

# Aplikasi Hidden Markov Model dalam Prediksi Harga Saham di Indonesia

Arfian Alimansyah<sup>1,a)</sup>, Acep Purqon<sup>2,b)</sup>

<sup>1</sup>Program Magister Sains Komputasi,  
Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Bandung,  
Jl. Ganesha no. 10 Bandung, Indonesia, 40132

<sup>2</sup>Kelompok Keilmuan Fisika Bumi dan Sistem Kompleks,  
Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Bandung,  
Jl. Ganesha no. 10 Bandung, Indonesia, 40132

<sup>a)</sup> arvianlimansyah@yahoo.com (corresponding author)

<sup>b)</sup> acep@fi.itb.ac.id

## Abstrak

*Hari ini, saham merupakan sebuah komoditas yang semakin sering ditransaksikan. Hal ini dapat dilihat dari peningkatan jumlah saham yang ditransaksikan maupun jumlah trader saham. Dengan alasan ini, maka diperlukan analisis saham yang dapat dipercaya. Salah satu cara untuk melakukan analisis saham adalah dengan melakukan pendekatan secara stokastik kepada dinamika perubahan nilai saham tersebut. Dalam sebuah sistem stokastik, pergerakan data tidak mengikuti sebuah persamaan tertentu sehingga untuk memperkirakan pergerakannya akan digunakan sistem probabilitas. Terdapat beberapa cara untuk memperkirakan pergerakan stokastik dan Hidden Markov Model (HMM) merupakan salah satunya. Pada penelitian ini, akan dilakukan simulasi HMM kepada 5 saham bluechip yang ada di Indonesia.*

*Kata-kata kunci: fluctuation rate, Hidden Markov Model, Return Price, Root Mean Square Error*

## PENDAHULUAN

Dikutip dari referensi [1], pada akhir tahun 2014 diketahui bahwa Indeks Harga Saham Gabungan (IHSG) mengalami pertumbuhan sebesar 22,23%. Jika data pertumbuhan pasar saham di Indonesia dibandingkan dengan beberapa negara lainnya di Asia dan Pasifik, Indonesia menempati peringkat keempat. Indonesia hanya kalah dari bursa dalam Shanghai, India dan Filipina. Dan melebihi peringkat negara-negara seperti Thailand, Malaysia, Jepang, dan Amerika Serikat. Bahkan dalam kurun waktu 6 tahun terakhir, IHSG mengalami peningkatan sebesar 282,6%. Karena perkembangannya yang sangat pesat, pasar saham di Indonesia akan menarik banyak trader yang ingin ikut berinvestasi didalamnya. Karena itu, analisis saham yang baik sangatlah diperlukan pada hari ini.

Analisis saham terdapat dua jenis, yaitu analisis fundamental dan analisis teknikal. Analisis fundamental merupakan sebuah analisis saham yang analisisnya didasarkan kepada indikator-indikator yang terkait kepada kondisi ekonomi suatu perusahaan yang menerbitkan saham. Dalam analisis fundamental hal-hal yang akan diperhatikan diantaranya pendapatan laba, pertumbuhan keuntungan dan penyerapan pasar. Selain faktor-faktor internal dari satu perusahaan, analisis ini juga bergantung pada kejadian tertentu yang mempengaruhi perusahaan tersebut. Sedangkan analisis teknikal merupakan sebuah analisis yang menggunakan data-data pergerakan saham untuk memprediksi bagaimana harga saham di hari mendatang. Pada analisis ini akan dianggap bahwa sifat manusia konstan, sehingga faktor-faktor seperti politik, kejenuhan pasar, dan kejadian eventual, yang mempengaruhi suatu perusahaan yang mengeluarkan saham, tidak akan diperhitungkan dalam

analisis ini. Selain itu, pada analisis teknikal, akan diperhatikan probabilitas suatu kejadian yang mungkin terjadi pada saham. Sehingga, kepastian dari hal yang akan terjadi bukanlah hal yang didapatkan.

