

Teori Dasar *Hidden Markov Model*

Muhammad Eko Budi Prasetyo 18209020¹

Program Studi Sistem dan teknologi Informasi

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia

¹18209020@std.stei.itb.ac.id

Abstract— *Hidden Markov Model (HMM)* adalah subset masalah dari metode statistika *Markov Model*. HMM cukup populer digunakan dalam berbagai bidang kecerdasan buatan dan bioinformatika. Hal ini diakibatkan oleh dua alasan kuat. Pertama, model ini sangat kaya dalam struktur matematika sehingga dapat digunakan sebagai teori dasar dibanyak bidang aplikasi. Kedua, jika model ini digunakan dengan benar, ia dapat bekerja dengan efektif dan efisien untuk implementasi beberapa aplikasi penting.

Index Terms— *Markov Model, Hidden Markov Model, Stokastik, Model Statistik.*

I. PENDAHULUAN

HMM merupakan pengembangan model statistik dari model Markov. Model ini dikembangkan pertama kali oleh Andreyevich Markov (Rusia: Андрей Андреевич Марков), seorang ilmuwan Rusia pada awal abad 20. Pada awalnya, model ini merupakan murni model teoritis namun mulai dikembangkan oleh kalangan akademisi dan engineer dalam paper-paper internasional seperti pada [1] – [4]. Aplikasi awal HMM adalah untuk program pengenalan suara (*speech recognizing*) yang dimulai pada pertengahan 1970. Pada pertengahan 1980, HMM mulai digunakan untuk analisis rangkaian biologis terutama pada DNA. Setelah itu, HMM menjadi salah satu model penting yang digunakan dalam bidang bioinformatika.

HMM telah dipelajari secara luas dalam bidang statistik. Model ini dipandang sebagai proses *bivariate parametric* dalam waktu diskrit. proses yang terjadi dalam HMM merupakan *finite-state* yang homogen dari *Markov Model* dan tidak dapat diamati. Proses kedua merupakan aliran variabel acak kondisional yang diberikan oleh *Markov Model*. Pada saat apapun, distribusi untuk setiap variabel acak dipengaruhi oleh nilai *Markov Model* pada waktu tersebut saja. Oleh karena itu, HMM merupakan bagian dari statistik prametri [5].

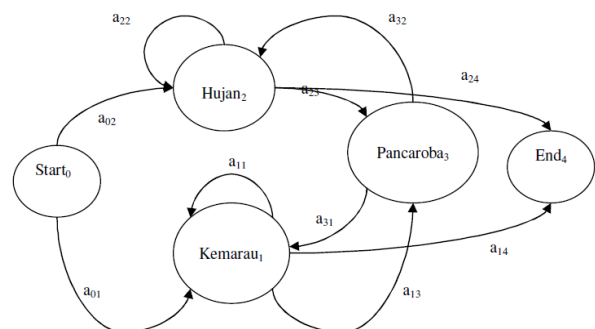
Tujuan dari pembuatan makalah ini adalah untuk menjelaskan teori-teori dasar HMM dan contoh-contoh aplikasinya. Untuk itu, makalah ini dibagi kedalam beberapa sub-bab. Pembahasan dimulai dari Bab II yang menjelaskan tentang *Markov Model*. Bab ini difokuskan kepada penjelasan superset dari HMM, Markov Model. Mulai dari definisi, contoh diagram, batasan-batasan,

sampai contoh kasus. Bab III menjelaskan tentang HMM. Pada bab ini, dijelaskan definisi dari HMM, sifat-sifatnya, contoh diagram, dan contoh kasus HMM. Bab IV menjelaskan tentang penyelesaian masalah menggunakan HMM. Pada bab ini dijelaskan mengenai evaluasi, inferensi, *learning*, dan implementasi HMM. Terakhir, pembahasan diakhiri dengan Bab V Kesimpulan.

