy the world each other do that to find inner peace be miserable is the same
for others wake up the same need calm minds to be happy peaceful and happy the way
the same others happy up to us need calm minds to be happy this goal the way
to us a meaningful life the goal up to us for others well-being the same be happy

Top-5 Generasi ke 10
you succeed the basis wake up for determination more peaceful the future in nature
a more can achieve our practice for determination more peaceful the future be happy
a more wake up of humanity for determination more peaceful the future the way
a more the basis of humanity for determination more peaceful the future of us
wake up the basis of humanity for determination more peaceful the future the way

Top-5 Generasi ke 15
the rest the way a matter the future peaceful and happy common sense to us
the rest the way a matter the future peaceful and happy common sense our potential
the rest the way a matter the future peaceful and happy common sense others well-being
the rest the way a matter in nature of us for positive makes us happy
the rest more peaceful a matter the goal others unhappy a more to us

Gambar 5.4 Hasil Generasi ke 5,10 dan 15 Tanpa Nilai *Semantic*

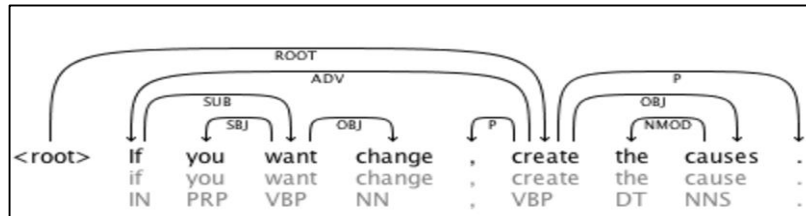
5.4. Hipotesis dan Perbaikan untuk Metode Evaluasi

5.4.1. Hipotesis

Hipotesis sementara untuk permasalahan pada bagian 5.2, adalah kalimat yang dihasilkan oleh sistem masih kurang baik, karena kalimat yang dihasilkan masih tidak memiliki struktur kalimat yang benar seperti pada Gambar 5.3 dan Gambar 5.4, sehingga membuat kalimat tersebut menjadi tidak bermakna. Hal tersebut diakibatkan karena dalam proses penilaian sistem tidak dapat menentukan bagaimana struktur kalimat yang baik atau tidak. Nilai *semantic* yang ditambahkan ke dalam penilaian tidak cukup untuk membuat kalimat yang memiliki struktur yang baik, karena kalimat yang memiliki kata-kata yang saling berhubungan tidak menjamin akan memiliki struktur kalimat yang bagus.

5.4.2. Perbaikan untuk Metode Evaluasi

Perbaikan yang dapat dilakukan untuk mengatasi masalah yang muncul pada bagian 5.2, adalah sistem harus bisa menentukan bagaimana struktur kalimat yang baik berdasarkan dari susunan kata dalam sebuah kalimat. Terdapat sebuah cara untuk menentukan peran kata dalam sebuah kalimat yang bernama *Semantic Role Labeling*. *Semantic Role Labeling* merupakan sebuah metode untuk menentukan peran setiap kata dalam sebuah kalimat (University of Illinois at Urbana-Champaign, 2014). Seperti pada Gambar 5.5, sebuah *tweet* dari akun @DalaiLama “*If you want change, create the cause.*” Setiap kata pada kalimat tersebut diberikan peran. Tidak hanya peran perkata yang diberikan, tetapi beberapa gabungan kata atau frasa juga diberikan peran dalam sebuah kalimat.



Gambar 5.5 Contoh *Semantic Role Labeling*

Dengan memanfaatkan *Semantic Role Labeling* maka diharapkan sistem dapat menilai sebuah kalimat memiliki struktur yang baik atau tidak. Metode tersebut juga dapat dimanfaatkan pada saat pembuatan individu awal atau pada saat melakukan proses kawin silang maupun mutasi.

