# PREDIKSI SAHAM MENGGUNAKAN DBN ( DEEP BELIEF NETWORK )

# STOCK PREDICTION USING DBN ( DEEP BELIEF NETWORK )

[1]Giali Ghazali, [2]Jondri, M.Si,
[1,2,3]Ilmu Komputasi Fakultas Informatika Universitas Telkom, Bandung
[1]gialighazali@gmail.com, [2]jondri@gmaill.com,

**ABSTRAK**

Dalam penelitian ini akan dibahas prediksi indeks harga saham dengan metode *Deep Belief Network* (DBN). Penelitian ini menggunakan indeks saham dari pasar saham Indonesia yaitu Indeks Harga Saham Gabungan (IHSG). Percobaan ini didasarkan pada data historis harian selama 5 tahun. Prediksi dilakukan dengan menggunakan Restricted Boltzman Machines (RBM) dua tahap dan RBM tiga tahap, dengan input data menggunakan (20 + ri) dan (40 + ri) dengan jumlah neuron (n = 10) dan (n = 50). Dengan menggunakan metode Deep Belief Network (DBN) didapatkan hasil *Root Mean Square Eror* (RMSE) sepuluh kali percobaan. Mendapatkan hasil data (20 + ri) dengan (n = 10) dengan RBM dua tahap 0,18291 dan RBM tiga tahap 0.17912. Pada data (40 + ri) dengan (n = 10) dengan RBM dua tahap 0,1804 dan tiga tahap RBM 0.17118. Pada data (20 + ri) dengan (n = 50) dengan RBM dua tahap 0,16996 dan tiga tahap RBM 0.05892. Pada data (40 + ri) dengan (n = 50) didapatkan RBM dua tahap 0,17781 dan tiga tahap RBM 0.16904. dapat disimpulkan bahwa RBM tiga tahap lebih baik dibandingkan RBM dua tahap, dan nilai neuron yang besar mempengaruhi hasil prediksi.

**Kata kunci : *prediksi saham, Indeks Harga Saham Gabungan (IHSG),* Deep Belief Network (DBN).**

## 1. Pendahuluan

### 1.1 Latar Belakang

Di dalam pasar modal, saham merupakan salah satu jenis investasi yang mampu memberikan keuntungan yang menarik sehingga banyak dipilih oleh investor sebagai alternatif investasi. Keuntungan yang diperoleh dengan investasi saham adalah pendapatan yang bersumber dari *deviden* (pembagian keuntungan perusahaan kepada pemegan saham) dan *capital gain* (selisih antara harga beli dan jual). Pada sisi lain, ada dua risiko investasi saham yaitu capital loss (kebalikan capital gain) dan Risiko Likuidasi (perusahaan dinyatakan bangkrut oleh pengadilan atau perusahaan tersebut dibubarkan).

Di pasar saham fluktuasi merupakan hal yang tidak terhindarkan. Para pemilik saham harus memiliki acuan dalam pengambilan keputusan agar terhindar dari risiko investasi. Indeks harga saham adalah indikator atau cerminan pergerakan harga saham yang dapat digunakan sebagai acuan oleh para pemilik saham untuk mengetahui trend pergerakan saham sehingga dapat mengambil keputusan apakah akan membeli, mempertahankan , atau menjual saham tersebut. Pergerakan pasar saham dipengaruhi oleh banyak faktor mikro-ekonomi seperti peristiwa politik, kebijakan perusahaan, kondisi ekonomi secara umum,komoditas harga, bunga bank, nilai tukar bank, psikologi, dll. Perubahan pada faktor-fatkor tersebut menyebabkan ketidakpastian harga saham sehingga dibutuhkan prediksi untuk melihat pergerakan pasar saham di masa depan.

Deep Learning merupakan salah satu metode yang dapat digunakan untuk memberikan solusi yang optimal untuk kasus prediksi harga saham. Deep learning merupakan perkembangan dari metode-metode prediksi lainnya seperti metode SVM (Support Vector Machine). Maka pada tugas akhir ini akan dibuat prediksi menggunakan metode Deep learning

### 1.2 Perumusan Masalah

Permasalahan yang akan dibahas dalam penelitian ini adalah sebagai berikut:

1. Bagaimana cara melakukan prediksi harga saham dengan metode *DBN (Deep Belief Network)* ?
2. Bagaimana menganalisa performansi hasil prediksi harga saham dengan menggunakan metode *DBN (Deep Belief Network)?*

### 1.3 Tujuan dari penelitian adalah :

1. Memprediksi harga saham dengan metode *DBN (Deep Belief Network)* berdasarkan data runtun waktu harga saham sebagai data input
2. Menganalisa performansi hasil prediksi harga saham dengan metode *DBN (Deeb Belief Network).*

## 2. Landasan Teori

### 2.1 Saham

Saham adalah surat berharga dengan satuan nilai sebagai tanda kepemilikan suatu perusahaan. Dengan menerbitkan saham, perusahaan perusahaan yang membutuhkan suntikan modal bisa menjual saham kepada inverstor (penanam modal) dengan kata lain saham bisa diperjual belikan dalam pasar modal untuk kepentingan pendanaan perusahaan. Saham bisa diperjual belikan dalam pasar primer (*primary market*) dan pasar sekunder (*secondary market*).

### 2.2 Prediksi Atau Peramalan

Prediksi adalah proses penggambaran sesuatu yang akan terjadi dimasa depan dengan menggunakan data dan informasi saat ini dan sebelumnya. Prediksi memiliki tujuan untuk memberikan gambaran sedekat mungkin dengan apa yang akan terjadi dan tidak memberikan suatu kondisi pasti tentang apa yang akan terjadi.Kuncidalam permalan adalah membaca sejumlah data untuk menemukan sejumlah karakteristik dan kecenderungan, dengan prinsip bahwa kejadian dimasa lalu akan berulang dimasa depan.

Dalam membuat peramalan atau menerapkan suatu peramalan maka ada beberapa hal yang harus diperitimbangkan yaitu:

1. Ramalan pasti mengandung kesalahan. Peramal hanya bisa mengurangi ketidakpastian yang akan terjadi, tetapi tidak dapat menghilangkan ketidakpastian tersebut.
2. Peramalan seharusnya memberikan informasi tentang beberapa ukuran kesalahan. Hal ini dikarenakan peramalan pasti mengandung kesalahan dan oleh karena itu, penting bagi peramal untuk menginformasikan seberapa besar kesalahan yang mungkin terjadi.
3. Peramalan jangka pendek lebih akurat dibandingkan peramalan jangka panjang. Hal ini disebabkan karena pada peramalan jangka panjang, memiliki kemungkinan terjadinya perubahan faktor-faktor yang mempengaruhi sesuatu yang diramalkan tersebut semakin besar. Berdasarkan sifat ramalan yang telah disusun.

#### 2.2.1 Peramalan Kualitatif

Peramalan kualitatif merupakan peramalan yang didasarkan atas kualitatif pada masa lalu. Hasil peramalan yang dibuat sangat tergantung pada orang yang menyusunnya. Hal ini penting karena hasil peramalan tersebut ditentukan berdasarkan pemikiran yang bersifat intuisi, judgement, pendapat, pengetahuan, dan pengalaman dari penyusunnya. Contoh dari peramalan kualitatif adalah *Panel Consensus. Panel Consensus* cocok digunakan dalam situasi yang sangat sensitif terhadap intuisi dari suatu atau sekelompok kecil orang yang karena pengalamannya mampu memberikan opini yang kritis dan relevan. Umumnya panel consensus dilakukan oleh manajemen senior suatu organisasi atau perusahaan.