II. MARKOV MODEL

Markov Model biasa disebut sebagai *Markov Chain* atau *Markov Process*. Model ini ditemukan oleh Andrey Markov dan merupakan bagian dari proses stokastik yang memiliki properti Markov. Dengan memiliki properti tersebut berarti, apabila diberikan inputan keadaan saat ini, keadaan akan datang dapat diprediksi dan ia lepas dari keadaan di masa lampau. Artinya, deskripsi kondisi saat ini menangkap semua informasi yang mempengaruhi evolusi dari suatu sistem dimasa depan. Dengan kata lain, Kondisi masa depan dituju dengan menggunakan probabilitas bukan dengan determinitas.

Model ini merupakan bagian dari finite state atau finite automaton. *Finite automaton* sendiri adalah kumpulan state yang transisi antar state-nya dilakukan berdasarkan masukan observasi. Pada *Markov Chain*, setiap busur antar state berisi probabilitas yang mengindikasikan kemungkinan jalur tersebut akan diambil. Jumlah probabilitas semua busur yang keluar dari sebuah simpul adalah satu. Gambar dibawah memperlihatkan contoh *Markov Chain* yang menggambarkan kondisi cuaca.



Gambar 1. Markov Chain yang menggambarkan perubahan cuaca

Pada Gambar 1, a_{ij} adalah probabilitas transisi dari state i ke state j . Misalkan, dari simpul $start_0$ keluar dua kemungkinan, a_{02} dan a_{01} . Maka jumlah probabilitas $a_{01} + a_{02}$ adalah satu. Hal ini juga berlaku bagi simpul-simpul yang lain. Markov Chain bermanfaat untuk menghitung probabilitas suatu kejadian teramati yang secara umum dapat dirumuskan sebagai berikut:

$$P(\sigma_t) = P(\sigma_t | \sigma_{t-1}, \sigma_{t-2}, \sigma_{t-3}, \dots)$$

σ_1 adalah kondisi saat ini, dan σ_t adalah kondisi pada waktu tertentu yang berhubungan dengan σ_1 . Sedangkan σ_{t-1} adalah kondisi sebelum σ_t . Kemudian kita dapat berasumsi bahwa sebelah kanan persamaan bersifat invariant, yaitu, dihipotesiskan dalam keseluruhan sistem, transisi diantara keadaan tertentu tetap sama dalam hubungan probabilitiknya.

Berdasarkan asumsi tersebut, kita dapat membuat suatu set keadaan probabilitas a_{ij} diantara dua keadaan S_i dan S_j :

$$a_{ij} = p(\sigma_t = s_i | \sigma_{t-1} = s_j), \quad 1 \leq i, j \leq N$$

Karena i dan j dapat sama, maka berlaku batasan berikut:

$$a_{ij} \geq 0, \text{ and } \sum_{i=1}^N a_{ij} = 1$$

Contoh permasalahan:

Cuaca dalam satu hari dimodelkan ke dalam tiga status: hujan (1), berawan (2) dan cerah (3). Aturan probabilitas dari setiap transisi status dideskripsikan sebagai berikut:

$$A = \{a_{ij}\} = \begin{matrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{matrix}$$

Misal, probabilitas cuaca untuk 8 hari berturut-turut "cerah-cerah-cerah-hujan-hujan-cerah-berawan-cerah".

O adalah tahap pengamatan dengan:

$O = \{\text{cerah, cerah, cerah, hujan, hujan, cerah, berawan, cerah}\} = \{3,3,3,1,1,3,2,3\}$

Maka pemodelan peluangnya dalam *Markov Model*:

$$\begin{aligned} P(O|Model) &= P[3,3,3,1,1,3,2,3|Model] \\ &= P[3] P[3|3]^2 P[1|3] P[1|1] \\ &\quad P[3|1] P[2|3] P[3|2] \\ &= \Pi_3 \cdot (a_{33})^2 a_{31} a_{11} a_{13} \\ &\quad a_{32} a_{23} \\ &= (1.0)(0.8)^2(0.1)(0.4)(0.3) \\ &\quad (0.1)(0.2) \\ &= 1.536 \times 10^{-4} \end{aligned}$$

Dari penjelasan diatas, dapat disimpulkan bahwa Markov Chain bermanfaat untuk menghitung probabilitas urutan kejadian yang dapat diamati. Masalahnya, terkadang ada urutan kejadian yang ingin diketahui tetapi tidak dapat diamati. Untuk menyelesaikan kasus tersebut, dikembangkanlah model baru yang memodelkan kejadian yang tersembunyi, *Hidden Markov Model*.

