

BAB III

PROSES KEPUTUSAN MARKOV DENGAN MODEL INFINITE STAGE

3.1 Proses Keputusan Markov

M. Iqbal Hasan (2004, hal: 9) mengatakan bahwa keputusan merupakan suatu pemecahan masalah sebagai suatu hukum yang dilakukan melalui pemilihan satu alternatif dari beberapa alternatif. P. Siagian (1987, hal: 317) menyatakan bahwa keputusan ialah suatu kesimpulan dari suatu proses untuk memilih tindakan yang terbaik dari sejumlah alternatif yang ada. Mengambil atau membuat keputusan adalah suatu proses yang dilaksanakan orang berdasarkan pengetahuan dan informasi yang ada padanya pada saat tersebut, dengan pengharapan bahwa sesuatu akan terjadi (Kuntoro Mangkusubroto et al, 1987). P. Siagian (1987, hal: 317) mengatakan bahwa pengambilan keputusan adalah proses yang mencakup semua pemikiran dan kegiatan yang diperlukan guna membuktikan dan memperlihatkan pilihan terbaik. M Iqbal Hasan (2004, hal: 10) menyatakan bahwa pengambilan keputusan merupakan suatu proses pemilihan alternatif terbaik dari beberapa alternatif secara sistematis untuk ditindaklanjuti atau digunakan sebagai suatu cara pemecahan masalah. Pengambilan keputusan dalam model-model probabilitas dapat diselesaikan dengan proses keputusan Markov (T. Tarlih Dimiyati et al, 2010). Proses keputusan Markov telah banyak dipakai dalam pemeliharaan peralatan, pengawasan persediaan dan bidang lainnya dalam bidang manajemen (W. K Ching et al, 2006). Serupa dengan rantai Markov, proses keputusan Markov adalah sebuah sistem yang dapat memindahkan satu keadaan yang khusus ke keadaan yang mungkin lainnya. Dalam langkah, pembuat keputusan harus mengambil suatu tindakan dari alternatif-alternatif yang ditetapkan. Tindakan ini mempengaruhi peluang transisi pada perpindahan selanjutnya dan mendatangkan sebuah keuntungan atau kerugian setelah itu. Permasalahan yang nyata adalah pengambilan keputusan dihadapi untuk

menentukan perencanaan yang sesuai. Sehingga keuntungan secara keseluruhan optimal, dengan kata lain proses keputusan Markov adalah suatu proses keputusan yang mengamati sistem pada waktu t dan menghasilkan keadaan i , kemudian dipilih sebuah keputusan k dari sejumlah keputusan yang mungkin, di mana sistem berkembang sepanjang waktu dengan probabilitas transisi yang diketahui.

3.2 Model Infinite Stage

Perilaku jangka panjang dari suatu proses Markov ditandai oleh perilaku ketidakbergantungannya pada state awal (initial state) dari sistemnya. Dalam hal ini, sistem tersebut dikatakan telah mencapai steady state (keadaan yang tetap). Ada dua metode yang dapat digunakan untuk menyelesaikan persoalan dengan stage yang tidak terbatas ini. Metode pertama disebut metode policy Iteration, yaitu metode yang melakukan iterasi untuk mendapatkan nilai policy optimumnya, metode ini umumnya bersifat efisien, dalam arti dapat mencapai solusi optimum dalam jumlah iterasi yang kecil, metode yang kedua adalah metode enumerasi sempurna (Exhaustive Enumeration) yang mengenumerasi seluruh stationary policy, hingga diperoleh solusi optimumnya. Metode ini hanya dapat digunakan apabila jumlah total stationary policy-nya tidak terlalu besar sehingga masih dapat dihitung. Namun, kedua metode tersebut akan menghasilkan solusi optimum yang sama. [61]

3.2.1 Metode Policy Iteration

Dalam bagian ini akan dikenalkan metode policy iteration yang merupakan salah satu metode yang dikenal dalam proses pengambilan keputusan Markov (Markov Decision Process) yang masuk dalam salah satu model infinite stage.

