BAB III

HIDDEN MARKOV MODELS

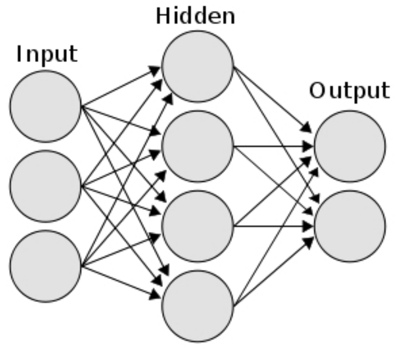
Rantai Markov bermanfaat untuk menghitung probabilitas urutan keadaan yang dapat diamati. Tetapi terkadang ada urutan dari suatu keadaan yang ingin diketahui tetapi tidak dapat diamati. Maka untuk kasus seperti itu dikembangkanlah suatu model baru yang dapat memodelkan keadaan yang tersembunyi, yaitu *Hiden Markov Models* (HMM). HMM semakin popular pada dekade terakhir ini karena model tersebut kaya akan struktur matematika dan mengandung teori dasar yang bisa digunakan untuk beberapa aplikasi yang penting.

HMM pada dasarnya perluasan dari rantai markov yang merupakan model stokastik. Biasanya dalam model Markov setiap keadaan dapat terlihat langsung oleh pengamat, sehingga kemungkinan transisi antar keadaan menjadi satu-satunya parameter yang teramati. Dalam HMM, keadaan tidak dapat terlihat langsung meskipun parameter model diketahui, model tersebut tetap tersembunyi, tetapi hasil keluaran (*output*) yang bergantung pada keadaan tersebut dapat terlihat.

3.1 Definisi *Hidden Markov Models*

HMM adalah sebuah proses stokastik ganda di mana salah satu prosesnya tidak dapat diobservasi (*hidden*). Proses yang tidak dapat diobservasi ini hanya dapat diobservasi melalui proses yang dapat diobservasi. Sebuah HMM menggabungkan dua atau lebih rantai Markov dengan hanya satu rantai yang terdiri dari *state* yang dapat diobservasi dan rantai lainnya membentuk *state* yang tidak dapat diobservasi (*hidden*), yang mempengaruhi hasil dari *state* yang dapat diobservasi.

Jika $X = \{X_1, X_2, \ldots, X_T\}$ adalah sebuah proses Markov, dan $O = \{O_1, O_2, \ldots, O_T\}$, adalah sebuah fungsi dari X, maka X adalah sebuah HMM yang dapat diobservasi melalui O, atau dapat ditulis $O = f(X)$ untuk suatu fungsi *f*. Parameter X menyatakan *state process* yang tersembunyi (*hidden*), sementara parameter O menyatakan *observation process* yang dapat diobservasi. Untuk ilustrasi HMM dapat dilihat gambar 3.1 berikut,



