Hidden Markov Model

Oleh :
Firdaniza, Nurul Gusriani dan Akmal

Dosen Jurusan Matematika FMIPA Universitas Padjadjaran
Jl. Raya Bandung Sumedang Km 21, Jatinangor, Jawa Barat
Telp. / Fax : 022-7794696

Abstrak

Hidden Markov Model (*HMM*) adalah peluasan dari rantai Markov di mana *state*nya tidak dapat diamati secara langsung (tersembunyi), tetapi hanya dapat diobservasi melalui suatu himpunan pengamatan lain. Pada *HMM* terdapat tiga permasalahan mendasar yang harus diselesaikan yakni *evaluation problem, decoding problem, dan learning problem.*

Dalam paper ini, akan dijelaskan tentang Hidden Markov Model(HMM) dan solusi dari ketiga masalah mendasar dalam HMM tersebut, yakni evaluation problem dengan algoritma forward, decoding problem dengan algoritma viterbi, dan learning problem dengan algoritma Baum-Welch.

Kata kunci : *Hidden Markov Model, evaluation problem, decoding problem, learning problem*

1. Pendahuluan

Proses stokastik adalah keluarga peubah acak $\{X_t, t \in T\}$, T disebut himpunan parameter. Himpunan nilai – nilai yang mungkin dari X_t disebut Ruang *State.*

Proses stokastik dengan ruang *state* (keadaan) diskrit, serta mempunyai sifat di mana *state* "saat ini" hanya tergantung pada *state* "sebelumnya" dan bebas dari histori yang lalu disebut dengan rantai markov.

Pada rantai Markov setiap *state* dapat diamati secara langsung. Seperti kasus cuaca, keadaan cuaca esok hari dapat kita prediksi melalui keadaan cuaca hari ini. Andaikan cuaca dikelompokkan menjadi tiga (cerah, hujan, berawan), maka dengan diberikan peluang perubahan cuaca, kita dapat menentukan peluang cuaca esok hari dan beberapa hari berikutnya. Akan tetapi bila seseorang berada dalam ruang tertutup dan dia tidak mengetahui cuaca di luar, kemudian dia disuruh menebak keadaan cuaca esok hari. Dalam hal ini *state* cuaca tidak dapat diamati secara langsung, namun pengamatan dapat

Dipresentasikan dalam Seminar Nasional Matematika dan Pendidikan Matematika 2006 dengan tema " **Trend Penelitian dan Pembelajaran Matematika di Era ICT "** yang diselenggarakan pada tanggal 24 Nopember 2006

dilakukannya dengan memperhatikan apakah orang masuk ke ruangan tersebut membawa payung atau tidak. Kasus ini dapat dimodelkan sebagai *Hidden Markov Model* (*HMM*).

Pada paper ini akan diberikan ulasan tentang *HMM* dan tiga masalah mendasar yang ada dalam HMM beserta solusinya, yakni *Evaluation problem,*

Decoding problem,

Learning problem.

HMM ini bermanfaat dalam menyelesaikan kasus-kasus dimana pengamatan tidak

dapat dilakukan terhadap suatu state, tetapi kita dapat menentukan peluang proses berada dalam stste tertentu. Pembahasan pada paper ini dibatasi hanya pada pengamatan diskrit saja.

2. Metode Penelitian

Pertama-tama dilakukan kajian ulang tentang rantai markov diskrit, agar dapat memberi gambaran tentang HMM, kemudian dipelajari tentang Teori *Hidden Markov Model* (HMM). Untuk lebih jelasnya diberikan suatu contoh. Kemudian dibahas tiga masalah mendasar dalam HMM, yakni *evaluation problem, decoding problem,* dan *learning problem*. Untuk *evaluation problem* diselesaikan dengan Algoritma *Forward* , untuk memecahkan *decoding problem* digunakan Algoritma *Viterbi,* serta untuk *learning problem* digunakan algorithm Baum-Welch. Sebagai gambaran, diberikan studi kasus tentang masalah cuaca.

3. Pembahasan

3.1 Rantai Markov Diskrit

Misalkan $\{Q_t,\ t = 0,1,2,3,...\}$ adalah proses stokastik parameter (waktu) diskrit dengan ruang state $\{s_1, s_2, s_3, ...\}$.

