

## 5. PENGUJIAN SISTEM

Pada bab ini, akan dibahas pengujian terhadap sistem NEAT dan *backpropagation* yang telah dibuat.

### 5.1 Sistem Pengujian

Waktu pengujian pada bab ini sangat ditentukan oleh spesifikasi yang digunakan. Semua pengujian yang dilakukan pada bab ini dilakukan menggunakan *notebook asus A412DA* dengan spesifikasi sebagai berikut:

Tabel 5.1 Spesifikasi sitem pengujian

|                         |                          |
|-------------------------|--------------------------|
| <i>Processor</i>        | <i>AMD Ryzen 5 3500u</i> |
| <i>RAM</i>              | <i>8GB DDR4</i>          |
| <i>Operating System</i> | <i>Windows 10</i>        |

### 5.2 Pengujian NEAT

Pengujian NEAT pada penelitian ini akan dibagi kedalam 3 tahap berdasarkan *feature* yang digunakan. Pembagian ini bertujuan untuk melihat pengaruh dari *feature* yang digunakan terhadap akurasi dari NEAT. Pada setiap tahap, akan dicoba berbagai konfigurasi untuk mendapatkan hasil yang optimal. Pengujian yang dilakukan meliputi proses *training*, yang bertujuan untuk mencari individu terbaik, dan proses *testing*, untuk mengukur akurasi dari individu yang dihasilkan oleh proses *training*.

Data yang digunakan berasal dari situs *whosocred.com*, yang merupakan situs penyedia data pertandingan sepak bola. Pada proses *training*, data berasal dari pertandingan liga inggris pada musim 2014/2015, 2015/2016, dan 2016/2017. Sedangkan untuk proses *testing*, data yang digunakan berasal dari pertandingan liga inggris pada musim 2017/2018. Pada setiap musim, terdapat 380 pertandingan.

#### 5.2.1 Tahap 1

Pada tahap ini, *feature* yang digunakan adalah *rating* pemain dari kedua *team*, sehingga ada 22 *feature* yang terdiri dari 11 pemain dari masing-masing *team*.

### 5.2.1.1 Tahap 1 Pengujian 1

|  |   |                               |
|--|---|-------------------------------|
| [NEAT]                                   | conn_delete_prob = 0.5                  | weight_max_value = 1000       |
| fitness_criterion = max                  |   | weight_min_value = -1000      |
| fitness_threshold = 1140                 | # connection enable options             | weight_mutate_power = 0.1     |
| pop_size = 1000                          | enabled_default = True                  | weight_mutate_rate = 0.8      |
| reset_on_extinction = True               | enabled_mutate_rate = 0.4               | weight_replace_rate = 0.4     |
| no_fitness_termination = True            |   |                               |
| [DefaultGenome]                          | feed_forward = True                     | [DefaultSpeciesSet]           |
| # node activation options                | initial_connection = partial_direct 0.7 | compatibility_threshold = 3.0 |
| activation_default = relu                |   |                               |
| activation_mutate_rate = 0.0             | # node add/remove rates                 | [DefaultStagnation]           |
| activation_options = relu                | node_add_prob = 0.8                     | species_fitness_func = max    |
|  | node_delete_prob = 0.5                  | max_stagnation = 1000         |
|  |   | species_elitism = 5           |
| # node aggregation options               | # network parameters                    |                               |
| aggregation_default = sum                | num_hidden = 1                          | [DefaultReproduction]         |
| aggregation_mutate_rate = 0.0            | num_inputs = 22                         | elitism = 40                  |
| aggregation_options = sum                | num_outputs = 2                         | survival_threshold = 0.5      |
|  |   |                               |
| # node bias options                      | # node response options                 |                               |
| bias_init_mean = 0.0                     | response_init_mean = 1.0                |                               |
| bias_init_stdev = 1.0                    | response_init_stdev = 0.0               |                               |
| bias_max_value = 100.0                   | response_max_value = 3000.0             |                               |
| bias_min_value = -100.0                  | response_min_value = -3000.0            |                               |
| bias_mutate_power = 0.1                  | response_mutate_power = 0.0             |                               |
| bias_mutate_rate = 0.7                   | response_mutate_rate = 0.0              |                               |
| bias_replace_rate = 0.3                  | response_replace_rate = 0.0             |                               |
|  |   |                               |
| # genome compatibility options           | single_structural_mutation = False      |                               |
| compatibility_disjoint_coefficient = 1.0 | structural_mutation_surfer = False      |                               |
| compatibility_weight_coefficient = 0.5   |   |                               |
|  |   |                               |
| # connection add/remove rates            | # connection weight options             |                               |
| conn_add_prob = 0.8                      | weight_init_mean = 0.0                  |                               |
|  | weight_init_stdev = 1.0                 |                               |

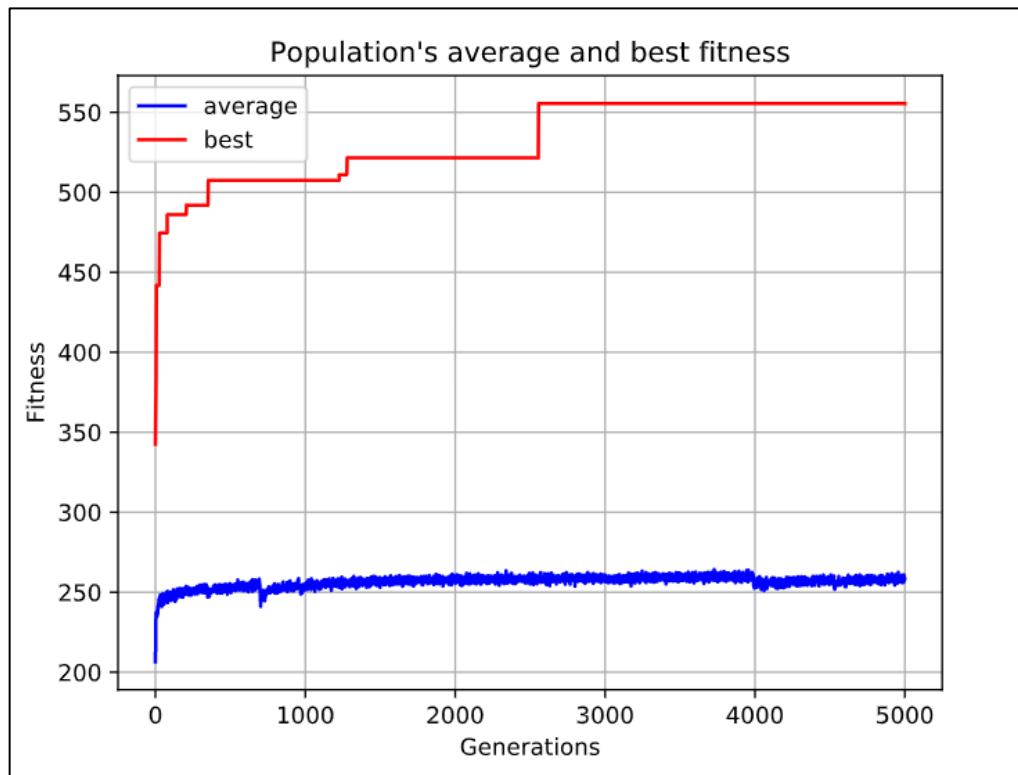
Gambar 5.1 Konfigurasi yang digunakan pada Tahap 1 Pengujian 1

Seperti yang sudah dijelaskan pada Bab 3 dan 4, tahap pertama pada *training* NEAT adalah pengaturan konfigurasi. Konfigurasi yang dilakukan pada pengujian ini dapat dilihat pada Gambar 5.1. Setelah konfigurasi, proses selanjutnya adalah pemilihan *fitness function*. Untuk pengujian pertama ini, *fitness function* yang digunakan adalah *fitness function* pertama, yang dapat dilihat pada Segmen Program 4.4.

```
Population's average fitness: 259.39311 stdev: 97.52384
Best fitness: 555.50000 - size: (11, 23) - species 2 - id 2249411
Average adjusted fitness: 0.360
Mean genetic distance 2.179, standard deviation 0.435
Population of 1000 members in 3 species:
  ID  age  size  fitness  adj fit  stag
  ===  ==  ===  =====  =====  =====
    1 4965  255   515.5    0.367  2410
    2 4965  431   555.5    0.347  2381
    3 4965  314   549.5    0.365  1273
Total extinctions: 0
Generation time: 6.861 sec
```

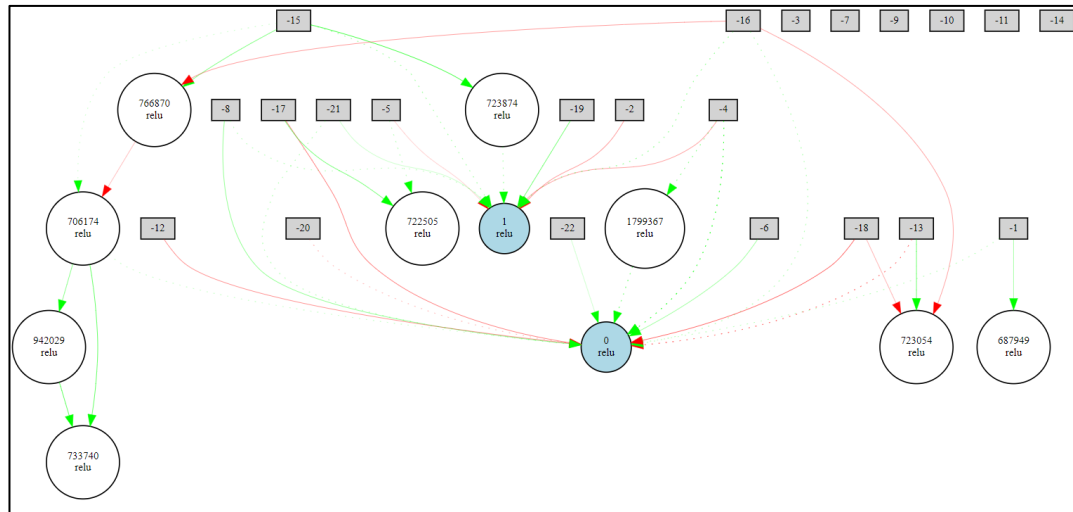
Gambar 5.2 Hasil *training* dari tahap 1 pengujian 1

Setelah dijalankan sebanyak 5000 generasi dengan waktu rata-rata 6.8 detik per generasi, total waktu yang dibutuhkan NEAT untuk menjalankan proses *training* dengan tahap 1 pengujian 1 adalah 571.75 menit. Hasil pengujian dapat dilihat pada Gambar 5.2. Dengan konfigurasi dan *fitness function* yang digunakan, individu terbaik yang dihasilkan NEAT memiliki nilai fitness sebesar 555.5, dengan rata-rata *fitness* pada populasi sebesar 259.39311. Grafik *fitness* terbaik selama pengujian dan rata-rata *fitness* dalam populasi dapat dilihat pada Gambar 5.3.



Gambar 5.3 Grafik nilai *fitness* terbaik dan nilai *fitness* rata-rata dalam populasi selama proses *training* pada Tahap 1 Pengujian 1

Individu terbaik yang dihasilkan oleh NEAT dapat dilihat pada Gambar 5.4. Individu ini memiliki kompleksitas jaringan berupa 22 *input nodes*, 9 *hidden nodes*, dan 2 *output nodes* dengan 23 koneksi aktif. Dimana *nodes* -1 sampai -22 merupakan *input nodes*, 0 dan 1 merupakan *output nodes*, dan sisanya merupakan *hidden nodes*. Garis merah dan hijau menandakan koneksi antar *nodes*, yang berarti *weight* positif untuk garis hijau dan negatif untuk garis merah. Sedangkan garis putus-putus menandakan koneksi yang tidak aktif.



Gambar 5.4 *Network* terbaik yang dihasilkan NEAT pada Tahap 1 Pengujian 1

Setelah *training* selesai, proses selanjutnya adalah *testing* yang bertujuan untuk mengukur tingkat akurasi dari *network* terbaik yang dihasilkan NEAT. Namun setelah dilakukannya proses *testing*, tingkat akurasi dari *network* terbaik ini masih sangat buruk. *Network* hanya mampu memprediksi hasil pertandingan dengan benar sebanyak 107 pertandingan, dimana 32 diantaranya dapat diprediksi dengan skor yang tepat. Ringkasan proses *training* dan proses *testing* dapat dilihat pada Tabel 5.2.

Tabel 5.2 Rangkuman proses *training* dan *testing* Tahap 1 Pengujian 1

|   |  |
|---|--|
| Generasi  | 5000   |
| Rata-rata waktu per generasi  | 6.861 detik                                    |
| Total waktu   | 571.75 menit                                   |
| <i>Fitness Function</i>   | <i>Fitness Function 1</i> (Segmen Program 4.4) |
| Rata-rata nilai <i>fitness</i> dalam populasi                                 | 259.39311                                      |
| Nilai <i>fitness</i> terbaik dalam populasi                                   | 555.5  |
| Kompleksitas <i>network</i> pada individu dengan nilai <i>fitness</i> terbaik | 33 <i>nodes</i> dan 23 koneksi aktif           |
| Skor pertandingan benar   | 32 (8%)  |
| Hasil pertandingan benar  | 107 (28%)                                      |

### 5.2.1.2 Tahap 1 Pengujian 2

Karena hasil Tahap 1 Pengujian 1 kurang memuaskan, pada pengujian ini *fitness function* yang digunakan diganti menjadi *fitness function 2*, yang dapat dilihat pada Segmen Program 4.2.