#### 2.2.2 Peramalan kuantitatif

Peramalan kuantitatif yaitu peramalan yang didasarkan atas data kuantitatif pada masa lalu. Hasil peramalan yang dibuat sangat tergantung pada metode yang dipergunakan dalam peramalan tersebut. Dengan metode yang berbeda meski studi kasusnya sama, akan diperoleh hasil peramalan yang berbeda. Baik tidaknya metode yang digunakan sangat ditentukan oleh perbedaan atau penyimpangan antara hasil ramalan dengan kenyataan yang terjadi. Metode yang baik adalah metode yang memberikan nilai-nilai perbedaan atau penyimpangan seminimal mungkin. Peramalan kuantitatif hanya dapat digunakan apabila terdapat tiga kondisi berikut:

a. Adanya informasi tentang keadaan masa lalu.
b. Informasi tersebut dapat pertanggung jawabkan dalam bentuk data.
c. Diasumsikan bahwa pola masa lalu akan berkelanjutan pada masa yang akan datang.

### 2.3 Prediksi *Time Series*

Prediksi secara *time series* merupakan peramalan yang dilakukan dengan menggunakan data tertentu yang disebut data *time series*. Data time series adalah nilai suatu variabel yang disusun berdasarkan urutan/deret waktu bisa dalam satuan harian, mingguan, bulanan, atau tahunan tergantung keperluan terhadap perilaku data yang akan diobservasi. Data *time series* yang digunakan adalah data historis yang diukur berdasarkan suatu pengamatan tertentu. Karena data yang digunakan adalah data yang terukur, maka peramalan secara *time series*, termasuk ke dalam peramalan kuantitatif. Metode prediksi *time series* beranggapan bahwa data atau kejadian masa lalu akan cenderung berulang dimasa yang akan datang. Fokus prediksi secara *time series* adalah apa yang akan terjadi, bukan mengapa hal itu terjadi. Hal yang harus diperhatikan sebelum melakukan prediksi *time series* adalah.

1. Adanya ketergantungan masa depan dengan masa lampau.
2. Aktivitas masa yang akan datang mengikuti pola masa lalu.
3. Hubungan keterkaitan masa lalu dan masa kini dapat ditentukan melalui pengamatan.

Dalam prediksi data time series, diasumsikan bahwa nilai variabel dependen dapat diketahui dari nilai-nilai variabel dependen yang telah lampau saja. Hal ini berbeda dengan *causal forecasting* yang menggunakan

nilai-nilai variable independen. Variabel dependen lampau yang digunakan didalam prediksi data time series ini biasa disebut dengan variabel lag. Salah satu contoh peramalan time series dengan variabel dependen yaitu prediksi harga saham dengan menggunakan data niali saham sebelumnya.

Kelebihan metode prediksi time series adalah sederhana, cepat dan murah. Metode prediksi ini jauh lebih murah dibandingkan dengan metode prediksi yang lain yang membutuhkan jauh lebih banyak data yang seringkali belum tersedia dan harus dikumpulkan terlebih dahulu. Selain itu, jika terjadi perubahan kondisi yang disebabkan oleh berbagai faktor, misalnya faktor alam, politik, ekonomi, dan lain sebagainya, maka prediksi menggunakan metode ini akan mengabaikannya .

Prediksi *time series* dengan menggunakan JST dapat dilakukan dengan menggunakan input berupa *2 series*, *3 series*, *4 series*, dan sebagainya. Jika digunakan *4 series*, berarti prediksi dilakukan dengan menggunakan 4 data sebelumnya, dan begitu juga untuk series yang lain. Misalkan akan dilakukan prediksi dengan menggunakan input *4 series* pada 10 buah data. Kesepuluh data tersebut terlebih dahulu dibagi menjadi dua atau tiga kelompok. Jika dibagi menjadi dua kelompok, data tersebut akan dibagi menjadi data untuk proses pelatihan dan data untuk proses pengujian. Namun, jika dibagi menjadi tiga kelompok, data tersebut akan dibagi menjadi data untuk proses pelatihan, data untuk proses *validasi*, dan data untuk proses pengujian. Proses *validasi* digunakan untuk menghindari JST agar tidak mengalami kondisi overfit dan overweight.

**a. Deep Learning**

Deep Learning (deep machine learning, or deep structured learning, or hierarchical learning, or sometimes DL) adalah cabang dari machine learning berdasarkan satu set algoritma yang digunakan untuk model abstraksi tingkat tinggi pada data dengan menggunakan beberapa lapisan implementasi dan menggunakan struktur yang kompleks atau sebaliknya, terdiri dari beberapa transformasi non-linear. Teknik dan algoritma dalam Deep Learning dapat digunakan baik untuk kebutuhan supervised learning, unsupervised learning dan semi-supervised learning dalam berbagai aplikasi seperti pengenalan citra, pengenalan suara, klasifikasi teks. Pada tugas akhir ini, Deep Learning digunakan untuk memprediksi churn pada perusahaan jaringan telekomunikasi. Model Deep pada dasarnya dibangun berdasarkan Jaringan saraf tiruan, yang risetnya sudah berlangsung sejak era 80an namun baru-baru ini kembali bangkit dengan adanya komputer yang semakin cepat apalagi ditambah dengan kemampuan grafis modern yang mampu melakukan kalkulasi berbasis matriks secara simultan. Berdasarkan riset yang baru-baru ini dilakukan, Deep Learning mampu melakukan pengenalan grafis, pola tulis tangan dan beberapa pola lainnya lebih akurat dibandingkan dengan algoritma pembelajaran mesin lainnya.

### 2.5 Jaringan Saraf Tiruan

Jaringan Syaraf Tiruan adalah salah satu representasi buatan dari otak manusia, yang diimplementasikan menggunakan program computer yang mampu menyelesaikan tugas-tugas dengan efektifitas yang tinggi seperti pengenalan pola atau klasifikasi karena proses pembelajaran. Otak manusia berisi berjuta-juta sel saraf yang bertugas untuk memproses informasi. Tiap-tiap sel bekerja seperti suatu processor sederhana. Masing-masing sel tersebut saling berinteraksi sehingga mendukung kemampuan kerja otak manusia. Jaringan Syaraf Tiruan dibentuk sebagai generalisasi model matematika dari jaringan syaraf biologi, dengan asumsi bahwa :

- Pemrosesan informasi terjadi pada banyak elemen sederhana (neuron).
- Sinyal dikirimkan antar neuron-neuron melalui penghubung-peghubung.
- Penghubung antar neuron memiliki bobot yang akan memperat atau memperlemah sinyal.
- Untuk menentukan output, setiap neuron menggunakan fungsi aktivasi yang dikenakan pada jumlah input yang diterima. Besar output ini kemudian dibandingkan dengan batas ambang.