Disajikan suatu penerapan program dinamis terhadap pemecahan suatu proses keputusan stokastik yang dapat dijelaskan oleh jumlah state yang tidak terbatas. Peluang transisi diantara state ini dijelaskan oleh suatu rantai Markov (Markov chain), sedangkan struktur biaya proses ini dijelaskan oleh suatu matriks yang elemennya menyatakan pendapatan yang dihasilkan dari pergerakan satu state ke state yang lain, dalam hal ini struktur biaya dikatakan reward diasumsikan sebagai jumlah kursi di parlemen yang didapat oleh suatu partai. Baik matriks transisi maupun matriks pendapatan ini sifatnya bergantung pada alternatif-alternatif keputusan yang dapat digunakan oleh pengambil keputusan. Tujuan persoalan ini adalah menentukan

keputusan optimum yang dapat memaksimumkan ekspektasi pendapatan dari proses yang mempunyai jumlah state tidak terbatas.

Diberikan berbagai notasi yang diperlukan untuk model matematika dari pengambilan keputusan policy iteration:

a. State (keadaan)

Suatu keadaan, akibat, atau kejadian (alamiah) pada suatu waktu di mana pengambilan keputusan hanya mempunyai sedikit kontrol atau bahkan tidak memiliki kontrol terhadapnya. State ditandai dengan $i, 0, 1, 2, N$ dan lokasi peralihan State $j, 0, 1, 2, \dots, N$. Himpunan semua state yang mungkin dilambangkan dengan I . [12]

Maka dari definisi di atas dapat disimpulkan bahwa State merupakan suatu keadaan yang dilambangkan dengan i dan j . Sehingga untuk prosedur penentuan simbol keadaan dapat dimisalkan pada studi kasus penentuan pilihan pada partai politik.

b. Alternatif keputusan

Suatu bagian dari aksi atau Strategi yang mungkin dipilih oleh seorang pengambil keputusan di setiap State i , dilambangkan dengan notasi $A(i) = \{1, 2, \dots, K\}$. Berdasarkan definisi ini dapat disimpulkan bahwa alternatif keputusan merupakan suatu pilihan aksi yang akan diambil pada setiap state, dan dilambangkan dengan k . [12]

c. Peluang Transisi

Peluang suatu proses bergerak dari suatu State ke State yang lain pada suatu alternatif keputusan ke- k . Peluang transisi dinotasikan dengan $P_{ij}^k, i, j \in I, k \in A(i)$. [12] Sehingga untuk prosedur pembentukan matriks peluang transisi satu langkah lebih dahulu akan dibuat diagram transisi state dan distribusi transisi state untuk tiap partai.

Banyaknya transisi dari state i ke State j adalah E_{ij} , dan E_{ii} menyatakan banyaknya elemen dalam ruang sampel, maka berdasarkan definisi peluang diperoleh rumus peluang transisi dari state i ke state j sebagai berikut:

$$P_{ij}^k = \frac{E_{ij}^k}{E_{ii}^k} \quad (3.1) \text{ untuk setiap } i = 0, 1, 2, 3 \text{ dan } j = 0, 1, 2, 3.$$

Berdasarkan persamaan tersebut, akan dihitung semua peluang transisi dari state yang satu ke state yang lain. Dari hasil perhitungan tersebut maka akan diperoleh matriks peluang transisi satu langkah sebagai berikut:

$$P^{(1)} = \begin{bmatrix} P_{00} & P_{01} & P_{02} & \dots & P_{0j} \\ P_{10} & P_{11} & P_{12} & \dots & P_{1j} \\ P_{20} & P_{21} & P_{22} & \dots & P_{2j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{i0} & P_{i1} & P_{i2} & \dots & P_{ij} \end{bmatrix}$$

