# **BAB IV**

**METODOLOGI PENELITIAN**

## Diagram Alur Metodologi Penelitian

Metodologi penelitian merupakan pendekatan yang sistematis untuk keseluruhan kegiatan penelitian. Permasalahan atau pertanyaan penelitian dapat diselesaikan dengan pendekatan metodologi tertentu. Dalam metodologi penelitian dipelajari bagaimana proses dan tahapan suatu kegiatan penelitian. Metodologi penelitian bertujuan untuk membuat berbagai tahap penelitian, mulai dari rencana kerja suatu penelitian sampai dengan publikasi[32]. Metodologi penelitian yang digunakan adalah CRISP-DM (*Cross Standart Industries for Data Mining*). Metodologi CRISP-DM adalah standarisasi *data mining* yang disusun oleh tiga penggagas *data mining* *market* yaitu Daimcler Chrysler (Daimler Benz), SPSS, dan NCR. Kemudian dikembangkan pada berbagai *workshop* (antara tahun 1997-1999). Lebih dari 30 organisasi yang berkontribusi dalam proses *modelling* ini dan akhirnya CRISP-DM dipublikasikan pada tahun 1999[33].



Gambar 4. 1 Diagram Alur CRISP-DM.

## Tahapan-Tahapan Diagram Alur Metodologi Penelitian

Terdapat beberapa tahapan dalam CRISP-DM (*Cross Industries Process for Data Mining*), namun pada penelitian ini tahapan yang dilakukan sampai pada tahapan *deployment* dikarenakan tujuan dari penelitian ini adalah untuk mengetahui pengaruh jumlah pelanggan dan jumlah produk terhadap nilai pendapatan serta untuk mengoptimasi nilai pendapatan tersebut[34].

### Bussiness Understanding

Fase ini meliputi penentuan tujuan bisnis, menilai situasi saat ini, menetapkan tujuan data mining, dan mengembangkan rencana penelitian. Tujuan bisnis yang akan dilakukan pada penelitian ini adalah bagaimana mengetahui pengaruh jumlah pelanggan dan jumlah produk terhadap nilai pendapatan menggunakan Regresi Linier Berganda dan Algoritma Genetika untuk melakukan optimasi nilai tersebut.

### Data Understanding

Fase ini dikenal sebagai fase pemahaman terhadap data yang diperoleh dan kemudian data awal yang dikumpulkan melalui observasi langsung, dengan melakukan analisa terhadap berkas data pelanggan pada objek penelitian. Kajian perlu dilakukan sehingga diketahui data yang akan digunakan. Data yang digunakan dalam penelitian berjumlah 2000 *records* sebagai populasi, dimana data tersebut diambil dari periode bulan Januari sampai dengan Oktober berdasarkan berkas yang ditunjukan oleh objek penelitian. Jika data yang diperoleh semakin banyak, maka hasil dari akurasi datanya juga akan maksimal. Dalam tahapan ini, perlu identifikasi terhadap masalah kualitas data yang baik sehingga diperoleh nilai subset yang sesuai, menarik dalam pembuatan hipotesa awal.

### Data Preparation

Tahapan ini merupakan tahapan pengolahan data atau dapat juga dikatakan sebagai tahapan persiapan data. Banyak persiapan yang dilakukan pada tahapan ini sehingga tak jarang fase ini juga disebut sebagai fase padat karya. Beberapa kegiatan seperti pemilihan tabel dan *field* terjadi pada fase ini. Pemilihan tabel dan *field* tersebut akan dimasukan atau ditransformasikan kedalam *database* yang lain atau *database* baru sebagai bahan atau *data mining* mentah.

### Modelling

Pada fase pemodelan dilakukan dengan menggunakan aplikasi R Studio, dan dimasukkan juga metode regresi linier berganda dan algoritma genetika. Data-data yang telah dipilih pada fase *data preparation* digunakan sebagai parameter untuk melakukan peramalan serta optimasi.