Ciri utama dari JST adalah pada kemampuannya dalam hal belajar. Agar JST berfungsi seperti yang diinginkan, jaringan tidak diprogam melainkan melalui proses belajar. Proses belajar dalam JST adalah proses penyesuain nilai-nilai bobot. Jadi bobot jaringan berubah sesuai dengan sasaran yang ditangkap dari keteraturan masukan.

#### 2.5.1 Bentuk dan Arsitektur Jaringan Saraf Tiruan

*neuron* yang terletak pada sebuah lapisan yang sama akan memiliki keadaan sama. Faktor terpenting dalam menentukan kelakuan suatu *neuron* adalah fungsi aktivasi dan pola bobotnya. Pada setiap lapisan yang sama, *neuron–neuron* akan memiliki fungsi aktivasi yang sama. Apabila *neuron–neuron* dalam suatu lapisan lapisan tersembunyi

akan dihubungkan dengan *neuron–neuron* pada lapisan yang lain lapisan *output*, maka setiap *neuron* pada lapisan

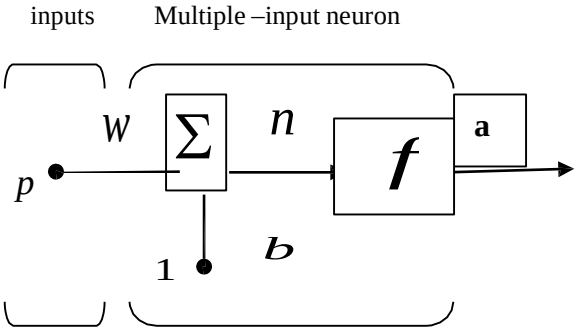
tersebut lapisan tersembunyi juga harus dihubungkan dengan lapisan lainnya lapisan *output*.

### 2.5.2 Jaringan Dengan Lapisan Tunggal (*single layer net)*

Jaringan *single layer net* hanya memiliki satu lapisan dengan bobot–bobot terhubung. Jaringan ini menerima *input* yang kemudian secara langsung akan mengolahnya menjadi *output* tanpa harus melalui lapisan tersembunyi. Untuk lebih jelas dapat dilihat pada gambar 2.1:

Sebuah neuron berinput tunggal diperlihatkan pada gambar 2.1. Input skalar *p* dikalikan dengan skalar *weight w* untuk kodisi bentuk *wp*, salah bentuk ini yang dikirim ke dalam penjumlah. Input yang lain, 1, adalah dikalikan dengan *bias b* dan dilewatkan ke dalam penjumlah. Output dari penjumlah *n* sering kali digunakan sebagai input *net*, yang berjalan ke dalam fungsi transfer *f* , yang menghasilkan output neuron skalar *a*. fungsi aktivasi disebut juga *fungsi transfer* dan "offset" disebut juga *bias*.

Bobot *w* berhubungan dengan sebuah kekuatan gabungan dari sel badan yang direpresentasikan dengan penjumlahandan fungsi transfer, dan output neuron *a* mengambarkan signal dalam axon




$$a = f(\mathrm{W}p + b)$$

*Gambar 2.1 : Neuron berinput tunggal*

Output neuron dihitung sebagai berikut

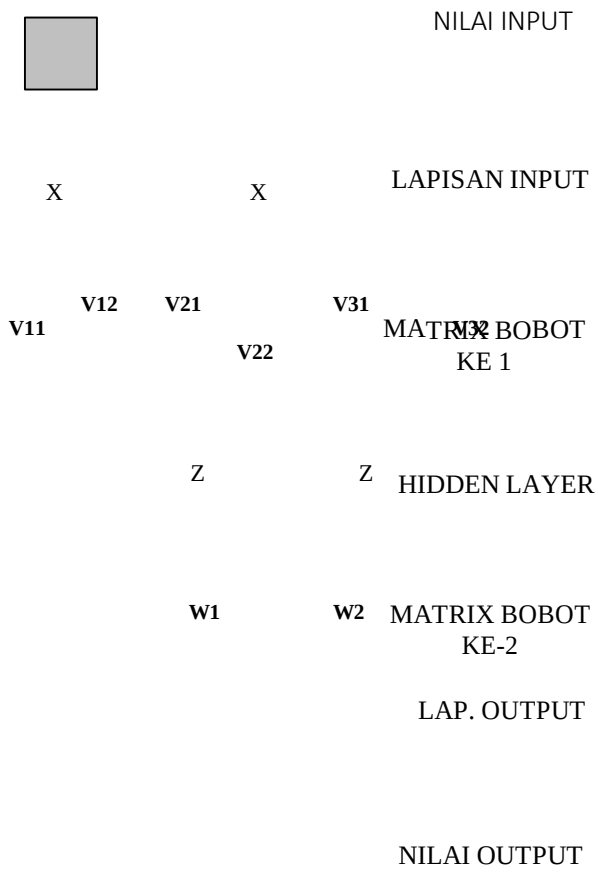
$$a = f(wp + b) \qquad (2\text{-}3)$$

$$a = f(3(2) - 1.5) = f(4.5) \qquad (2\text{-}4)$$

Output nyata bergantung pada fakta-fakta fungsi transfer yang dipilih. Kita akan mendiskusikan transfer fungsi dalam bagian berikutnya.

### 2.5.3 Jaringan Dengan Banyak Lapisan (*multilayer*)

Jaringan ini memiliki satu atau lebih lapisan yang terletak diantara lapisan *input* dan lapisan *output*nya (memiliki satu atau lebih lapisan tersembunyi). Jaringan dengan banyak lapisan ini dapat menyelesaikan permasalahan yang lebih sulit dari pada jaringan dengan lapisan tunggal, tetapi pembelajarannya lebih rumit. Jaringan ini pada banyak kasus lebih sukses dalam menyelesaikan masalah.




*Gambar 2.2 : Jaringan saraf multilayer.*

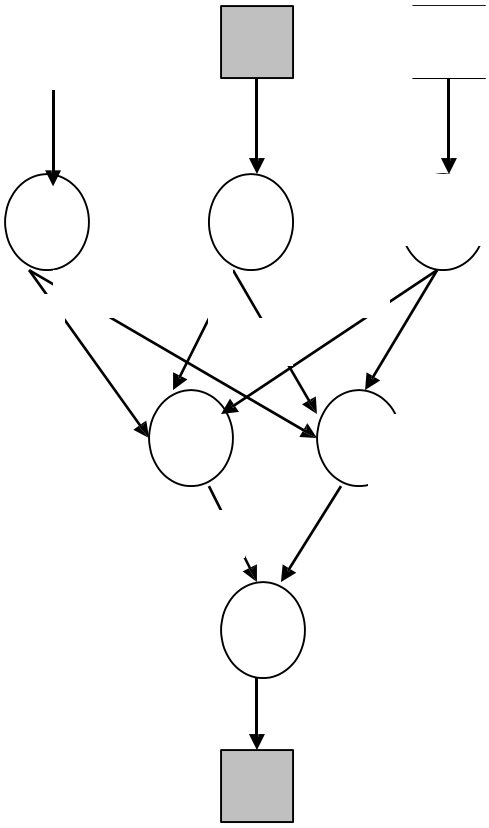
### 2.6 Deep Belief Network

DBN dibangun oleh tumpukan RBM. Urutan training dan pembuatan layer, kita dapat membangun sebuah kedalaman neural network. Sekali stack dari RBM yang ditrain, itu dapat digunakan untuk menginiliasisasi sebuah multilayer neural network untuk klasifikasi. Untuk

