

## GENETIC ALGORITHM DALAM MENENTUKAN DECISION TREE

### I. Pendahuluan

*Genetic Algoritim* (GA) ialah algoritma komputasi yang diinspirasi teori evolusi Darwin yang menyatakan bahwa kelangsungan hidup suatu makhluk dipengaruhi aturan “yang kuat adalah yang menang”. Darwin juga menyatakan bahwa kelangsungan hidup suatu makhluk dapat dipertahankan melalui proses reproduksi, crossover, dan mutasi. Konsep dalam teori evolusi Darwin tersebut kemudian diadopsi menjadi algoritma komputasi untuk mencari solusi suatu permasalahan dengan cara yang lebih “alamiah”. Sebuah solusi yang dibangkitkan dalam algoritma genetika disebut sebagai chromosome, sedangkan kumpulan chromosome-chromosome tersebut disebut sebagai populasi. Sebuah chromosome dibentuk dari komponen-komponen penyusun yang disebut sebagai gen dan nilainya dapat berupa bilangan numerik, biner, simbol ataupun karakter tergantung dari permasalahan yang ingin diselesaikan.

Decision Tree (Pohon Keputusan) adalah pohon dimana setiap cabangnya menunjukkan pilihan diantara sejumlah alternatif pilihan yang ada, dan setiapdaunnya menunjukkan keputusan yang dipilih. Decision tree biasa digunakan untuk mendapatkan informasi untuk tujuan pengambilan sebuah keputusan. Decision tree dimulai dengan sebuah root node (titik awal) yang dipakai oleh user untuk mengambil tindakan. Dari node root ini, user memecahnya sesuai dengan algoritma decision tree. Hasil akhirnya adalah sebuah decision tree dengan setiap cabangnya menunjukkan kemungkinan skenario dari keputusan yang diambil serta hasilnya

### II. Analisa Masalah

Kasus yang harus dipecahkan ialah *Decision Tree* dengan menggunakan data cuaca seperti berikut:

```
Normal Siang Berawan Normal Tidak
Normal Siang Cerah Normal Tidak
Normal Pagi Hujan Normal Tidak
Normal Sore Rintik Rendah Ya
Normal Malam Berawan Tinggi Tidak
Normal Malam Cerah Normal Ya
Normal Pagi Cerah Rendah Ya
Normal Siang Berawan Rendah Tidak
Normal Sore Hujan Normal Tidak
Normal Siang Rintik Rendah Ya
Normal Sore Berawan Tinggi Tidak
Normal Sore Cerah Rendah Tidak
Normal Pagi Berawan Normal Ya
Normal Malam Hujan Rendah Ya
```

### III. Penyelesaian

- Langkah pertama yang saya ambil ialah memanggil data yang sudah disediakan untuk dianalisa

```
def loadData(fileName):
    f = open(fileName, "r");
    i = 0;
    dataTest = [];
    line = f.readline();
    while( line != ""):
        i = i + 1;
        line = line.replace("\n", "");
        print(i, line);
        dataTest.append(decode(line));
        line = f.readline();
    f.close();
    return dataTest;
```

- Setelah itu melakukan decode yang tertera secara satu persatu agar mendapat rule.

```
def decode(code):
    arrCode = code.split(" ");
    print(arrCode);
    rule = [0]*15;

    if(arrCode[0] == "Rendah"):
        rule[8] = 1;
    elif(arrCode[0] == "Normal"):
        rule[1] = 1;
    elif(arrCode[0] == "Tinggi"):
        rule[2] = 1;

    if(arrCode[1] == "Pagi"):
        rule[3] = 1;
    elif(arrCode[1] == "Siang"):
        rule[4] = 1;

    elif(arrCode[1] == "Sore"):
        rule[5] = 1;
    elif(arrCode[1] == "Malam"):
        rule[6] = 1;
```

```
if(arrCode[2] == "Berawan"):
    rule[7] = 1;
elif(arrCode[2] == "Rintik"):
    rule[8] = 1;
elif(arrCode[2] == "Hujan"):
    rule[9] = 1;
elif(arrCode[2] == "Cerah"):
    rule[10] = 1;

if(arrCode[3] == "Rendah"):
    rule[11] = 1;
elif(arrCode[3] == "Normal"):
    rule[12] = 1;
elif(arrCode[3] == "Tinggi"):
    rule[13] = 1;

if(arrCode[4] == "Ya"):
    rule[14] = 1;
else:
    rule[14] = 0;
print(rule);
return rule;
```