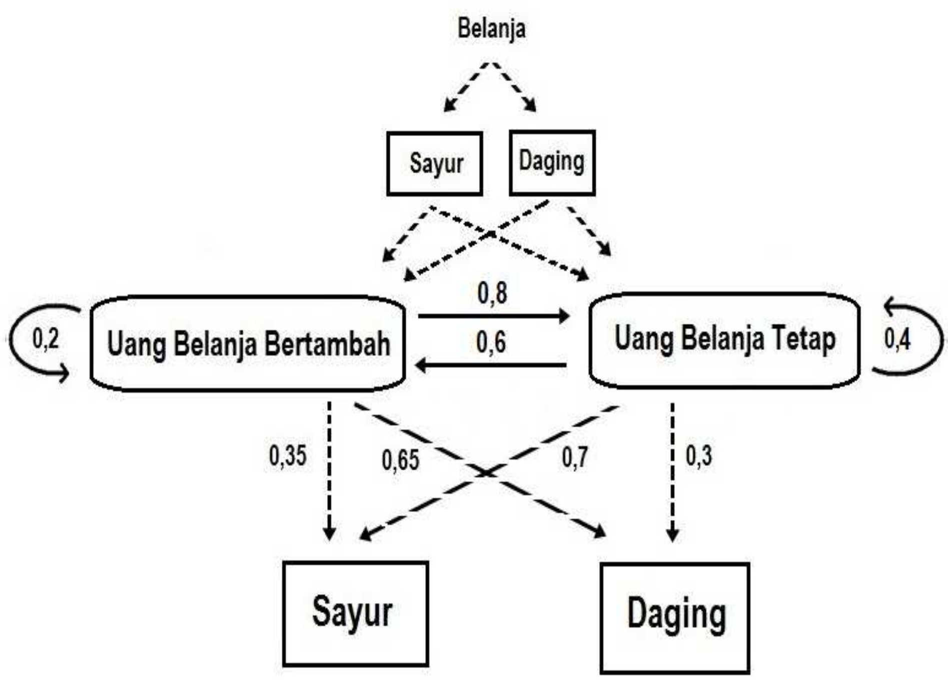
Gambar 3.1 Ilustrasi HMM

Contoh 3.1

Seorang ibu rumah tangga setiap minggunya berbelanja di suatu pasar tradisional yang beranggapan setiap berbelanja dipengaruhi waktu pergi ke pasar tersebut, yaitu pagi dan siang. Jika diketahui minggu ini ibu membeli 0,40 sayuran dan 0,60 daging, maka ibu ingin memperkirakan apa yang akan ibu beli untuk minggu selanjutnya. Dengan pengalaman yang diketahui , peluang minggu depan ibu berbelanja sayur jika minggu ini berbelanja daging adalah 0,7 dan peluang minggu depan berbelanja daging jika minggu ini berbelanja sayur adalah 0,3.

Tetapi barang yang dibeli ibu tersebut juga dipengaruhi oleh faktor yang tersembunyi (*hidden*) yang memepengaruhi perpindahan keadaan terobservasi, yaitu faktor uang belanja. Jika ibu tersebut mendapatkan uang belanja tambahan maka ibu tersebut akan membeli sayuran dengan probabilitas 0,35 dan daging dengan probabilitas 0,65. Sementara jika tidak mendapatkan uang belanja maka ibu tersebut akan membeli sayuran dengan probabilitas 0,70 dan daging dengan probabilitas 0,30.

Jika minggu ini ibu mendapat tambahan uang belanja, maka minggu depan ibu akan mendapatkan lagi tambahan uang belanja dengan probabilitas 0,2 dan tidak mendapatkan tambahan uang belanja dengan probabilitas 0,8. Sementara jika minggu ini ibu tidak mendapatkan tambahan uang belanja, maka minggu depan ibu akan memdapatkan tambahan uang belanja dengan probabilitas 0,6 dan tidak mendapatkan tambahan uang belanja dengan probabilitas 0,4.



Gambar 3.2 Ilustrasi Contoh Soal

Suatu HMM secara umum memiliki parameter sebagai berikut :

1) N, yaitu jumlah elemen keadaan tersembunyi *(hidden state)* pada model HMM. Sebagai contoh, dapat dilihat pada contoh 3.1, dimana keadaaan yang tersembunyi adalah penambahan uang belanja dan tidak adanya penambahan uang belanja, maka N = 2 atau dapat ditulis $X_t = i, i = 1$ (penambahan uang belanja), 2 (tidak adanya penambahan uang belanja).

2) M, yaitu jumlah elemen keadaan yang terobservasi. Pada contoh 3.1, untuk masalah pilihan jenis barang yang akan dibeli oleh ibu, banyaknya elemen hanya 2, yaitu sayur dan daging, sehingga $k = 1, 2$.