*Hidden Markov Model* (HMM) merupakan sebuah model pengembangan dari *Markov chain*. Model ini pada awalnya digunakan oleh Lawrence Rabiner sebagai metode untuk pengenalan suara. Pada HMM, terdapat dua jenis *state* yang memiliki probabilitasnya masing-masing, yaitu *observable* dan *hidden*. Dengan menggunakan dua *state* itu, kita dapat memperkirakan baik *observable state* maupun *hidden state*. Saat ini HMM dipakai dalam pemodelan proses-proses yang bersifat stokastik seperti pada pemodelan DNA, perkiraan cuaca hingga aplikasi seperti deteksi suara, deteksi wajah dan pemodelan harga saham. Pemodelan harga saham dengan menggunakan HMM sudah dilakukan dalam beberapa penelitian, diantaranya pada referensi [2],[3], dan [4]. Hal yang membedakan HMM dengan metode pemodelan lain adalah penggunaan probabilitas. Simulasi HMM akan menggunakan matriks probabilitas, yang didapatkan dari data saham, sebagai masukan. Sedangkan model lain, seperti ANFIS atau *moving average*, data yang diolah adalah data harga saham yang akan di modelkan.

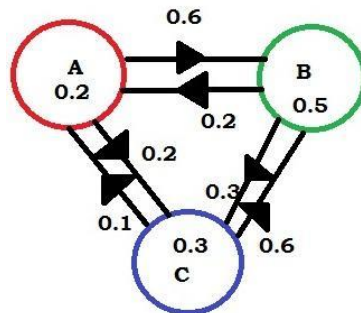
Dalam penelitian ini, saham akan dilihat sebagai sebuah sistem stokastik dan akan dimodelkan pergerakannya dengan menggunakan HMM. 5 saham *bluechip* dari sektor yang berbeda akan dimodelkan pergerakannya dengan HMM. Saham *bluechip* digunakan agar data yang dimodelkan lebih stabil, sedangkan penggunaan saham dengan sektor berbeda bertujuan untuk melihat performa HMM pada jenis saham yang berbeda-beda.

## IMPLEMENTASI HIDDEN MARKOV MODEL PADA PASAR SAHAM

### Hidden Markov Model

Analisis teknikal saham akan didekati sebagai sebuah sistem yang stokastik. Dimana pada sebuah sistem stokastik, pergerakan data dari waktu ke waktu tidak dapat diprediksi oleh sebuah persamaan tertentu. Pergerakan dalam sistem stokastik akan terjadi berdasarkan probabilitas tertentu. Salah satu sistem yang memanfaatkan probabilitas adalah *Markov chain*.

Pada *Markov chain*, sebuah kejadian pada waktu  $t+1$  hanya akan dipengaruhi oleh kejadian pada waktu  $t$ . Kejadian dalam *Markov chain* akan digambarkan sebagai beberapa *state* tertentu yang membentuk suatu sistem *state*. Disini, akan dianggap yang mempengaruhi *state* tersebut hanyalah *state* yang berada pada waktu sebelumnya. Faktor eksternal akan dihilangkan sehingga yang mempengaruhi sistem tersebut hanyalah sistem itu sendiri.



Gambar 1. Skema proses *Markov chain*

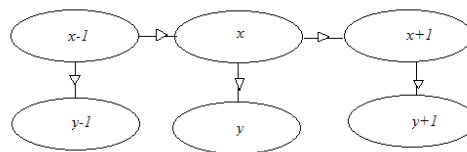
Dilihat dari gambar sebelumnya, terdapat sebuah sistem *Markov chain* yang memiliki 3 *state*: A, B, dan C. Misalkan sistem tersebut menggambarkan gerak partikel. Ketika  $t=0$ , partikel tersebut akan berada pada *state* a, ketika  $t=1$  partikel tersebut memiliki probabilitas untuk berpindah ke B atau C atau tetap berada di A. Ketika  $t=2$ , partikel memiliki probabilitas lagi dan seterusnya. Dalam *Markov chain*, setiap waktu, sebuah *state* dapat berubah dan perubahan tersebut akan digambarkan dengan probabilitas. Probabilitas tersebut dapat disebut dengan probabilitas transisi.

Untuk setiap  $N$  *state*, akan dimiliki probabilitas untuk berpindah sebanyak  $N$  buah yang berarti bahwa sebuah *state* dapat berpindah kepada semua *state* lain atau tetap berada pada *state* tersebut untuk waktu berikutnya. Hal ini menyebabkan untuk setiap anggota  $N$  *state* memiliki probabilitas sebanyak  $N$  buah. Dengan itu akan dimiliki sebuah matriks transisi  $N \times N$  yang menggambarkan probabilitas transisi sebuah sistem *Markov chain*.