Sedangkan untuk konfigurasi yang digunakan, masih sama dengan konfigurasi yang digunakan pada Tahap 1 Pengujian 1.

|  |   |                               |
|--|---|-------------------------------|
| [NEAT]                                   | conn_delete_prob = 0.5                  | weight_max_value = 1000       |
| fitness_criterion = max                  |   | weight_min_value = -1000      |
| fitness_threshold = 1140                 | # connection enable options             | weight_mutate_power = 0.1     |
| pop_size = 1000                          | enabled_default = True                  | weight_mutate_rate = 0.8      |
| reset_on_extinction = True               | enabled_mutate_rate = 0.4               | weight_replace_rate = 0.4     |
| no_fitness_termination = True            |   |                               |
| [DefaultGenome]                          | feed_forward = True                     | [DefaultSpeciesSet]           |
| # node activation options                | initial_connection = partial_direct 0.7 | compatibility_threshold = 3.0 |
| activation_default = relu                |   |                               |
| activation_mutate_rate = 0.0             | # node add/remove rates                 | [DefaultStagnation]           |
| activation_options = relu                | node_add_prob = 0.8                     | species_fitness_func = max    |
|  | node_delete_prob = 0.5                  | max_stagnation = 1000         |
| # node aggregation options               |   | species_elitism = 5           |
| aggregation_default = sum                | # network parameters                    |                               |
| aggregation_mutate_rate = 0.0            | num_hidden = 1                          | [DefaultReproduction]         |
| aggregation_options = sum                | num_inputs = 22                         | elitism = 40                  |
|  | num_outputs = 2                         | survival_threshold = 0.5      |
| # node bias options                      | # node response options                 |                               |
| bias_init_mean = 0.0                     | response_init_mean = 1.0                |                               |
| bias_init_stdev = 1.0                    | response_init_stdev = 0.0               |                               |
| bias_max_value = 100.0                   | response_max_value = 3000.0             |                               |
| bias_min_value = -100.0                  | response_min_value = -3000.0            |                               |
| bias_mutate_power = 0.1                  | response_mutate_power = 0.0             |                               |
| bias_mutate_rate = 0.7                   | response_mutate_rate = 0.0              |                               |
| bias_replace_rate = 0.3                  | response_replace_rate = 0.0             |                               |
| # genome compatibility options           | single_structural_mutation = False      |                               |
| compatibility_disjoint_coefficient = 1.0 | structural_mutation_surrer = False      |                               |
| compatibility_weight_coefficient = 0.5   |   |                               |
| # connection add/remove rates            | # connection weight options             |                               |
| conn_add_prob = 0.8                      | weight_init_mean = 0.0                  |                               |
|  | weight_init_stdev = 1.0                 |                               |

Gambar 5.5 Konfigurasi yang digunakan pada Tahap 1 Pengujian 2

Proses *training* pada pengujian ini juga dijalankan sebanyak 5000 generasi. Namun, pengujian ini memakan waktu yang lebih lama dibanding Tahap 1 Pengujian 1. Rata-rata waktu per generasi yang dibutuhkan pada pengujian ini adalah 8.124 detik, sehingga total waktu yang dibutuhkan untuk menyelesaikan proses *training* adalah 677 menit. Hasil *training* dapat dilihat pada Gambar 5.6

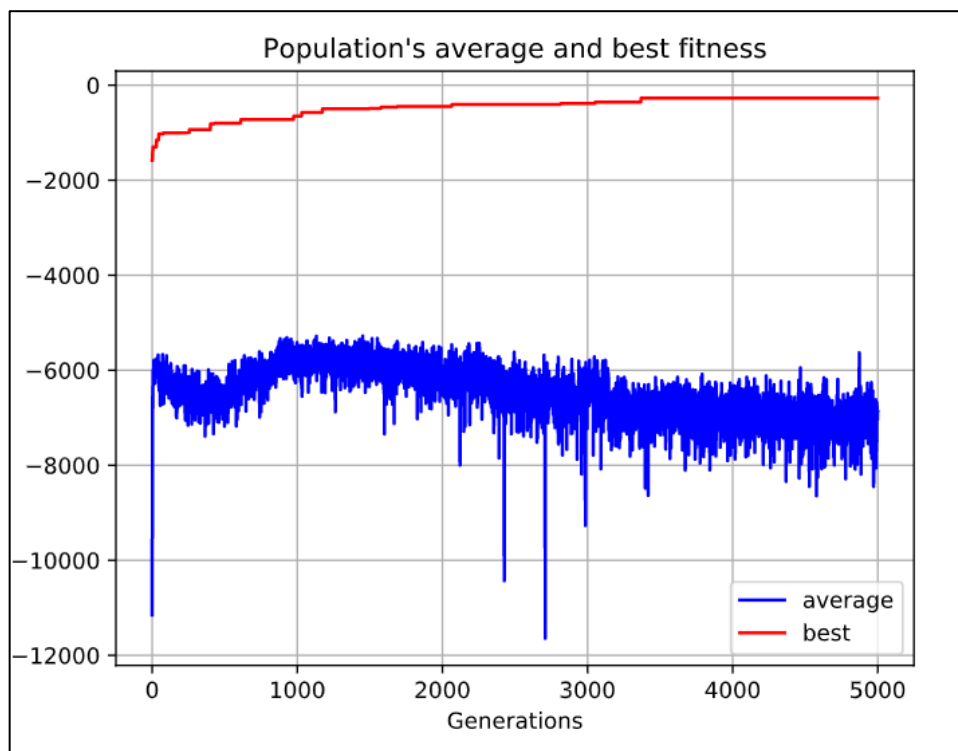
```

Population's average fitness: -7509.48097 stdev: 11127.63908
Best fitness: -272.27539 - size: (7, 29) - species 5 - id 2554238
Average adjusted fitness: 0.963
Mean genetic distance 2.429, standard deviation 0.498
Population of 1000 members in 5 species:
  ID   age  size  fitness  adj fit  stag
  ===  ===  ===  =====  =====  ===
    3  4990   164   -691.2    0.963   3414
    4  4990   346   -328.9    0.962    988
    5  4388   149   -272.3    0.971   1622
    6  4378   165   -372.7    0.961    153
    7  4181   176   -381.0    0.960   2175
Total extinctions: 0
Generation time: 8.124 sec

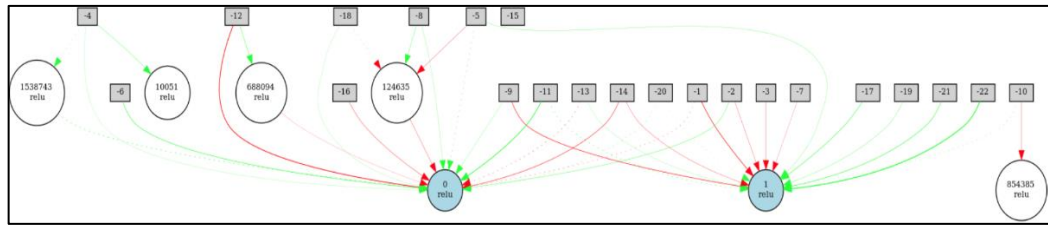
```

Gambar 5.6 Hasil training dari Tahap 1 Pengujian 2

Setelah proses *training* selesai, individu yang keluar sebagai pemenang memiliki nilai *fitness* sebesar -272.275. Perlu diingat bahwa *fitness function* yang digunakan pada penelitian ini berbeda dengan Tahap 1 Pengujian 1, sehingga nilai *fitness* pada penelitian ini tidak bisa dibandingkan secara langsung dengan nilai *fitness* terbaik yang ada pada Tahap 1 Pengujian 1. Grafik *fitness* terbaik selama pengujian dan rata-rata *fitness* dalam populasi dapat dilihat pada Gambar 5.7.



Gambar 5.7 Grafik nilai *fitness* terbaik dan nilai *fitness* rata-rata dalam populasi selama proses *training* pada Tahap 1 Pengujian 2



Gambar 5.8 *Network* terbaik yang dihasilkan NEAT pada Tahap 1 Pengujian 2

Setelah dilakukan proses testing, *network* yang dihasilkan NEAT ternyata memiliki tingkat akurasi yang jauh lebih tinggi jika dibandingkan dengan Tahap 1 Pengujian 1. *Network* pada pengujian ini mampu memprediksi hasil pertandingan sebanyak 273 dari 380 pertandingan, dimana 85 diantaranya dapat diprediksi dengan skor yang benar. Hal ini membuktikan bahwa *fitness function 2* mampu menilai sebuah *network* lebih baik daripada *fitness function 1* karena konfigurasi yang digunakan pada pengujian ini sama dengan konfigurasi yang digunakan pada Tahap 1 Pengujian 1. Rangkuman Tahap 1 Pengujian 2 dapat dilihat pada Tabel 5.3.

Tabel 5.3 Rangkuman proses *training* dan *testing* Tahap 1 Pengujian 2

|   |  |
|---|--|
| Generasi  | 5000   |
| Rata-rata waktu per generasi  | 8.124 detik                                    |
| Total waktu   | 677 menit                                      |
| <i>Fitness Function</i>   | <i>Fitness Function 2</i> (Segmen Program 4.5) |
| Rata-rata nilai <i>fitness</i> dalam populasi                                 | -7509.48                                       |
| Nilai <i>fitness</i> terbaik dalam populasi                                   | -272.275                                       |
| Kompleksitas <i>network</i> pada individu dengan nilai <i>fitness</i> terbaik | 29 <i>nodes</i> dan 29 koneksi aktif           |
| Skor pertandingan benar   | 85 (22%)                                       |
| Hasil pertandingan benar  | 273 (71%)                                      |

### 5.2.1.3 Tahap 1 Pengujian 3

Berdasarkan hasil dari Tahap 1 Pengujian 2, *fitness function* yang akan digunakan pada pengujian ini adalah *fitness function* 2 karena sudah terbukti lebih baik daripada *fitness function* 1.

Konfigurasi pada pengujian ini sedikit berbeda dengan 2 pengujian sebelumnya. Pada pengujian ini, `conn_delete_prob` dan `node_delete_prob` yang mengatur kemungkinan untuk hilangnya koneksi dan *nodes* dalam proses *mutation* diubah menjadi 0 dari 0.5 seperti pada 2 pengujian sebelumnya. Dengan dilakukannya perubahan konfigurasi ini, diharapkan individu dapat berkembang lebih cepat sehingga mampu menghasilkan akurasi yang lebih baik.

|   |  |  |
|---|--|--|
| [NEAT]  | <code>conn_delete_prob = 0.0</code>                  | <code>weight_max_value = 1000</code>       |
| <code>fitness_criterion = max</code>                  |  | <code>weight_min_value = -1000</code>      |
| <code>fitness_threshold = 1140</code>                 | # connection enable options                          | <code>weight_mutate_power = 0.1</code>     |
| <code>pop_size = 1000</code>                          | <code>enabled_default = True</code>                  | <code>weight_mutate_rate = 0.8</code>      |
| <code>reset_on_extinction = True</code>               | <code>enabled_mutate_rate = 0.0</code>               | <code>weight_replace_rate = 0.4</code>     |
| <code>no_fitness_termination = True</code>            |  |  |
| [DefaultGenome]                                       | <code>feed_forward = True</code>                     | [DefaultSpeciesSet]                        |
| # node activation options                             | <code>initial_connection = partial_direct 0.7</code> | <code>compatibility_threshold = 3.0</code> |
| <code>activation_default = relu</code>                | # node add/remove rates                              | [DefaultStagnation]                        |
| <code>activation_mutate_rate = 0.0</code>             | <code>node_add_prob = 0.5</code>                     | <code>species_fitness_func = max</code>    |
| <code>activation_options = relu</code>                | <code>node_delete_prob = 0.0</code>                  | <code>max_stagnation = 1000</code>         |
| # node aggregation options                            | # network parameters                                 | <code>species_elitism = 5</code>           |
| <code>aggregation_default = sum</code>                | <code>num_hidden = 1</code>                          | [DefaultReproduction]                      |
| <code>aggregation_mutate_rate = 0.0</code>            | <code>num_inputs = 22</code>                         | <code>elitism = 40</code>                  |
| <code>aggregation_options = sum</code>                | <code>num_outputs = 2</code>                         | <code>survival_threshold = 0.5</code>      |
| # node bias options                                   | # node response options                              |  |
| <code>bias_init_mean = 0.0</code>                     | <code>response_init_mean = 1.0</code>                |  |
| <code>bias_init_stdev = 1.0</code>                    | <code>response_init_stdev = 0.0</code>               |  |
| <code>bias_max_value = 100.0</code>                   | <code>response_max_value = 3000.0</code>             |  |
| <code>bias_min_value = -100.0</code>                  | <code>response_min_value = -3000.0</code>            |  |
| <code>bias_mutate_power = 0.1</code>                  | <code>response_mutate_power = 0.0</code>             |  |
| <code>bias_mutate_rate = 0.7</code>                   | <code>response_mutate_rate = 0.0</code>              |  |
| <code>bias_replace_rate = 0.3</code>                  | <code>response_replace_rate = 0.0</code>             |  |
| # genome compatibility options                        | <code>single_structural_mutation = False</code>      |  |
| <code>compatibility_disjoint_coefficient = 1.0</code> | <code>structural_mutation_surfer = False</code>      |  |
| <code>compatibility_weight_coefficient = 0.5</code>   | # connection weight options                          |  |
| # connection add/remove rates                         | <code>weight_init_mean = 0.0</code>                  |  |
| <code>conn_add_prob = 0.5</code>                      | <code>weight_init_stdev = 1.0</code>                 |  |