5.5. Perbandingan Kalimat Hasil Genetic Algorithm dan Information Retrieval (*Lucene.Net*)

Gambar 5.6 merupakan perbandingan antara kalimat yang dihasilkan oleh algoritma genetik dan kalimat yang dihasilkan oleh *information retrieval* menggunakan *Lucene.Net*. Masing-masing kalimat tersebut memiliki kekurangan dan kelebihan masing-masing. Untuk kalimat yang dihasilkan oleh *Lucene.Net* memiliki kelebihan dibandingkan dengan kalimat hasil algoritma genetik adalah kalimat yang dihasilkan memiliki struktur kalimat yang bagus dan sesuai dengan tema yang diberikan. Sedangkan kekurangannya adalah kalimat yang dihasilkan terpaku pada data yang ada, jadi tidak bisa menghasilkan kalimat lain selain kalimat yang sudah disediakan sebelumnya. Begitu juga sebaliknya kalimat yang dihasilkan oleh algoritma genetik memiliki kekurangan pada struktur kalimatnya yang kurang bagus, akan tetapi algoritma genetik dapat menghasilkan kalimat yang lebih beragam tidak terpaku pada data yang ada.

```
--Kalimat hasil IR (Lucene.Net)--
1. the future is in your hands, but for it to be peaceful and happy requires a sense of compassion and concern for others' well-being.
2. if we cherish non-violence and concern for others' well-being it is possible to make this a more peaceful century.
3. everyone can practise non-violence, it only calls for determination. if you succeed, it will open the way to a far more peaceful world.
4. changing the world is up to us. if each of us tries, the next generation may see a happier, more peaceful world emerge.
5. the goal of happier human beings living together, supporting each other, in a more peaceful world, is, i believe, something we can achieve.

--Kalimat Hasil GA--
1. the goal successful and happy of humanity to us wake up the basis each other
2. just money the same is the same to us the way successful and happy the future
3. a matter for positive is our special quality to be happy the way a more just money
4. be miserable more peaceful you succeed to us the way successful and happy the future
5. the future the way your hands to us our own interest successful and happy the same
```

Gambar 5.6 Perbandingan Kalimat *Genetic Algorithm* dan *Information Retrieval*

BAB VI

SIMPULAN DAN SARAN

6.1. Simpulan

Simpulan yang diperoleh dari hasil analisis, pengambilan data, implementasi, dan pengujian aplikasi ini adalah:

1. Pembuatan aplikasi untuk membangun kalimat bernuansa positif tidak memberikan hasil yang maksimal dikarenakan kalimat yang dihasilkan tidak memiliki struktur kalimat yang baik.
2. Kalimat yang dihasilkan oleh aplikasi ini memiliki struktur kalimat yang tidak benar sehingga kalimat menjadi tidak bermakna.
3. Besaran individu maksimal pada sebuah proses algoritma genetika dapat mempengaruhi banyaknya kemungkinan individu baru yang dapat dihasilkan.
4. Semakin tinggi jumlah generasi yang dilakukan akan menghasilkan kalimat yang semakin sejenis atau serupa.
5. Proses evaluasi yang hanya melihat nilai probabilitas, nilai *semantic*, dan kandungan kata dalam kalimat tidak dapat menghasilkan kalimat yang memiliki struktur yang baik.

6.2. Saran

Berikut ini adalah saran-saran yang bisa dipakai untuk mengembangkan aplikasi:

1. Manfaatkan metode *Semantic Role Labeling* untuk proses evaluasi kalimat.
2. Menambah jumlah sumber data *tweet*, agar kalimat yang dihasilkan lebih beragam.
3. Mencari solusi yang lebih baik pada saat pemilihan frasa yang tepat untuk dijadikan sebuah kalimat.
4. Mencari solusi yang lebih baik pada saat pembangunan individu awal agar lebih cepat dan baik.
5. Struktur kalimat yang kurang baik dapat mengakibatkan sebuah kalimat menjadi tidak memiliki makna.
6. *Semantic role labeling* diharapkan bisa memperbaiki proses evaluasi pembuatan kalimat menjadi lebih baik.