

ANALISIS PREDIKSI MAHASISWA YANG TIDAK MENYELESAIKAN PENDIDIKAN DENGAN MENGGUNAKAN METODE DATA MININ

LINK DOWNLOAD [1.14 MB]

ANALISIS PREDIKSI MAHASISWA YANG TIDAK MENYELESAIKAN PENDIDIKAN DENGAN MENGGUNAKAN METODE DATA MINING

Hotler Manurung

Dosen Tetap Akademi Teknologi Industri Immanuel Medan

NIDN : 01180268

Abstrak

Data mining merupakan penambangan atau penemuan informasi baru dengan memprediksi, mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar yang diharapkan dapat mengatasi kondisi tersebut. Dalam hal ini penggunaan jaringan syaraf tiruan merupakan salah satu dari sistem proses informasi yang dirancang dengan meniru cara kerja otak manusia dalam menyelesaikan masalah dengan melakukan memproses belajar melalui perubahan bobot sinaps. jaringan syaraf tiruan dapat melakukan aktivitas pengakuan dasar pada data masa lalu. Data masa lalu akan dipelajari oleh jaringan syaraf tiruan, sehingga memiliki kemampuan untuk memberikan keputusan terhadap data yang belum pernah dipelajari. Sasaran yang ingin dicapai adalah mampu memprediksi mahasiswa yang tidak menyelesaikan pendidikan secara tepat waktu, dengan mengacu pada data diawal semester, lalu melihat data mahasiswa semester berikutnya selama 3 semester berturut, diharapkan dapat menghasilkan informasi tentang akumulasi jumlah mahasiswa diawal, di pertengahan semester berjalan maupun diakhir. Proses dengan menggunakan 9 variabel sebagai variabel masukan selama 3 semester berturut yang masing-masing berisi 3 variabel input data hadir, 3 variabel data nilai dan 3 variabel data bayar. Kemudian variabel dilatih dan diuji oleh jaringan syaraf tiruan yang dapat memprediksi hasil analisis dengan baik dengan menggunakan algoritma backpropagation. Setelah dilakukan pelatihan dengan algoritma perambatan galat mundur (backpropagation) mendapat hasil yang baik dengan parameter belajar jaringan konstanta 0,001 dan konvergen, yaitu pada epoch ke-50000.

Kata-kata Kunci: Prediksi Mahasiswa, Data Mining, Backpropagation.

Pendahuluan

Metode data mining yang dituangkan ke dalam jaringan syaraf tiruan merupakan penambangan atau penemuan informasi baru dengan memprediksi, mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar yang diharapkan dapat mengatasi kondisi tersebut.

Dalam hal ini penggunaan jaringan syaraf tiruan merupakan salah satu dari sistem proses informasi yang dirancang dengan meniru cara kerja otak manusia dalam menyelesaikan masalah dengan melakukan memproses belajar melalui perubahan bobot sinaps. Jaringan syaraf tiruan dapat melakukan aktivitas pengakuan dasar pada data masa lalu.

Pembahasan

Data dilatihkan dengan dipengaruhi oleh model algoritma jaringan yang digunakan. Jumlah lapisan tersembunyi, data hadir, data nilai dan data pembayaran. Untuk menghasilkan iterasi tercepat, nilai-nilai dari jumlah lapisan tersembunyi diubah-ubah. Perubahan dilakukan dengan memberi nilai tetap pada salah satu item. Pelatihan data yang paling konvergen terjadi pada epoch ke-50000 dengan konfigurasi JST-PB sebagai berikut :

Jumlah sel lapisan tersembunyi : 100

Konstanta Belajar : 0,001

Besar Galat : 0,006

Hasil pelatihan dilakukan melalui 2 tahap, yaitu pelatihan terhadap data yang dilatihkan dan pengujian pada data yang baru yang belum pernah dilatihkan. Adapun data yang dilatihkan terdiri dari 50 data, dimana terdiri dari 25 data mahasiswa yang tamat dan 25 data mahasiswa yang tidak tamat. Setelah itu, jaringan akan diuji dengan 25 data baru setelah data yang telah dilatihkan tersebut dikenali. Hal tersebut berfungsi untuk menguji seberapa besar JST mengenali data yang baru.

