

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Perkembangan ilmu pengetahuan dan teknologi yang sangat pesat, menjadikan statistika memegang peranan penting dalam kehidupan. Hampir semua fenomena yang terjadi di alam ini bersifat probabilistik, sehingga statistika dapat digunakan untuk membantu kegiatan-kegiatan manusia.

Salah satu bagian penting dalam statistika adalah rantai Markov. Analisis rantai Markov adalah suatu teknik probabilitas yang menganalisis pergerakan probabilitas dari suatu kondisi ke kondisi lainnya. Rantai Markov baru diperkenalkan sekitar tahun 1906, oleh seorang matematikawan Rusia Andrei A. Markov (1856-1922). Rantai Markov merupakan suatu bentuk khusus dari model probabilistik yang lebih umum yaitu proses stokastik atau proses perubahan-perubahan probabilistik yang bersifat terus menerus.

Rantai Markov adalah sebuah proses Markov dengan ruang parameter yang diskrit (dapat dihitung) yang berada pada suatu ruang keadaan yang diskrit (Ross, 1996). Analisis Markov hampir sama dengan *decision analysis*, bedanya adalah analisis rantai Markov tidak memberikan keputusan rekomendasi, melainkan hanya informasi probabilitas mengenai situasi keputusan yang dapat

membantu pengambil keputusan. Dengan demikian, analisis rantai Markov bukanlah teknik optimisasi, tetapi adalah teknik deskriptif yang menghasilkan informasi probabilitas di masa mendatang.

Untuk dapat menerapkan analisis rantai Markov ke dalam suatu kasus, ada beberapa syarat yang harus dipenuhi :

1. Jumlah probabilitas transisi untuk suatu keadaan awal dari sistem bernilai satu.
2. Probabilitas-probabilitas tersebut berlaku untuk semua partisipan dalam sistem.
3. Probabilitas transisi konstan sepanjang waktu.
4. Keadaannya independen terhadap waktu.

Rantai Markov juga merupakan suatu proses random (proses stokastik) di mana keadaan (*state*) saat ini, keadaan (*state*) yang akan datang bersifat independen terhadap keadaan (*state*) yang lampau dan hanya tergantung pada keadaan yang terdekat sebelumnya. Sifat ini dinamakan Markov *Property* dan dapat dirumuskan sebagai berikut:

$$P\{X_{n+1} = j | X_n = i, X_n = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0\} = P_{ij}$$

untuk semua keadaan  $i_0, i_1, \dots, i_{n-1}, i, j$  dan  $n \geq 0$ , di mana nilai yang memungkinkan dari  $i$  adalah suatu himpunan terbatas yang sering dinamakan sebagai Ruang keadaan (Ibe, 2009). Probabilitas di atas umumnya dinamakan dengan nama *state transition probability* yang sering dilambangkan dengan simbol  $P_{ij}$  di mana nilai  $P_{ij} \geq 0, \forall i, j$  dan nilai  $\sum_{j=0}^{\infty} P_{ij} = 1, \forall i = 0, 1, \dots$

Penerapan rantai Markov mula-mula digunakan untuk menganalisis dan memperkirakan perilaku partikel-partikel gas di ruang tertutup serta meramalkan keadaan cuaca. Sekarang ini rantai Markov juga telah banyak digunakan untuk menganalisis tentang perpindahan merek (*brands switching*) dalam pemasaran (Pramuditya, 2010), perhitungan rekening, jasa penyewaan mobil, perencanaan penjualan, masalah-masalah persediaan, pemeliharaan mesin, antrian, perubahan harga pasar saham, penentuan premi asuransi kendaraan, dan administrasi rumah sakit. Hongki (2004) menggunakan rantai Markov untuk penelusuran sifat dengan menghitung peluang kromosom seorang anak berasal dari leluhur ayah (*grandparental*) atau dari leluhur ibu (*grandmathernal*). Penelusuran sifat ini berguna dalam mengurangi kemungkinan penyakit bawaan pada seorang anak.

Meskipun aplikasi rantai Markov telah digunakan secara luas dalam kehidupan nyata, namun pada kenyataannya masih ada beberapa kekurangan dalam penerapan rantai Markov. Contohnya pada kasus penelusuran sifat yang dilakukan oleh Hongki (2004), penelusuran sifat hanya dilakukan secara fisik, struktur dan susunan DNA anak maupun leluhurnya tidak dipelajari dengan mendalam. Padahal diperlukan informasi apakah suatu gen berasal dari suatu mutasi atau bukan, karena jika terjadi mutasi pada gen tersebut, maka susunan DNA-nya akan berubah. Perubahan susunan DNA itu sendiri nantinya akan mengubah sebagian dari sifat atau karakter individu yang dikendalikan oleh gen tersebut.

Kelemahan-kelemahan tersebut pada akhirnya menjadi pemicu pengembangan rantai Markov. Para ahli matematika terus melakukan

pengembangan pada rantai Markov, salah satunya adalah *Hidden Markov Models*(HMM). HMM merupakan salah satu model stokastik yang banyak menarik perhatian akhir-akhir ini.