	A'	B'	C'
A	0.2	0.6	0.2
B	0.2	0.5	0.3
C	0.1	0.6	0.3

Gambar 2. Contoh matriks transisi dari gambar 1

*Hidden Markov model* (HMM) merupakan pengembangan dari *Markov chain*. Jika pada *Markov chain* sebuah kejadian pada waktu  $n+1$  hanya akan dipengaruhi oleh kejadian pada waktu ke  $n$ , maka pada HMM kejadian pada  $n+1$  dipengaruhi kejadian lain yang berbeda pada waktu  $n$ . Dengan kata lain, dalam HMM terdapat dua set kejadian, atau dapat disebut sebagai sistem *state*. Set pertama merupakan *state* yang dapat diamati. Kelompok *state* ini disebut *observable state*. Pada *observable state*, kejadian pada waktu  $n$ , tidak akan mempengaruhi kejadian pada waktu  $n+1$ . *State observable* hanya akan dipengaruhi oleh set kejadian yang lain yang tidak teramati. Set kejadian ini akan disebut *hidden state*.



Gambar 3. Skema proses HMM

Seperti gambar diatas, *hidden state* ( $x$ ) akan mempengaruhi *observable state* ( $y$ ) dan *hidden state* ( $x$ ) dalam bentuk probabilitas. Hal ini menyebabkan pada sistem akan terdapat dua buah matriks transisi, yaitu matriks transisi yang menggambarkan hubungan *state* pada *hidden state* yang disebut matriks transisi dan matriks transisi yang akan menggambarkan hubungan diantara *hidden state* dan *observable state* yang disebut matriks observasi.

Dari penjelasan diatas dapat disimpulkan dalam sebuah sistem HMM terdapat :

- $S$  sebagai *hidden state* dimana  $S$  merupakan set dari  $N$  buah *state* tersembunyi yang mungkin didalam satu sistem. *Hidden state* pada sebuah waktu  $t$  akan disimbolkan sebagai  $qt$
- $V$  sebagai *observable state* dimana  $V$  merupakan set dari  $M$  buah *state* yang diamati dalam suatu sistem. *Observable state* pada waktu  $t$  akan disimbolkan sebagai  $ot$
- $A$  sebagai matriks transisi  $N \times N$
- $B$  sebagai matriks observasi  $N \times M$

### Implementasi HMM pada Pasar Saham

Dalam peneltian ini akan digunakan data saham dari lima perusahaan penerbit saham bluechip dengan jumlah data sebanyak 2000 data. Dimana 1900 data akan digunakan sebagai data untuk membangun sistem HMM dan 100 data digunakan sebagai data yang akan digunakan untuk menguji hasil prediksi HMM. Setelah mendapatkan data yang akan digunakan, maka data yang ada perlu diolah terlebih dahulu sebelum nantinya diproses oleh MATLAB untuk dilakukan pemodelan dengan HMM. Dalam penelitian ini return price akan dilihat sebagai *observable state* dan fluctuation rate yang akan dilihat sebagai *hidden state*.

Berikut merupakan persamaan return price (RP) dan fluctuation rate (FR) yang digunakan dalam penelitian ini :

$$RP = \frac{close_t - close_{t-1}}{close_t} \tag{1}$$

$$FR = \frac{High - Low}{Low} \tag{2}$$

Dari dua persamaan diatas dapat dilihat bahwa *return price* (RP) dan *fluctuation rate* (FR) yang digunakan memiliki sifat diskrit. Dimana kedua data tersebut akan dianggap berdiri sendiri dan tidak bersambung dengan data pada  $t-1$  dan  $t+1$ . Hal ini dilakukan karena pada return price dia merupakan *observable state* yang tidak saling mempengaruhi sehingga dapat didekati sebagai data diskrit. Sedangkan fluctuation rate bisa dianggap diskrit karena data yang digunakan merupakan data pada satu  $t$  dan tidak melihat  $t+1$  atau  $t-1$ .