3) Matriks peluang tansisi $A = \{a_{ij}\}$ di mana a_{ij} adalah elemen dari A yang merupakan himpunan distribusi kemungkinan perpindahan *state (transition probability*) pada saat $t + 1$, jika diketahui keadaan X pada saat t, atau $a_{ij} = P(X_{t+1} = j|X_t = i)$ di mana $1 \leq i, j \leq N$. Pada contoh 3.1, untuk masalah pilihan jenis barang yang akan dibeli oleh ibu, dengan keadaan $X_t = i, i = 1$ (penambahan uang belanja) dan 2 (tidak adanya penambahan uang belanja), akan ada matriks transisi berukuran 2×2.

$$A = \begin{pmatrix} 0{,}2 & 0{,}8 \\ 0{,}6 & 0{,}4 \end{pmatrix}$$

4) $B = \{b_j(k)\}$ merupakan himpunan distribusi kemungkinan simbol pengamatan pada *state j*. Distribusi peluang observasi pada saat *t*, pada keadaan *i*, yang biasa dikenal dengan matriks emisi $B = \{b_i(k)\}$, di mana

$$b_i^k = b_i(k) = P(O_t = k|X_t = i), 1 \leq i \leq N, 1 \leq k \leq M$$

K adalah observasi pada waktu ke-*t* bernilai *k*, jadi B adalah matriks berukuran $N \times M$.

Pada contoh 3.1, mengenai pilihan jenis barang yang akan dibeli oleh ibu akan ada matriks B yang berukuran 2×2,

$$B = \begin{pmatrix} 0{,}35 & 0{,}65 \\ 0{,}7 & 0{,}3 \end{pmatrix}$$

5) $\pi = \{\pi_i\}$ dimana $\pi_i = P(q_1 = S_i)$, merupakan himpunan distribusi probabilitas awal, dengan $\pi(i) = P(X_1 = i), 1 \leq i \leq N$ Contoh mengenai pilihan jenis barang yang akan dibeli oleh ibu, $\pi(1)$ = (penambahan uang belanja), $\pi(2)$ = (tidak adanya penambahan uang belanja).

Misal untuk keadaan awal pada contoh 3.1 ada atau tidaknya tambahan uang belanja, sehingga ada atau tidaknya tambahan uang belanja tersebut diasumsikan memiliki peluang yang sama atau, $\pi = \begin{pmatrix} 0{,}5 \\ 0{,}5 \end{pmatrix}$

Spesifikasi lengkap dari sebuah HMM yakni mencakup dua ukuran tetap parameter N dan M, juga terdapat tiga komponen yang merupakan ukuran (probabilitas), yaitu A, B, dan π yang merupakan bentuk ringkas dari HMM. Akibatnya HMM lebih dikenal dengan notasi $\lambda = (A, B, \pi)$ dengan A berukuran $N \times N$ dan B berukuran $N \times M$.

3.2 Asumsi pada HMM

Ada tiga asumsi pokok yang dibutuhkan dalam analisis HMM, yaitu:

1. Asumsi Markov

   Asumsi ini menyatakan bahwa keadaan berikutnya hanya dipengaruhi oleh keadaan saat ini. Model yang dihasilkan adalah HMM orde pertama. Pada beberapa kasus di kehidupan nyata, keadaan selanjutnya mungkin dipengaruhi oleh k keadaan sebelumnya, yang akan menghasilkan HMM orde ke-k yang lebih sulit untuk dianalisa daripada HMM orde pertama.

2. Asumsi Stasioneritas

   Asumsi ini menyatakan bahwa peluang transisi dari suatu keadaan ke keadaan lainnya independen dengan waktu saat transisi itu terjadi. Sehingga untuk sembarang t_1 dan t_2 berlaku :

$$P\left(X_{t_1+1} = j \middle| X_{t_1} = i\right) = P\left(X_{t_2+1} = j \middle| X_{t_2} = i\right) = P_{ij} \qquad (3.1)$$