d. RewardTransisi

Reward Transisi adalah pendapatan yang diperoleh sebagai implikasi terjadinya transisi antar state pada alternatif keputusan ke-k. Reward dinotasikan dengan Berdasarkan definisi di atas dapat disimpulkan bahwa reward transisi merupakan suatu pendapatan yang diperoleh dari hasil perpindahan state i ke State j, yang dilambangkan dengan notasi $r_i(j^k)$ diasumsikan sebagai jumlah kursi di parlemen Yang didapat Oleh setiap partai. Bentuk matriks transisi reward dalam berbagai state sebagai berikut:

$$R = \begin{bmatrix} r_{00} & r_{01} & r_{02} & \dots & r_{0j} \\ r_{10} & r_{11} & r_{12} & \dots & r_{1j} \\ r_{20} & r_{21} & r_{22} & \dots & r_{2j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{i0} & r_{i1} & r_{i2} & \dots & r_{ij} \end{bmatrix}$$

e. Poltcy

Suatu langkah mengambil suatu aksi sebagai kebijakan secara prosedural untuk mencapai langkah strategis yaitu mendapatkan keputusan optimal Secara matematis dinyatakan sebagai himpunan semua keputusan yang diambil dalam setiap state. Dinotasikan dengan s dan t dan dikarakterisasikan oleh matriks. [12]

f. Policy optimal

Suatu kebijakan yang terbaik dari sekian banyak aksi yang mungkin sebagai hasil menjalankan serangkaian proses pengambilan keputusan. Secara matematis dinyatakan sebagai himpunan semua keputusan di setiap state yang memberikan reward maksimal atau cost minimal. Mendapatkan suatu keputusan inilah yang sebenarnya menjadi tujuan kita menggunakan teknik-teknik pengambilan keputusan. Berdasarkan definisi diatas dapat disimpulkan bahwa policy optimal merupakan suatu hasil akhir dari keputusan yang akan

diambil pada setiap state yang memberikan reward maksimal atau cost minimal [12]

g. Faktor Koreksi

Suatu nilai yang dihasilkan dari selisih antara pendapatan optimum kumulatif dan banyaknya stage yang diamati dikalikan dengan ekspektasi pendapatan per stage [6]

Secara matematis persamaan ini dapat dituliskan:

$$f(i) = f(i) - T \quad (3.2) \quad h. \text{ Pendapatan}$$

Optimum Kumulatif

Suatu nilai yang menjadi patokan untuk mendapatkan policy optimal yang merupakan hasil penjumlahan antara ekspektasi dengan faktor koreksi yang dikalikan dengan matriks peluang transisi suatu state yang diketahui. Nilai pendapatan optimum ini terletak pada langkah perbaikan policy, di mana sebelumnya ditentukan langkah penentuan nilai terlebih dahulu. [6]

Didefinisikan $v_n(i)$ = ekspektasi pendapatan optimum dari stage n , $n+1, \dots, N$, jika state dari sistem pada awal bulan ke- n adalah i . [6] Maka perhitungan mundur yang menghubungkan $v_n(i)$ dan $v_{n+1}(j)$ dapat dituliskan sebagai berikut:

$$v_n(i) = \max_k \left\{ \sum_{j=0}^m P_{ij}^k [r_{ij}^k + v_{n+1}(j)] \right\}, \quad i = 0, 1, 2, \dots, m$$

di mana

$v_{n+1}(j) = 0$ untuk seluruh j
Sebagai bukti kebenaran persamaan diatas adalah pendapatan kumulatif, yaitu $r_{ij}^k + v_{n+1}(j)$, yang diperoleh dengan mencapai state j pada stage $n+1$ dari state i pada stage n , terjadi dengan peluang P_{ij}^k . [6]