Hasil Pembahasan

Adapun hasil pelatihan I terlihat pada Gambar 1 sebagai berikut :

Gambar 1. Hasil pelatihan ke-I terhadap data

latih pada JST

Gambar 1, menunjukkan bahwa proses pelatihan berhenti pada epoch ke 50000 dengan sintaks sebagai berikut :

```
tp=[50 50000 0.001 0.006]
```

```
[w1,b1,w2,b2]=initff(p,100,'logsig',t,'logsig')
```

```
[w1,b1,w2,b2]=trainbp (w1,b1,'logsig',w2,b2,'logsig',p,t,tp)
```

Durasi : 3 jam 10 menit (20.05-23.15 WIB)

```
a=simuff(r,w1,b1,'logsig',w2,b2,'logsig')
```

```
a=a'
```

hasil konvergen = 58 %

Setelah itu dilanjutkan dengan pelatihan data ke 2 yang terlihat pada Gambar 2.

Gambar 2. Hasil pelatihan ke-2 terhadap data latih pada JST

Gambar 2, menunjukkan bahwa proses pelatihan berhenti pada epoch ke 5228 dengan sintaks sebagai berikut :

```
tp=[50 100000 0.006 0.009]
```

```
[w1,b1,w2,b2]=initff(p,200,'logsig',t,'logsig')
```

```
[w1,b1,w2,b2]=trainbp (w1,b1,'logsig',w2,b2,'logsig',p,t,tp)
```

Durasi : 3 jam 10 menit (20.05-23.15 WIB)

```
a=simuff(r,w1,b1,'logsig',w2,b2,'logsig')
```

```
a=a'
```

hasil konvergen = 46 %

Setelah itu dilanjutkan dengan pelatihan data ke 3 yang terlihat pada Gambar 3.

Gambar 3. Hasil Pelatihan ke-3 Terhadap Data Latih Pada JST

Gambar 3, menunjukkan bahwa proses pelatihan berhenti pada epoch ke 51345 dengan sintaks sebagai berikut :

```
tp=[50 100000 0.004 0.001]
[w1,b1,w2,b2]=initff(p,200,'logsig',t,'logsig')
[w1,b1,w2,b2]=trainbp (w1,b1,'logsig',w2,b2,'logsig',p,t,tp)
Durasi : 3 jam 20 menit (20.05-23.25 WIB)
a=simuff(r,w1,b1,'logsig',w2,b2,'logsig')
a=a'
```

hasil konvergen = 54 %

Kemudian dilanjutkan dengan pelatihan data ke 4 yang terlihat pada Gambar 4.

Gambar 4. Hasil pelatihan ke-4 terhadap data latih pada JST

Gambar 4, menunjukkan bahwa proses pelatihan berhenti pada epoch ke 6097 dengan sintaks sebagai berikut :

```
tp=[50 50000 0.06 0.06]
[w1,b1,w2,b2]=initff(p,100,'logsig',t,'logsig')
[w1,b1,w2,b2]=trainbp (w1,b1,'logsig',w2,b2,'logsig',p,t,tp)
Durasi : 1 jam 5 menit (07.05-08.00 WIB)
a=simuff(r,w1,b1,'logsig',w2,b2,'logsig')
```

a=a'

hasil konvergen = 50 %

Kemudian diteruskan dengan pelatihan data ke 5 yang terlihat pada Gambar 5.

Gambar 5. Hasil pelatihan ke-5 terhadap data latih pada JST

Gambar 5, menunjukkan bahwa proses pelatihan berhenti pada epoch ke 110 dengan sintaks sebagai berikut :

```
tp=[30 3000 0.1 0.1]
```

```
[w1,b1,w2,b2]=initff(p,20,'logsig',t,'logsig')
```

```
[w1,b1,w2,b2]=trainbp (w1,b1,'logsig',w2,b2,'logsig',p,t,tp)
```

Durasi : 5 menit (08.05-08.10 WIB)

```
a=simuff(r,w1,b1,'logsig',w2,b2,'logsig')
```

```
a=a'
```

hasil konvergen = 50 %

Kemudian diteruskan dengan pelatihan data ke 6 yang terlihat pada Gambar 6.

Gambar 6. Hasil pelatihan ke-6 terhadap data latih pada JST

Gambar 6, menunjukkan bahwa proses pelatihan berhenti pada epoch ke 6307 dengan sintaks sebagai berikut :

```
tp=[40 40000 0.1 0.1]
[w1,b1,w2,b2]=initff(p,50,'logsig',t,'logsig')
[w1,b1,w2,b2]=trainbp (w1,b1,'logsig',w2,b2,'logsig',p,t,tp)
Durasi : 30 menit (08.05-08.35 WIB)
a=simuff(r,w1,b1,'logsig',w2,b2,'logsig')
a=a'
```