- Selanjutnya saya melakukan penentuan populasi yang dapat dibentuk dari decode diatas dengan menentukan populasi numbernya adalah 100.
- Dalam menentukan fitness saya menentukan populasi yang digunakan untuk membentuk fitness sebanyak 10 berdasarkan datatest.txt.
- Menhitung dalam pemilihan orang tua, gunakan hasil random lalu di kalikan total fitness, setelah itu lakukan crossever antara parent dan child.

```
def gen_population(count):
    populations = []
    for i in range(count):
        chromosome = []
        for b in range(bit_count):
            chromosome.append(random.randint(0,1))
        populations.append(chromosome)
    return populations

def getresult(chromosome, data):
    ruleCount = len(chromosome)/15
    elseCount = int(chromosome[len(chromosome)-1] == 0)
    valid = False
    rule = 0
    while(valid==False):
        if(rule >= ruleCount):
            break
        crPos = rule*15
        sesi = data[0] and chromosome[crPos + 0]
        sesi = sesi or (data[1] and chromosome[crPos + 1])
        sesi = sesi or (data[2] and chromosome[crPos + 2])
        if(sesi==0):
            rule = rule + 1
        continue
    sesi = data[1] and chromosome[crPos + 3]
    sesi = sesi or (data[4] and chromosome[crPos + 8])
    sesi = sesi or (data[5] and chromosome[crPos + 9])
    sesi = sesi or (data[10] and chromosome[crPos + 10])
    if(sesi==0):
        rule = rule + 1
    continue
    sesi = data[11] and chromosome[crPos + 11]
    sesi = sesi or (data[12] and chromosome[crPos + 12])
    sesi = sesi or (data[13] and chromosome[crPos + 13])
    if(sesi==0):
        rule = rule + 1
    continue
    valid = True;
    return rule;

def getlowestFitnessIndex():
    min = 1;
    idx = 0;
    for i in range(0, len(populasi)):
        if(populasi[i].fitness < min):
            min = populasi[i].fitness;
            idx=i;
        elif(populasi[i].fitness == min and random() < 0.2):
            idx=i;
    return idx;

def gethighestFitnessIndex():
    min = 0;
    idx = 0;
    for i in range(0, len(populasi)):
        if(populasi[i].fitness > min):
            min = populasi[i].fitness;
            idx=i;
        elif(populasi[i].fitness == min and random() < 0.2):
            idx=i;
    return idx;

def get_parent(populations, fit_res):
    sorted_fit = sorted(fit_res, key=lambda x: x['fit'], reverse=True);
    parent1 = populations[sorted_fit[0]['index']];
    parent2 = populations[sorted_fit[1]['index']];
    return parent1, parent2;
```

```
def crossTypeA(parentA, parentB, pointA, pointB):
    child = []
    maximum = (pointB[0] + (pointA[1]-pointA[0]+1)*(len(parentB) - 1 - pointB[1]));
    print(maximum);
    maximum = maximum % (maximum%15);

    for i in range(0, pointB[0]):
        print("B",i)
        child.append(parentB[i]);
        maximum -= 1;
        if(maximum==0):
            print(1, len(child));
            return child;

    for i in range(pointA[0], pointA[1]+1):
        print("A",i)
        child.append(parentA[i]);
        maximum -= 1;
        if(maximum==0):
            print(2, len(child));
            return child;
```

```
def gen_population():
    i = 1;
    crPos = 0;
    crPos = crPos + 15;
    sesi = data[0] and chromosome[crPos + 0]
    sesi = sesi or (data[1] and chromosome[crPos + 1])
    sesi = sesi or (data[2] and chromosome[crPos + 2])
    if(sesi==0):
        rule = rule + 1
    continue
    sesi = data[1] and chromosome[crPos + 3]
    sesi = sesi or (data[4] and chromosome[crPos + 8])
    sesi = sesi or (data[5] and chromosome[crPos + 9])
    sesi = sesi or (data[10] and chromosome[crPos + 10])
    if(sesi==0):
        rule = rule + 1
    continue
    sesi = data[11] and chromosome[crPos + 11]
    sesi = sesi or (data[12] and chromosome[crPos + 12])
    sesi = sesi or (data[13] and chromosome[crPos + 13])
    if(sesi==0):
        rule = rule + 1
    continue
    valid = True;
    return rule;
```