Jika q_i menyatakan ekspektasi pendapatan yang dihasilkan dari suatu transisi tunggal dari state i pada alternatif keputusan ke- k [6], maka q_i dapat dinyatakan sebagai

$$q_i = \sum_{j=0}^m P_{ij}^k r_{ij}^k \quad (3.3)$$

Sehingga hubungan rekursifnya dapat dinyatakan sebagai:

$$v_n(i) = \max_k (q_i^k + v_{n+1}(j)), \quad i = 0, 1, 2, 3, \dots, m \quad (3.4)$$

Persamaan rekursif ini adalah dasar untuk mengembangkan metode policy Iteration tersebut. Namun, bentuknya saat ini harus dimodifikasi lebih dahulu untuk memungkinkan mempelajari perilaku asimptotik dari prosesnya.

Definisikan sebagai banyaknya stage yang akan diamati. Im berbeda dai n dari persamaan diatas, karena menyatakan stage ke-n [6]. Maka persamaan rekursifnya ditulis sebagai:

$$v'(i) = \max_j \{ r_{ij} + \gamma \sum_{k=0}^m \pi_{ik} v(k) \}, \text{ di mana } i = 0, 1, 2, 3, \dots, m. \text{ Perhatikan bahwa}$$

ini adalah ekspektasi pendapatan kumulatif dengan adalah banyak stage yang diamati. Dengan definisi yang baru ini, maka perilaku asimptotik dari prosesnya dapat dipelajari dengan menetapkan 00.

Diketahui bahwa $r = (r_1, r_2, \dots, r_m)$ adalah vektor peluang pada keadaan

$$\pi_1 q_1 + \dots$$

steadystate dari matriks transisi $P = [P_{ij}]$ [6], dan $g = \sum_{i=0}^m \pi_i q_i + \sum_{i=0}^m \pi_i m q_m$ adalah ekspektasi pendapatan per-tahap [6] Untuk yang sangat besar, maka:

$$v_r(i) = ng + v(i).$$

karena $v_r(i)$ adalah pendapatan optimum kumulatif untuk stage berdasarkan state i , dan g adalah ekspektasi pendapatan per-stage, maka secara intuitif dapat dilihat mengapa $v(i)$ sama dengan ng ditambah dengan suatu faktor korelasi $v(i)$ yang dihitung untuk state i tertentu. Hasil ini tentu saja berasumsi bahwa sangat besar.

Dengan menggunakan informasi diatas, maka persamaan rekursifnya dapat ditulis sebagai:

$$ng + v(i) = q_i + \sum_{j=0}^m P_{ij} (ng + v(j)), \quad i = 0, 1, 2, \dots, n.$$

Dengan menyederhanakan persamaan ini dapat diperoleh $g = q_i + v(i) - v(i)$.

$$j = 0, 1, 2, \dots, m$$

3.2.2 Metode Exhaustive Enumeration

Definisi 3.1: [16]

Metode Exhaustive Enumeration merupakan salah satu metode untuk mencari nilai optimum suatu tindakan pada proses keputusan markov dengan cara kerjanya adalah mengenumerasi secara keseluruhan dari setiap stationary policy yang ada untuk mencari nilai ekspektasi optimum suatu tindakan

Diberikan berbagai notasi yang diperlukan untuk model matematika dari pengambilan keputusan Metode Exhaustive Enumeration:

a. Stationary policy

Suatu ketetapan policy yang didefinisikan sebagai bentuk kausalitas yang merupakan sebab terpenuhinya syarat salah satu policy s yang mengakibatkan diambilnya suatu tindakan yang bersesuaian dengan policy s tersebut [6], Stationary policy selalu didefinisikan diawal saat proses enumerasi akan dilakukan, sehingga nilai policy optimum merupakan salah satu policy pada stationary policy yang maksimal.

b. Probabilitas Stationary

Suatu nilai probabilitas stationary policy s yang merupakan taraf jangka panjang dari matriks transisi P yang berkaitan dengan policy s, dilambangkan dengan π^s . [6]