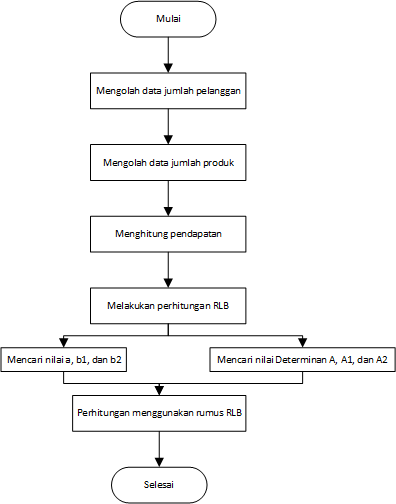
### Evaluation

Fase ini merupakan tahapan analisa yang merupakan hasil dari pengolahan fase sebelumnya dengan menginterpretasikan data yang kemudian diperoleh peramalan dan optimasi dari nilai tersebut.

### Deployment

Fase ini merupakan pengetahuan atau informasi yang telah diperoleh akan diatur dan dipresentasikan dalam bentuk khusus sehingga dapat digunakan oleh pengguna. Fase *deployment* dapat berupa pembuatan laporan sederhana atau mengimplementasikan proses *data mining* yang berulang dalam perusahaan.

## Diagram Alur Metode Regresi Linier Berganda

****

Gambar 4. 2 Diagram Alur Metode Regresi Linier Berganda.

## Tahapan – Tahapan Diagram Alur Metode Regresi Linier Berganda

Tahapan – tahapan diagram alur metode regresi linier berganda yang digunakan peneliti adalah sebagai berikut:

### Mengolah Data Jumlah Pelanggana

Untuk mengolah data jumlah pelanggan ini dapat dilakukan oleh pegawai maupun admin, cara yang dapat dilakukan untuk mengolah data jumlah pelanggan ini dengan mengakumulasikan jumlah pelanggan untuk setiap bulannya dari setiap produk yang berhasil dijual kepada pelanggan. Lalu hasil dari setiap bulannya dijumlahkan kembali untuk mengetahui keseluruhan pelanggan.

### Mengolah Data Jumlah Produk

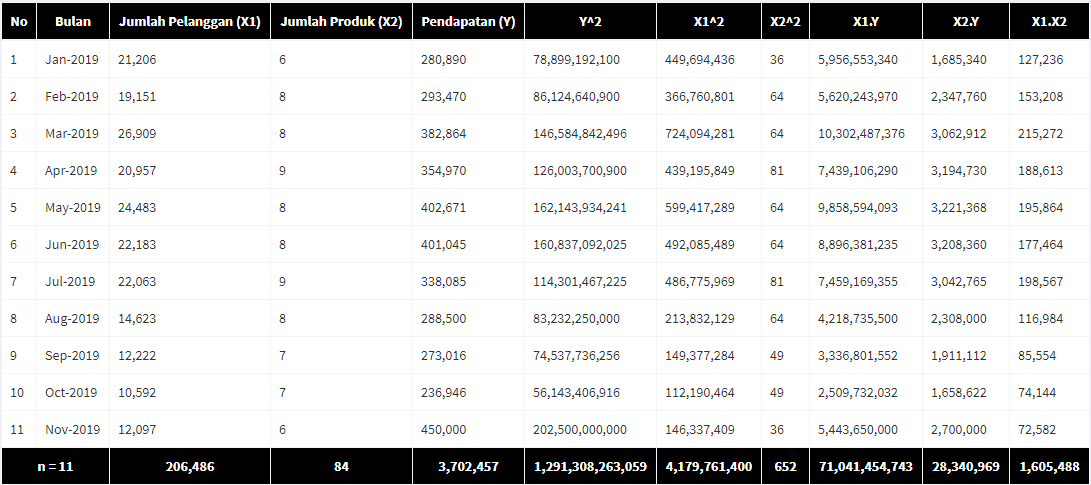
Dalam mengolah jumlah produk, langkah yang dapat dilakukan yakni dengan cara menghitung setiap produk yang terjual setiap bulannya, terhitung bulan Januari.