Jika $P(Q_{t+1} = s_j \mid Q_0 = s_{i_0}, Q_{i_1} = s_{i_1}, ..., Q_t = s_i)$

$= P(Q_{t+1} = s_j \mid Q_t = s_i)$

$= a_{ij}$

(1)

maka proses disebut rantai markov waktu diskrit dan a_{ij} disebut peluang transisi dari state s_i ke state s_j. Apabila a_{ij} tidak tergantung pada n maka a_{ij} disebut peluang transisi stasioner. Matriks $\mathbf{A} = \left[a_{ij}\right]$, dengan $\sum_{j=0}^{\infty} a_{ij} = 1$ disebut matriks peluang transisi.

Dengan mengetahui distribusi inisial dan peluang transisi, suatu rantai markov dapat dikenali secara lengkap, seperti terlihat pada persamaan berikut

$$P(Q_0 = s_{i_0}, Q_1 = s_{i_1}, ..., Q_t = s_{i_t})$$

$$= P(Q_t = s_{i_t} \mid Q_0 = s_{i_0}, ..., Q_{t-1} = s_{i_{t-1}}) \,.\, P(Q_0 = s_{i_0}, ..., Q_{t-1} = s_{i_{t-1}})$$

$$= P(Q_t = s_{i_t} \mid Q_{t-1} = s_{i_{t-1}}) \,.\, P(Q_0 = s_{i_0}, ..., Q_{t-1} = s_{i_{t-1}})$$

$$= a_{i_{t-1}i_t} \, P(Q_0 = s_{i_0}, ..., Q_{t-1} = s_{i_{t-1}})$$

$$= ......$$

$$= a_{i_{t-1}i_t} a_{i_{t-2}i_{t-1}} ... a_{i_0 i_1} P(Q_0 = i_0)$$

(2)

$P(Q_0 = i_0)$ disebut peluang inisial . $\pi(0) = (\pi_0(0), \pi_1(0), ...)$dimana $\pi_j(0) = P(Q_0 = s_j)$ disebut distribusi inisial. Sebagai gambaran, perhatikan contoh berikut:

Contoh 1

Misalkan cuaca dalam satu hari dapat dikelompokkan menjadi cerah, hujan, dan berawan. Perubahan cuaca dalam hari yang berurutan dinyatakan dalam peluang pada tabel berikut :

Cuaca hari ini	Cuaca besok		
	cerah	hujan	berawan
cerah	0.8	0.05	0.15
hujan	0.2	0.6	0.2
berawan	0.2	0.3	0.5

Tabel 1. Peluang perubahan cuaca pada hari yang berurutan

Jika cuaca pada hari pertama (t = 1) adalah cerah, berapa peluang cuaca pada 5 hari berikutnya cerah-hujan-berawan-hujan-hujan.

Jawab: Misalkan State 1 (s_1) : cerah, state 2 (s_2) : hujan, state 3 (s_3) : berawan.

Dengan menggunakan persamaan (2) diperoleh :

$$P(Q_0 = s_1, Q_1 = s_1, Q_2 = s_2, Q_3 = s_3, Q_4 = s_2, Q_5 = s_2)$$

$$= P(Q_0 = s_1)P(Q_1 = s_1 \mid Q_0 = s_1)P(Q_2 = s_2 \mid Q_1 = s_1)P(Q_3 = s_3 \mid Q_2 = s_2)P(Q_4 = s_2 \mid Q_3 = s_3)P(Q_5 = s_2 \mid Q_4 = s_2)$$

$$= \pi_1(0).a_{11}.a_{12}.a_{23}.a_{32}a_{22}$$

= (1)(0.8)(0.05)(0,2)(0,3)(0,6) = 0,00144

3.2 Hidden Markov Model

Dari contoh 1 di atas terlihat bahwa state (cuaca) dapat diamati secara langsung. Akan tetapi, jika seseorang dikunci dalam satu ruangan tertutup sehingga dia tidak dapat mengetahui keadaan cuaca diluar, kemudian orang tersebut disuruh menerka keadaan cuaca, maka pengamatan yang dapat dilakukan hanyalah dengan melihat apakah orang yang masuk ke ruangan terkunci tersebut membawa payung atau tidak. Masalah seperti ini dapat dimodelkan dalam bentuk *Hidden Markov Model* (*HMM*).