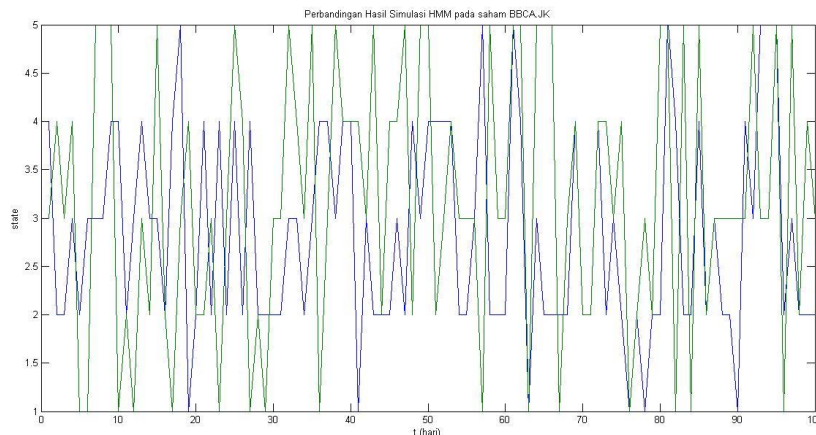
Nilai FR dan RP yang didapat akan dicari distribusinya untuk dilihat persebarannya. Dari persebaran tu akan dibagi menjadi lima *state*. Untuk mendapatkan nilai batas untuk setiap *state*, maka data akan dibagi menjadi lima bagian dimana setiap bagian memiliki populasi data yang sama. Dengan populasi data yang sama, maka range untuk tiap *state* menjadi berbeda, sehingga data yang memiliki frekuensi kemunculan yang sedikit akan terfokus pada dua *state*. Dan *state* yang memiliki frekuensi kemunculan yang tinggi akan memiliki range yang sempit. Dari lima *state* tersebut akan dicari matriks transisi FR-FR sebagai matriks *hidden* dan FR-RP sebagai matriks *observable*. Dari kedua matriks ini akan digunakan MATLAB untuk melakukan simulasi HMM

Hasil simulasi HMM akan mengeluarkan data sebanyak data tes yang nantinya akan dibandingkan dengan data tes untuk mengetahui akurasi dari metode yang digunakan. Dalam penelitian ini akan digunakan RMSE sebagai parameter ketepatan simulasi. *Root Mean Squared Error* (biasa disingkat RMSE) merupakan sebuah metode yang digunakan untuk menghitung galat dari data yang diprediksi. Dengan mengetahui galat yang dihasilkan, maka kita dapat mengetahui kinerja dari program yang digunakan. Untuk menghitung nilai RMSE, akan digunakan persamaan sebagai berikut :

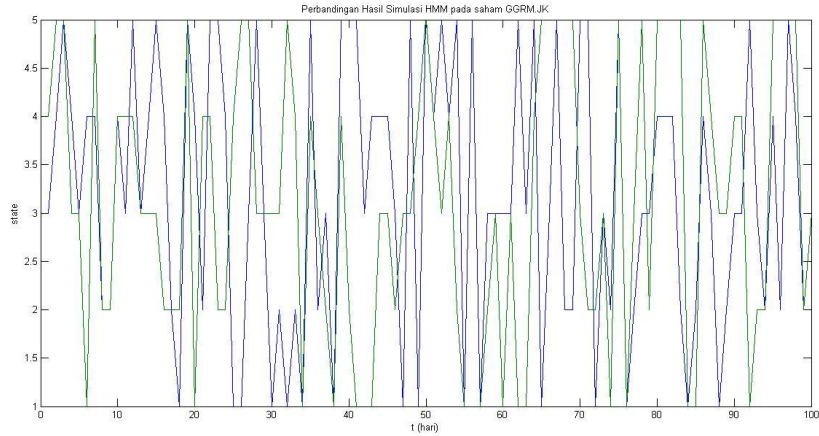
$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{t_i} - x_{h_i})^2}{n}} \tag{3}$$

Dimana  $n$  merupakan panjang data,  $x_{t_i}$  merupakan data nyata yang kita jadikan pembanding terhadap data yang dihasilkan program (disebut juga data tes) pada waktu  $i$ , dan  $x_{h_i}$  merupakan data yang dihasilkan oleh program pada waktu ke  $i$ . Pada RMSE diharapkan nilai yang kecil. Semakin kecil nilai dari RMSE, maka eror yang dihasilkan oleh sebuah program semakin jarang terjadi.