Jika untuk sesaat, $w = 3, p = 2$ *dan* $b = -1.5$, kemudian

mengerti bagaimana pengerjaan mesin learning ini kita memerlukan teori pengdukung dari *boltzman mecine* dan *rbm.*

**2.6.1 Boltzmann Machines**



Pendekatan ini diambil dalam teori dari Boltzmann machine dari stastic pysic dan untuk sebuah akun dari metode2 tambahan. Mesin boltzman stochastic binary learning yang tersusun dari *stochastic neuron* dan bobot *simtris* . stochastic neuron dapat diamabil pada dua stage +1 on stage tapi kalau -1 of state ( seperti skring lampu) tapi +1 untuk on state dan 0 of state juga pelajaran dan aplikasi

Neuron dari mesin boltzman di bagi menjadi 2 group atau layer yaitu viible layer dan hidenlayer ini dapat dinyatakan dengan $\mathbf{v} \in \{0,1\}^{v}$ dan $\mathbf{h} \in \{0,1\}^{h}$ dengan **v** adalah visible layer dan **h** adalah hiden layer .visible neuron adalah penghubung ke satu sama lain pada lingkungan – lingkungan neuron itu menjadi unit visible bahwa kita mengirim sinyal input dibawah training visible neuron adalah dijepit pada keadaan tetentu oleh lingkungan hiden neuron menjalankan dengan bebas dan mengextrac sifat dari sinyal input oleh penangkap pesan oleh statistic yang lebih tinggi di vector penjepit.

Pada mesin boltzman visible neuron adalah penghubung yang baik ke setiap neuron pada hiden layer tipe yang sama dari pemakain koneksi ke neuron hiden layer. Pringatan bahwa koneksi antara kedua layer adalah presentasi oleh dua panah dimana dalam gambar di implementasikan oleh garis tanpa kepala panah secara sederhana ini mengartikan bahwa koneksi berbentuk simentris yang diamana itu juga kasus untuk koneksi antara neuron pada layer yang sama .

Pembelajaran jaringan mendasari distribusi probalitas dari data oleh proses pengiriman pola penjepit ke visible neuron. Jaringan kemudian dapat melengkapi pola jika bagian dari informasi asumsi nya tersedia bahwa jaringan telah tepat

detraining

Asumsi mesin boltzman adalah

1. Vector input cukup jauh berlanjut untuk jaringan pencapaian keseimbagan thermal.
2. Tidak ada struktur berada pada pesanan squensial yang dimana input vector adalah menjepit ke visible neuron

Sebuah konfigurasi berat tertentu dari mesin boltzman dikatakan menjadi model sempurna dari input data jika itu

menyebabkan distribusi probalitas sama dari wilayah. Unit terlihat ketika mereka menjalankan dengan bebas ketika

unti ini dijepit oleh input vector. Untuk mencapai sebuah

model yang sempurna dibutuhkahkan sebuah angka exponent yang besar dari hiden unit sebagai perbandingan untuk angka dari visible unit tetapi pada mesin ini dapat mencapai perfomansi bagus jika ada sebuah structure yang regular didalam data khususnya ini benar ketika penggunan jaringan hiden unit untuk menangkap regulasi ini dibawah

1. Tahap positive. Pada tahap ini mesin beroprasi pada kondisi terjepit, itu mengatakan bahwa dibawah ini mempengaruhi dari sample training
2. Tahap negative. Pada tahap ini mesin diijinkan untuk menjalankan bebas, ini artinya bahwa disana tidak ada lingkungan hidup.

### 2.6.2 Restricted Boltzmann Machines

Boltzman mesin adalah penghubung antara input neuron dan hiden layer di samping itu menghubungkan antara input dan hiden layer . rbm hanya koneksi antara input dan hiden layer dan unit yang tidak terkonkneksi pada layer yang sama dan tidak berkoneksi diantara unit pada layer yang sama. Mesin rbm adalah sebuah model generative dan kemudian kita melihat bahwa itu digunakan sebagai penggebangun block pada dbn.rbm juga mempunyai penggunanan sebagai pembangun block pada kedalaman pada auto-encoders. Rbm terdiri dari sebuah visible dan hiden layer dari unit penghubung binary unit oleh berat simetris. Pada rbm hiden unit dapat dilihat sebagai sifat detector jaringan menetapkan sebuah probalitas kesetiap pasangan dari visble dan vector hiden neuron direkam untuk distribusi .

$$P(v,h;\theta) = \frac{1}{Z\theta} e^{-E(v,h;\theta)} \qquad (2\text{-}5)$$

Yang dimana fungsi partisi diberikan oleh $Z(\theta) = \sum_v \sum_h \exp(-E(v,h;\theta))$. Energi dari system ini adalah:

$$E(v,h) = \sum_i b_i h_i - \sum_j a_j v_j - \sum_{i,j} v_i w_{ij} h_j$$

$$= -\sum_i b_i h_i - \sum_j a_j v_j - \sum_{i,j} v_i w_{ij} h_j \qquad (2\text{-}6)$$

yang mana $b_i$ dan $a_j$ adalah bias dari variable input $v_j$ dan variable tersembunyi $h_i$, dan $w_{ij}$ adalah bobot dari interaksi antara unit i dan j, di visible dan hiden layer. *Marginal probability* distribusi P dari data vektor V diberikan oleh :

$$P(V;\theta) = \sum_h \frac{e^{-E(v,h;\theta)}}{\sum_{v,h} e^{-E(v,h;\theta)}}$$

$$= \frac{1}{Z(\theta)} \sum_h e^{-E(v,h;\theta)} \qquad (2\text{-}7)$$

keadaan mesin dapat merecontruksi mendasari distribusi dengan sebuah angka yang dikelola dari hiden unit.

Mesin boltzman mengatakan mempunyai keadaan dari oprasi :

$$= \sum_{h} exp\,( \bullet wh + \bullet v + \bullet h)$$

dalam persamaan ini, v, adalah vektor input dan h adalah vektor hiden unit. Ini adalah fungsi distribusi marginal visible unit yang lebih.Sebuah sistem (v, h) dengan energi rendah diberikan sebuah probabilitas tinggi dan satu dengan energi tinggi diberikan probalitas rendah. Dengan bantuan fungsi energi yang kita dapat tentukan berikut probabilitas.

$$= \frac{1}{\Omega(\Omega)} \exp(\Omega v) \prod_{\Omega=1}^{\Omega} \sum \exp(\Omega h_{\Omega\Omega} + \sum_{\Omega=1}^{\Omega} \Omega_{\Omega\Omega} \Omega h_{\Omega\Omega}) \quad (2\text{-}8)$$

$$= \frac{1}{\Omega(\Omega)} \exp(\Omega v) \prod_{\Omega=1}^{\Omega} \sum (1 + \exp(\Omega + \sum_{\Omega=1}^{\Omega} \Omega_{\Omega\Omega} \Omega)$$