III. HIDDEN MARKOV MODEL

HMM merupakan model statistik dimana suatu sistem yang dimodelkan diasumsikan sebagai markov proses dengan kondisi yang tidak terobservasi. Suatu HMM dapat dianggap sebagai jaringan Bayesian dinamis yang sederhana (simplest dynamic Bayesian network).

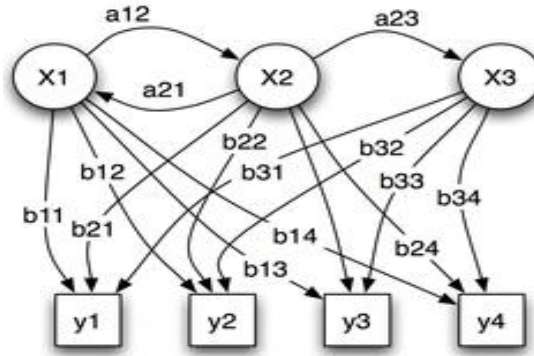
Dalam *Markov Model* biasa, setiap keadaan dapat terlihat langsung oleh pengamat. Oleh karena itu, kemungkinan dari transisi antar kondisi menjadi satu-satunya parameter teramati. Dalam HMM, keadaan tidak terlihat secara langsung. Tetapi output yang bergantung terhadap keadaan tersebut terlihat.

Setiap kondisi memiliki distribusi kemungkinan disetiap output yang mungkin. Oleh karena itu, urutan langkah yang dibuat oleh HMM memberikan suatu informasi tentang urutan dari keadaan. Perlu dipahami, bahwa sifat *hidden* 'tersembunyi' menunjuk kepada kondisi langkah yang dilewati model, bukan kepada parameter dari model tersebut. Walaupun parameter model diketahui, model tersebut tetap tersembunyi.

Hidden Markov Model dapat digunakan untuk aplikasi dibidang *temporal pattern recognition* 'pengenalan pola temporal' seperti pengenalan suara, tulisan, gestur, bioinformatika, kompresi kalimat, *computer vision*, ekonomi, finansial, dan pengenalan not balok.

HMM adalah variasi dari *finite state machine* yang memiliki kondisi tersembunyi Q , suatu nilai output O (observasi), kemungkinan transisi A , kemungkinan output B , sebuah kondisi awal Π . Kondisi saat ini tidak terobservasi. Tetapi, setiap keadaan menghasilkan output kemungkinan B . biasanya, Q dan O dimengerti, jadi HMM disebut triple (A, B, Π) .

1. Himpunan observed state: $O = o_1, o_2, \dots, o_N$.
2. Himpunan hidden state: $Q = q_1, q_2, \dots, q_N$.
3. Probabilitas transisi: $A = a_{01}, a_{02}, \dots, a_{n1}, \dots, a_{nm}$; a_{ij} adalah probabilitas untuk pindah dari state i ke state j .
4. Probabilitas emisi atau observation likelihood: $B = b_i(O_t)$, merupakan probabilitas observasi O_t dibangkitkan oleh state i .
5. State awal dan akhir: q_0, q_{end} , yang tidak terkait dengan observasi.