Elemen-elemen dari *HMM* adalah:

1. N, yaitu jumlah *state*, dengan ruang *state* $S = \{s_1, s_2, ..., s_N\}$ dan *state* pada waktu t dinyatakan dengan Q_t. Untuk kasus di atas adalah keadaan cuaca.

2. *M*, yaitu jumlah pengamatan (observasi) tiap state, dengan ruang observasi $V = \{v_1, v_2, ..., v_M\}$, untuk kasus di atas, orang datang membawa payung atau tidak membawa payung.
3. $\mathbf{A} = \left[a_{ij}\right]$, yaitu matriks peluang transisi.
4. $\mathbf{B} = \left[b_{jm}\right]$, yaitu matriks peluang bersyarat observasi v_m jika proses berada pada *state j*, dimana:

   $b_{jm} = b_j\left(O_t\right) = P\left(O_t = v_m \middle| Q_t = s_j\right)$, $1 \le j \le N$ dan $1 \le m \le M$.

   (3)
5. π_i yaitu distribusi *state* awal.

Sehingga *Hidden Markov Model* dapat dituliskan dalam notasi $\lambda = \left(\mathbf{A}, \mathbf{B}, \pi\right)$. Jika diberikan *N*, *M*, **A**, **B**, dan π, *HMM* dapat digunakan sebagai pembangkit barisan observasi: $O = O_1, O_2, O_3, ..., O_T$

Jika seseorang dikurung dalam ruangan tertutup dan ia tidak mengetahui keadaan cuaca di luar. Peluang cuaca pada hari ke-*t* $\left(Q_t\right)$, dengan kemungkinan $\left\{s_1 = \text{cerah}, s_2 = \text{hujan}, s_3 = \text{berawan}\right\}$ hanya bisa diperoleh berdasarkan observasi O_t, dengan $O_t = \text{membawa payung}$ atau $O_t = \text{tidak membawa payung}$.

Untuk menyimpulkan cuaca di luar berdasarkan observasi bahwa orang masuk membawa payung atau tidak, digunakan ukuran untuk peluang, yang disebut kemungkinan (*L*):

$$L\left(Q_1 = s_{i_1}, ..., Q_T = s_{i_T} \middle| O_1 = v_{m_1}, ..., O_T = v_{m_T}\right)$$

$$= \prod_{t=1}^{T} P\left(O_t = v_{m_t} \middle| Q_t = s_{i_t}\right) \cdot P\left(Q_1 = s_{i_1}\right) \cdot \prod_{t=2}^{T} P\left(Q_t = s_{i_t} \middle| Q_{t-1} = s_{i_{t-1}}\right) \qquad (4)$$

3.3 Tiga Masalah Dasar HMM

Agar HMM dapat diaplikasikan ke berbagai masalah nyata, ada tiga masalah mendasar dalam HMM yang harus diselesaikan, yakni:

a. ***Evaluation problem***

yakni akan dicari $P\left(O|\lambda\right)$ atau peluang dari barisan observasi

$O=\left\{O_1=v_{m_1},O_2=v_{m_2},\ldots,O_T=v_{m_T}\right\}$ jika diberikan *HMM;* $\lambda=\left(\mathbf{A},\mathbf{B},\pi\right)$.

Peluang ini dapat ditentukan secara induksi dengan menggunakan algoritma *forward*.

Definisikan variabel *forward* $\alpha_t(i)$:

$$\alpha_t(i)=P\left(O_1=v_{m_1},\ldots,O_t=v_{m_t},Q_t=s_i\,|\lambda\right)$$

(5)

yaitu peluang barisan observasi $O_1,O_2,\ldots,O_t$ dan *state* s_i pada waktu *t* jika diberikan λ.