*(2-9)*

di mana φ adalah fungsi sigmoid φ (x) = 1 / (1 + exp (x)). Fungsi energy didefinisikan lebih vektor biner dan tidak cocok untuk data kontinu, tetapi dengan memodifikasi fungsi energi untuk menentukan *Gaussian - Bernoulli* RBM dengan memasukkan istilah kuadrat dari unit yang terlihat kami juga dapat menggunakan RBM pada data kontinu.

$$E(\mathbf{v}, \mathbf{h}; \theta) = \sum_{\Omega} \frac{\Omega - b_{\Omega}}{2\Omega^{\Omega}} - \sum_{\Omega} \Omega h - \sum_{\Omega} \frac{\Omega}{\Omega} hj \quad (2\text{-}10)$$

Vektor θ = {W, a, b, $\sigma^2$} dan $\sigma_{\Omega}$ repesentasi varian dari variable input merupakan $\Omega$. Vektor adalah bias dari unit terlihat dan b adalah bias dari unit tersembunyi. Distribusi marjinal atas vektor terlihat di v.

$$P(\mathbf{v}; \theta) = \sum_h \frac{\exp(-\Omega(\Omega h; \theta))}{\int \Omega \sum_{\Omega} \exp(-\Omega(\Omega h; \Omega)) \Omega} \quad (2\text{-}11)$$

P (v | h) menjadi *Gaussian multivariat* dengan *mean* $\Omega + \Omega \sum_{\Omega} \Omega h_{\Omega}$ dan diagonal kovarians matriks. Distribusi bersyarat untuk terlihat dan unit tersembunyi menjadi.

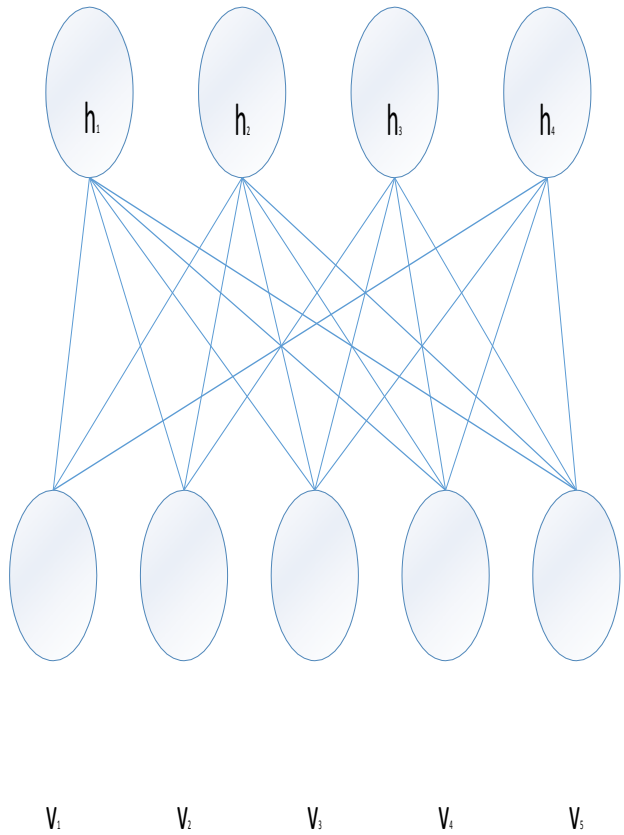
$$P(\Omega = x|h) = \frac{1}{\Omega\sqrt{2\pi}} \exp\left(-\frac{(\Omega - \Omega - \Omega \sum_{\Omega} \Omega h_{\Omega})^2}{2\Omega^2}\right),$$

$$P(h_{\Omega} = 1|v) = \phi\left(\Omega + \sum \frac{\Omega}{\Omega} wij\right), \quad (2\text{-}13)$$

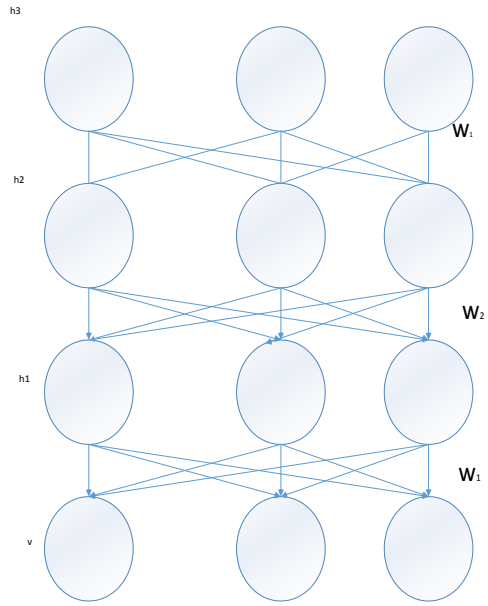
di mana φ (x) = 1 / (1 + exp (-x)) adalah fungsi aktivasi sigmoid. menyebutkan bahwa *Binomial Bernoulli* RBM juga bekerja untuk data kontinu jika data dinormalisasi untuk interval [0,1]. Ini mempunyai tes pada percobaan dan pengliahatan untuk bekerja baik.

### 2.6.3 Deep Belief Networks

Deep Belief Network di dua lapisan atas dimodelkan sebagai arahan *bipartite associative memory*, yaitu, RBM. Lapisan bawah merupakan :




*Gambar 2.3 : Gambar dari Restricted Boltzmann machine*




*Gambar 2.4 : Gambar dari deep belief network*

Sebuah model grapic tepat disebut juga dengan jaringan sigmoid Belief Network, Perbedaan antara jaringan sigmoid

Belief Network dan DBN adalah di parameterisasi lapisan tersembunyi.

$$P(v, h^1, \ldots, h^l) = P(h^{l-1}, h^l) \left(\prod_{k=0}^{l-2} P(h^k | h^{k+1})\right), \quad (2\text{-}15)$$

DBN dibuat untuk menumpuk RBM di atas satu sama lain. lapisan masing-masing RBM dalam tumpukan diatur ke lapisan tersembunyi dari RBM. Ketika belajar model untuk

satu set data, kita ingin mencari model $Q(h^1|h^{\hat{\diamond}-1})$ untuk
Pertama untuk tingkat atas $P(h^{\hat{\diamond}-1}|h^{\hat{\diamond}})$ adalah sebuah perkiraan
benar ke posterior benar, $P(h^{\hat{\diamond}-1}|h^{\hat{\diamond}})$ dimana ditingkat atas
tumpukan RBM memungkinkan kita untuk membuat kesimpulan yang tepat.

V adalah vektor unit yang dapat terlihat, P $(h^{\hat{\diamond}-1}|h^{\hat{\diamond}})$ adalah probabilitas bersyarat dari unit yang terlihat diberikan dan yang tersembunyi dalam RBM di tingkat k. dari distribusi gabungan di tingkat atas, P $(h^{\hat{\diamond}-1}|h^{\hat{\diamond}})$ , adalah RBM. Cara lain untuk menggambarkan DBN dengan model sederhana pada gambar 2.4 menjelaskan mengapa DBN adalah model generatif.

Garis Panah mewakili fitur *learning* sementara garis tanpa panah yang berarti bahwa DBN adalah model generatif. Garis Panah mewakili proses *learning* dan bukan bagian dari model. Garis tanpa panah menunjukkan bagaimana generasi data mengalir dalam jaringan. Generative Model tidak termasuk panah ke atas di lapisan bawah.