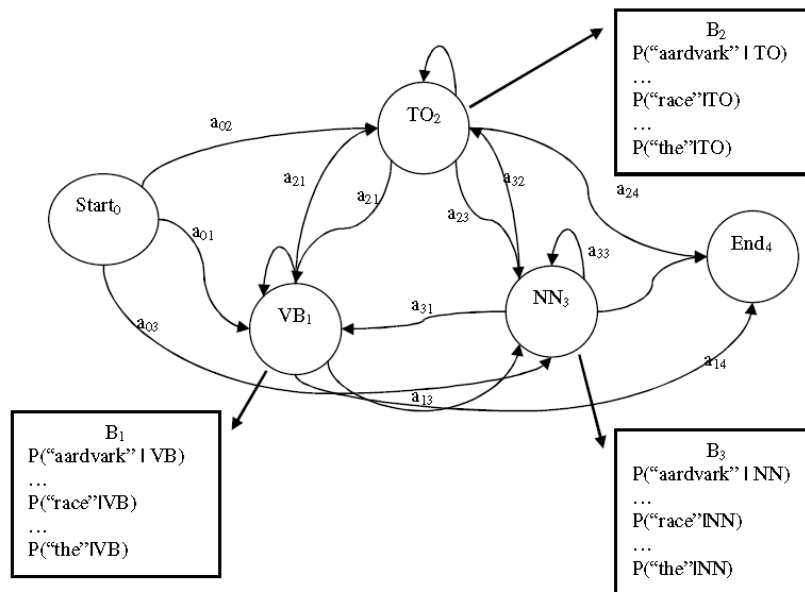


Gambar 2. Representasi Parameter HMM

Penjelasan Gambar 2:

- x = kondisi
- y = observasi yang mungkin
- a = kemungkinan keadaan transisi
- b = kemungkinan output

contoh kasus HMM:



Gambar 3. Contoh HMM untuk part of speech tagger [7]

Pada gambar 3, kata yang dicari menggunakan metode HMM adalah "aardvark", "race", dan "the". Ketiga kata tersebut menjadi *observed state*. Sedangkan *hidden state* adalah "TO" (*to infinitive*), "VB" (*verb base form*), dan "NN" (*mass noun*). B_i himpunan semua probabilitas emisi untuk *hidden state* i , sedangkan a_{ij} adalah probabilitas transisi dari *state* i ke *state* j [12].

IV. PENYELESAIAN MASALAH DENGAN HMM

Ada tiga permasalahan khusus yang dapat diselesaikan oleh metode *Hidden Markov Model*. Diantaranya :

1. Evaluation (evaluasi)
2. Inference (penarikan kesimpulan)
3. Learning (pembelajaran)

Karakter masalah seperti di ataslah yang mampu diselesaikan oleh metode *Hidden Markov Model*.

A. Evaluasi

Pengertian dari operasi evaluasi dalam *Hidden Markov Model* adalah perhitungan probabilitas dari urutan nilai observasi yang di berikan oleh *Hidden Markov Model*. Masalah ini dapat diselesaikan dengan Algoritma Forward dan Backward.

Adapun perbedaan algoritma Forward dan Backward adalah pada runutan observasi dari nilai probabilitasnya. Algoritma Forward seperti namanya bergerak maju, dengan observasi sesuai dengan urutan, dan sebaliknya Algoritma Backward bergerak secara mundur.

Nilai probabilitas pada setiap langkah observasi digunakan dalam perhitungan selanjutnya.

B. Inferensi

Pengertian dari operasi inference dalam *Hidden Markov Model* adalah penarikan kesimpulan berdasarkan asumsi yang diperoleh dari nilai probabilitas observasi yang didapat sebelumnya pada operasi evaluasi. Operasi ini juga sering kali digunakan untuk mencari nilai optimum. Masalah ini dapat diselesaikan dengan Algoritma Viterbi.

Algoritma Viterbi adalah sebuah algoritma pemrograman dinamis yang berfungsi untuk menemukan urutan statement yang disembunyikan. Dalam hal ini Algoritma Viterbi erat kaitannya dengan Algoritma Forward.

C. Learning

Pengertian dari operasi learning dalam *Hidden Markov Model* adalah melatih parameter HMM jika diberikan dataset barisan-barisan tertentu agar dapat menemukan himpunan transisi state yang paling mungkin beserta probabilitas outputnya.