Gambar 5.9 Konfigurasi yang digunakan pada Tahap 1 Pengujian 3



```

Population's average fitness: -12051.08154 stdev: 31720.80446
Best fitness: 75.79604 - size: (14, 70) - species 6 - id 2718253
Average adjusted fitness: 0.981
Mean genetic distance 2.847, standard deviation 0.700
Population of 1000 members in 6 species:
  ID   age  size  fitness  adj fit  stag
  ----  ---  ----  -
  1  4991  167   34.0    0.975   25
  5  4915  166  -78.6    0.977  2386
  6  4915  167   75.8    0.985  1015
  7  4915  166    3.8    0.985   405
  8  4915  167 -100.9    0.979   755
  9  4915  167    4.2    0.985   615
Total extinctions: 0
Generation time: 18.803 sec

```

Gambar 5.10 Hasil *training* dari Tahap 1 Pengujian 3

Efek perubahan konfigurasi dapat dilihat pada hasil *training* pengujian ini pada Gambar 5.10. Dengan mengatur kemungkinan hilangnya koneksi dan *nodes* menjadi 0, koneksi yang dimiliki *network* terbaik pada pengujian ini menjadi jauh lebih banyak dan nilai *fitness* terbaik juga lebih tinggi dari 2 pengujian sebelumnya. Grafik *fitness* terbaik selama pengujian dan rata-rata *fitness* dalam populasi dapat dilihat pada Gambar 5.3.



Gambar 5.11 Grafik nilai *fitness* terbaik dan nilai *fitness* rata-rata dalam populasi selama proses *training* pada Tahap 1 Pengujian 1

Tidak hanya efek positif, perubahan konfigurasi ini juga disertai efek negatif. Pada 2 pengujian sebelumnya, 1 generasi dapat diselesaikan dalam waktu kurang dari 10 detik. Tetapi pada pengujian ini, dibutuhkan rata-rata 18.8 detik untuk menyelesaikan 1 generasi. Sehingga total waktu yang dibutuhkan untuk menyelesaikan 5000 generasi adalah sekitar 1566 menit, jauh lebih lama dari 2 pengujian sebelumnya.

Efek positif yang ditemukan pada proses *training* juga berpengaruh pada proses *testing*. Tingkat akurasi dari *network* yang dihasilkan pada pengujian ini mampu melebihi tingkat akurasi pada 2 pengujian sebelumnya. *Network* pada pengujian ini mampu memprediksi 300 dari 380 pertandingan dengan benar, dimana 125 diantaranya dapat diprediksi dengan skor yang benar. Rangkuman proses *training* dan *testing* dapat dilihat pada Tabel 5.4.

Tabel 5.4 Rangkuman proses *training* dan *testing* pada Tahap 1 Pengujian 3

|   |  |
|---|--|
| Generasi  | 5000   |
| Rata-rata waktu per generasi  | 18.803 detik                                   |
| Total waktu   | 1566 menit                                     |
| <i>Fitness Function</i>   | <i>Fitness Function 2</i> (Segmen Program 4.5) |
| Rata-rata nilai <i>fitness</i> dalam populasi                                 | -12051.081                                     |
| Nilai <i>fitness</i> terbaik dalam populasi                                   | 75.796   |
| Kompleksitas <i>network</i> pada individu dengan nilai <i>fitness</i> terbaik | 36 <i>nodes</i> dan 70 koneksi aktif           |
| Skor pertandingan benar   | 125 (30%)                                      |
| Hasil pertandingan benar  | 300 (78%)                                      |

### 5.2.2 Tahap 2

Pada tahap ini, *feature* yang digunakan adalah *rating* pemain dari kedua *team*, dan *rating* kedua team itu sendiri. Sehingga ada total 24 *feature* yang terdiri dari 11 pemain dari masing-masing *team* dan *rating* dari kedua team.

### 5.2.2.1 Tahap 2 Pengujian 1

Pada pengujian ini, konfigurasi dan *fitness function* yang digunakan sama seperti Tahap 1 Pengujian 3, hanya saja `num_inputs` berubah menjadi 24 dari yang sebelumnya 22 karena adanya penambahan 2 *feature* baru.

```
[NEAT]
fitness_criterion = max
fitness_threshold = 1140
pop_size = 1000
reset_on_extinction = True
no_fitness_termination = True

[DefaultGenome]
# node activation options
activation_default = relu
activation_mutate_rate = 0.0
activation_options = relu

# node aggregation options
aggregation_default = sum
aggregation_mutate_rate = 0.0
aggregation_options = sum

# node bias options
bias_init_mean = 0.0
bias_init_stddev = 1.0
bias_max_value = 100.0
bias_min_value = -100.0
bias_mutate_power = 0.1
bias_mutate_rate = 0.7
bias_replace_rate = 0.3

# genome compatibility options
compatibility_disjoint_coefficient = 1.0
compatibility_weight_coefficient = 0.5

# connection add/remove rates
conn_add_prob = 0.5
conn_delete_prob = 0.0

# connection enable options
enabled_default = True
enabled_mutate_rate = 0.0

feed_forward = True
initial_connection = partial_direct 0.7

# node add/remove rates
node_add_prob = 0.5
node_delete_prob = 0.0

# network parameters
num_hidden = 1
num_inputs = 24
num_outputs = 2

# node response options
response_init_mean = 1.0
response_init_stddev = 0.0
response_max_value = 3000.0
response_min_value = -3000.0
response_mutate_power = 0.0
response_mutate_rate = 0.0
response_replace_rate = 0.0

single_structural_mutation = False
structural_mutation_surer = False

# connection weight options
weight_init_mean = 0.0
weight_init_stddev = 1.0
weight_max_value = 1000
weight_min_value = -1000
weight_mutate_power = 0.1
weight_mutate_rate = 0.8
weight_replace_rate = 0.4

[DefaultSpeciesSet]
compatibility_threshold = 3.0

[DefaultStagnation]
species_fitness_func = max
max_stagnation = 1000
species_elitism = 5

[DefaultReproduction]
elitism = 40
survival_threshold = 0.5
```

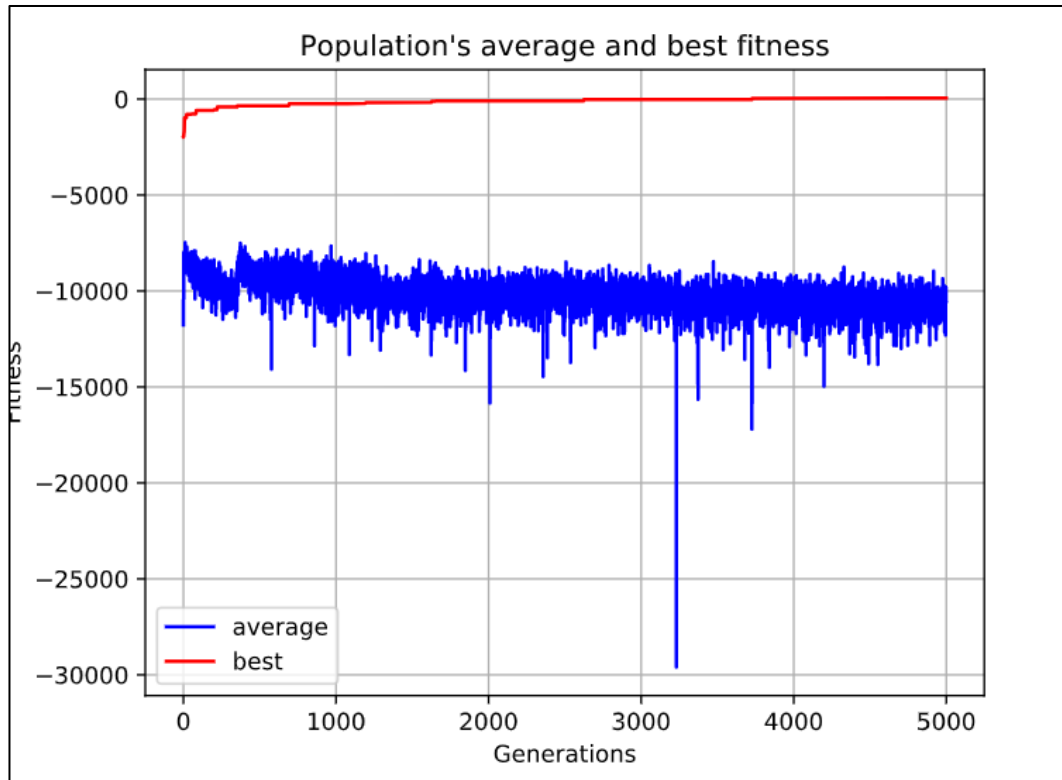
Gambar 5.12 Konfigurasi yang digunakan pada Tahap 2 Pengujian 1

```
Population's average fitness: -11662.11922 stdev: 24262.91578
Best fitness: 50.41736 - size: (9, 61) - species 6 - id 3614032
Average adjusted fitness: 0.971
Mean genetic distance 2.773, standard deviation 0.650
Population of 1000 members in 5 species:
  ID  age  size  fitness  adj fit  stag
  ====  ===  ====  =====  =====  =====
    1  4999   189   -155.7    0.983    48
    3  4999   199   -25.2    0.967   2375
    5  4975   200  -107.9    0.970   3372
    6  4653   213    50.4    0.974    267
    8  4635   199  -211.3    0.962   2201
Total extinctions: 0
Generation time: 16.769 sec
```

Gambar 5.13 Hasil *training* dari Tahap 2 Pengujian 1

Sama seperti proses *training* pada pengujian sebelumnya, *training* pada pengujian ini juga dijalankan selama 5000 generasi. Penambahan *feature* baru pada

pengujian ini ternyata tidak memiliki banyak efek positif pada proses *training*. Dapat dilihat pada Gambar 5.13., nilai *fitness* terbaik yang dihasilkan ternyata lebih kecil daripada nilai *fitness* yang dihasilkan oleh Tahap 1 Pengujian 3. Grafik *fitness* terbaik dan rata-rata *fitness* dalam populasi dapat dilihat pada Gambar 5.14.



Gambar 5.14 Grafik nilai *fitness* terbaik dan nilai *fitness* rata-rata dalam populasi selama proses *training* pada Tahap 2 Pengujian 1

Efek positif penambahan 2 *feature* pada pengujian ini baru terlihat pada setelah proses *testing* selesai dijalankan. Tingkat akurasi dari *network* yang dihasilkan pengujian ini sedikit lebih tinggi daripada Tahap 1 Pengujian 3, yang merupakan *network* dengan tingkat akurasi sejauh ini. *Network* yang dihasilkan pada pengujian mampu memprediksi pertandingan dengan benar sebanyak 302 dari 380 pertandingan, dimana 136 diantaranya dapat diprediksi dengan skor yang tepat. *Network* pada pengujian ini juga memiliki tingkat kompleksitas sedikit lebih rendah dari Tahap 1 Pengujian 3. Rangkuman pengujian ini dapat dilihat pada Tabel 5.5.