## 3.1 Deskripsi Umum

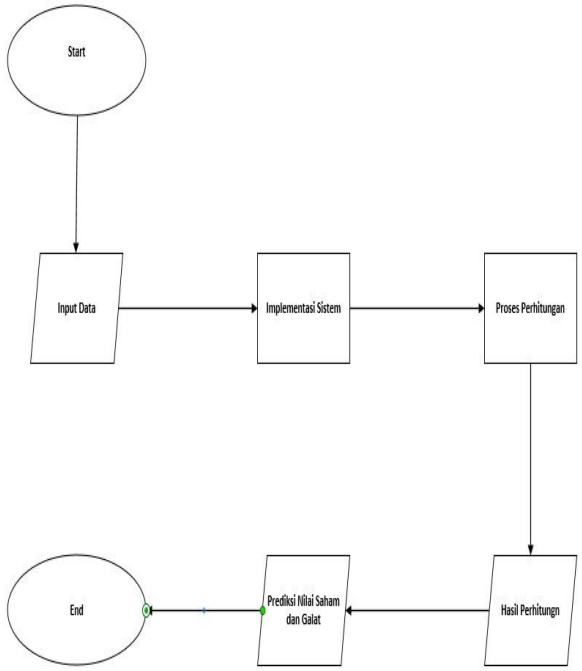
Dalam tugas akhir ini akan dirancang sebuah sistem untuk memprediksi indeks harga saham menggunakan metode DBN (*Deep Belief network*) yang teridi dari tumpukan RBM (*Restricted Boltzman Machines*), pada tahap pertama akan mencoba kombinasi dari banyak nya pengunaan RBM (*Restricted Boltzman Machines*) Dengan nilai yang berbeda. Tahap kedua *Output* dari system adalah nilai prediksi indeks harga saham. Dalam setiap proses RBM (*Restricted Boltzman Machines*) untuk menentukan perbandingan prediksi yang digunakan pada algoritma DBN yang mendapatkan nilai yang optimal untuk fungsi RBM

## 3.2 Pengumpulan Data

Untuk menyelesaikan permasalahan yang dibahas sebelumnya dikumpulkan data yang dibutuhkan antara lain data Saham IHSG historis per hari selama 5 tahun yang di ambil dari www.duniainvestasi.com/bei.

## 3.3 Perancangan Sistem

Sistem prediksi yang akan dibangun dibagi menjadi tiga tahap proses utama. Tahap pertama adalah perhitungan dengan variasi RBM (*Restricted Boltzman Machines*) yang akan dijelaskan pada subbab 3.4 Implementasi Sistem, yang terdiri dari perubahan perhitungan dengan macam variasi RBM. Kemudian pada tahap kedua yaitu proses perhitungan dan hasil prediksi dari masing-masing macam implementasi sitem yang akan menghasilkan nilai prediksi harga saham.pada tahap ketiga adalah mengupulkan dan membandingkan hasil RBM yang lebih baik untuk prediksi.




**Gambar 3.1 : perancangan system prediksi harga saham**

a.) Input Data : Data historis saham IHSG
Data yang digunakan adalah data pembukan, tertinggi, terendah dan penutup saham selama 5 tahun terhitung tanggal 4 Januari 2010 – 31 Agustus 2015.

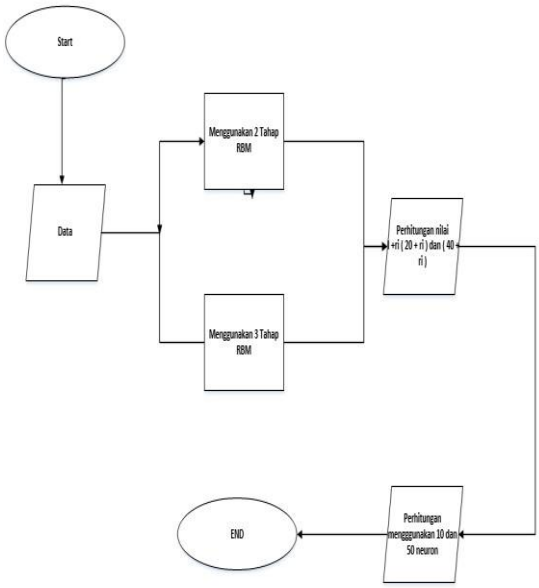
b.) Implementasi Sistem
akan dijabarkan pada subbab selanjutnya (3.4).

c.) Proses Perhitungan
Proses ini menghitung masing masing RBM pada implementasi sistem.

d.) Hasil Prediksi
Proses ini menghasilkan prediksi dari proses perhitungan implementasi system.

e.) Prediksi Saham dan nilai galat
Ini akan didapatkan dari hasil implementasi sistem dan akan membandingkan hasil RBM terbaik dalam prediksi saham.

## 3.4 Implementasi Sistem




**Gambar 3.2 : implementasi sistem prediksi harga saham**

### 3.5 Prediksi Data Menggunakan Training

Pada tahap *training* menggunakan tiga tahap yaitu inputnum, hiddennum, dan outputnum, pada tahap pertama yaitu intputnum yaitu membuat matriks, matriks pada input dipengaruhi oleh niali RI, nilai yang dipakai yaitu ( ri (i-1) = log10 (dataT(i)./dataT(i-1)). Pada tahap kedua adalah hiddennum yaitu jumlah RBM (*Restricted Boltzman Machines*) yang dipakai dengan nilai neuron pada masing-masing hiddennum atau RBM. Pada tahap ketiga yaitu nilai outputnum yaitu nilai prediksi saham dengan hasil yang dipengaruhi oleh inputnum dan hiddennum.

## 4 Hasil Dan Pengujian

### 4.1 Pengujian Sistem

Seperti yang telah dijelaskan pada bab sebelumnya, Tugas Akhir ini bertujuan untuk menerapkan metode DBN dengan metode menumpuk RBM pada prediksi saham IHSG. Pada tahap pertama yaitu memprediksi nilai saham menggunakan (𝑖 + 𝑟𝑖) nilai i dengan 20 dan 40 data sebelumnya, kemudian setiap tahap (𝑖 + 𝑟𝑖) yaitu I akan memprediksi dengan 2 RBM dan 3 RBM dengan neuron yang berbeda yaitu ($n = 20$) dan

($n = 40$) hasil prediksi tahap pertama dan kedua akan di bandingkan yang mengahasilkan nilai lebih baik.

### 4.2 Tujuan Pengujian Sistem

Tujuan dilakukannya pengujian pada tugas akhir ini adalah sebagai berikut:

- Menganalisis nilai pediksi saham yang didapat dari sistem dengan menguji nilai harga penutup IHSG.
- Mendapatkan nilai prediksi yang mendekati nilai aslinya

### 4.3 Skenario Pengujian

Pengujian pada Tugas Akhir ini menggunakan data historis IHSG. Selanjutnya akan digunakan beberapa skenario yang akan menguji sistem yang dibentuk.