hasil konvergen = 52 %

Hasil pelatihan data latih dapat dilihat pada Tabel 1 dari 50 data baru yang telah diujikan, adapun pengujian terhadap data latih dilakukan sebanyak 6 kali pelatihan dan pengujian dengan angka input, besar galat dan konstanta belajar yang berbeda-beda. Tampak bahwa hasil pengujian data latih yang pertama 29 data (58 %) sesuai dengan target sedangkan 21 data (42 %) tidak sesuai dengan target. Setelah itu dilakukan kembali pengujian data kedua, dimana hasilnya menunjukkan bahwa 23 data (46 %) sesuai dengan target sedangkan 27 data (54 %) tidak sesuai dengan target. Setelah itu dilakukan kembali pengujian data ketiga, dimana hasilnya menunjukkan bahwa 27 data (54 %) sesuai dengan target sedangkan 23 data (46 %) tidak sesuai dengan target. Kemudian dilanjutkan dengan pengujian data keempat, dimana 25 data (50 %) sesuai target dan 25 (50 %) tidak sesuai. Adapun hasil pengujian data kelima hasilnya masih sama dengan pengujian data keempat dimana hasilnya menunjukkan 50 % data sesuai target dan 50 % lagi tidak sesuai. Selanjutnya dilakukan pengujian data latih keenam dimana hasilnya menunjukkan 26 data (52 %) sesuai target dan 24 data (48 %) tidak sesuai target.

Tabel 1. Hasil pengujian data latih ke dalam JST

Pengujian

Data Sesuai Target

Persentase

(%)

Data Tidak Sesuai Target

Persentase

(%)

I

29

58 %

21

42 %

II

23

46 %

27

54 %

III

27

54 %

23

46 %

IV

25

50 %

25

50 %

IV

25

50 %

25

50 %

VI

26

52 %

24

48 %

Berdasarkan hasil pelatihan dan uji coba terhadap data hadir, data nilai dan data pembayaran mahasiswa, menunjukkan data pada pengujian pertama merupakan data yang tertinggi dengan tingkat konvergen 58 %. Adapun data kehadiran dan nilai yang bagus tidak selalu menjamin bahwasanya mahasiswa tersebut akan menyelesaikan pendidikannya atau tidak. Hal ini dapat disebabkan karena banyak faktor lain yang mempengaruhi proses belajar di STMIK Kaputama itu sendiri antara lain dapat disebabkan faktor lingkungan, dosen, staf/karyawan, sarana dan prasarana dan lain-lain.

Adapun hasil pengujian terhadap data yang baru yang paling konvergen dapat dilihat pada Tabel 2. (Lampiran 1).

Setelah ke 50 data latih tersebut diuji selama 6 kali pengujian dan didapat hasilnya 58% konvergen atau sesuai dengan target, selanjutnya, jaringan akan diuji dengan 25 data baru setelah data yang telah dilatihkan tersebut dikenali. Hasil pengujian terhadap 25 data yang baru tertuang dalam Tabel 3 (Lampiran 2)

Berdasarkan pada Tabel 3 dapat dianalisis bahwa hasil ketepatan jaringan syaraf mendekati 100%, yaitu 0,92% data prediksi telah konvergen sesuai target.

Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan dapat diambil beberapa kesimpulan antara lain :

Metode data mining yang dituangkan kedalam jaringan syaraf tiruan dengan penggunaan algoritma backpropagation pada penelitian ini mampu untuk memprediksi mahasiswa apakah akan menyelesaikan pendidikannya atau tidak.

Untuk mendapatkan hasil prediksi mahasiswa yang tidak menyelesaikan pendidikan pada jaringan syaraf tiruan yang terbaik, perlu memperhatikan jumlah simpul lapisan tersembunyi, laju pembelajaran, toleransi galat dan jumlah epoch.

- Target output pada proses pengujian terhadap data baru yang akan diprediksi yang diinginkan telah mendekati ketepatan pada hasil karena telah terjadi penyesuaian nilai bobot dan bias yang semakin baik pada proses pelatihan

Daftar Pustaka

Kusrini dan Luthfi Emha Taufiq, 2009, Algoritma Data Mining, Andi Offset, Yogyakarta

Sri Kusumadewi, 2003, Artificial Inteligenci, Graha Ilmu, Yogyakarta

Wibowo, 2007: Performance,

<http://www.scrib.com>, tahun, 2012

Lampiran 1.

Tabel 2. Hasil analisa prediksi terhadap data yang telah dilatihkan sebelumnya

NPM

Convert Data Hadir

Convert Data

Nilai

Convert Data

Bayar

Target

HPL

KSB

Sem.

I

Sem.

II

Sem. III

Sem.

I

Sem. II

Sem. III

Sem.

I

Sem.

II

Sem.