Untuk menyelesaikan permasalahan *learning*, digunakan algoritma Baum-Welch. Algoritma ini secara umum berfungsi untuk menentukan nilai harapan dan maksimalisasi.

Algoritma ini mempunyai dua langkah dalam penyelesaian masalah, yaitu:

1. Menghitung nilai probabilitas forward dan backward untuk setiap statement.
2. Menentukan frekuensi dari pasangan transisi emisi dan membaginya dengan nilai

probabilitas semua observasi.

D. Implementasi HMM

Kompresi Kalimat

Kompresi kalimat adalah pemilihan kata atau frasa yang penting dalam satu kalimat dan menyusun ulang kata tersebut menjadi kalimat yang lebih ringkas. Sebaliknya, kompresi kalimat juga dapat dipandang sebagai proses pembuangan kata atau frasa yang tidak penting. Kompresi kalimat erat kaitannya dengan peringkasan dokumen, keduanya memiliki tujuan yang sama yaitu menghasilkan teks yang lebih pendek tanpa kehilangan informasi penting. Perbedaannya terletak pada masukan yang diberikan. Peringkasan dokumen menerima masukan berupa teks dokumen sedangkan kompresi kalimat berupa kalimat [12].

Speech Recognition

Pengenalan ucapan atau pengenalan wicara—dalam istilah bahasa Inggrisnya, *automatic speech recognition (ASR)*—adalah suatu sistem yang memungkinkan komputer untuk menerima masukan berupa kata yang diucapkan. Teknologi ini memungkinkan suatu perangkat untuk mengenali dan memahami kata-kata yang diucapkan dengan cara digitalisasi kata dan mencocokkan sinyal digital tersebut dengan suatu pola tertentu yang tersimpan dalam suatu perangkat. Kata-kata yang diucapkan diubah bentuknya menjadi sinyal digital dengan cara mengubah gelombang suara menjadi sekumpulan angka yang kemudian disesuaikan dengan kode-kode tertentu untuk mengidentifikasi kata-kata tersebut. Hasil dari identifikasi kata yang diucapkan dapat ditampilkan dalam bentuk tulisan atau dapat dibaca oleh perangkat teknologi sebagai sebuah komando untuk melakukan suatu pekerjaan, misalnya penekanan tombol pada telepon genggam yang dilakukan secara otomatis dengan komando suara.

Face Recognition

Merupakan aplikasi komputer untuk identifikasi identitas seseorang menggunakan gambar digital atau video digital. Identifikasi tersebut dimungkinkan dengan membandingkan ekspresi wajah dari gambar atau video dan membandingkannya dengan basisdata dalam sistem. Sistem ini biasa digunakan dalam untuk autentifikasi pengamanan secara digital.

Penerapan di bidang bioinformatika

Dalam bidang bioinforamtika, HMM dapat digunakan untuk:

- prediksi daerah produsen protein dalam barisan genome
- pemodelan famili DNA terkait atau barisan protein
- prediksi elemen struktur sekunder dari barisan protein primer

Handwriting Recognition

Merupakan suatu sistem yang memungkinkan komputer untuk menerima dan menginterpretasi input tulisan tangan dari sumber-sumber seperti kertas, foto, *touch-screen*, dan *device* lainnya. Proses interpretasi biasanya dilakukan dengan mengenali huruf-huruf yang terdapat dalam input tulisan tangan dan mengubahnya menjadi huruf digital.

Gesture Recognition

Merupakan suatu sistem yang memungkinkan komputer untuk dapat menginterpretasi gesture gerakan dari manusia. Biasanya, gesture tersebut berasal dari gerakan badan atau sentuhan jari. Setelah gesture dapat dibaca, komputer akan menginterpretasinya menjadi sinyal input yang dapat digunakan sebagai tanda terjadinya suatu kejadian tertentu dalam komputer. Sebagai contoh, monitor dengan layar sentuh. Pengembangan terbaru dari sistem ini adalah kemampuan untuk mengenali emosi seseorang dari mimik wajahnya.