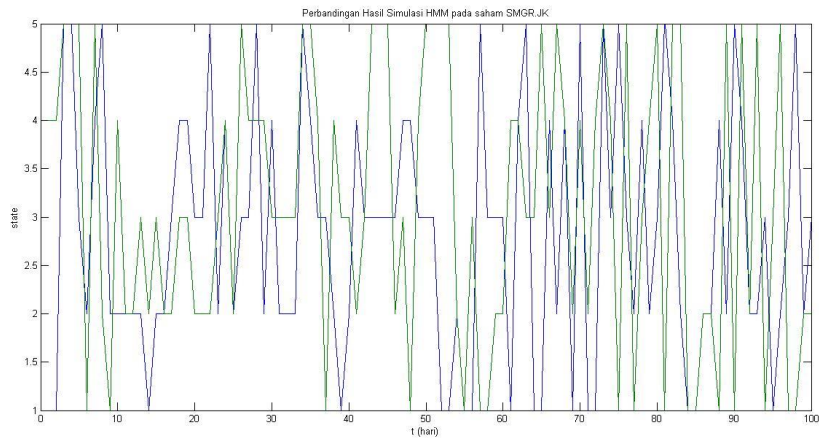
### HASIL SIMULASI DAN PEMBAHASAN



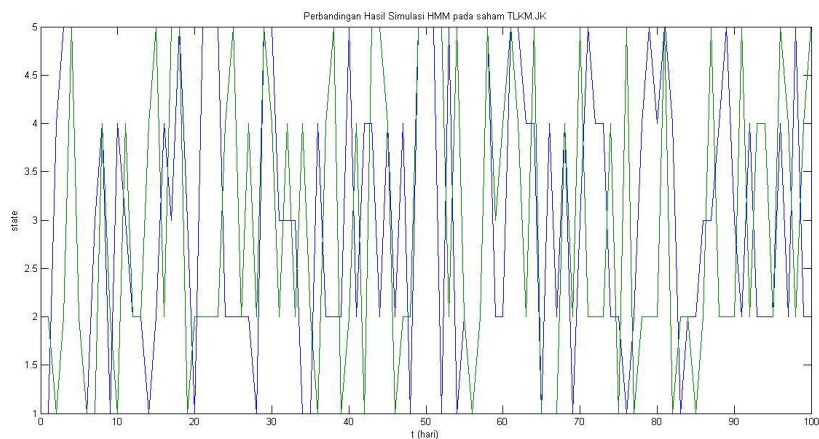
Gambar 4. Perbandingan hasil simulasi HMM (hijau) dengan referensi (biru) pada saham BBKA.JK



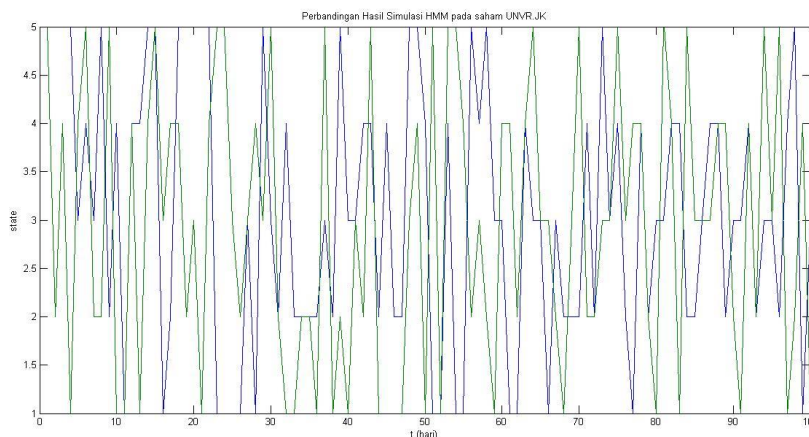
Gambar 5. Perbandingan hasil simulasi HMM (hijau) dengan referensi (biru) pada saham GGRM.JK



Gambar 6. Perbandingan hasil simulasi HMM (hijau) dengan referensi (biru) pada saham SMGR.JK



Gambar 7. Perbandingan hasil simulasi HMM (hijau) dengan referensi (biru) pada saham TLKM.JK



Gambar 8. Perbandingan hasil simulasi HMM (hijau) dengan referensi (biru) pada saham UNVR.JK