Tabel 5.5 Rangkuman proses *training* dan *testing* Tahap 2 Pengujian 1

|   |  |
|---|--|
| Generasi  | 5000   |
| Rata-rata waktu per generasi  | 16.769 detik                                   |
| Total waktu   | 1397 menit                                     |
| <i>Fitness Function</i>   | <i>Fitness Function 2</i> (Segmen Program 4.5) |
| Rata-rata nilai <i>fitness</i> dalam populasi                                 | -11662.11922                                   |
| Nilai <i>fitness</i> terbaik dalam populasi                                   | 50.41736                                       |
| Kompleksitas <i>network</i> pada individu dengan nilai <i>fitness</i> terbaik | 33 <i>nodes</i> dan 61 koneksi aktif           |
| Skor pertandingan benar   | 136 (35%)                                      |
| Hasil pertandingan benar  | 302 (79%)                                      |

### 5.2.2.2 Tahap 2 Pengujian 2

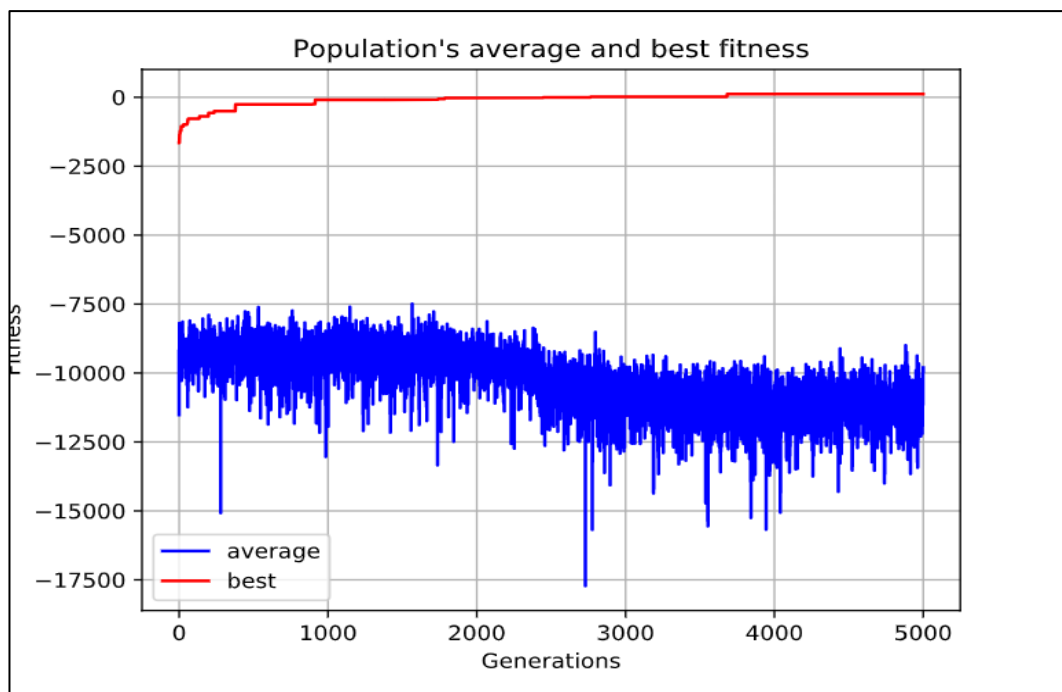
|  |   |   |
|--|---|---|
| [NEAT]<br>fitness_criterion = max<br>fitness_threshold = 1140<br>pop_size = 1000<br>reset_on_extinction = True<br>no_fitness_termination = True<br><br>[DefaultGenome]<br># node activation options<br>activation_default = relu<br>activation_mutate_rate = 0.0<br>activation_options = relu<br><br># node aggregation options<br>aggregation_default = sum<br>aggregation_mutate_rate = 0.0<br>aggregation_options = sum<br><br># node bias options<br>bias_init_mean = 0.0<br>bias_init_stddev = 1.0<br>bias_max_value = 100.0<br>bias_min_value = -100.0<br>bias_mutate_power = 0.1<br>bias_mutate_rate = 0.7<br>bias_replace_rate = 0.3<br><br># genome compatibility options<br>compatibility_disjoint_coefficient = 1.0<br>compatibility_weight_coefficient = 0.5<br><br># connection add/remove rates<br>conn_add_prob = 0.8 | conn_delete_prob = 0.0<br><br># connection enable options<br>enabled_default = True<br>enabled_mutate_rate = 0.0<br><br>feed_forward = True<br>initial_connection = partial_direct 0.7<br><br># node add/remove rates<br>node_add_prob = 0.8<br>node_delete_prob = 0.0<br><br># network parameters<br>num_hidden = 1<br>num_inputs = 24<br>num_outputs = 2<br><br># node response options<br>response_init_mean = 1.0<br>response_init_stddev = 0.0<br>response_max_value = 3000.0<br>response_min_value = -3000.0<br>response_mutate_power = 0.0<br>response_mutate_rate = 0.0<br>response_replace_rate = 0.0<br><br>single_structural_mutation = False<br>structural_mutation_surfer = False<br><br># connection weight options<br>weight_init_mean = 0.0<br>weight_init_stddev = 1.0 | weight_max_value = 1000<br>weight_min_value = -1000<br>weight_mutate_power = 0.1<br>weight_mutate_rate = 0.8<br>weight_replace_rate = 0.4<br><br>[DefaultSpeciesSet]<br>compatibility_threshold = 3.0<br><br>[DefaultStagnation]<br>species_fitness_func = max<br>max_stagnation = 1000<br>species_elitism = 5<br><br>[DefaultReproduction]<br>elitism = 40<br>survival_threshold = 0.5 |
|--|---|---|

Gambar 5.15 Konfigurasi yang digunakan pada Tahap 2 Pengujian 2

Pada 2 pengujian sebelumnya, menghilangkan kemampuan individu untuk kehilangan *nodes* dan koneksi pada saat terjadinya *mutation* terbukti memberikan efek positif pada tingkat akurasi yang dihasilkan *network*. Pada pengujian ini, *node\_add\_prob* dan *conn\_add\_prop* yang mengatur kemungkinan terjadinya penambahan koneksi *nodes* pada proses *mutation* ditingkatkan dari 0.5 menjadi 0.8, dengan harapan *network* dapat berkembang lebih cepat lagi.

```
Population's average fitness: -11531.72723 stdev: 32019.12033
Best fitness: 116.81931 - size: (20, 77) - species 8 - id 2696919
Average adjusted fitness: 0.981
Mean genetic distance 2.868, standard deviation 0.614
Population of 1000 members in 5 species:
  ID   age  size  fitness  adj fit  stag
  ----  ---  ----  -
  1  4987   201   -49.5    0.983   433
  5  4951   200    14.5    0.975  1190
  6  4920   199    34.8    0.981   359
  7  4672   199   -69.8    0.978  1669
  8  4562   201   116.8    0.986  1305
Total extinctions: 0
Generation time: 18.802 sec
```

Gambar 5.16 Hasil *training* dari Tahap 2 Pengujian 2



Gambar 5.17 Grafik nilai *fitness* terbaik dan nilai *fitness* rata-rata dalam populasi selama proses *training* pada Tahap 2 Pengujian 2

Dan ternyata benar, dapat dilihat pada Gambar 5.16., peningkatan kemungkinan terjadinya penambahan koneksi dan *nodes* saat *mutation* memberikan hasil *training* yang lebih baik. Nilai *fitness* yang dihasilkan pada pengujian ini mampu mengungguli nilai *fitness* yang dihasilkan Tahap 2 Pengujian 1 dan Tahap 1 Pengujian 3, yang merupakan 2 *network* terbaik dalam hal nilai *fitness*. Grafik nilai *fitness* terbaik dan rata-rata *fitness* dalam populasi selama proses *training* dapat dilihat pada Gambar 5.17.

Tidak hanya pada proses *training*, efek positif dari konfigurasi yang digunakan juga dapat lihat pada hasil proses *testing*. *Network* yang dihasilkan pengujian ini kembali mampu menghasilkan tingkat akurasi yang lebih tinggi dari pengujian-pengujian sebelumnya. *Network* yang dihasilkan pada pengujian ini mampu memprediksi hasil pertandingan dengan benar sebanyak 308 dari 380 pertandingan, dimana 136 diantaranya dapat diprediksi dengan skor yang tepat.

Tentu saja dengan meningkatkan kemungkinan terjadinya penambahan koneksi dan *nodes* pada saat *mutation* membuat kompleksitas *network* yang dihasilkan menjadi lebih tinggi. Rangkuman proses *training* dan *testing* pada pengujian ini dapat dilihat pada Tabel 5.6

Tabel 5.6 Rangkuman proses *training* dan *testing* dari Tahap 2 Pengujian 2

|   |  |
|---|--|
| Generasi  | 5000   |
| Rata-rata waktu per generasi  | 18.802 detik                                   |
| Total waktu   | 1566 menit                                     |
| <i>Fitness Function</i>   | <i>Fitness Function</i> 2 (Segmen Program 4.5) |
| Rata-rata nilai <i>fitness</i> dalam populasi                                 | -11531.72723                                   |
| Nilai <i>fitness</i> terbaik dalam populasi                                   | 116.81931                                      |
| Kompleksitas <i>network</i> pada individu dengan nilai <i>fitness</i> terbaik | 44 <i>nodes</i> dan 71 koneksi aktif           |
| Skor pertandingan benar   | 136 (35%)                                      |
| Hasil pertandingan benar  | 302(80%)                                       |

### 5.2.3 Tahap 3

Pada tahap ini, *feature* yang digunakan kembali ditambah. Selain *rating* dari masing-masing *team* dan pemain, posisi pemain juga digunakan sebagai *feature*. Terdapat 16 posisi pemain pada *dataset* yang digunakan, dimana masing-masing pemain memiliki 1 posisi. Karena posisi pemain tidak berbentuk angka, harus dilakukan *encoding* terlebih dulu sebelum posisi dapat digunakan. Banyaknya *feature* yang digunakan pada tahap ini bergantung pada jenis *encoding* yang digunakan.

#### 5.2.3.1 Tahap 3 Pengujian 1

Pada pengujian pertama dalam tahap 3, jenis *encoding* yang digunakan adalah label *encoding*. Label *encoding* merubah posisi pemain menjadi sebuah nilai tunggal. Dengan menggunakan *encoding* jenis ini, jumlah *feature* yang digunakan adalah 46, yang terdiri dari 22 *rating* pemain, 22 posisi yang telah diencoding, dan 2 *rating* dari masing-masing *team*.

|  |   |                               |
|--|---|-------------------------------|
| [NEAT]                                   | conn_delete_prob = 0.0                  | weight_max_value = 1000       |
| fitness_criterion = max                  |   | weight_min_value = -1000      |
| fitness_threshold = 1140                 | # connection enable options             | weight_mutate_power = 0.1     |
| pop_size = 1000                          | enabled_default = True                  | weight_mutate_rate = 0.8      |
| reset_on_extinction = True               | enabled_mutate_rate = 0.0               | weight_replace_rate = 0.4     |
| no_fitness_termination = True            |   |                               |
| [DefaultGenome]                          | feed_forward = True                     | [DefaultSpeciesSet]           |
| # node activation options                | initial_connection = partial_direct 0.7 | compatibility_threshold = 3.0 |
| activation_default = relu                |   |                               |
| activation_mutate_rate = 0.0             | # node add/remove rates                 | [DefaultStagnation]           |
| activation_options = relu                | node_add_prob = 0.8                     | species_fitness_func = max    |
|  | node_delete_prob = 0.0                  | max_stagnation = 1000         |
| # node aggregation options               |   | species_elitism = 5           |
| aggregation_default = sum                | # network parameters                    |                               |
| aggregation_mutate_rate = 0.0            | num_hidden = 1                          | [DefaultReproduction]         |
| aggregation_options = sum                | num_inputs = 46                         | elitism = 40                  |
|  | num_outputs = 2                         | survival_threshold = 0.5      |
| # node bias options                      | # node response options                 |                               |
| bias_init_mean = 0.0                     | response_init_mean = 1.0                |                               |
| bias_init_stddev = 1.0                   | response_init_stddev = 0.0              |                               |
| bias_max_value = 100.0                   | response_max_value = 3000.0             |                               |
| bias_min_value = -100.0                  | response_min_value = -3000.0            |                               |
| bias_mutate_power = 0.1                  | response_mutate_power = 0.0             |                               |
| bias_mutate_rate = 0.7                   | response_mutate_rate = 0.0              |                               |
| bias_replace_rate = 0.3                  | response_replace_rate = 0.0             |                               |
| # genome compatibility options           | single_structural_mutation = False      |                               |
| compatibility_disjoint_coefficient = 1.0 | structural_mutation_surfer = False      |                               |
| compatibility_weight_coefficient = 0.5   |   |                               |
| # connection add/remove rates            | # connection weight options             |                               |
| conn_add_prob = 0.8                      | weight_init_mean = 0.0                  |                               |
|  | weight_init_stddev = 1.0                |                               |