### Menghitung Pendapatan

Untuk melakukan perhitungan pendapatan hal yang perlu diketahui adalah harga untuk setiap produknya. Jika harga sudah diketahui maka tinggal dikalikan saja antara harga produk dikali dengan jumlah pelanggan pada bulan tersebut. Hasilnya diakumulasikan dengan bulan-bulan yang lain, yang selanjutnya digunakan untuk melakukan prediksi pada bulan berikutnya.

### Melakukan Perhitungan Regresi Linier Berganda

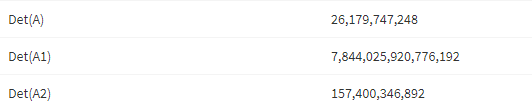
Perhitungan ini dilakukan dengan cara kuadaratkan, dan kalikan dengan nilai pada tabel regresi linier berganda atau dapat dilihat seperti berikut.



Gambar 4. 3 Hasil Perhitungan Tabel RLB.

### Mencari Nilai Determinan A, A1, dan A2

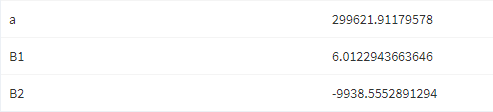
Setelah melakukan perhitungan regresi linier berganda, hal selanjutnya yang perlu kita lakukan ialah mencari determinan A, A1, dan A2 agar mendapatkan nilai untuk melakukan prediksi. Berikut adalah hasil dari determinan A, A1, dan A2.



Gambar 4. 4 Hasil Nilai Determinan.

### Mencari Nilai A, B1, dan B2

Untuk mencari nilai a, b1, dan b2. Persyaratan yang perlu dipenuhi adalah nilai determinan A, A1, dan A2 sudah didapatkan, nilai a, b1, dan b2 ini akan digunakan untuk tahap selanjutnya dalam menentukan prediksi atau peramalan, berikut merupakan hasil nilai dari a, b1, dan b2.

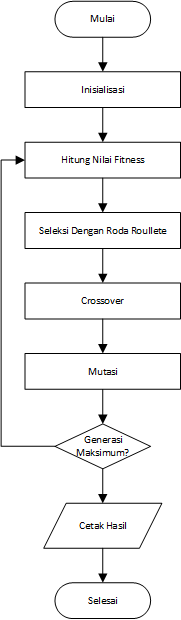


Gambar 4. 5 Hasil Nilai A, B1, dan B2.

### Perhitungan Menggunakan Rumus Regresi Linier Berganda

Proses perhitungan menggunakan rumus regresi linier berganda dilakukan jika nilai a, b1, dan b2 sudah didapatkan, karena rumus yang digunakan pada metode regresi linier berganda ini memerlukan nilai a, b1, dan b2 dalam proses peramalan.

## Diagram Alur Algoritma Genetika



Gambar 4. 6 Diagram Alur Algoritma Genetika.

## Tahapan – Tahapan Diagram Alur Algoritma Genetika

Tahapan – tahapan diagram alur algoritma genetika yang digunakan peneliti adalah sebagai berikut:

### Inisialisasi

Proses inisialisasi dilakukan dengan cara memberikan nilai awal gen-gen dengan nilai acak sesuai batasan yang telah ditentukan. Sebagai contoh dalam tahap inisialisasi ini jumlah produk dan jumlah pelanggan diberikan inisialisasi identitas atau ID berupa angka 1 sampai jumlah data yang akan diprediksi, sehingga sebuah inisialisasi identitas atau ID berupa angka 1 sampai dengan jumlah data yang akan digunakan adalah gen.