Secara induktif $P\left(O|\lambda\right)$ dapat dihitung sebagai berikut :

(i) Inisialisasi

$$\alpha_1\left(i\right)=\pi_i b_i\left(O_1\right) \qquad 1\le i\le N$$

(6)

(ii) Induksi

$$\alpha_{t+1}\left(j\right)=\left[\sum_{i=1}^{N}\alpha_t\left(i\right)a_{ij}\right]b_j\left(O_{t+1}\right) \qquad ,1\le t\le T-1 \quad 1\le j\le N$$

(7)

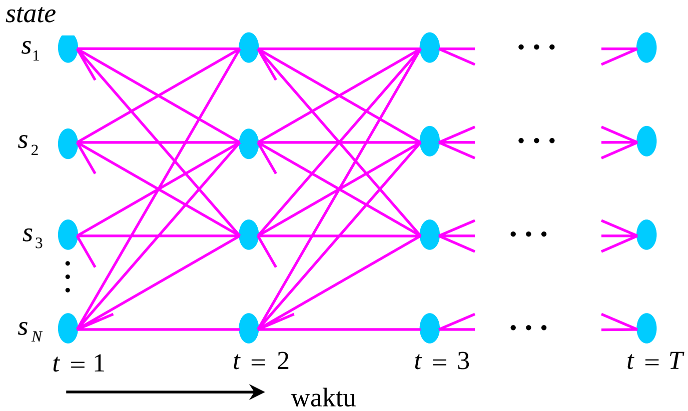
(iii) Akhir ; $P\left(O|\lambda\right)=\sum_{i=1}^{N}\alpha_T\left(i\right)$

(8)

b. ***Decoding problem***

Akan dicari barisan *state* yang optimal $Q^*=\left\{Q^*{}_1,Q^*{}_2,\ldots,Q^*{}_T\right\}$ jika diberikan barisan observasi $O=\{O_1,O_2,\ldots,O_T\}$ dan model $\lambda=(\mathbf{A},\mathbf{B},\pi)$. Barisan *state* terbaik yang akan ditentukan yaitu berupa lintasan tunggal yang terhubung dari $t=1,2,\ldots,T$.

Seperti yang terlihat pada gambar 1, begitu banyak lintasan tunggal yang mungkin. Kemudian akan dipilih satu lintasan tunggal yang memiliki peluang tertinggi diantara semua lintasan yang mungkin.



Gambar 1. Barisan *state* (lintasan) yang mungkin

untuk $1 \le i \le N$ dan $1 \le t \le T$

Untuk menyelesaikan *decoding problem* ini digunakan algoritma Viterbi.

Dalam algoritma Viterbi digunakan dua variabel bantu, yaitu:

1. $\delta_t(i) = \max\limits_{Q_1, Q_2, \ldots, Q_{t-1}} P\left(Q_1 = s_{i_1}, \ldots, Q_t = s_i, O_1 = v_{m_1}, \ldots, O_t = v_{m_t} \middle| \lambda\right)$

   Dengan induksi diperoleh:

   $\delta_{t+1}(j) = \max\limits_{1 \le i \le N} \left[\delta_t(i) \cdot a_{ij}\right] \cdot b_j(O_{t+1})$

2 $\psi_t(i) = \arg\max\limits_{Q_1, Q_2, \ldots, Q_{t-1}} P\left(Q_1 = s_{i_1}, \ldots, Q_t = s_i, O_1 = v_{m_1}, \ldots, O_t = v_{m_t} \middle| \lambda\right)$

Langkah-langkah dalam algoritma Viterbi untuk menentukan barisan *state* terbaik yaitu:

1. Inisialisasi

   $\delta_1(i) = \pi_i \cdot b_i(O_1)$, $\qquad 1 \le i \le N$

   (9)

$$\psi_1(i) = 0$$

(10)

2. Rekursi

$$\delta_t(j) = \max_{1 \le i \le N}\left[\delta_{t-1}(i)a_{ij}\right] \cdot b_j(O_t), \qquad 2 \le t \le T, \qquad 1 \le j \le N$$

(11)

$$\psi_t(j) = \arg\max_{1 \le i \le N}\left[\delta_{t-1}(i)a_{ij}\right], \qquad 2 \le t \le T, \qquad 1 \le j \le N$$

(12)

3. *State* terbaik pada waktu $T\ (Q_T)$

$$P^* = \max_{1 \le i \le N}\left[\delta_T(i)\right]$$

(13)

$$Q^*_T = \arg\max_{1 \le i \le N}\left[\delta_T(i)\right]$$

(14)

4. Barisan *state* terbaik pada $t = T-1, T-2, \ldots, 1$

$$Q^*_t = \psi_{t+1}(Q^*_{t+1}), \quad t = T-1, T-2, \ldots, 1 \tag{15}$$

c. Learning problem

Jika diberikan sebuah *HMM*, dan barisan observasi $O = O_1, O_2, \ldots, O_T$, bagaimana mengatur parameter model $\lambda = (\mathbf{A}, \mathbf{B}, \pi)$ agar $P(O|\lambda)$ maksimum.