Gambar 5.18 Konfigurasi yang digunakan pada Tahap 3 Pengujian 1



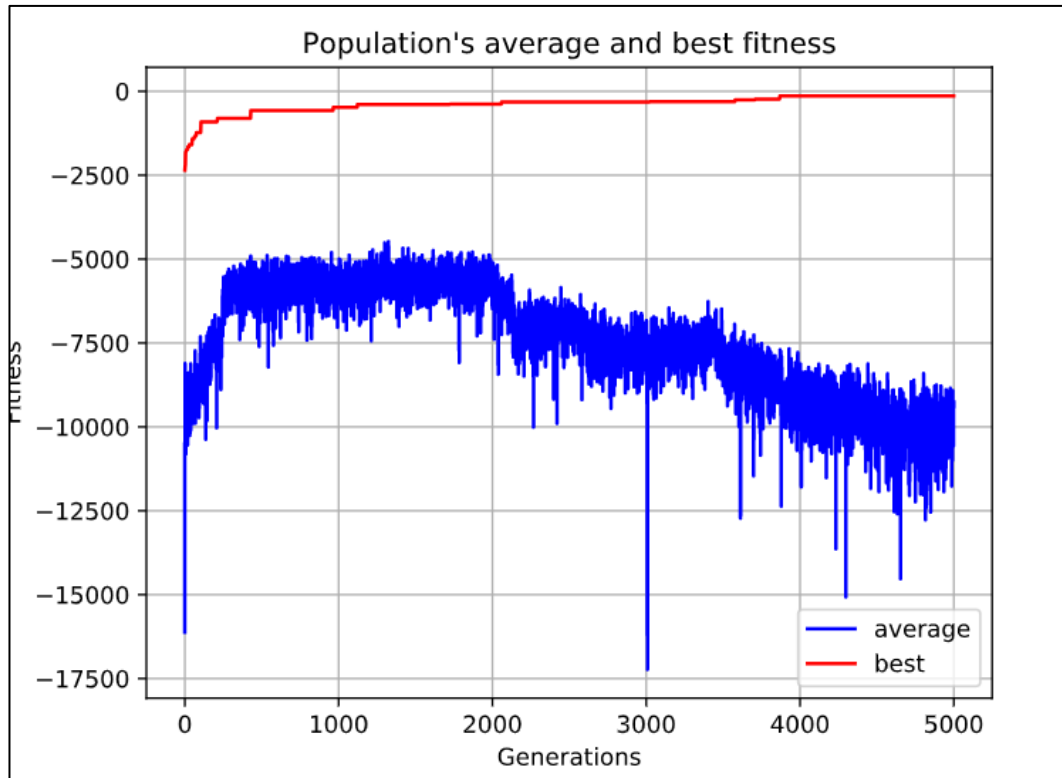
Konfigurasi yang digunakan pada pengujian ini masih sama seperti yang digunakan pada Tahap 2 Pengujian 3, yang merupakan pengujian dengan tingkat akurasi terbaik sejauh ini. Hanya saja, `num_inputs` berubah menjadi 46 dari yang sebelumnya 24. Konfigurasi pada pengujian ini dapat dilihat pada Gambar 5.18.

```
Population's average fitness: -11009.71893 stdev: 24462.22540
Best fitness: -138.71074 - size: (25, 137) - species 8 - id 1892881
Average adjusted fitness: 0.974
Mean genetic distance 2.969, standard deviation 0.608
Population of 1001 members in 7 species:
  ID   age  size  fitness  adj fit  stag
  ===  ===  ===  =====  =====  =====
    2  4988   142   -480.1    0.970   119
    4  4980   142   -232.8    0.974  1252
    7  4974   142   -405.4    0.969   214
    8  4915   144   -138.7    0.976  1121
    9  4895   144   -192.6    0.979    33
   10  4855   143   -307.6    0.970  1966
   12  4751   144   -227.4    0.979   426
Total extinctions: 0
Generation time: 26.809 sec
```

Gambar 5.19 Hasil *training* dari Tahap 3 Pengujian 1

Penambahan *feature* pada pengujian ini membawa banyak efek negatif. Seperti yang dapat dilihat pada Gambar 5.19., waktu yang dibutuhkan untuk menyelesaikan 1 generasi melebihi 20 detik, yang mana tidak pernah terjadi pada penelitian sebelumnya. Untuk menyelesaikan proses *training* sebanyak 5000 generasi, dibutuhkan kurang lebih 2234.08333 menit.

Efek negatif juga dirasakan pada nilai *fitness* yang dihasilkan. Nilai *fitness* terbaik yang dihasilkan pada proses *training* sangat rendah, hanya -138.71074. Hal ini menjadikan pengujian ini sebagai pengujian dengan nilai *fitness* terendah kedua sejauh ini yang menggunakan *fitness function 2*, hanya Tahap 1 Pengujian 2 yang memiliki nilai *fitness* lebih rendah. Belum lagi, kompleksitas *network* juga menjadi tinggi. Grafik nilai *fitness* dan rata-rata nilai *fitness* dalam populasi pada pengujian ini dapat dilihat pada Gambar 5.20



Gambar 5.20 Grafik nilai *fitness* terbaik dan nilai *fitness* rata-rata dalam populasi selama proses *training* pada Tahap 3 Pengujian 1

Efek negatif juga masih terasa pada saat proses testing. Network yang dihasilkan pada pengujian ini hanya mampu memprediksi hasil pertandingan dengan benar sebanyak 260 dari 380 pertandingan, dimana hanya 90 pertandingan yang skornya dapat ditebak dengan benar. Hasil yang kurang memuaskan pada pengujian ini dapat disebabkan oleh *encoding* yang tidak cocok, atau posisi pemain sulit diproses oleh network sehingga akurasi prediksi menjadi rendah. Rangkuman proses *training* dan *testing* dari pengujian ini dapat dilihat pada Tabel 5.7.

Tabel 5.7 Rangkuman proses *training* dan *testing* pada Tahap 3 Pengujian 1

|                              |                       |
|------------------------------|-----------------------|
| Generasi                     | 5000                  |
| Rata-rata waktu per generasi | 26.809 detik          |
| Total waktu                  | 2234.08333 menit      |
| Jenis <i>Encoding</i>        | Label <i>Encoding</i> |

|   |  |
|---|--|
| <i>Fitness Function</i>   | <i>Fitness Function 2</i> (Segmen Program 4.5) |
| Rata-rata nilai <i>fitness</i> dalam populasi                                 | -11009.71893                                   |
| Nilai <i>fitness</i> terbaik dalam populasi                                   | -138.71074                                     |
| Kompleksitas <i>network</i> pada individu dengan nilai <i>fitness</i> terbaik | 71 <i>nodes</i> dan 137 koneksi aktif          |
| Skor pertandingan benar   | 90 (23.6%)                                     |
| Hasil pertandingan benar  | 260 (68.4%)                                    |

### 5.2.3.2 Tahap 3 Pengujian 2

Karena hasil yang tidak memuaskan pada Tahap 3 Pengujian 2, pada pengujian ini jenis *encoding* akan diganti. Jenis *encoding* yang akan digunakan pada pengujian ini adalah *onehot encoding*. Terdapat 16 posisi berbeda yang dapat ditempati pemain, yang menjadikan jumlah *feature* pada pengujian ini menjadi sangat banyak. Total jumlah *feature* yang digunakan pada pengujian ini adalah 376, yang terdiri dari 22 *rating* pemain, 2 *rating team*, dan 22 posisi pemain yang setelah diencode menjadi 352 ( $22 * 16$ ).

|  |  |   |
|--|--|---|
| [NEAT]<br>fitness_criterion = max<br>fitness_threshold = 1140<br>pop_size = 700<br>reset_on_extinction = True<br>no_fitness_termination = True<br><br>[DefaultGenome]<br># node activation options<br>activation_default = relu<br>activation_mutate_rate = 0.0<br>activation_options = relu<br><br># node aggregation options<br>aggregation_default = sum<br>aggregation_mutate_rate = 0.0<br>aggregation_options = sum<br><br># node bias options<br>bias_init_mean = 0.0<br>bias_init_stdev = 1.0<br>bias_max_value = 100.0<br>bias_min_value = -100.0<br>bias_mutate_power = 0.1<br>bias_mutate_rate = 0.7<br>bias_replace_rate = 0.3<br><br># genome compatibility options<br>compatibility_disjoint_coefficient = 1.0<br>compatibility_weight_coefficient = 0.5<br><br># connection add/remove rates<br>conn_add_prob = 0.8 | conn_delete_prob = 0.0<br><br># connection enable options<br>enabled_default = True<br>enabled_mutate_rate = 0.0<br><br>feed_forward = True<br>initial_connection = partial_direct 0.4<br><br># node add/remove rates<br>node_add_prob = 0.8<br>node_delete_prob = 0.0<br><br># network parameters<br>num_hidden = 1<br>num_inputs = 376<br>num_outputs = 2<br><br># node response options<br>response_init_mean = 1.0<br>response_init_stdev = 0.0<br>response_max_value = 3000.0<br>response_min_value = -3000.0<br>response_mutate_power = 0.0<br>response_mutate_rate = 0.0<br>response_replace_rate = 0.0<br><br>single_structural_mutation = False<br>structural_mutation_surfer = False<br><br># connection weight options<br>weight_init_mean = 0.0<br>weight_init_stdev = 1.0 | weight_max_value = 1000<br>weight_min_value = -1000<br>weight_mutate_power = 0.1<br>weight_mutate_rate = 0.8<br>weight_replace_rate = 0.4<br><br>[DefaultSpeciesSet]<br>compatibility_threshold = 3.0<br><br>[DefaultStagnation]<br>species_fitness_func = max<br>max_stagnation = 1000<br>species_elitism = 5<br><br>[DefaultReproduction]<br>elitism = 40<br>survival_threshold = 0.5 |
|--|--|---|

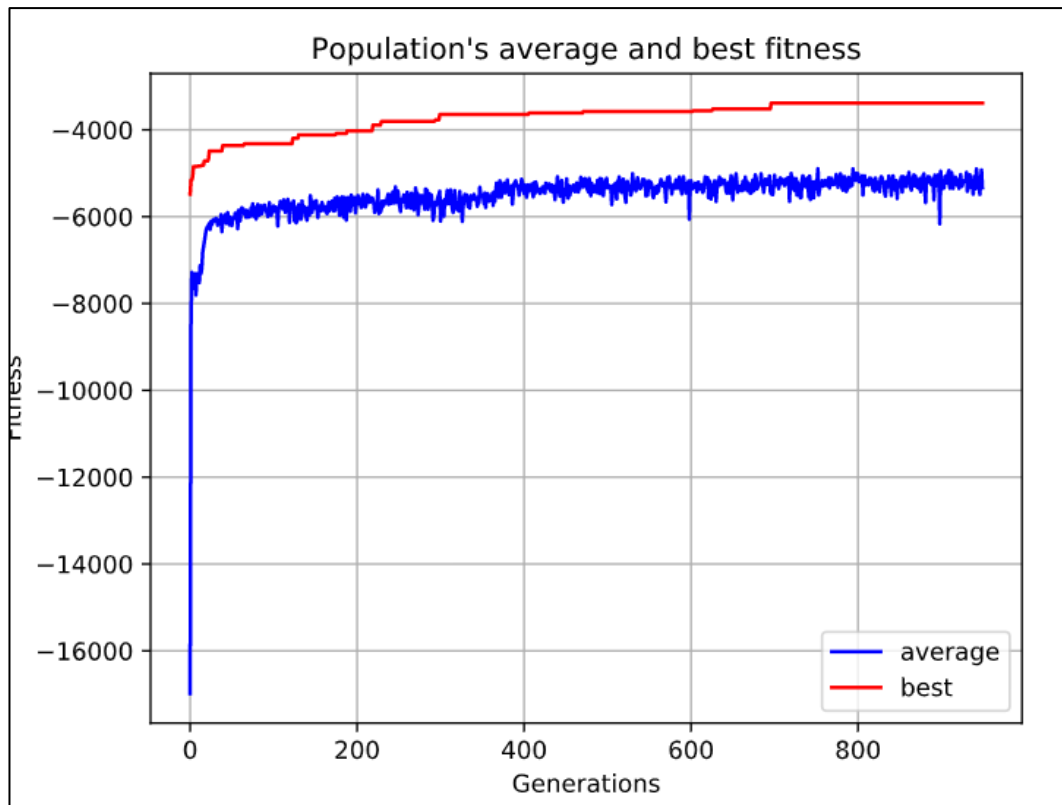
Gambar 5.21 Konfigurasi yang digunakan pada Tahap 3 Pengujian 2

Konfigurasi yang digunakan juga mengalami penyesuaian karena jumlah *feature* yang terlalu besar. Seperti *pop\_size* yang mengatur besarnya jumlah populasi diturunkan menjadi 700 dari 1000, dan *initial\_connection* *partial\_direct*, yang mengatur koneksi awal pada masing-masing individu juga diturunkan menjadi 0.4 dari yang sebelumnya 0.7.

```
Population's average fitness: -5281.77021 stdev: 4672.45793
Best fitness: -3384.64161 - size: (17, 478) - species 10 - id 117370
Average adjusted fitness: 0.968
Mean genetic distance 3.034, standard deviation 0.530
Population of 698 members in 14 species:
  ID  age  size  fitness  adj fit  stag
  ===  ==  ===  =====  =====  ===
    1  995   50  -4192.2    0.973   364
    2  995   50  -3839.4    0.972    57
    3  995   50  -3667.3    0.972   560
    4  993   50  -3798.0    0.973    15
    5  993   49  -4206.4    0.963   147
    6  993   50  -3519.2    0.973   369
    7  984   51  -3869.0    0.970   157
    8  984   50  -3709.3    0.974   353
    9  984   49  -3663.6    0.953   167
   10  983   50  -3384.6    0.972   299
   11  983   51  -3557.6    0.982   392
   12  982   48  -4040.4    0.949    70
   13  982   50  -3842.8    0.955   125
   14  639   50  -3895.3    0.975   205
Total extinctions: 0
Generation time: 55.592 sec
```

Gambar 5.22 Hasil *training* dari Tahap 3 Pengujian 2

Berbeda dari pengujian-pengujian sebelumnya, proses *training* pada pengujian ini dihentikan pada 1000 generasi. Pengujian dihentikan karena *progress* dari nilai *fitness* dan rata-rata nilai *fitness* dalam populasi sangat lambat, dan waktu yang dibutuhkan untuk menjalankan 1 generasi sangat lama. Dapat dilihat pada Gambar 5.22., nilai *fitness* terbaik yang dihasilkan sangat rendah, yaitu -3384.64161 dan waktu yang untuk menyelesaikan 1 generasi lebih dari 50 detik. Untuk menjalankan 1000 generasi saja, total waktu yang dibutuhkan adalah 926.53333 menit. Selain itu, kompleksitas *network*, terutama koneksi menjadi sangat banyak karena banyaknya *feature* yang digunakan. Grafik nilai *fitness* terbaik dan nilai *fitness* rata-rata dalam populasi dapat dilihat pada Gambar 5.23.