3. Asumsi Kebebasan

   Jika diketahui suatu barisan observasi $O = O_1, O_2, \dots, O_T$ dan suatu barisan keadaan $X = X_1, X_2, \dots, X_T$. Maka pengamatan saat ini bersifat independen secara statistik dengan pengamatan sebelumnya. Atau dapat dinyatakan:

$$P(O|X, \lambda) = \prod_{t=1}^{T} P(O_t|X_t, \lambda) \qquad (3.2)$$

3.3 Masalah-masalah Utama Dalam HMM

Terdapat tiga masalah utama yang dapat diselesaikan oleh *Hidden Markov Models,* yaitu:

1. Evaluasi (Menghitung Peluang Observasi)

   Operasi evaluasi dalam HMM adalah perhitungan probabilitas dari urutan nilai observasi yang diberikan oleh *Hidden Markov Models*. Untuk menghitung peluang terjadinya suatu barisan observasi dibutuhkan algoritma yang lebih efisien untuk menyelesaikan masalah evaluasi. Algoritma yang banyak digunakan untuk penyelesaian masalah evaluasi adalah algoritma maju *(forward algorithm)* yang keadaannya mengalir kedepan, algoritma mundur *(backward algorithm)* yang keadaannya mengalir kebelakang dari observasi terakhir pada saat *T*, dan algoritma maju-mundur *(forward-backward algorithm)* yang merupakan gabungan dari algoritma *forward-backward*.

2. Decoding (Menentukan Barisan Keadaan Tersembunyi)

   Permasalahan *decoding* ini yaitu menemukan barisan *state* terbaik (optimal) yang berasosiasi dengan barisan observasi dari sebuah model yang juga telah diketahui. Barisan *state* yang optimal didefinisikan sebagai barisan *state* yang mempunyai probabilitas tertinggi dalam menghasilkan barisan observasi yang telah diketahui sebelumnya. Untuk menentukan keadaan tersembunyi dari suatu barisan observasi perlu digunakan suatu metode yang mempertimbangkan probabilitas transisi *state* pada proses pencarian barisan *state* yang paling optimal. Metode yang banyak digunakan untuk penyelesaian masalah ini antara lain algoritma Viterbi dan Entropi.

3. *Learning* (Menaksir Parameter-parameter HMM)

   Operasi *learning* dalam HMM adalah melatih parameter HMM jika diberikan data set barisan-barisan tertentu agar dapat menemukan himpunan transisi *state* yang paling mungkin beserta probabilitas output. Untuk menyelesaikan permasalahan terakhir pada HMM ini, biasanya digunakan algoritma Baum-Welch yang merupakan kasus khusus dari algoritma EM (Ekspektasi Maksimum). Algoritma EM sendiri merupakan algoritma yang digunakan untuk mempelajari model-model probabilistik dalam suatu kasus yang melibatkan keadaan-keadaan tersembunyi.

3.4 Menghitung Peluang Observasi dengan Algoritma Maju (*The Forward Algorithm*)

3.4.1 Definisi Algoritma *Forward*

Algoritma *Forward* adalah proses iterasi yang didasarkan pada perhitungan peluang bersyarat melalui sifat-sifat pada peluang. Algoritma Forward digunakan untuk menghitung probabilitas barisan yang terobservasi.

Bila diketahui sebuah model $\lambda = (A, B, \pi)$ dan sebuah barisan observasi $O = \{O_1, O_2, \dots, O_T\}$, kemudian akan dihitung $P(O|\lambda)$ yang dapat ditulis sebagai,

$$P(O|\lambda) = \sum_X P(O|X, \lambda) P(X|\lambda) \tag{3.3}$$

Di mana $X = \{X_1, X_2, \dots, X_T\}$ adalah suatu barisan, $P(O|X, \lambda)$ adalah probabilitas barisan observasi O untuk suatu barisan *state* X , dan $P(X|\lambda)$ merupakan probabilitas dari X bila diberikan sebuah model. Karena pada HMM barisan observasi diasumsikan independen, maka

$$P(O|X, \lambda) = \prod_{t=1}^{T} P(O_t|X_t, \lambda) = b_1(O_1).b_2(O_2).\dots.b_T(O_T).$$

$$P(X|\lambda) = \pi(1) a_{12} a_{23} \dots . a_{\mathrm{T-1,T}} \tag{3.4}$$

Sehingga diperoleh,

$$P(O|\lambda) = \sum_X P(O|X, \lambda) P(X|\lambda)$$

$$= \sum_{1,2,\dots,T} \pi(1) b_1(o_1)\, a_{12} b_2(o_2) a_{23} \dots a_{\mathrm{T-1,T}} b_T(o_T) \tag{3.5}$$