Part-Of-Speech Tagging

Adalah suatu sistem yang memungkinkan komputer untuk menandai kata pada suatu teks kedalam suatu kelas kata tertentu berdasarkan definisi dan maknanya. Dengan kata lain, komputer dapat membedakan hubungan kata yang mendampingi atau yang terkait dengan kata yang ditandai pada suatu frasa, kalimat, atau paragraf. Bentuk sederhana proses ini adalah identifikasi kata sebagai nomina, verba, adjektiva, adverbial, dll.

Musical Score Following

Adalah suatu sistem yang memungkinkan komputer untuk secara otomatis mendengarkan pagelaran musik dan mengikuti pergerakannya dalam partitur. Sistem ini merupakan interseksi dari kecerdasan buatan, pengenalan pola, pemroses sinyal, dan musikologi.

Selain implementasi yang disebutkan diatas, masih banyak contoh implementasi dari *Hidden Markov Model* yang bervariasi, sesuai dengan karakteristik dan operasi operasinya.

V. KESIMPULAN

Dalam makalah ini kami membahas teori dasar dan implementasi algoritma dari *Hidden Markov Model*. Pembahasan ini meliputi teori dasar *Markov Model* dan *Hidden Markov Model* beserta aplikasi algoritma Forward-Backward, Viterbi, dan Baum-Welch. Pembahasan dalam makalah ini terutama berupa penjelasan dasar dari teori HMM dan penjabaran implementasi HMM dalam bidang teknologi informasi.

REFERENCES

- [1] L.E Baum dan T. Petrie, "Statistical Inference for Probabilistic function of finite state Markov Chains," *Ann. Math. Stat.*, vol. 37, pp. 1554—1536, 1966.

- [2] L. E. Baum and J. A. Egon, "An Inequality with application to statistic estimation for probabilistic function of Markov Process and to a model for Ecology," *Bull. Amer. Meteorol. Soc.*, vol. 73, pp 360-363, 1967.
- [3] L. E. Baum dan G. R. Sell, "Growth functions for Transformation on Manifolds," *Pac. J. Math.*, vol. 27, no.2, pp. 211-227, 1968.
- [4] L. E. Baum, T. Petrie, G. Soules, dan N. Weiss, "A Maximization Technique Occurring in Statistical Analysis of Probabilistic function of Markov Chains," *Ann. Math. Stat.*, vol. 41, no. 1, pp. 164-171, 1970.
- [5] Ephraim Y, Merhav N (June 2002). "Hidden Markov processes". *IEEE Trans. Inform. Theory* 48: 1518-1569.
- [6] For a more advanced approach cf: *Markov Processes and Semi-groups*, Ch. X, § 8, Vol II *Introduction to Probability Theory and Its Applications* (2nd edition), William Feller, Wiley 1971, LCCCN 57-10805, ISBN 0-471-25709-5
- [7] Jurafsky, D., Martin, J.H. (2006) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*.
- [8] Kristie Seymore, Andrew McCallum, and Roni Rosenfeld. *Learning Hidden Markov Model Structure for Information Extraction*. AAAI 99 Workshop on Machine Learning for Information Extraction, 1999.
- [9] Lawrence R. Rabiner (February 1989). "A tutorial on *Hidden Markov Models* and selected applications in speech recognition.
- [10] Munir, Rinaldi.2006. "Strategi Algoritmik". Departemen Teknik Informatika, Institut Teknologi Bandung
- [11] Olivier Cappé, Eric Moulines, Tobias Rydén (2005). *Inference in Hidden Markov Models*. Springer.
- [12] Wibisono, Y. (2008). *Penggunaan Hidden Markov Model untuk Kompresi Kalimat*. Bandung: Institut Teknologi Bandung.

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 29 April 2010

ttd

Muhammad Eko Bui Prasetyo