Untuk menyelesaikan masalah ini digunakan Algoritma *Baum-Welch.* Dalam algoritma *Baum-Welch*, didefinisikan empat variabel, yaitu : variabel *forward*, variabel *backward*, variabel $\xi_t(i,j)$, dan variabel $\gamma_t(i)$. Variabel *forward* dan variabel *backward* akan digunakan dalam perhitungan variabel $\xi_t(i,j)$ dan variabel $\gamma_t(i)$.

Variabel pertama, variabel *forward* telah didefinisikan pada persamaan (5), serta tahapan induksi diberikan oleh persamaan (6) – (8).

Variabel kedua, yakni variabel *backward* didefinisikan sebagai,

$$\beta_t(i) = P(O_{t+1}, O_{t+2}, ..., O_T \mid Q_t = s_i, \lambda)$$

(16)

yaitu, peluang dari barisan observasi parsial $O_{t+1}, O_{t+2}, ..., O_T$, diberikan *state* s_i pada waktu t dan model λ. Selanjutnya, $\beta_t(i)$ dapat diselesaikan secara induksi sebagai berikut:

1. Inisialisasi

$$\beta_T(i) = 1, \qquad 1 \le i \le N$$

(17)

2. Induksi

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j), \;\; t = T-1, T-2, ..., 1, \; 1 \le i \le N$$

(18)

Variabel ketiga, yakni variabel $\xi_t(i, j)$ didefinisikan sebagai,

$$\xi_t(i, j) = P(Q_t = s_i, Q_{t+1} = s_j \mid O, \lambda)$$

(19)

yaitu, peluang proses pada saat t berada pada *state* s_i dan pada saat t + 1 berada pada state s_j, jika diberikan barisan O dan model λ.

Persamaan (19) dapat ditulis

$$\xi_t(i, j) = \frac{P(Q_t = s_i, Q_{t+1} = s_j, O \mid \lambda)}{P(O \mid \lambda)}$$

(20)

atau dapat diekspresikan dengan variabel *forward-backward* sebagai

$$\xi_t(i, j) = \frac{\alpha_t(i).\beta_{t+1}(j).b_j(O_{t+1}).a_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t(i).a_{ij}.b_j(O_{t+1}).\beta_{t+1}(j)}$$

(21)

Variabel keempat, yakni variabel $\gamma_t(i)$ didefinisikan sebagai:

$$\gamma_t(i) = P(Q_t = s_i \mid O, \lambda)$$

(22)

yaitu, peluang proses berada pada *state* s_i saat waktu *t*, jika diberikan barisan observasi O, dan model λ. Persamaan (22) dapat diekspresikan dengan variabel *forward-backward* sebagai berikut,

$$\gamma_t(i) = \frac{\alpha_t(i).\beta_t(i)}{\sum_{i=1}^{N} \alpha_t(i).\beta_t(i)}$$

(23)

Peluang proses berada pada *state* s_i saat *t*, jika diberikan barisan pengamatan *O*,

$$\gamma_t(i) = \sum_{j=1}^{N} \xi_t(i, j)$$

(24)

Sehingga,

$$\sum_{t=1}^{T-1} \gamma_t(i) = \text{Ekspektasi jumlah transisi dari } s_i$$

(25)

Sama halnya, jika $\xi_t(i, j)$ dijumlahkan atas indeks waktu t (dari $t = 1$ ke $t = T - 1$) dapat diartikan sebagai perkiraan banyaknya transisi dari *state* s_i ke *state* s_j. Jadi,

$$\sum_{t=1}^{T-1} \xi_t(i, j) = \text{Ekspektasi jumlah transisi dari } s_i \text{ ke } s_j \tag{26}$$

Dengan menggunakan persamaan (25) dan (26) di atas, maka diperoleh rumus re-estimasi sebagai berikut:

$$\begin{aligned} \overline{\pi}_i &= \text{Ekspektasi frekuensi dalam } \textit{state } s_i \text{ ketika } t = 1 \\ &= \gamma_1(i), \qquad\qquad 1 \le i \le N \end{aligned}$$