Dengan menggunakan definisi peluang bersyarat $P(O|\lambda)$ dapat dihitung dari urutan observasi O dengan model λ, namun operasi perhitungan yang dibutuhkan akan bertambah banyak karena operasinya akan naik secara eksponensial, seiring dengan bertambah panjangnya barisan observasi yang ada. Algoritma *forward* menyimpan nilai yang telah dihitung pada iterasi sebelumnya. Algoritma ini akan sangat efisien ketika panjang barisan observasinya cukup besar.

Jika variabel $\alpha_t(i)$ sebagai variabel maju, dimana:

$$\alpha_t(i) = P(O_1, O_2, \dots, O_t, X_t = i|\lambda) \qquad (3.6)$$

dengan $\alpha_t(i)$ mendeskripsikan probabilitas dari sebagian urutan observasi $\{O_1, O_2, \dots, O_t\}$ dan berakhir pada *state* i pada saat t dimana $t = 1,2, \dots, T$ jika diketahui suatu barisan observasi $\{O_1, O_2, \dots, O_t\}$.

3.4.2 Menghitung Peluang Menggunakan Algoritma *Forward*

Menurut Rabiner (1989), untuk menghitung peluang dalam algoritma *forward* dapat terdiri dari tiga bagian, yaitu:

1. Inisialisasi

$$\alpha_1(i) = \pi(i) b_i(O_1) \text{ dimana } 1 \leq i \leq N \qquad (3.7)$$

Persamaan tersebut diperoleh dari definisi variabel maju dengan cara mensubstitusikan dua definisi parameter HMM yaitu $\pi(i) = P(X_t = i)$ dan $b_i(k) = P(O_t = k|X_t = i)$:

$$\alpha_1(i) = P(O_1, X_1 = i|\lambda)$$

$$= P(X_1 = i, \lambda)P(O_1|X_1 = i, \lambda)$$

$$= \pi(i)P(O_1|X_1 = i, \lambda)$$

$$= \pi(i)b_i(O_1)$$

2. Induksi

$$\alpha_{t+1}(j) = \{\sum_{i=1}^{N} \alpha_t(i)a_{ij}\}b_j(O_{t+1}) \quad (3.8)$$

, dengan $j = 1, \dots, N, t = 1, \dots, T-1$

Pada tahap ini akan dihitung nilai α pada saat $t \geq 1$, sama seperti pada tahap inisialisasi, pembuktian dilakukan dengan mensubstitusikan dua parameter HMM yaitu $b_i(k) = P(O_t = k|X_t = i)$ dan a_{ij} sehingga diperoleh:

$$\alpha_{t+1}(j) = P(O_1, O_2, \dots, O_t, O_{t+1}, X_{t+1} = j|\lambda)$$

$$= P(O_1, O_2, \dots, O_t, O_{t+1}|X_{t+1} = j, \lambda)P(X_{t+1} = j|\lambda)$$

$$= P(O_1, O_2, \dots, O_t|X_{t+1} = j, \lambda)P(O_{t+1}|X_{t+1} = j, \lambda)P(X_{t+1} = j|\lambda)$$

$$= P(O_1, O_2, \dots, O_t, X_{t+1} = j|\lambda)\, P(O_{t+1}|X_{t+1} = j, \lambda)$$

$$= P(O_1, O_2, \dots, O_t, X_{t+1} = j|\lambda)b_j(O_{t+1})$$

$$= \left[\sum_{i=1}^{N} P(O_1, O_2, \dots, O_t, X_t = i, X_{t+1} = j|\lambda)\right] b_j(O_{t+1})$$