ID untuk nilai X1

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 10 | 6 | 8 | 2 | 3 | 5 | 7 | 9 | 1 | 4 |

ID untuk nilai X2

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2 | 6 | 10 | 8 | 1 | 3 | 9 | 7 | 5 | 4 |

Setelah itu bentuk awal populasi dibangkitkan secara acak, pada penelitian ini kromosom yang dibangkitkan ada 6 kromosom sebagai berikut:

* Kromosom[1] = [1 4]
* Kromosom[2] = [2 5]
* Kromosom[3] = [3 6]
* Kromosom[4] = [6 3]
* Kromosom[5] = [3 2]
* Kromosom[6] = [4 1]

Selanjutnya kita lakukan evaluasi nilai kromosom dengan memasukkan nilai tersebut pada rumus regresi linier berganda.

* Kromosom[1] = (27 + 8.15\*1 + 17\*4)

= (27 + 8.15 + 68)

= 103.15

* Kromosom[2] = (27 + 8.15\*2 + 17\*5)

= (27 + 17 + 85)

= 129

* Kromosom[3] = (27 + 8.15\*3 + 17\*6)

= (27 + 24.45 + 102)

= 153.45

* Kromosom[4] = (27 + 8.15\*6 + 17\*3)

= (27 + 48.9 + 51)

= 126.9

* Kromosom[5] = (27 + 8.15\*3 + 17\*2)

= (27 + 24.45 + 34)

= 85.45

* Kromosom[6] = (27 + 8.15\*4 + 17\*1)

= (27 + 32.6 + 17)

= 76.6

* Rata-rata nilai dari evaluasi

= (103.15 + 129 + 153.45 + 126.9 + 85.45 + 76.6) / 6

= 112.425

Rata-rata nilai tersebut merupakan nilai dari kromosom sebelum dilakukan optimasi.

### Hitung Nilai *Fitness*

Fungsi dari nilai *fitness* suatu kromosom adalah nilai kecocokan kromosom terhadap permasalahan. Semakin tinggi nilai *fitness*, seharusnya solusi tersebut semakin optimal. Berikut merupakan faktor-faktor penunjang nilai *fitness* pada penelitian ini, diantaranya:

1. Jumlah pelanggan pada periode bulan tertentu.
2. Jumlah produk pada periode bulan tertentu.

Pada proses seleksi ini dilakukan dengan cara membuat kromosom yang mempunyai nilai evaluasi kecil, mendapatkan kemungkinan terpilih lebih besar dengan nilai probabilitasnya yang tinggi. Untuk mencari *fitness* rumus yang digunakan sebagai berikut: F = 1 / hasil\_evaluasi.

* Fitness[1] = 1 / 104.15

= 0.0096

* Fitness[2] = 1 / 130

= 0.0076

* Fitness[3] = 1 / 154.45

= 0.0064

* Fitness[4] = 1 / 127.9

= 0.0078

* Fitness[5] = 1 / 86.45

= 0.0115

* Fitness[6] = 1 / 77.6

= 0.0128

* Total Nilai *Fitness*

= 0.0096 + 0.0076 + 0.0064 + 0.0078 + 0.0115 + 0.0128

= 0.0557

### Seleksi Dengan Roda *Roullete*

Seleksi dengan *roullete wheel* memilih anggota populasi tertentu untuk menjadi *parents* dengan probabilitas sama dengan *fitness* dibagi dengan total *fitness* populasi. Pada seleksi ini, *parents* dipilih berdasarkan *fitness*. Jika suatu individu lebih baik, maka akan lebih besar kesempatan terpilih. Probabilitas suatu individu terpilih untuk *crossover* sebanding dengan *fitness*-nya.