Gambar 5.23 Grafik nilai *fitness* terbaik dan nilai *fitness* rata-rata dalam populasi selama proses *training* pada Tahap 3 Pengujian 2

Selaras dengan proses training, tingkat akurasi yang dihasilkan proses testing juga tidak memuaskan. Network hanya mampu memprediksi hasil pertandingan dengan benar sebanyak 194 dari 380 pertandingan, dimana 47 diantaranya dapat diprediksi dengan skor yang benar. Rangkuman proses *training* dan *testing* pada pengujian ini dapat dilihat pada Tabel

Tabel 5.8 Rangkuman proses *training* dan *testing* pada Tahap 3 Pengujian 2

|                              |  |
|------------------------------|--|
| Generasi                     | 1000   |
| Rata-rata waktu per generasi | 55.592 detik                                   |
| Total waktu                  | 926.53333 menit                                |
| Jenis <i>Encoding</i>        | <i>OneHot Encoding</i>                         |
| <i>Fitness Function</i>      | <i>Fitness Function 2</i> (Segmen Program 4.5) |

|   |  |
|---|--|
| Rata-rata nilai <i>fitness</i> dalam populasi                                 | -5281.77021                            |
| Nilai <i>fitness</i> terbaik dalam populasi                                   | -3384.64161                            |
| Kompleksitas <i>network</i> pada individu dengan nilai <i>fitness</i> terbaik | 393 <i>nodes</i> dan 478 koneksi aktif |
| Skor pertandingan benar   | 47 (12 %)                              |
| Hasil pertandingan benar  | 194 (51 %)                             |

### 5.2.3.3 Tahap 3 Pengujian 3

Karena hasil dari Tahap 3 Pengujian 2 masih belum memuaskan juga, pada pengujian kali ini jenis *encoding* akan kembali diganti. Jenis *encoding* yang akan digunakan pada pengujian ini adalah *binary encoding*. *Encoding* jenis ini akan merubah merubah nilai posisi pemain menjadi sebuah nilai *binary*. Keuntungan *encoding* jenis ini dibanding *onehot encoding* adalah dimensi *output* yang dihasilkan lebih sedikit. Karena ada 16 posisi pemain, maka *output* dari *binary encoding* hanya memiliki 4 dimensi, dibandingkan 16 dimensi dari *output* yang dihasilkan *onehot encoding*. Total ada 112 *feature* yang digunakan pada pengujian ini, yang terdiri dari 22 *rating* pemain, 2 *rating* team, dan 88 nilai posisi pemain (22 \* 4).

|   |  |   |
|---|--|---|
| [NEAT]<br>fitness_criterion = max<br>fitness_threshold = 1140<br>pop_size = 1000<br>reset_on_extinction = True<br>no_fitness_termination = True<br><br>[DefaultGenome]<br># node activation options<br>activation_default = relu<br>activation_mutate_rate = 0.0<br>activation_options = relu<br><br># node aggregation options<br>aggregation_default = sum<br>aggregation_mutate_rate = 0.0<br>aggregation_options = sum<br><br># node bias options<br>bias_init_mean = 0.0<br>bias_init_stdev = 1.0<br>bias_max_value = 100.0<br>bias_min_value = -100.0<br>bias_mutate_power = 0.1<br>bias_mutate_rate = 0.7<br>bias_replace_rate = 0.3<br><br># genome compatibility options<br>compatibility_disjoint_coefficient = 1.0<br>compatibility_weight_coefficient = 0.5<br><br># connection add/remove rates<br>conn_add_prob = 0.8 | conn_delete_prob = 0.4<br><br># connection enable options<br>enabled_default = True<br>enabled_mutate_rate = 0.0<br><br>feed_forward = True<br>initial_connection = partial_direct 0.7<br><br># node add/remove rates<br>node_add_prob = 0.8<br>node_delete_prob = 0.4<br><br># network parameters<br>num_hidden = 1<br>num_inputs = 112<br>num_outputs = 2<br><br># node response options<br>response_init_mean = 1.0<br>response_init_stdev = 0.0<br>response_max_value = 3000.0<br>response_min_value = -3000.0<br>response_mutate_power = 0.0<br>response_mutate_rate = 0.0<br>response_replace_rate = 0.0<br><br>single_structural_mutation = False<br>structural_mutation_survival = False<br><br># connection weight options<br>weight_init_mean = 0.0<br>weight_init_stdev = 1.0 | weight_max_value = 1000<br>weight_min_value = -1000<br>weight_mutate_power = 0.1<br>weight_mutate_rate = 0.8<br>weight_replace_rate = 0.4<br><br>[DefaultSpeciesSet]<br>compatibility_threshold = 3.0<br><br>[DefaultStagnation]<br>species_fitness_func = max<br>max_stagnation = 1000<br>species_elitism = 5<br><br>[DefaultReproduction]<br>elitism = 40<br>survival_threshold = 0.5 |
|---|--|---|

Gambar 5.24 Konfigurasi yang digunakan pada Tahap 3 Pengujian 3

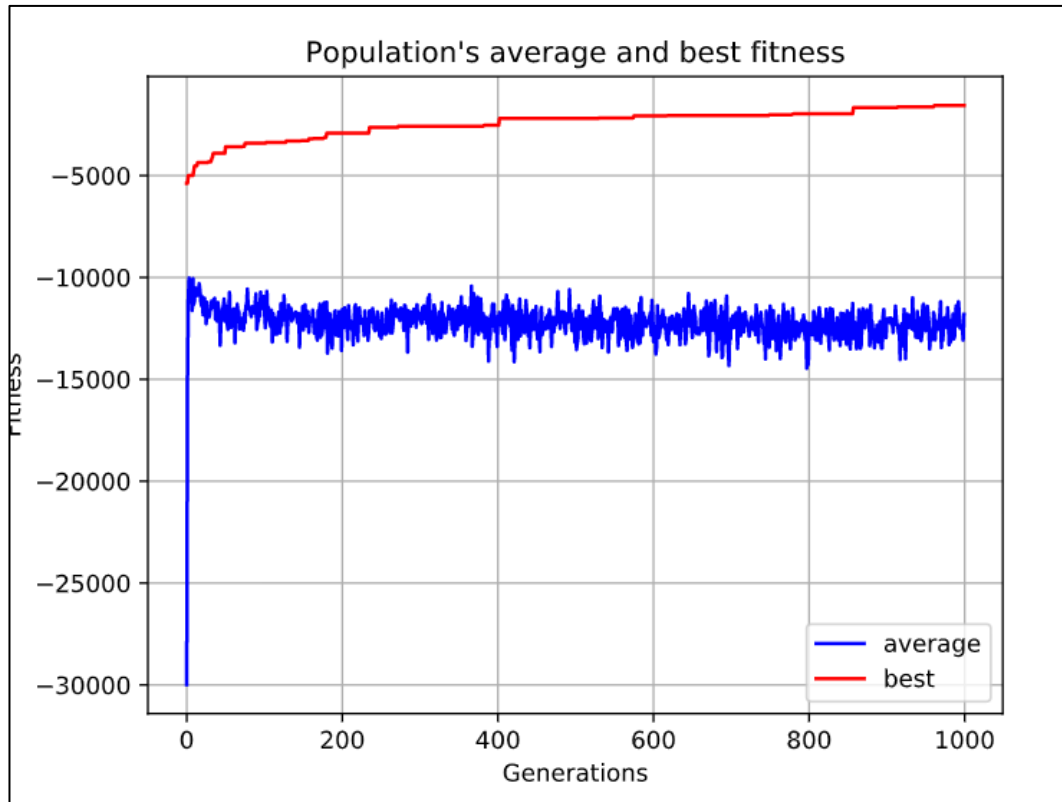
Penyesuaian konfigurasi juga kembali dilakukan pada penelitian ini. Selain `num_inputs` yang berganti menjadi 112, `pop_size` yang mengatur jumlah populasi juga kembali ditingkatkan menjadi 1000 karena jumlah *feature* tidak sebanyak Tahap 3 Pengujian 2. Kemampuan *network* untuk kehilangan koneksi dan *nodes* yang diatur oleh `conn_delete_prob` dan `nodes_delete_prop` juga dikembali diaktifkan, dan diatur ke 0.4.

```
Population's average fitness: -12364.50885 stdev: 18992.98633
Best fitness: -1560.54523 - size: (16, 114) - species 3 - id 860255
Average adjusted fitness: 0.945
Mean genetic distance 2.747, standard deviation 0.443
Population of 999 members in 3 species:
  ID  age  size  fitness  adj fit  stag
  ===  ==  ===  =====  =====  =====
    1  995   332  -1954.1    0.942    41
    2  949   335  -1836.1    0.947   123
    3  697   332  -1560.5    0.947    34
Total extinctions: 0
Generation time: 29.952 sec
```

Gambar 5.25 Hasil *training* dari Tahap 3 Pengujian 3

Sama seperti pengujian sebelumnya, proses *training* pengujian ini juga dihentikan pada 1000 generasi. Alasan dihentikannya proses *training* ini juga sama, yaitu waktu yang terlalu lama disertai *progress* dari nilai *fitness* sangat lambat. Waktu yang dibutuhkan untuk menyelesaikan 1 generasi pada proses *training* ini sekitar 29 detik. Nilai *fitness* yang dihasilkan proses *training* pada pengujian ini juga terbilang kecil, hanya -1560.5453 dengan rata-rata nilai *fitness* pada populasi sebesar -12364.50885. Grafik *fitness* terbaik selama pengujian dan rata-rata *fitness* dalam populasi dapat dilihat pada Gambar 5.26.

Hasil dari proses *testing* pada pengujian ini juga kurang memuaskan. *Network* hanya mampu menebak hasil pertandingan dengan benar sebanyak 223 dari 380 pertandingan, dimana 46 pertandingan diantaranya dapat ditebak dengan skor yang benar. Tingkat akurasi yang dihasilkan pengujian ini hanya sedikit lebih baik daripada Tahap 3 Pengujian 2 dan masih tertinggal cukup jauh daripada pengujian dengan tingkat akurasi terbaik yang pernah dilakukan, yaitu pada Tahap 2 Pengujian 2. Rangkuman proses *training* dan *testing* pada pengujian ini dapat dilihat pada Tabel 5.9.