Sekenario yang akan diuji yaitu membandingkan ( i + ri ) dengan 2 dan 3 RBM dengan neuron yang berbeda yaitu ($n = 10$) dan ($n = 50$)

#### 4.3.1 Skenario 1 ( Prediksi (n=10) (20+ri) vs Prediksi (n=10) (40+ri) )

Pada skenario ini akan membandingkan hasil prediksi untuk ( 20 + ri ) dengan ( 40+ ri ), menggunakan ($N = 10$) antara RBM dua tahap dan RBM tiga tahap. Jika nilai galat RBM ( 20 + ri ) lebih kecil berarti hasil prediksi ( 20 + ri ) lebih baik daripada ( 40 + ri ), jika hasilnya sebaliknya yaitu RBM ( 40 + ri ) lebih kecil maka untuk ( 40 + ri )memprediksi lebih baik.

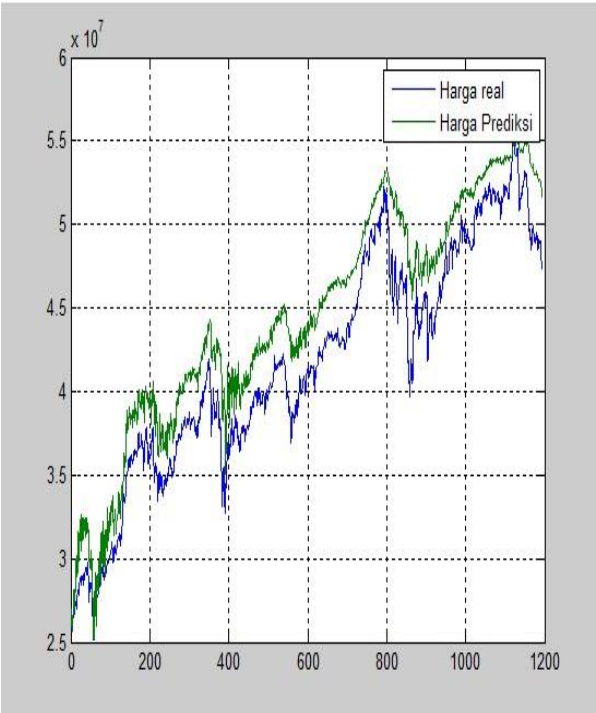
#### 4.3.2 Skenario 2 ( Prediksi (n=50) (20+ri ) vs Prediksi (n=50) (40+ri) )

Pada skenario ini akan membandingkan hasil prediksi untuk ( 20 + ri ) dengan ( 40+ ri ), menggunakan ($N = 50$) antara RBM dua tahap dan RBM tiga tahap. Jika nilai galat RBM ( 20 + ri ) lebih kecil berarti hasil prediksi ( 20 + ri ) lebih baik daripada ( 40 + ri ), jika hasilnya sebaliknya yaitu RBM ( 40 + ri ) lebih kecil maka untuk ( 40 + ri )memprediksi lebih baik.

### 4.4 Hasil dan Analisis

#### 4.4.1 Hasil Prediksi Untuk Skenario ($n = 10$).

Proses perhitungan nilai prediksi ($n = 10$) dilakukan dengan RBM dua layer dan RBM tiga layer untuk untuk setiap ($20 + ri$) dan ($40 + ri$) sebagai membandingkan data setiap nilai ($20 + ri$) akan mengahasilkan RBM dua tahap dan RBM tiga tahap dengan nilai N yang sama, nilai N yang digunakan adalah ($n = 10$) digunakan pada ($20 + ri$) dan ($40 + ri$) pada setiap prediksi.




*Gambar 4.1 : Grafik Hasil dari RBM ( 40 + ri ) tiga layer ( n = 10 )*

**Tabel 4.1 : Hasil prediksi skenario ( 40 + ri ) dengan ( n + 10 )**

| Skanario | RBM ( 20 + ri ) | | | | | |
|---|---|---|---|---|---|---|
| | RBM Dua Layer | | | RBM Tiga layer | | |
| | No | RMSE | MAPE | No | RMSE | MAPE |
| N = 10 | 1 | 0,1827 | | 1 | 0,1794 | |
| | 2 | 0,1828 | | 2 | 0,1813 | |
| | 3 | 0,1839 | | 3 | 0,1803 | |
| | 4 | 0,1827 | | 4 | 0,1783 | |
| | 5 | 0,1827 | | 5 | 0,1785 | |
| | 6 | 0,1826 | | 6 | 0,1763 | |
| | 7 | 0,1834 | | 7 | 0,1821 | |
| | 8 | 0,1827 | | 8 | 0,1763 | |
| | 9 | 0,1829 | | 9 | 0,1823 | |
| | 10 | 0,1828 | | 10 | 0,1726 | |
| | Rata – rata | 0,18291 | | Rata - rata | 0,17912 | |

Setelah pengujian maka di dapatkan Hasil table (4.1) dengan RBM ( 20 + ri ) dengan ( n = 10 ), pada hasil pertama didapatkan RBM dua layer dengan 10 kali percobaan diapatkan rata- rata RMSE nya adalah 0,18291 dengan nilai terkecil 0,1826, sedangkan pada tahap kedua menggunakan RBM tiga layer didapatkan rata-rata 0,17912 dengan nilai terkecil 0,1763.

Tabel 4.2 : Hasil prediksi skenario ( 40 + ri ) dengan ( n + 10 )

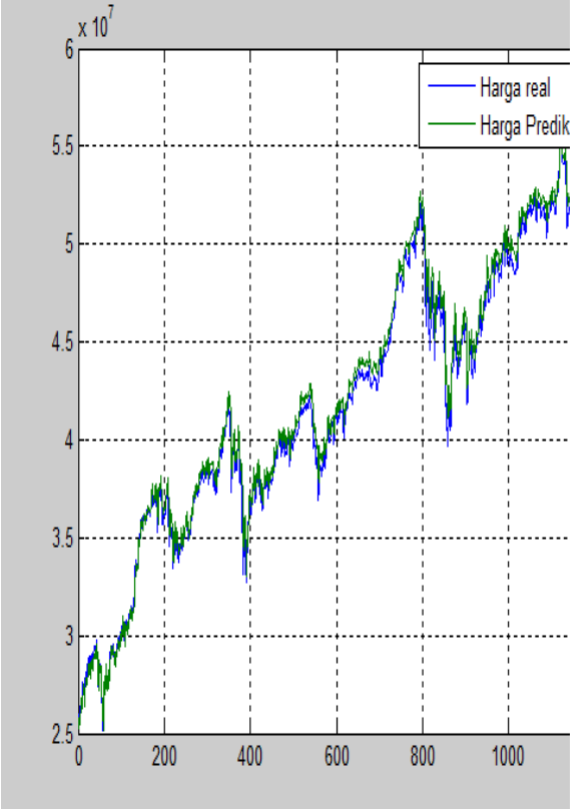
| Skanario | RBM ( 40 + ri ) | | | | | |
|---|---|---|---|---|---|---|
| | RBM dua layer | | | RBM Tiga Layer | | |
| | No | RMSE | MAPE | No | RMSE | MAPE |
| N = 10 | 1 | 0,1805 | | 1 | 0,1723 | |
| | 2 | 0,1801 | | 2 | 0,1739 | |
| | 3 | 0,1802 | | 3 | 0,1763 | |
| | 4 | 0,1809 | | 4 | 0.1562 | |
| | 5 | 0,1802 | | 5 | 0.1743 | |
| | 6 | 0,1801 | | 6 | 0.1638 | |
| | 7 | 0,1803 | | 7 | 0.1775 | |
| | 8 | 0,1806 | | 8 | 0.1757 | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 9 | 0,1802 | | 9 | 0.1699 | |
| | 10 | 0,1809 | | 10 | 0.1719 | |
| | Rata – rata | 0,1804 | | Rata – rata | 1,7118 | |

Setelah pengujian maka di dapatkan Hasil table (4.2) dengan RBM ( 20 + ri ) dengan ( n = 10 ), pada hasil pertama didapatkan RBM dua layer dengan 10 kali percobaan diapatkan rata- rata RMSE nya adalah 0,1804 dengan nilai terkecil 0,1801, sedangkan pada tahap kedua menggunakan RBM tiga layer didapatkan rata-rata 0,17118 dengan nilai terkecil 0,1562.