(27)

$$\bar{a}_{ij} = \frac{\text{Ekspektasi jumlah transisi dari } \textit{state } s_i \text{ ke } \textit{state } s_j}{\text{Ekspektasi jumlah transisi dari } \textit{state } s_i}$$

$$= \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \qquad 1 \le i \le N,\ 1 \le j \le N \tag{28}$$

$$\bar{b}_j(k) = \frac{\text{Ekspektasi lamanya dalam } \textit{state } j \text{ dan simbol pengamatannya } v_k}{\text{Ekspektasi lamanya dalam } \textit{state } j}$$

$$= \frac{\sum_{\substack{t=1 \\ \text{s.t } O_t = v_k}}^{T} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)}, \qquad 1 \le j \le N,\ 1 \le k \le M$$

(29)

Persamaan (27) sampai (29) dikenal dengan rumus re-estimasi, yang dapat memperbaharui (mengatur kembali) parameter model $\lambda = (\mathbf{A}, \mathbf{B}, \pi)$. Aspek yang sangat penting dari prosedur re-estimasi adalah batasan stokastik parameter *HMM* yang selalu dipenuhi pada setiap iterasi, yakni:

$$\sum_{i=1}^{N} \bar{\pi} = 1 \tag{30}$$

$$\sum_{j=1}^{N} \bar{a}_{ij} = 1 \quad ,\ 1 \le i \le N \tag{31}$$

$$\sum_{k=1}^{M} \bar{b}_j(k) = 1 \quad ,\ 1 \le j \ge N \tag{32}$$

Jika model awal adalah $\lambda = (A, B, \pi)$ dan proses dilaksanakan sehingga diperoleh taksiran parameter $\bar{\lambda} = (\bar{A}, \bar{B}, \bar{\pi})$. Maka dengan algoritma Baum Welch tadi diperoleh $\bar{\lambda}$ lebih “optimal” dibandingkan dengan λ dalam artian $P(O \mid \bar{\lambda}) > P(O \mid \lambda)$, yaitu diperoleh model baru $\bar{\lambda}$ sehingga barisan pengamatan lebih mirip untuk dihasilkan. Jika proses tersebut dilakukan berulang-ulang sampai dipenuhinya syarat tertentu maka peluang barisan pengamatan dapat di observasi dari model dapat ditingkatkan. Langkah yang dilakukan diatas adalah Algoritma Baum-Welch.

3.4 Studi Kasus

Misalkan seseorang dikurung di dalam ruangan tertutup selama beberapa hari, dan tidak mengetahui cuaca yang sedang terjadi di luar. Jika ia diminta menebak keadaan cuaca di luar, maka pengamatan yang ia lakukan adalah dengan melihat apakah orang yang datang ke ruangan tersebut membawa payung atau tidak. Asumsikan keadaan cuaca; cerah, hujan
dan berawan.

Andaikan diberikan matriks peluang transisi $\mathbf{A} = \left[a_{ij}\right] = \begin{bmatrix} 0,8 & 0,05 & 0,15 \\ 0,2 & 0,6 & 0,2 \\ 0,2 & 0,3 & 0,5 \end{bmatrix}$ dan

matriks peluang observasi $\mathbf{B} = \left[b_{jm}\right] = \begin{bmatrix} 0,1 & 0,9 \\ 0,8 & 0,2 \\ 0,3 & 0,7 \end{bmatrix}$.

Bagaimana barisan cuaca yang paling mungkin untuk tiga hari pertama ?