Seleksi dengan menggunakan metode *roullete wheel* dilakukan dengan tahapan-tahapan sebagai berikut:

1. Menghitung nilai *fitness* tiap individu.
2. Menghitung nilai probabilitas seleksi setiap individu.

Setelah nilai *fitness* ditemukan, selanjutnya mencari nilai probabilitas dengan rumus; P[i] = fitness[i] / total\_fitness

* P[1] = 0.0096 / 0.0557

= 0.1723

* P[2] = 0.0076 / 0.0557

= 0.1364

* P[3] = 0.0064 / 0.0557

= 0.1149

* P[4] = 0.0078 / 0.0557

= 0.1400

* P[5] = 0.0115 / 0.0557

= 0.2064

* P[6] = 0.0128 / 0.0557

= 0.2298

Dari probabilitas tersebut terlihat bahwa kromosom ke-6 memiliki *fitness* paling besar, maka kromosom tersebut mempunyai probabilitas untuk terpilih pada generasi selanjutnya lebih besar dari kromosom lainnya.

1. Mencari nilai komutatif

Untuk proses seleksi digunakan *roullete-wheel*, untuk itu terlebih dahulu mencari nilai komutatif dari probabilitasnya.

C[1] = 0.1723

C[2] = 0.1723 + 0.1364 = 0.3087

C[3] = 0.3087 + 0.1149 = 0.4236

C[4] = 0.4236 + 0.1400 = 0.5636

C[5] = 0.5636 + 0.2064 = 0.77

C[6] = 0.77 + 0.2298 = 1

Proses *roullete-wheel* adalah membangkitkan nilai acak R antara 0-1. Jika R[k] < C[k], maka kromosom ke-k tersebut sebagai induk. Setelah itu putar *roullete-wheel* sebanyak jumlah kromosom.

1. Membangkitkan nilai acak R

R[1] = 0.201

R[2] = 0.284

R[3] = 0.009

R[4] = 0.822

R[5] = 0.398

R[6] = 0.501

1. Populasi baru terbentuk

Angka acak R[1] nilainya lebih besar dari C[1] dan lebih kecil dari C[2], maka pilih kromosom[2] sebagai kromosom pada populasi baru. Sehingga kromosom baru dari hasil proses seleksi tersebut, sebagai berikut:

Kromosom[1] = [2] = [2 5]

Kromosom[2] = [2] = [2 5]

Kromosom[3] = [1] = [1 4]

Kromosom[4] = [5] = [3 2]

Kromosom[5] = [2] = [2 5]

Kromosom[6] = [3] = [3 6]

### Crossover

Proses *crossover* adalah proses persilangan, yaitu membentuk dua *offspring* (kromosom anak) baru dari dua *parent*. Hanya *offspring* yang memenuhi syarat pada permasalahan yang akan ditambahkan ke populasi yang ada. Proses mutasi yaitu mengganti suatu gen dengan gen yang baru. Proses *crossover* dan mutasi biasanya dilakukan secara acak.

Kromosom yang dijadikan *parent* dipilih secara acak dari jumlah kromosom yang di *crossover* dipengaruhi oleh parameter *crossover* probabilitas, misalnya = 25% = 0.25

1. Membangkitkan nilai acak R

R[1] = 0.191

R[2] = 0.259

R[3] = 0.760

R[4] = 0.006

R[5] = 0.159

R[6] = 0.340

\*kromosom ke-k yang dipilih sebagai induk. Jika R[k] < nc, maka yang akan dijadikan induk adalah kromosom [1], [4], dan [5].

1. Bilangan acak untuk 3 kromosom induk yang akan di *crossover*

C[1] = 1

C[4] = 1

C[5] = 2

1. Proses *crossover*

Kromosom[1] = kromosom[1] >< kromosom[4]

= [1 4] >< [6 3]

= [1 6]

Kromosom[4] = kromosom[4] >< kromosom[5]

= [6 3] >< [3 2]

= [3 2]

Kromosom[5] = kromosom[5] >< kromosom[1]

= [3 2] >< [1 4]

= [2 1]