Sama seperti halnya ekspektasi pendapatan satu langkah, nilai dari probabilitas stationary pun nantinya akan menjadi salah satu syarat yang harus dipenuhi untuk mencari nilai ekspektasi pendapatan dari policy s untuk setiap langkah transisi.

c. Ekspektasi Pendapatan dari policy s untuk Setiap Transisi

Suatu harapan yang diperoleh dengan mencari nilai kumulatif perkalian antara matriks reward dan matriks peluang transisi di setiap langkah transisi dari policy s, dilambangkan dengan E^s . [6]

Mencari policy optimum, dilakukan dengan mengambil nilai E^s maksimum.

Misalkan suatu persoalan keputusan mempunyai sejumlah S policy, dan asumsikan bahwa P dan R adalah matriks transisi (satu langkah) dan matriks pendapatan yang berkaitan dengan policy ke-k, $s = 1, 2, \dots, S$. maka langkah-langkah enumerasinya adalah sebagai berikut [6]:

Langkah 1.

Hitung harga V_i^s yaitu ekspektasi pendapatan satu langkah (satu periode) dari policy s, pada state i, $i=1,2,\dots,m$.

Langkah 2•

Hitung yaitu probabilitas stationary jangka panjang dari matriks transisi P^s yang berkaitan dengan policy s. probabilitas ini, jika ada, dihitung dengan persamaan:

$$\pi^s P^s = \pi^s \quad (3.5)$$

$$\pi_1^s + \pi_2^s + \pi_3^s + \dots + \pi_m^s = 1$$

dimana

$$lts = (Th, TC2, \dots, rrm).$$

Langkah 3

Tentukan E^s , yaitu ekspektasi pendapatan dari policy s untuk setiap langkah transisi (periode) dengan menggunakan persamaan:

$$E^s = \sum_{i=1}^m \pi_i^s v_i^s. \quad (3.6)$$

Langkah 4:

Policy optimum s^* ditentukan dengan:

$$E^{s^*} = \max_s E^s.$$

3.3 Model Infinite Stage pada Kasus Politik

Misalkan terdapat seorang calon member tetap yang akan bergabung terhadap suatu partai politik, tujuan dari calon member ini adalah mendapatkan Pilihan partai politik yang paling tepat yang bisa dia jadikan sebagai batu loncatan untuk mendapatkan jabatan di parlemen, yang jadi permasalahannya adalah calon member ini kebingungan untuk menentukan pilihan partai tersebut mengingat dinamika perpolitikan yang tidak bisa ditentukan arahnya, atau bisa juga kasus lain seperti suatu partai yang harus menentukan Dapil terpotensial untuk menentukan basis daerahnya. Markov decision Processes Model Infinite Stage bisa menjawab kebingungan ini. Seperti yang dijelaskan pada awal bab ini bahwa proses keputusan Markov dengan model infinite stage terbagi menjadi dua yaitu metode exhaustive enumeration dan policy Iteration. Berikut langkah-langkah untuk menentukan keputusan dalam politik dengan dua metode tersebut:

3.3.1 Metode Policy Iteration pada Keputusan Pilihan Politik

- a. Hal pertama yang harus dilakukan adalah melakukan pendefinisian calon matriks peluang transisi dan matriks reward, matriks peluang transisi dilambangkan dengan P dan matriks reward dilambangkan dengan R , pada kasus politik partai politik bisa dijadikan sebagai state yang dalam prosesnya

dapat menghasilkan matriks peluang transisi karena ada perpindahan pilihan (bisa berupa perpindahan partai atau perpindahan keadaan jumlah suara) dan jumlah kursi yang didapat oleh partai pada pemilu didefinisikan sebagai reward.