$$= \left[\sum_{i=1}^{N} P(O_1, O_2, \dots, O_t, X_t = i|\lambda)\, P(X_{t+1} = j|O_1, O_2, \dots, O_t, X_t = i, \lambda)\right] b_j(O_{t+1})$$

$$= \left[\sum_{i=1}^{N} P(O_1, O_2, \dots, O_t, X_t = i|\lambda)\, P(X_{t+1} = j|X_t = i, \lambda)\right] b_j(O_{t+1})$$

$$= \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1})$$

3. Terminasi

Pada tahap ini adalah menjumlahkan semua peluang gabungan dari observasi dan *hidden state* bila diketahui sebuah model sehingga diketahui peluang marjinal dari observasi tersebut atau ditulis:

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i) \qquad (3.9)$$

Contoh 3.2

Perhatikan kembali contoh 3.1, untuk permasalahan pertama pada HMM adalah menghitung peluang bahwa model $\lambda = (A, B, \pi)$ menghasilkan barisan observasi O = *{sayur, daging}*

Diketahui : $A = \begin{pmatrix} 0{,}2 & 0{,}8 \\ 0{,}6 & 0{,}4 \end{pmatrix}$, $B = \begin{pmatrix} 0{,}35 & 0{,}65 \\ 0{,}7 & 0{,}3 \end{pmatrix}$

Dengan memisalkan proses dapat dimulai dengan belanja sayur dan belanja daging, dengan peluang yang sama yaitu $\pi = \begin{pmatrix} 0{,}5 \\ 0{,}5 \end{pmatrix}$

Penyelesaian:

Dari contoh 3.1 terdapat dua keadaan untuk “Sayur” (1), dan “Daging” (2), dan untuk keadaan yang tersembunyi memiliki dua keadaan, yaitu “Penambahan Uang Belanja” (1) dan “Tidak Ada Penambahan Uang Belanja” (2). Permasalahan tersebut akan diselesaikan dengan menggunakan algoritma maju, di mana panjang barisan observasi $T = 2$.

- Tahap Inisialisasi

  Pada tahap inisialisasi akan ada 2 buah α masing-masing untuk penambahan uang belanja (1), tidak adanya penambahan uang belanja (2) , sehingga probabilitasnya :

  $\alpha_1(i) = \pi(i)b_i(O_1)$

  $\alpha_1(1) = \pi(1)b_1(sayur) = (0{,}5)(0{,}35) = 0{,}175$

  $\alpha_1(2) = \pi(2)b_2(sayur) = (0{,}5)(0{,}7) = 0{,}35$

- Tahap Induksi

  Pada tahap ini akan dihitung $\alpha_2(j)\ dan\ \alpha_3(j)$

  Probabilitas pada $t = 2$ untuk keadaan $j = 1, 2$ dihitung sebagai berikut :

  $$\alpha_2(j) = \left[\sum_{i=1}^{N} \alpha_1(i)a_{ij}\right]b_j(O_2)$$

$$\alpha_2(1) = [(\alpha_1(1)a_{11}) + (\alpha_1(2)a_{21})]b_1(O_2 = daging)$$

$$= [(0{,}175)(0{,}2) + (0{,}35)(0{,}6)](0{,}65) \ = 0{,}1592$$

$$\alpha_2(2) = [(\alpha_1(1)a_{12}) + (\alpha_1(2)a_{22})]b_2(O_2 = daging)$$

$$= [(0{,}175)(0{,}8) + (0{,}35)(0{,}4)](0{,}3) \ = 0{,}084$$

- Tahap Terminasi

  Dari tahap terminasi maka algoritma menghasilkan penyelesaian sebagai berikut:

$$P(O = sayur, daging|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

$$= \alpha_1(1) + \alpha_1(2)$$

$$= 0{,}175 + 0{,}35$$

$$= 0{,}525$$