Populasi setelah dilakukan *crossover*

* Kromosom[1] = [1 6]
* Kromosom[2] = [2 5]
* Kromosom[3] = [3 6]
* Kromosom[4] = [3 2]
* Kromosom[5] = [2 1]
* Kromosom[6] = [4 1]

### Mutasi

Proses mutasi ini dilakukan setelah proses persilangan atau *crossover* dengan cara memilih kromosom yang akan dimutasi secara acak, dan kemudian menentukan titik mutasi pada kromosom tersebut secara acak. Mutasi diperlukan untuk mengembalikan informasi bit yang hilang akibat pindah silang atau *crossover*. Mutasi diterapkan dengan probabilitas yang sangat kecil. Jika mutasi dilakukan terlalu sering, maka akan menghasilkan individu yang lemah karena konfigurasi gen pada individu yang unggul akan dirusak. Untuk melakukan proses mutasi dapat dilakukan dengan cara sebagai berikut:

Menghitung panjang total gen

Panjang total gen = jumlah gen dalam 1 kromosom \* jumlah kromosom

= 2 \* 6 = 12

Untuk memilih posisi gen yang mengalami mutasi dilakukan dengan memberikan bilangan acak 1 – 12. Misal kita tentukan nilai nm = 20%, maka jumlah gen yang dimutasi adalah 0.2 \* 12 = 2.4 = 2. 2 buah posisi gen yang akan dimutasi, setelah diacak adalah posisi gen 4 dan 7. Lalu gen pada posisi tersebut nilainya diganti dengan ID yang telah ditentukan sebelumnya.

1. Proses mutasi

Kromosom[1] = [1 6]

Kromosom[2] = [2 6]

Kromosom[3] = [3 6]

Kromosom[4] = [1 2]

Kromosom[5] = [2 1]

Kromosom[6] = [4 1]

1. Menghitung nilai fungsi objektif / evaluasi kromosom

* Kromosom[1] = (27 + 8.15\*1 + 17\*6)

= (27 + 8.15 + 102)

= 137.15

* Kromosom[2] = (27 + 8.15\*2 + 17\*6)

= (27 + 16.3 + 102)

= 145.3

* Kromosom[3] = (27 + 8.15\*3 + 17\*6)

= (27 + 24.45 + 102)

= 153.45

* Kromosom[4] = (27 + 8.15\*1 + 17\*2)

= (27 + 8.15 + 34)

= 69.15

* Kromosom[5] = (27 + 8.15\*2 + 17\*1)

= (27 + 16.3 + 17)

= 60.3

* Kromosom[6] = (27 + 8.15\*4 + 17\*1)

= (27 + 32.6 + 17)

= 76.6

* Rata-rata nilai dari evaluasi

= (137.15 + 145.3 + 153.45 + 69.15 + 60.3 + 76.6) / 6

= 106.991

Dapat dilihat dari hasil perhitungan evaluasi tersebut bahwa setelah satu generasi nilai hasil rata-rata evaluasi menurun dibandingkan dengan rata-rata sebelumnya. Hal ini menunjukkan bahwa kromosom atau solusi yang dihasilkan setelah satu generasi lebih baik dibandingkan generasi sebelumnya.

### Generasi Maksimum

Proses generasi maksimum ini akan melakukan pengecekan terhadap populasi baru yang telah dibentuk dengan melalui tahapan-tahapan yang cukup panjang. Jika populasi baru yang dihasilkan terbaik, maka selanjutnya akan mengulangi tahapan-tahapan dalam pembentukan populasi baru seperti evaluasi nilai *fitness*, proses seleksi, proses *crossover*, dan mutasi untuk membentuk populasi baru selanjutnya.

### Cetak Hasil

Setelah melewati tahapan-tahapan yang cukup panjang, langkah selanjutnya adalah melakukan cetak hasil dari populasi baru yang terbentuk dengan solusi yang optimal untuk menyelesaikan permasalahan optimasi nilai pendapatan pada penlitian ini.