Gambar 5.26 Grafik nilai *fitness* terbaik dan nilai *fitness* rata-rata dalam populasi selama proses *training* pada Tahap 3 Pengujian 3

Tabel 5.9 Rangkuman proses *training* dan *testing* pada Tahap 3 Pengujian 3

|   |  |
|---|--|
| Generasi  | 1000   |
| Rata-rata waktu per generasi  | 28.952 detik                                   |
| Total waktu   | 482.53333                                      |
| Jenis <i>Encoding</i>   | <i>Binary Encoding</i>                         |
| <i>Fitness Function</i>   | <i>Fitness Function 2</i> (Segmen Program 4.5) |
| Rata-rata nilai <i>fitness</i> dalam populasi                                 | -12364.50885                                   |
| Nilai <i>fitness</i> terbaik dalam populasi                                   | -1560.54523                                    |
| Kompleksitas <i>network</i> pada individu dengan nilai <i>fitness</i> terbaik | 128 <i>nodes</i> dan 114 koneksi aktif         |
| Skor pertandingan benar   | 46 (12 %)                                      |
| Hasil pertandingan benar  | 223 (58 %)                                     |



#### 5.2.3.4 Tahap 3 Pengujian 4

Berbeda dengan pengujian lainnya yang dilakukan pada Tahap 3., posisi setiap pemain pada pengujian ini tidak diencode satu persatu. Jumlah dari masing-masing *defender*, *midfielder*, dan *striker*lah yang menjadi parameter tambahan. Dengan begitu, jumlah *features* yang digunakan pada pengujian ini hanya 30, yang terdiri dari 11 *rating* pemain, 1 *rating team*, 1 jumlah *defender*, 1 jumlah *midfielder*, 1 jumlah *striker* dari masing masing *team*.

|  |   |                               |
|--|---|-------------------------------|
| [NEAT]                                   | conn_delete_prob = 0.0                  | weight_max_value = 1000       |
| fitness_criterion = max                  |   | weight_min_value = -1000      |
| fitness_threshold = 1140                 | # connection enable options             | weight_mutate_power = 0.1     |
| pop_size = 1000                          | enabled_default = True                  | weight_mutate_rate = 0.8      |
| reset_on_extinction = True               | enabled_mutate_rate = 0.0               | weight_replace_rate = 0.4     |
| no_fitness_termination = True            |   |                               |
| [DefaultGenome]                          | feed_forward = True                     | [DefaultSpeciesSet]           |
| # node activation options                | initial_connection = partial_direct 0.8 | compatibility_threshold = 3.0 |
| activation_default = relu                | # node add/remove rates                 | [DefaultStagnation]           |
| activation_mutate_rate = 0.0             | node_add_prob = 0.8                     | species_fitness_func = max    |
| activation_options = relu                | node_delete_prob = 0.0                  | max_stagnation = 1000         |
| # node aggregation options               | # network parameters                    | species_elitism = 5           |
| aggregation_default = sum                | num_hidden = 2                          | [DefaultReproduction]         |
| aggregation_mutate_rate = 0.0            | num_inputs = 30                         | elitism = 40                  |
| aggregation_options = sum                | num_outputs = 2                         | survival_threshold = 0.5      |
| # node bias options                      | # node response options                 |                               |
| bias_init_mean = 0.0                     | response_init_mean = 1.0                |                               |
| bias_init_stddev = 1.0                   | response_init_stddev = 0.0              |                               |
| bias_max_value = 100.0                   | response_max_value = 3000.0             |                               |
| bias_min_value = -100.0                  | response_min_value = -3000.0            |                               |
| bias_mutate_power = 0.1                  | response_mutate_power = 0.0             |                               |
| bias_mutate_rate = 0.7                   | response_mutate_rate = 0.0              |                               |
| bias_replace_rate = 0.3                  | response_replace_rate = 0.0             |                               |
| # genome compatibility options           | single_structural_mutation = False      |                               |
| compatibility_disjoint_coefficient = 1.0 | structural_mutation_surfer = False      |                               |
| compatibility_weight_coefficient = 0.5   | # connection weight options             |                               |
| # connection add/remove rates            | weight_init_mean = 0.0                  |                               |
| conn_add_prob = 0.8                      | weight_init_stddev = 1.0                |                               |

Gambar 5.27 Konfigurasi yang digunakan pada Tahap 3 Pengujian 4

Konfigurasi yang digunakan pada pengujian ini masih sama dengan konfigurasi yang digunakan pada Tahap 3 Pengujian 3., hanya saja jumlah input berkurang menjadi 30, yang menjadikan pengujian ini pengujian dengan jumlah *features* paling sedikit pada Tahap 3. Konfigurasi dapat dilihat pada Gambar 5.27.

```

Population's average fitness: -9819.25758 stdev: 25635.21339
Best fitness: 27.67263 - size: (12, 114) - species 9 - id 895896
Average adjusted fitness: 0.981
Mean genetic distance 2.942, standard deviation 0.576
Population of 1000 members in 8 species:
  ID   age  size  fitness  adj fit  stag
  ===  ===  ===  =====  =====  =====
    2  4990   125   -122.0    0.976   1586
    3  4990   124   -229.2    0.980    586
    8  4971   125    -66.7    0.985   1209
    9  4969   125    27.7    0.980   2673
   10  4969   126   -43.8    0.983    920
   11  4957   125  -184.7    0.984    722
   14  4857   125  -186.2    0.976    108
   16  4752   125   -38.1    0.984    620
Total extinctions: 0
Generation time: 25.464 sec

```

Gambar 5.28 Hasil *training* dari Tahap 3 Pengujian 4

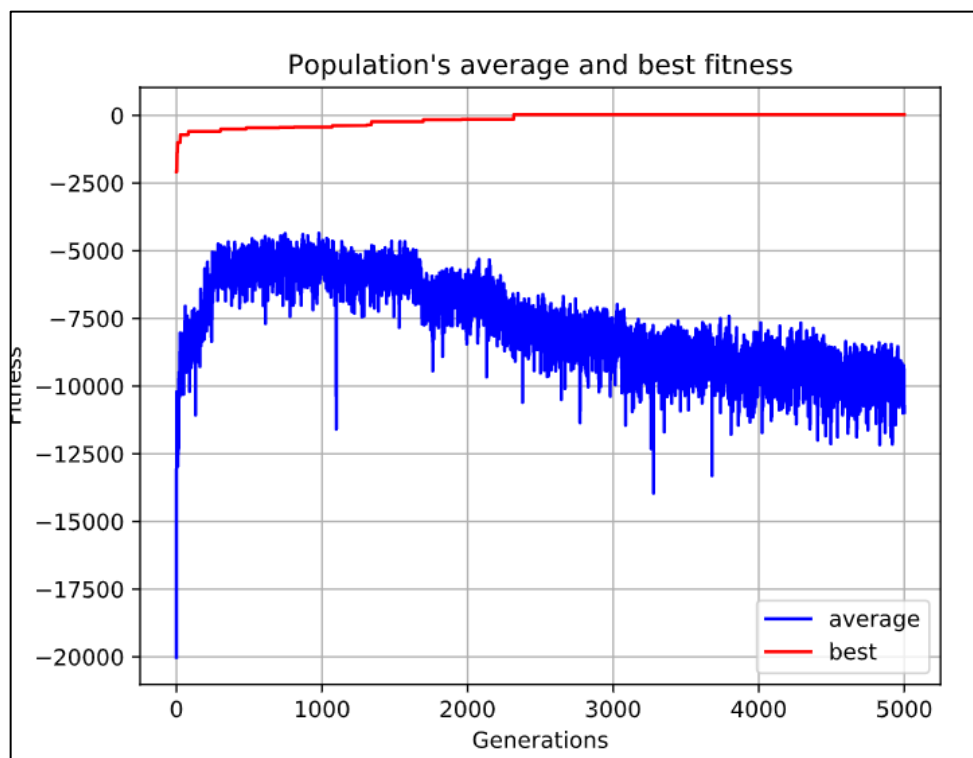
Ternyata, penggunaan *encoding* posisi pemain secara keseluruhan membawa hasil positif pada proses *training*. Selain lebih singkatnya waktu yang dibutuhkan untuk menyelesaikan satu generasi, nilai fitness terbaik yang dihasilkan, 27.67263, juga melebihi seluruh nilai fitness terbaik yang dihasilkan oleh pengujian pada Tahap 3. Proses training ini dijalankan sebanyak 5000 generasi dengan dengan rata-rata waktu yang dibutuhkan sebanyak 25 detik pergenerasi. Grafik *fitness* terbaik selama pengujian dan rata-rata *fitness* dalam populasi dapat dilihat pada Gambar 5.29.

Hasil positif juga berlanjut pada proses *testing*. Akurasi yang dihasilkan pengujian ini juga melampaui semua akurasi yang dihasilkan oleh pengujian lain pada tahap ini, baik dalam akurasi hasil maupun akurasi skor pertandingan. *Network* yang dihasilkan oleh pengujian ini mampu menebak 299 dari 380 pertandingan dengan benar, dimana 131 diantaranya dapat ditebak dengan skor yang tepat.

Tabel 5.10 Rangkuman proses *training* dan *testing* pada Tahap 3 Pengujian 4

|                              |  |
|------------------------------|--|
| Generasi                     | 5000   |
| Rata-rata waktu per generasi | 25.464   |
| Total waktu                  | 2122   |
| <i>Fitness Function</i>      | <i>Fitness Function 2</i> (Segmen Program 4.5) |

|   |                                       |
|---|---------------------------------------|
| Rata-rata nilai <i>fitness</i> dalam populasi                                 | -9819.25758                           |
| Nilai <i>fitness</i> terbaik dalam populasi                                   | 27.67263                              |
| Kompleksitas <i>network</i> pada individu dengan nilai <i>fitness</i> terbaik | 42 <i>nodes</i> dan 114 koneksi aktif |
| Skor pertandingan benar   | 131 (34 %)                            |
| Hasil pertandingan benar  | 299 (78.6 %)                          |



Gambar 5.29 Grafik nilai *fitness* terbaik dan nilai *fitness* rata-rata dalam populasi selama proses *training* pada Tahap 3 Pengujian 4

#### 5.2.4 Kesimpulan Pengujian NEAT

Selain *feature*, *fitness function* juga memiliki peranan yang sangat penting pada NEAT. Dapat dilihat pada Tahap 1 Pengujian 1 dan Tahap 1 Pengujian 2, terjadi peningkatan akurasi yang signifikan. Pengujian dilakukan dengan *feature* dan konfigurasi yang sama, namun *fitness function* yang digunakan berbeda. Tingkat akurasi hasil pertandingan dan akurasi skor yang dihasilkan Tahap 1

Pengujian 1, yaitu 28% dan 8% meningkat menjadi 71% dan 22% pada Tahap 1 Pengujian 2.

Selain *fitness function* yang digunakan, konfigurasi juga memiliki peranan penting ketika proses NEAT berjalan. Pentingnya konfigurasi dapat dilihat dari Tahap 1 Pengujian 2 dan Tahap 1 Pengujian 3. Dengan menggunakan *feature* yang sama, yaitu *rating* pemain, terjadi peningkatan akurasi dari 71% untuk akurasi prediksi hasil pertandingan dan 22% untuk akurasi prediksi skor yang dihasilkan oleh Tahap 1 Pengujian 2, menjadi 78% dan 32% yang dihasilkan oleh Tahap 1 Pengujian 3.

Penambahan 2 *feature* baru, yaitu *rating team* juga ikut meningkatkan tingkat akurasi dari sebuah *network*, walaupun tidak signifikan. Dengan menggunakan konfigurasi dan *fitness function* yang sama, akurasi prediksi hasil pertandingan dan akurasi prediksi skor meningkat dari 78% dan 32% yang dihasilkan oleh Tahap 1 Pengujian 3, menjadi 79% dan 35% pada Tahap 2 dan Pengujian 1.

Tingkat akurasi terbaik pada pengujian NEAT dihasilkan pada Tahap 2 Pengujian 2. Dengan sedikit melakukan perubahan konfigurasi yang digunakan pada Tahap 2 Pengujian 1, Tahap 2 Pengujian 2 mampu menghasilkan akurasi yang lebih baik, yaitu sebesar 81% untuk akurasi prediksi hasil pertandingan, dan 42% untuk akurasi prediksi skor.

Pada Tahap 3, posisi pemain ikut menjadi *feature* pada proses NEAT. Dari 3 pengujian yang mengaplikasikan *encoding* kepada posisi masing-masing pemain dan 1 pengujian yang menggunakan total dari posisi pemain pada setiap sektor, hasil terbaik diperoleh pada Tahap 3 Pengujian 4 yang menggunakan total posisi pemain pada tiap sektor. Pengujian ini menghasilkan tingkat akurasi sebesar 78% untuk akurasi hasil pertandingan dan 34% untuk akurasi skor.

Selain tingkat akurasi yang dihasilkan tidak sebaik Tahap 2, waktu yang dibutuhkan untuk menyelesaikan proses *training* juga jauh lebih lama. Pada Tahap 2, waktu yang dibutuhkan untuk menyelesaikan 1 generasi dalam proses *training* tidak pernah melebihi 20 detik. Sedangkan pada Tahap 3, waktu tercepat yang dibutuhkan untuk menyelesaikan 1 generasi adalah sekitar 25 detik.

Berdasarkan beberapa pengujian yang dilakukan dengan berbagai *features*, *fitness function*, dan konfigurasi yang berbeda, dapat disimpulkan bahwa pengujian yang menghasilkan tingkat akurasi terbaik adalah Tahap 2 Pengujian 3 dengan *features* berupa *rating* pemain dan *rating team*. Rangkuman pengujian dapat dilihat pada Tabel 5.11.

Tabel 5.11 Rangkuman seluruh pengujian NEAT

| Pengujian           | Akurasi Hasil<br>Pertandingan | Akurasi Skor<br>Pertandingan |
|---------------------|-------------------------------|------------------------------|
| Tahap 1 Pengujian 1 | 107 / 380 (28%)               | 32 / 380 (8%)                |
| Tahap 1 Pengujian 2 | 273 / 380 (71%)               | 85 / 380 (22%)               |
| Tahap 1 Pengujian 3 | 300 / 380 (78%)               | 125 / 300 (32%)              |
| Tahap 2 Pengujian 1 | 302 / 380 (79%)               | 136 / 380 (35%)              |
| Tahap 2 Pengujian 2 | 305 / 380 (80%)               | 136 / 380 (35%)              |
| Tahap 3 Pengujian 1 | 260 / 380 (68%)               | 90 / 380 (23%)               |
| Tahap 3 Pengujian 2 | 194 / 380 (51%)               | 47 / 380 (12%)               |
| Tahap 3 Pengujian 3 | 223 / 380 (52 %)              | 46 / 380 (12%)               |
| Tahap 3 Pengujian 4 | 299 / 380 (78.6 %)            | 131 / 380 (34 %)             |

### 5.3 Pengujian *Backpropagation*

Setelah proses NEAT selesai dijalankan, proses selanjutnya adalah *backpropagation*. Proses *backpropagation* bertujuan untuk mengoptimasi *weight* dari *network* yang dihasilkan oleh NEAT. Hasil dari proses *backpropagation* dapat dilihat pada Tabel 5.12.