Tabel 1. Hasil perhitungan RMSE pada 5 saham yang disimulasikan

Kode Saham	RMSE (%)
BBCA.JK	25,059928
SMGR.JK	27,45906
UNVR.JK	28,94823
TLKM.JK	27,784888
GGRM.JK	25,922963

Dari gambar 4,5,6,7 dan 8 dapat dilihat bahwa perbandingan antara hasil simulasi dengan data referensi untuk kelima saham yang diteliti semuanya mirip. Hal ini juga diperlihatkan dari nilai RMSE dari simulasi kelima saham tersebut yang memiliki nilai yang mirip. Dan memiliki rata-rata sebesar 27,035%. Dapat disimpulkan bahwa penggunaan HMM kepada 5 saham dengan sektor yang berbeda tidak memberikan hasil yang berbeda-beda, sehingga HMM dapat diaplikasikan untuk memodelkan saham yang berbeda-beda. Hal ini dikuatkan dengan referansi [2],[3], dan [4] dimana HMM juga diaplikasikan untuk memodelkan saham.

Salah satu kekurangan dari simulasi ini adalah hasil yang dihasilkan dari masih berupa perkiraan yang sangat kasar. Sehingga hasil dari simulasi ini tidak dapat dibandingkan dengan grafik *closing price* yang biasa dijadikan referensi. Untuk memperbaiki kekurangan ini, metode monte carlo dapat digunakan untuk mengeluarkan nilai *return price* yang nantinya bisa diolah lagi untuk menghasilkan nilai *closing price* yang diinginkan.

Penggunaan sistem pembagian *state* yang digunakan dalam penelitian ini dapat berefek baik jika dilakukan simulasi lanjutan dengan menggunakan metode monte carlo. Dengan menggunakan sistem ini, maka nilai yang memiliki frekuensi kemunculan yang tinggi akan terfokus pada 3 *state* sedangkan nilai yang frekuensi kemunculannya rendah akan fokus pada 2 *state*. Sehingga 3 *state* akan memiliki *range* nilai yang kecil dan dapat memperkecil batas nilai acak yang akan dihasilkan oleh simulasi monte carlo.

Dari nilai galat yang didapatkan, maka seharusnya simulasi HMM yang dilakukan masih dapat diperbaiki. Terdapat beberapa cara yang dapat dilakukan untuk memperbaiki simulasi HMM. Pertama, memperbanyak *state* ketika melakukan simulasi HMM. Dengan meperbanyak *state*, maka *range* nilai untuk setiap *state* akan menjadi lebih kecil sehingga hasil simulasi dapat menjadi lebih tepat. Kedua, memperbanyak data yang digunakan. Dengan memperbanyak data yang digunakan dalam membangun sistem matriks transisi pada HMM, diharapkan nilai matriks transisi akan lebih tepat dalam memberikan probabilitas yang nantinya akan digunakan untuk simulasi HMM.

## KESIMPULAN

Dari penelitian ini didapatkan rata-rata galat RMSE dari simulasi HMM pada 5 saham bluechip di Indonesia adalah 27,035%. Untuk memperbaiki nilai ini, dapat dilakukan beberapa cara, diantaranya dengan menambahkan *state* dan menambahkan data yang digunakan.

Selain itu, kekurangan dari simulasi HMM untuk memprediksi harga saham adalah hasil yang didapatkan masih berupa perkiraan kasar. Karena itu untuk memperbaikinya dapat digunakan metode monte carlo sehingga bisa didapatkan harga prediksi sebagai keluaran simulasi yang dilakukan.

## REFERENSI

1. [http://www.rri.co.id/post/berita/129845/ekonomi/pasar\\_modal\\_2015\\_di\\_indonesia\\_semakin\\_bergairah.html](http://www.rri.co.id/post/berita/129845/ekonomi/pasar_modal_2015_di_indonesia_semakin_bergairah.html)
2. Jaroslav Lajos et.al; A Six *State* HMM for the S&P 500 Stock Market Index, Computer Science Departement, Oklahoma *State* University
3. Md. Rafiul Hassan & Baikun Nath, Stock Market Forecasting Using *Hidden* Markov Model: A New Approach, Computer Science and Software Engineering, The University of Melbourne, 2005
4. Aditya Gupta & Bhuwan Dhingra, Stock Market prediction using *Hidden* Markov Model, Departement of Computer Science and Engineering, Indian Institute of Technology
5. Taylor Sauder, The Implementation of a *Hidden* Markov Model in MATLAB for the Prediction of Commodity Prices, Foster College of Business Administration, Bradley University, 2011