- b. Selanjutnya, hitung ekspektasi pendapatan kursi satu langkah dari setiap partai, yang dilambangkan dengan V_i^k , dengan $V_i^k = Y_i^m - \text{opt } r^k$ dengan P_i^k adalah elemen matriks transisi state i , langkah ke- k , dan r^k adalah elemen matriks kursi parlemen state i , langkah ke- k .
- c. Setelah didapat nilai ekspektasi reward satu langkah, lakukan langkah penentuan nilai dengan mencari nilai ekspektasi reward satu langkah yang maksimal.
- d. Langkah penentuan nilai ini akan menghasilkan beberapa persamaan, $E + f(i) - \sum_{l=0}^m p_l^k f(l) = V_i$, cari $f(i)$ untuk masing-masing partai.
- e. Selanjutnya, cari langkah perbaikan policy
Langkah perbaikan policy ini dilambangkan dengan $v_\eta(i)$, dengan
$$v_\eta(i) = \max_k (r_i^k + P_i \{ f(S_j) \})$$
- f. Hasil dari langkah perbaikan policy di atas kemudian dicari nilai k optimum, yaitu k yang menghasilkan langkah perbaikan policy sama dengan langkah penentuan nilai V_i^k pada masing-masing partai. Dilambangkan dengan k^* , di mana $k^* = \max_i v_\eta(i)$, $k^* = V_t$.
- g. Jika langkah perbaikan policy tidak sama atau tidak sesuai dengan langkah penentuan nilai V_i pada masing-masing partai, maka proses harus diulang dari langkah penentuan nilai sampai nilai langkah perbaikan policy sama dengan atau sesuai dengan langkah penentuan nilai.
- h. selanjutnya calon member bisa menentukan pilihan politiknya pada partai dengan policy optimum ini guna mendapatkan partai yang potensial yang bisa dijadikan "tunggangan" politiknya atau suatu partai untuk menentukan Dapil terpotensial.

3.3.2 Metode Exhaustive Enumeration pada Keputusan Pilihan Politik

Hal pertama yang harus dilakukan adalah melakukan pendefinisian calon matriks peluang transisi dan matriks reward, matriks peluang transisi dilambangkan dengan P dan matriks reward dilambangkan dengan R , pada kasus politik partai politik bisa dijadikan sebagai state yang dalam prosesnya dapat

menghasilkan matriks peluang transisi karena ada perpindahan pilihan (bisa berupa perpindahan partai atau perpindahan keadaan jumlah suara) dan jumlah kursi yang didapat oleh partai pada pemilu didefinisikan sebagai reward.