Tabel 5.12 Hasil pengujian *backpropagation*

| Pengujian              | <i>Error</i><br>Terakhir | <i>Epoch</i> | Akurasi<br>Hasil<br>Pertandingan | Akurasi<br>Skor<br>Pertandingan | Total<br>Waktu |
|------------------------|--------------------------|--------------|----------------------------------|---------------------------------|----------------|
| Tahap 1<br>Pengujian 1 | 759.55                   | 100.000      | 291 / 380<br>(76%)               | 94 / 380<br>(24%)               | 264<br>detik   |

|                        |         |         |                      |                     |               |
|------------------------|---------|---------|----------------------|---------------------|---------------|
| Tahap 1<br>Pengujian 2 | 460.110 | 100.000 | 289 / 380<br>(76%)   | 134 / 380<br>(35%)  | 234<br>detik  |
| Tahap 1<br>Pengujian 3 | 433.55  | 44.000  | 301 / 380<br>(79%)   | 155 / 380<br>(40 %) | 181<br>detik  |
| Tahap 2<br>Pengujian 1 | 351.01  | 100.000 | 303 / 380<br>(79 %)  | 167 / 380<br>(44 %) | 355<br>detik  |
| Tahap 2<br>Pengujian 2 | 395.55  | 100.000 | 308 / 380<br>(81%)   | 161 / 380<br>(42%)  | 482<br>detik  |
| Tahap 3<br>Pengujian 1 | 386.29  | 100.000 | 299 / 380<br>(78%)   | 153 / 380<br>(40%)  | 789<br>detik  |
| Tahap 3<br>Pengujian 2 | 492.47  | 100.000 | 306 / 380<br>(80%)   | 122 / 380<br>(32%)  | 2269<br>detik |
| Tahap 3<br>Pengujian 3 | 469.00  | 100.000 | 304 / 380<br>(80%)   | 162 / 380<br>(42%)  | 1199<br>detik |
| Tahap 3<br>Pengujian 4 | 326.57  | 100.000 | 310 / 380<br>(81.5%) | 185 / 380<br>(48%)  | 656<br>detik  |

Semua pengujian *backpropagation* dijalankan dengan *learning rate* sebesar 0.0001 dan selama 100.000 *epoch*, kecuali untuk Tahap 1 Pengujian 3 karena terjadinya *overfitting* pada *epoch* 44.000 – 45.000 yang menyebabkan proses *backpropagation* terpaksa dihentikan.

Berdasarkan pengujian-pengujian yang telah dilakukan, proses *backpropagation* terbukti mampu meningkatkan tingkat akurasi pada *network* yang dihasilkan oleh NEAT pada penelitian ini dengan rata-rata peningkatan akurasi sebesar 15,25% pada akurasi hasil pertandingan, dan 12,625% pada akurasi skor pertandingan.

Peningkatan paling signifikan pada *network* setelah dilakukannya *backpropagation* terjadi pada Tahap 1 Pengujian 1. *Network* yang dihasilkan oleh Tahap 1 Pengujian 1 pada awalnya hanya memiliki akurasi sebesar 28% dan 8%. Namun setelah proses *backpropagation* diaplikasikan pada pengujian tersebut, tingkat akurasi mengalami kenaikan menjadi 76% dan 24%. Hal ini membuktikan

bahwa walaupun tingkat akurasi yang dihasilkan oleh *network* dari proses NEAT rendah, struktur yang dimiliki *network* itu sudah cukup baik.

Perbandingan tingkat akurasi sebelum dan sesudah dilakukannya *backpropagation* dapat dilihat pada Tabel 5.13.

Tabel 5.13 Perbandingan akurasi dari *network* sebelum dan sesudah *backpropagation*

| Pengujian              | Sebelum<br><i>Backpropagation</i> |                     | Sesudah<br><i>Backpropagation</i> |                     |
|------------------------|-----------------------------------|---------------------|-----------------------------------|---------------------|
|                        | Akurasi Hasil                     | Akurasi Skor        | Akurasi Hasil                     | Akurasi Skor        |
| Tahap 1<br>Pengujian 1 | 107 / 380<br>(28%)                | 32 / 380<br>(8%)    | 291 / 380<br>(76%)                | 94 / 380<br>(24%)   |
| Tahap 1<br>Pengujian 2 | 273 / 380<br>(71%)                | 85 / 380<br>(22%)   | 289 / 380<br>(76%)                | 134 / 380<br>(35%)  |
| Tahap 1<br>Pengujian 3 | 300 / 380<br>(78%)                | 125 / 300<br>(32%)  | 301 / 380<br>(79%)                | 155 / 380<br>(40 %) |
| Tahap 2<br>Pengujian 1 | 302 / 380<br>(79%)                | 136 / 380<br>(35%)  | 303 / 380<br>(79 %)               | 167 / 380<br>(44 %) |
| Tahap 2<br>Pengujian 2 | 305 / 380<br>(80%)                | 136 / 380<br>(35%)  | 308 / 380<br>(81%)                | 161 / 380<br>(42%)  |
| Tahap 3<br>Pengujian 1 | 260 / 380<br>(68%)                | 90 / 380<br>(23%)   | 299 / 380<br>(78%)                | 153 / 380<br>(40%)  |
| Tahap 3<br>Pengujian 2 | 194 / 380<br>(51%)                | 47 / 380<br>(12%)   | 306 / 380<br>(80%)                | 122 / 380<br>(32%)  |
| Tahap 3<br>Pengujian 3 | 223 / 380<br>(52 %)               | 46 / 380<br>(12%)   | 304 / 380<br>(80%)                | 162 / 360<br>(42%)  |
| Tahap 3<br>Pengujian 4 | 299 / 380<br>(78.6 %)             | 131 / 380<br>(34 %) | 310 / 380<br>(81.5%)              | 185 / 380<br>(48%)  |

Setelah *backpropagation* selesai dijalankan, pengujian yang menghasilkan *network* dengan akurasi terbaik berubah. Tahap 3 Pengujian 4 mengalahkan Tahap 2 Pengujian 2 yang sebelumnya menjadi pengujian dengan akurasi terbaik.

Akurasi yang dihasilkan oleh Tahap 3 Pengujian 4 meningkat dari yang sebelumnya 78% untuk akurasi hasil pertandingan dan 34% akurasi skor pertandingan, menjadi 81% dan 48%. Akurasi yang dihasilkan oleh pengujian ini mengalahkan semua akurasi dari semua pengujian yang telah dilakukan.

#### **5.4 Pengujian *Real Life***

Pada data pertandingan *real*, tidak diketahui berapa *rating* pemain yang akan bermain pada pertandingan tersebut. Untuk memprediksi *rating* pemain yang akan bermain, akan dicoba beberapa cara, yaitu rata-rata 5 pertandingan terakhir, rata-rata 10 pertandingan terakhir.

Untuk rata-rata 5 dan 10 pertandingan terakhir, akan diambil *rating* dari pemain dan *team* pada 5 dan 10 pertandingan terakhir sebelum pertandingan akan dicoba untuk diprediksi. Sebagai contoh, pada pertandingan ke 20, akan dihitung rata-rata pada pertandingan ke-14 sampai ke-19 atau pertandingan ke-9 sampai pertandingan ke-19. Jika data yang tersedia kurang dari 5 atau 10, akan dihitung rata-rata dari jumlah pertandingan yang tersedia dalam rentang 5 atau 10 pertandingan terakhir. Jika seorang pemain tidak pernah bermain pada 5 atau 10 pertandingan terakhir, akan diberi *rating* 0.55.

Selain itu, akan diuji juga rata-rata 5 dan 10 pertandingan terakhir yang sudah dinormalisasi. Normalisasi dilakukan dengan cara mengalikan *rating* pemain dengan *rating team* lawan, lalu dibagi dengan *rating team* dari pemain itu sendiri.

Agar tersedianya cukup data, pengujian *real life* dimulai dengan pertandingan dari pertandingan ke-10 untuk rata-rata 5 pertandingan terakhir, dan pertandingan ke-15 untuk rata-rata 10 pertandingan terakhir.

Pengujian *real life* dilakukan dengan *network* hasil pengujian *backpropagation* yang memiliki tingkat akurasi tertinggi, yaitu *network* dari Tahap 3 Pengujian 4.



Tabel 5.14 Hasil pengujian *real life*

| Metode Prediksi<br>Rating                                      | Akurasi Hasil<br>Pertandingan | Akurasi Skor<br>Pertandingan |
|--|-------------------------------|------------------------------|
| Rata-rata 5<br>pertandingan<br>terakhir                        | 124 / 280<br>(44%)            | 28 / 280<br>(10%)            |
| Rata-rata 5<br>pertandingan<br>terakhir dengan<br>normalisasi  | 78 / 280<br>(27.8%)           | 25 / 280<br>(8.9%)           |
| Rata-rata 10<br>pertandingan<br>terakhir                       | 92 / 230<br>(40%)             | 21 / 230<br>(9%)             |
| Rata-rata 10<br>pertandingan<br>terakhir dengan<br>normalisasi | 63 / 230<br>(26.5%)           | 20 / 230<br>(8.6%)           |

Seperti yang dapat dilihat pada Tabel 5.14., hasil pengujian *real life* sangat memiliki akurasi yang sangat rendah jika dibandingkan dengan hasil pengujian *backpropagation* maupun hasil pengujian NEAT yang menggunakan *rating* yang sebenarnya. Kedua metode yang digunakan untuk melakukan pengujian ini memperoleh tingkat akurasi yang hampir sama, yaitu 44% dan 10% untuk rata-rata 5 pertandingan terakhir, 27.8% dan 8.9% jika dilakukan dengan normalisasi dan 40% dan 9% untuk rata-rata 10 pertandingan terakhir, 26.5% dan 8.6% jika dilakukan dengan normalisasi. Hal ini disebabkan oleh *rating* pemain yang tidak konsisten pada setiap pertandingan sehingga sulit untuk diprediksi menggunakan rata-rata.

## 5.5 Kesimpulan Pengujian

Setelah semua pengujian selesai dijalankan, dapat disimpulkan bahwa *features* terbaik pada pengujian NEAT adalah *rating* pemain disertai dengan *rating team* yang terjadi pada Tahap 2 Pengujian 2 dengan tingkat akurasi sebesar 80% untuk akurasi hasil pertandingan, dan 35% untuk akurasi skor. Penggunaan *rating* pemain dan *rating team* sedikit meningkatkan tingkat akurasi jika *feature* yang digunakan hanya *rating* pemain.

Penambahan posisi pemain sebagai *features* pada NEAT justru menimbulkan efek negatif. Selain tingkat akurasi yang lebih rendah jika dibandingkan dengan penggunaan *features* lainnya, waktu *training* yang dibutuhkan juga jauh lebih lama.

Dalam beberapa kasus, struktur yang dihasilkan oleh NEAT sudah cukup baik, tetapi *weight* yang tidak optimal menyebabkan tingkat akurasi pada pengujian menjadi rendah. Hal ini dibuktikan dengan peningkatan akurasi yang signifikan pada *network* yang memiliki tingkat akurasi rendah ketika proses *backpropagation* dijalankan.

*Backpropagation* terbukti mampu meningkatkan tingkat akurasi pada semua *network* yang dihasilkan oleh NEAT. Pada *network* yang sudah memiliki tingkat akurasi tinggi, memang optimasi *backpropagation* tidak terlalu signifikan. Tetapi pada *network* yang memiliki akurasi rendah, *backpropagation* mampu mengoptimasi *weight* sehingga terjadi peningkatan akurasi yang signifikan. Peningkatan akurasi terbesar terjadi pada Tahap 1 Pengujian 1, dimana akurasi hasil pertandingan meningkat dari 28% menjadi 76%.

Setelah proses *backpropagation* selesai, Tahap 3 Pengujian 4 yang menambahkan total posisi tiap pemain pada berbagai sektor keluar sebagai pengujian yang menghasilkan tingkat akurasi terbaik, mengalahkan Tahap 2 Pengujian 2 yang sebelumnya memiliki tingkat akurasi terbaik. Hal ini kembali membuktikan bahwa adanya kemungkinan *weight* yang dihasilkan NEAT tidak optimal.

Sayangnya, akurasi yang didapat pada pengujian *real life* tidak sebaik akurasi yang dihasilkan pada proses *testing*. Karena pada pengujian *real life rating* pemain maupun *rating team* tidak diketahui, sehingga harus diprediksi terlebih

dulu. Penggunaan metode rata-rata untuk menentukan *rating* pemain dan *rating* team gagal menghasilkan akurasi yang baik. Inkonsistensi dari pemain pada satu pertandingan ke pertandingan lainnya diyakini menjadi penyebab rendahnya akurasi yang dihasilkan.