### 4.4.2 Hasil Prediksi Untuk Skenario ( $n = 50$ ).

Proses perhitungan nilai prediksi ($n = 50$) dilakukan dengan RBM dua layer dan RBM tiga layer untuk pembanding data, setiap nilai ( $20 + ri$ ) dan ( $40 + ri$ ) sebagai mengahasilkan RBM dua tahap dan RBM tiga tahap dengan nilai N yang sama, nilai N yang digunakan adalah ($n = 10$) digunakan pada ( $20 + ri$ ) dan ( $40 + ri$ ) pada setiap prediksi.




Gambar 4.2 : Grafik Hasil dari RBM (40 +ri) tiga layer (n = 50) Tabel 4.1 : Hasil prediksi skenario ( 20 + ri ) dengan ( n + 50 )

| Skanario | RBM ( 20 + ri ) | | | | | |
|---|---|---|---|---|---|---|
| | RBM Dua Layer | | | RBM Tiga Layer | | |
| | No | RMSE | MAPE | No | RMSE | MAPE |
| N = 50 | 1 | 0,1739 | | 1 | 0,0717 | |
| | 2 | 0,1717 | | 2 | 0,0587 | |
| | 3 | 0,1607 | | 3 | 0,0477 | |
| | 4 | 0,1695 | | 4 | 0,0472 | |
| | 5 | 0,1743 | | 5 | 0,0596 | |
| | 6 | 0,1626 | | 6 | 0,0674 | |
| | 7 | 0,1744 | | 7 | 0,0558 | |
| | 8 | 0,1684 | | 8 | 0,0547 | |
| | 9 | 0,1747 | | 9 | 0,0536 | |
| | 10 | 0,1694 | | 10 | 0,0728 | |
| | Rata – rata | 0,16996 | | Rata - rata | 0,05892 | |

Setelah pengujian maka di dapatkan Hasil table (4.3) dengan RBM ( 20 + ri ) dengan ( n = 50 ), pada hasil pertama didapatkan RBM dua layer dengan 10 kali percobaan diapatkan rata- rata RMSE nya adalah 0,16996 dengan nilai terkecil 0,1607, sedangkan pada tahap kedua menggunakan RBM tiga layer didapatkan rata-rata 0,05892 dengan nilai terkecil 0,0472.

Tabel 4.1 : Hasil prediksi skenario ( 40 + ri ) dengan ( n + 50 )

| | RBM ( 40 + ri ) |
|---|---|
| | |

| Skanario | RBM dua layer | | | RBM Tiga Layer | | |
|---|---|---|---|---|---|---|
| | No | RMSE | MAPE | No | RMSE | MAPE |
| N = 50 | 1 | 0,1781 | | 1 | 0,1761 | |
| | 2 | 0,1775 | | 2 | 0,1762 | |
| | 3 | 0,1775 | | 3 | 0,1621 | |
| | 4 | 0,1772 | | 4 | 0,1754 | |
| | 5 | 0,1777 | | 5 | 0,6609 | |
| | 6 | 0,1782 | | 6 | 0,1623 | |
| | 7 | 0,1785 | | 7 | 0,1724 | |
| | 8 | 0,1784 | | 8 | 0,1697 | |
| | 9 | 0,1781 | | 9 | 0,1730 | |
| | 10 | 0,1769 | | 10 | 0,1617 | |
| | Rata – rata | 0,17781 | | Rata - rata | 0,16904 | |

Setelah pengujian maka di dapatkan Hasil table (4.4) dengan RBM ( 40 + ri ) dengan ( n = 50 ), pada hasil pertama didapatkan RBM dua layer dengan 10 kali percobaan diapatkan rata- rata RMSE nya adalah 0,17781 dengan nilai terkecil 0,1769, sedangkan pada tahap kedua menggunakan RBM tiga layer didapatkan rata-rata 0,16904 dengan nilai terkecil 0,1621.

## 5. Kesimpulan dan Saran

### 5.1 Kesimpulan

Berdasarkan analisis terhadap implementasi sistem dan pengujian sistem prediksi indeks harga saham IHSG dengan metode *Deep Belief Network* (DBN), maka dapat ditarik kesimpulan sebagai berikut :

1. Metode RBM tiga layer menunjukan akurasi yang lebih baik dibandingkan dengan akurasi RBM tiga layer.
2. Berdasarkan hasil prediksi indeks harga saham dengan metode RBM yang mempengaruhinya adalah jumlah RBM layer dengan menggunakan lebih 2 RBM maka hasil akan baik, selain RBM yang mempengaruhi adalah jumlah *Neuron* semakin besar nilai *Neuron* semakin baik. Dan untuk ( i + ri ) tidak terlalu mempengaruhi karana tergantung dengan data yang digunakan.

### 5.2 Saran

Setelah dilakukan beberapa percobaan, saran untuk penelitian tugas akhir ini adalah metode *Deep Belief Network* (DBN) menggunakan satu skanario saja pada tahap ( i + ri ) karana tidak mempengaruhi hasil prediksi.

## Daftar Pustaka


[1] Batres-Estrada, G, “A tutorial on Support Vector Reggresion,” *Statistics and Computing 14*, 2004.

[2] Bursa Efek Indonesia, “IDX,” [Online]. Avaliable : http://www.idx.co.id/id-id/beranda/informasi/bagiinvestor/indeks.aspx. [Diakses 2 3 2015].

[3] Bursa Efek Indonesia, “IDX”, [Online]. Avaliable : http://www.idx.co.id/id-id/beranda/informasi/bagiinvestor/saham.aspx. [Diakses 2 3 2015].

[4] C. Cheng, W. Xu, dan J.Wang, “A Comparsion of Ensemble Methods in Financial Market Prediction,” *In 2012 Fifth International Joint Conference on Computational Sciences and Optimation (CSO)*, 2012.

[5] K. Miao, F. Chen, dan Z.-g. Zhao, “ Stock Price Forecast Based on Bacterial Colony RBF neural network,” *Journal of Qiangdao University*, 2007

[6] “Repositori Univesitas Sumatera Utara : Prediksi,” [Online]. Available: http://repository.usu.ac.id/bitstream/123456789/42955/8/Chapter%20I.pdf. [Diakses 2 3 2014].

[7] “Repository Universitas Sumatera Utara : Saham,” [Online]. Available: http://repository.usu.ac.id/bitstream/123456789/27690/4/Chapter%20II.pdf. [Diakses 6 3 2014].


.