Teori dasar HMM telah dipublikasikan oleh Baum dan Petrie pada awal tahun 1970-an. Tetapi perkembangan pesat dari HMM baru terjadi beberapa tahun terakhir ini. Metode statistik HMM semakin populer pada dekade terakhir ini karena model tersebut kaya akan struktur matematika dan mengandung teori dasar yang bisa digunakan untuk beberapa aplikasi yang penting. Pada awalnya, penerapan dari HMM meliputi pengenalan ucapan (*speech recognition*) seperti yang dilakukan Rabiner (1989). Kemudian Gupta (2004) mulai mengaplikasikan HMM pada bidang bioinformatika. Sementara Rahma (2007) menerapkan HMM pada masalah *Multiple Sequence Alignment*, yaitu suatu masalah pengenalan dan perbandingan pola *sequence* protein. Selain pada bidang-bidang tersebut, HMM juga telah banyak dimanfaatkan dalam bidang lain seperti analisis sistem jaringan *wireless*, data akustik, *human motion* dan lain-lain.

Dalam sebuah model Markov, semua *state* dalam suatu *sequence* yang linear dapat secara langsung diobservasi. Dalam situasi tertentu, beberapa faktor yang tidak dapat diobservasi mempengaruhi perhitungan perpindahan atau transisi *state*. Faktor-faktor tersebut dinamakan dengan *hidden states* (keadaan-keadaan tersembunyi). Untuk memasukkan *hidden states* dalam proses perhitungan, dibutuhkan model yang lebih “pintar”, yaitu *Hidden Markov Model* (HMM). Sebuah HMM, menggabungkan dua atau lebih rantai Markov dengan hanya satu

rantai yang terdiri dari *state* yang dapat diobservasi dan rantai lainnya yang membentuk *state* yang tidak dapat diobservasi (*hidden*).

Dalam sebuah HMM, seperti halnya dalam sebuah rantai Markov, probabilitas perpindahan dari suatu *state* ke *state* lainnya dinamakan *transition probability*. Setiap *state* mungkin dibentuk oleh sejumlah elemen atau simbol. Untuk *sequence* nukleotida, terdapat empat buah kemungkinan simbol (A, T, G, dan C) dalam setiap *state*. Untuk *sequence* asam amino, terdapat dua puluh buah simbol. Nilai probabilitas yang berasosiasi dengan setiap simbol dalam *state* dinamakan *emission probability*. Untuk menghitung probabilitas total dari suatu jalur dalam model, baik *transition probability* ataupun *emission probability* yang menghubungkan semua keadaan terobservasi dan keadaan yang tidak terobservasi harus dimasukkan ke dalam perhitungan.

HMM didefinisikan sebagai kumpulan lima parameter ( $N, M, A, B, \pi$ ). Jika dianggap  $\lambda = \{A, B, \pi\}$  maka HMM mempunyai parameter tertentu di mana:

$N$  = banyaknya keadaan yang tidak terobservasi

$M$  = banyaknya keadaan terobservasi

$A$  = matriks peluang transisi

$B$  = distribusi peluang observasi pada saat  $t$  berada pada keadaan  $i$  (matriks emisi)

$\pi$  = keadaan awal.

Ada tiga masalah utama dalam *Hidden Markov Models*, yaitu:

- 1) Menghitung peluang terjadinya suatu barisan observasi.
- 2) Menentukan keadaan tersembunyi dari suatu barisan observasi.
- 3) Penaksiran parameter-parameter HMM.

Untuk memecahkan masalah pertama dapat digunakan algoritma Maju-Mundur. Masalah kedua dipecahkan dengan menggunakan algoritma Viterbi. Sedangkan masalah ketiga dipecahkan dengan algoritma Baum-welch.

Pada karya tulis ini, akan dibahas mengenai penyelesaian masalah kedua pada HMM, yaitu penentuan barisan keadaan tersembunyi (*hidden states*) yang paling optimal pada suatu HMM dengan menggunakan algoritma Viterbi.

Oleh karena itu, tugas akhir ini diberi judul “**APLIKASI ALGORITMA VITERBI UNTUK MENENTUKAN BARISAN KEADAAN TERSEMBUNYI PADA *HIDDEN MARKOV MODELS***”.

## 1.2 Rumusan Masalah

Permasalahan yang akan dikaji dalam tugas akhir ini adalah sebagai berikut :

1. Asumsi- asumsi apa saja yang diperlukan dalam HMM?
2. Permasalahan utama apa saja yang ada pada HMM?
3. Bagaimana cara menentukan barisan keadaan tersembunyi dari suatu barisan observasi yang diketahui pada suatu HMM?

4. Bagaimana akurasi algoritma Viterbi dalam meramalkan pergerakan saham PT. Astra *International* Indonesia, melalui pergerakan nilai tukar rupiah?

### 1.3 Tujuan Penulisan

Berdasarkan rumusan masalah di atas, maka tujuan pembuatan tugas akhir ini adalah sebagai berikut :

1. Mengetahui asumsi-asumsi apa saja yang diperlukan dalam HMM.
2. Mengetahui permasalahan-permasalahan utama pada HMM.
3. Mengetahui cara menentukan barisan keadaan tersembunyi dari suatu barisan observasi yang diketahui pada HMM.
4. Mengetahui akurasi algoritma Viterbi dalam meramalkan pergerakan saham PT. Astra *International* Indonesia, melalui pergerakan nilai tukar rupiah.

### 1.4 Manfaat Penulisan

*Hidden Markov Models* (HMM) adalah suatu metode penyelesaian masalah secara efektif dan efisien di bidang-bidang yang melibatkan HMM seperti asuransi, kedokteran, elektro dan bioinformatika. Melalui tugas akhir ini, akan diperoleh hasil peramalan *hidden states* melalui suatu barisan keadaan yang terobservasi. Selain itu, penulisan karya tulis ini diharapkan dapat menambah wawasan dan menambah literatur dalam bidang Statistika, khususnya tentang HMM.