- Saat melakukan mutasi dengan menentukan mutase probabilitas adalah 0.03. saat hasil random lebih besar daripada hasil mutase probabilitas maka nilai child akan diinisialisasi 0

```
def generate_population():
    print("Generating first generation in population");
    generatePopulasi();
    print("Complete...");
    print("Calculating fitness value");
    for i in range(0, populasiCount):
        populasi[i].fitness = getFitnessValue(populasi[i].gen);
        print(populasi[i].fitness);
    print("Complete...");
    loop = 0;
    while(loop < limit):
        totalFitness = 0;
        for i in range(0, populasiCount):
            totalFitness = totalFitness + populasi[i].fitness;
        print("Selecting Parent");
        parent = [0,0];
        while(parent[0] == parent[1]):
            parent[0] = getParent(totalFitness);
            parent[1] = getParent(totalFitness);
        print("parent",parent);
        print("Complete...");
        print("Get 2 Random Point");
        shortestParent = getShortestParent(parent);
        point = getRandom2Point(populasi[parent[shortestParent]]);
        print("Primary Point",point,"Other Point",otherPoints);
        print("Complete...");
        print("Duplicate Crossover and Generate Child");
        child = crossOver(populasi[parent[0]].gen, populasi[parent[1]].gen, point, otherPoints);
        print("Child A len : ", len(child[0]));
        print("Child B len : ", len(child[1]));
        print("Complete...");
        print("Mutation");
        child[0] = mutate(child[0]);
        child[1] = mutate(child[1]);
        print("Complete...");
        print("Survivor");
        insertToPopulation(child[0]);
        insertToPopulation(child[1]);
        print("Complete...");
        loop += 1;

    idx = getHighestFitnessIndex();
    print("Best Result : ", idx);
    print("Fitness : ", populasi[idx].fitness);
    print("Gen : ");
    print(populasi[idx].gen);
    print();
    print("Memasukkan data ke file");
```

```
def generate_population():
    print("Generating first generation in population");
    generatePopulasi();
    print("Complete...");
    print("Calculating fitness value");
    for i in range(0, populasiCount):
        populasi[i].fitness = getFitnessValue(populasi[i].gen);
        print(populasi[i].fitness);
    print("Complete...");
    loop = 0;
    while(loop < limit):
        totalFitness = 0;
        for i in range(0, populasiCount):
            totalFitness = totalFitness + populasi[i].fitness;
        print("Selecting Parent");
        parent = [0,0];
        while(parent[0] == parent[1]):
            parent[0] = getParent(totalFitness);
            parent[1] = getParent(totalFitness);
        print("parent",parent);
        print("Complete...");
        print("Get 2 Random Point");
        shortestParent = getShortestParent(parent);
        point = getRandom2Point(populasi[parent[shortestParent]]);
        print("Primary Point",point,"Other Point",otherPoints);
        print("Complete...");
        print("Duplicate Crossover and Generate Child");
        child = crossOver(populasi[parent[0]].gen, populasi[parent[1]].gen, point, otherPoints);
        print("Child A len : ", len(child[0]));
        print("Child B len : ", len(child[1]));
        print("Complete...");
        print("Mutation");
        child[0] = mutate(child[0]);
        child[1] = mutate(child[1]);
        print("Complete...");
        print("Survivor");
        insertToPopulation(child[0]);
        insertToPopulation(child[1]);
        print("Complete...");
        loop += 1;

    idx = getHighestFitnessIndex();
    print("Best Result : ", idx);
    print("Fitness : ", populasi[idx].fitness);
    print("Gen : ");
    print(populasi[idx].gen);
    print();
    print("Memasukkan data ke file");
```

#### IV. Kesimpulan

Hasil dari program ini

```
[0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0]
76 Tinggi Pagi Hujan Rendah Tidak
['Tinggi', 'Pagi', 'Hujan', 'Rendah', 'Tidak']
[0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0]
77 Tinggi Pagi Berawan Rendah Ya
['Tinggi', 'Pagi', 'Berawan', 'Rendah', 'Ya']
[0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1]
78 Tinggi Malam Berawan Normal Tidak
['Tinggi', 'Malam', 'Berawan', 'Normal', 'Tidak']
[0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0]
79 Tinggi Malam Rintik Rendah Ya
['Tinggi', 'Malam', 'Rintik', 'Rendah', 'Ya']
[0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1]
80 Tinggi Malam Berawan Tinggi Tidak
Loading data complete, data test have been saved
Generating first generation in population
```