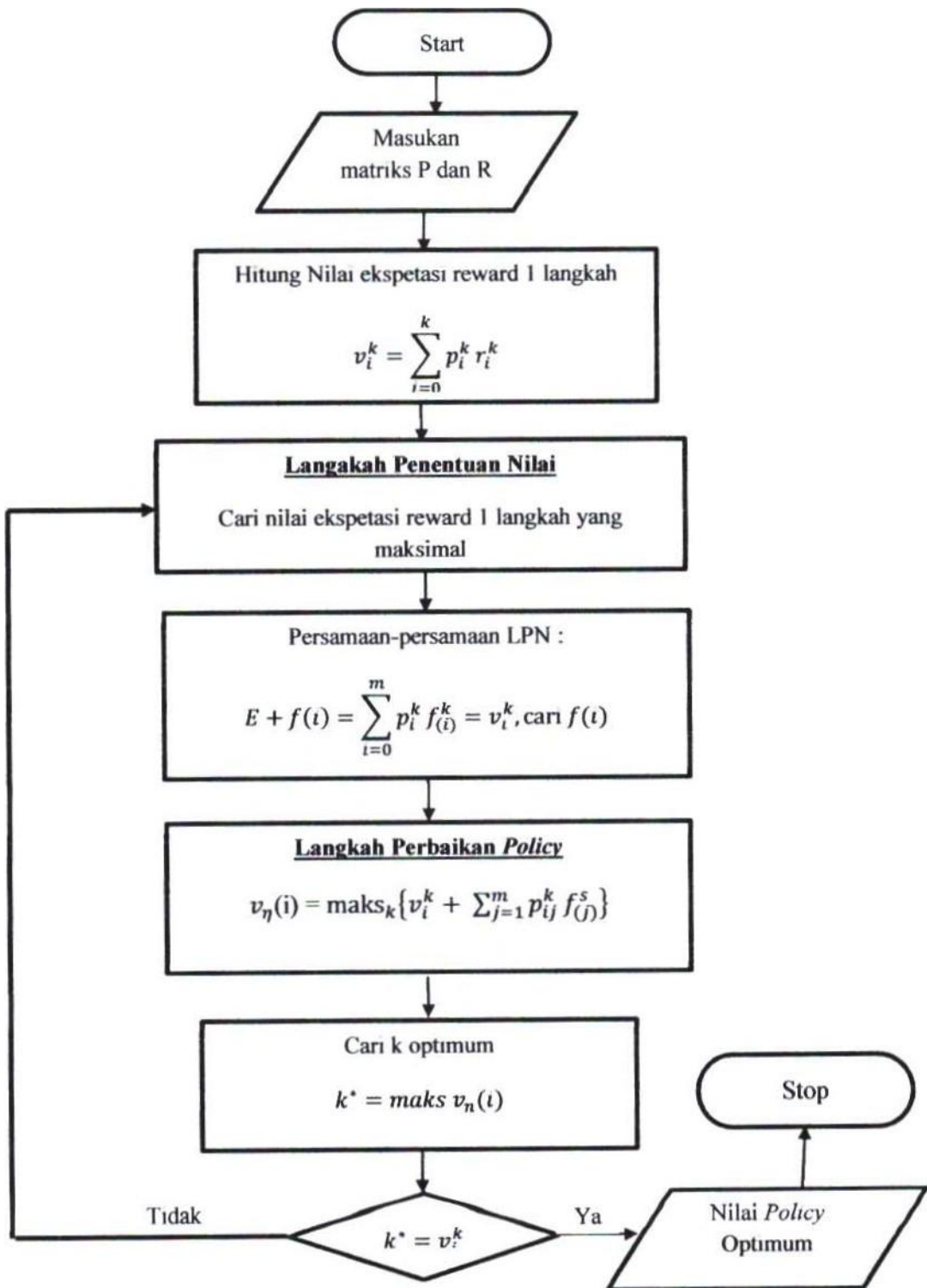
- 4 Setelah didapatkan P dan R dari data pemilihan umum yang diketahui selanjutnya ditentukan stationary policy yaitu suatu ketetapan policy yang didefinisikan sebagai bentuk kausalitas yang merupakan sebab terpenuhinya syarat salah satu policy s yang mengakibatkan diambilnya suatu tindakan yang bersesuaian dengan policy s tersebut, misal dalam pengamatan, ada tujuh partai politik yang dijadikan sasaran calon member, maka bisa ditetapkan tujuh policy stationary yang masing-masing policy menentukan tindakan pilihan partai, atau lima policy untuk lima Dapil
- 5 Selanjutnya, hitung ekspektasi pendapatan kursi satu langkah dari policy s setiap partai, yang dilambangkan dengan V_i^s , dengan

$$V_i^s = \sum_{j \in S} P_{ij}^s (R_{ij}^s + \gamma V_j^s)$$
 dengan P_{ij}^s adalah elemen matriks transisi state i, policy s, dan R_{ij}^s adalah elemen matriks reward state i, policy s.
- 6 Cari nilai probabilitas stationary policy setiap partai
 Nilai probabilitas stationary' setiap partai dilambangkan dengan T_i^s , dengan