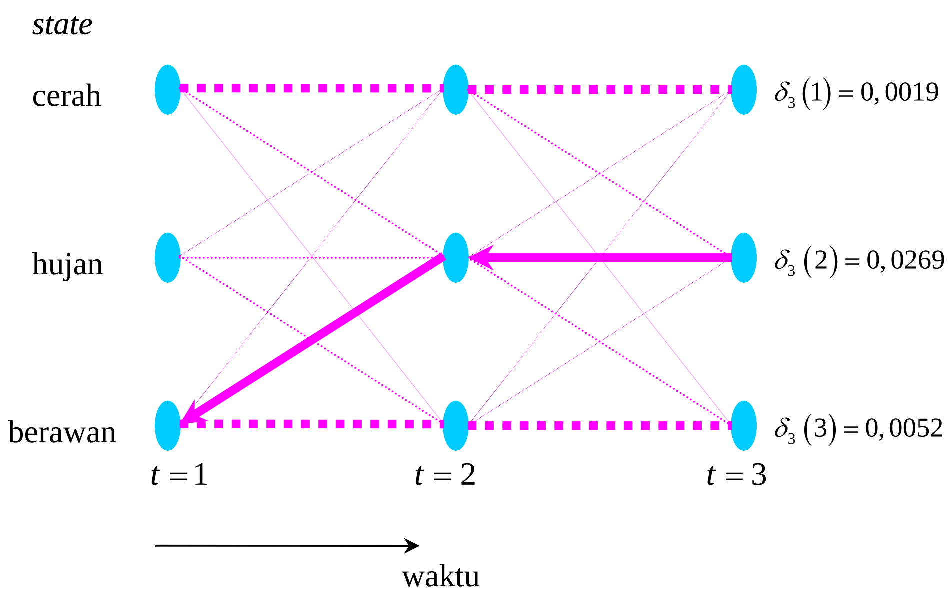
Dengan menggunakan algoritma Viterbi (matlab), diperoleh

Q(1) = 3

Q(2) = 2

Q(3) = 2

seperti digambarkan oleh gambar 2.



Gambar 2 Barisan *state* terbaik untuk kasus 1

Artinya, barisan cuaca yang paling mungkin untuk tiga hari berturut-turut adalah

berawan, hujan, hujan.

4. Kesimpulan dan saran

Dari pembahasan di atas, dapat disimpulkan:

1. *Hidden Markov Model* merupakan perluasan dari Rantai Markov Waktu diskrit, dimana state tidak dapat diamati secara langsung.
2. Untuk menyelesaikan *evaluation problem* dalam *HMM,* yakni mencari $P(O|\lambda)$ atau peluang dari barisan observasi, dapat digunakan algoritma *forward*.
3. Untuk menyelesaikan *decoding problem* pada *HMM*, yakni mencari barisan *state* yang optimal $Q^* = \{Q^*_1, Q^*_2, ..., Q^*_T\}$ jika diberikan barisan observasi $O = \{O_1, O_2, ..., O_T\}$ dan model $\lambda = (\mathbf{A}, \mathbf{B}, \pi)$ dapat menggunakan algoritma Viterbi.

4. Untuk menyelesaikan *learning problem* pada *HMM* yakni mengatur parameter model $\lambda = (\mathbf{A}, \mathbf{B}, \pi)$ sehingga $P(O \mid \lambda)$ maksimum, dapat menggunakan algoritma *Baum-Welch*.
5. Konsep *HMM* dapat digunakan dalam berbagai masalah nyata, dalam hal ingin mengetahui peluang suatu proses berada dalam keadaan tertentu, sementara statenya tidak dapat diamati secara langsung, seperti, kasus cuaca dengan orang terkurung, pengenalan pola, pengenalan suara, pengenalan DNA, dan lain-lain.
6. Dapat dilakukan studi lanjut *HMM* untuk waktu kontinu.

5. Daftar Pustaka

[1] Abdulla,W.H. and Kasabov,N.K., 1999. "*The Concepts of Hidden Markov Model in Speech Recognition*",Technical Report TR99/09,New Zealand.

[2] Barbara Resch. "*Hidden Markov Models*" , A tutorial for Course Computational Intelligence. http://www.igi.tugraz.at/lehre/CI. diakses 20 Januari 2006.

[3] Jedlik.phy.bme.hu/~gerjanos/HMM/node4.html

[4] Osaki,Sunji., 1992. "*Applied Stochastic System Modeling*". Springer Verlag,New york.

[5] Rabiner,L.R., 1989. "*A tutorial on Hidden Markov Models and Selected Aplications in Speech Recognition*",Proceedings of The IEEE,Vol.77,no.2. pp.257-286.

[6] Ross,S.M.,1996. "*Stochastic Processes*", second edition, John Wiley and Sons,New York.

[7] Young,S. etc., 2002. "*HMM Toolkit (HTK) Book* (for version 3.2.1)". Cambridge University Engineering Departement.