$$\sum_{s \in S} T_i^s = 1 \quad \text{dan} \quad \sum_{i \in S} T_i^s = 1.$$
- 7 Cari nilai ekspektasi reward jumlah kursi setiap langkah transisi nilai ekspektasi reward jumlah kursi setiap langkah transisi dilambangkan dengan E^s , di mana $E^s = \sum_i V_i^s$.
- 8 Tentukan policy optimum dari hasil enumerasi policy s setiap partai
 Policy optimum dilambangkan dengan s^* , di mana $E^{s^*} = \max_s E^s$.
 Policy optimum ini didapat dari nilai ekspektasi reward jumlah kursi suatu partai yang tertinggi.
- 9 Nilai policy optimum pasti terletak pada suatu partai tertentu, selanjutnya calon member bisa menentukan pilihan politiknya pada partai dengan policy optimum ini guna mendapatkan partai yang potensial yang bisa dijadikan "tunggangan" politiknya, atau Dapil terpotensial untuk partai politik tertentu.

3.3 Flow Chart Metode Policy Iteration

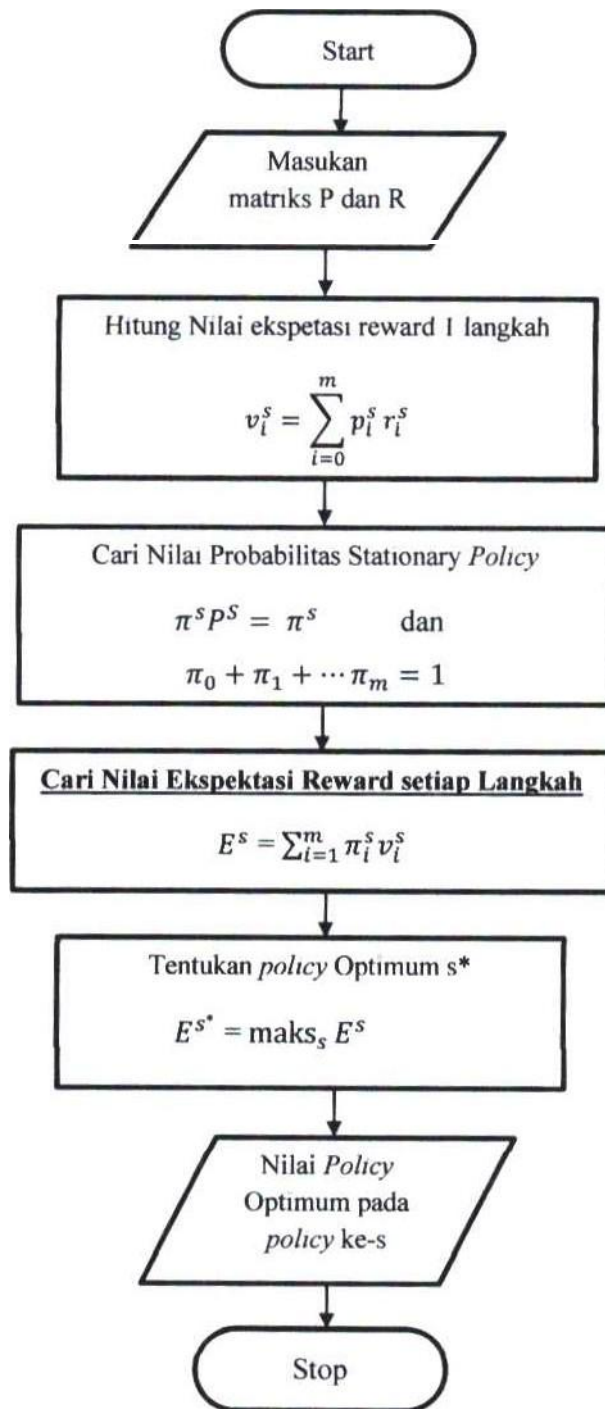
Berikut flow chart dari metode policy Iteration



Gambar 3. I. Flow Chart Metode policy iteration

3.4 Flow Chart Metode Exhaustive Enumeration

Berikut flow chart dari metode exhaustive enumeration



Gambar 3.2 Flow Chart Metode Exhaustive Enumeration