III

0631003

4

4

3

2.75

2.63

2.72

2

4

4

3

2.75

2.63

0631022

4

3

4

3.20

2.95

3.06

2

2

1

1

0.0

tamat

0631024

3

4

4

2.80

2.68

3.11

2

2
2
1
1.0
tamat

0631027
4
4
4
3.20
3.16
3.22
2
2
2
1
0.0
tamat

0631028
4
4
4
3.05
3.16
3.17
2
2
2
1
0.9
tamat

0631029
2
4
4
2.95
3.05
3.22
2
2
2
1
0.0
tamat

0631049

4

4

4

2.95

3.00

3.17

2

2

2

1

0.2

tamat

0632001

2

3

4

3.00

2.75

3.45

2

2

2

1

1.0

tamat

0632002

2

4

4

2.80

3.10

3.30

2

2

2

1

1.0

tamat

0632003

3

4

4
3.00
3.10
3.15
2
2
2
1
0.1
tamat

0632007
3
4
4
3.70
3.55
3.75
2
2
2
1
1.0
tamat

0632016
3
4
4
3.60
3.65
3.65
2
2
2
1
1.0
tamat

0632024
3
3
3
3.40
3.35
3.65
1

1
2
1
1.0
tamat

0633002

3
4
4
2.85
2.50
2.84
2
2
2
1
0.1
tamat

0633003

4
4
4
2.60
3.50
2.79
2
2
2
1
1.0
tamat

0633011

4
4
4
3.10
2.95
3.05
2
2
2
1
0.0
tamat

0641007

3

4

3

3.35

3.05

2.83

2

2

2

1

0.0

tamat

0642002

4

4

4

3.00

3.60

3.26

2

2

2

1

0.3

tamat

0642005

3

4

4

2.40

2.90

2.95

2

2

2

1

0.1

tamat

0642018

3

4

4
2.60
3.20
3.26
2
2
2
1
0.8
tamat

0731034

3
4
4
3.00
3.42
3.11
2
2
2
1
1.0
tamat

0731040

3
4
4
2.70
3.16
3.06
2
2
2
1
0.7
tamat

0731041

4
3
4
2.75
2.79
3.00
2

2
2
1
1.0
tamat

0732001
3
4
4
2.89
3.05
3.20
2
2
2
1
0.0
tamat

0732002
3
4
4
2.78
2.95
2.85
2
2
2
1
0.0
tamat

0631042
1
1
1
0.00
0.00
0.00
2
1
1
0
0.0
tdk tmt

0631047

1
1
1
0.00
0.00
0.00
2
2
1
0
0.0
tdk tmt

0631048

1
1
1
3.05
2.42
2.28
2
2
2
0
1.0
tdk tmt

0632004

3
2
3
2.89
2.84
2.90
2
2
2
0
0.0
tdk tmt

0633007

1
1

1
1.2
0.00
0.42
1
1
1
0
0.9
tdk tmt

0633009
1
1
1
1.50
0.60
0.53
2
2
2
0
0.0
tdk tmt

0641001
2
1
2
1.85
1.53
2.56
2
2
1
0
0.2
tdk tmt

0641012
1
1
1
0.3
0.47
0.17
2

2
1
0
1.0
tdk tmt

0641018
1
2
4
2.65
2.95
3.11
2
2
2
0
1.0
tdk tmt

0641020
1
3
3
2.30
2.26
2.78
2
2
1
0
0.1
tdk tmt

0731001
1
4
1
3.15
3.37
0.33
2
2
2
0
1.0
tdk tmt

0731004

1

3

2

2.25

2.42

1.06

2

2

1

0

1.0

tdk tmt

0731010

1

1

4

2.55

2.21

2.39

2

2

2

0

1.0

tdk tmt

0732016

1

1

1

0.56

0.00

0.30

1

1

1

0

0.1

tdk tmt

0732018

1

1

1
2.61
2.20
0.80
1
1
1
0
1.0
tdk tmt

0732019
1
1
1
3.17
0.60
0.80
1
1
1
0
0.0
tdk tmt

0732021
1
1
1
0.22
0.00
0.10
1
1
1
0
0.0
tdk tmt

0733005
1
1
1
1.50
0.35
0.11
2

1
1
0
0.3
tdk tmt

0733010
1
1
1
0.00
0.00
0.00
1
1
1
0
0.1
tdk tmt

0733012
1
1
1
0.00
0.00
0.00
1
1
1
0
0.8
tdk tmt

0733013
1
1
1
0.10
0.00
0.00
1
1
1
0
1.0
tdk tmt

0733014

1
1
1
0.10
0.00
0.00
1
1
1
0
0.7
tdk tmt

0733016

1
1
1
0.00
0.00
0.00
2
1
1
0
1.0
tdk tmt

0733017

1
1
1
1.55
0.00
0.00
2
1
1
0
0.0
tdk tmt

0741003

1
1

1
0.00
0.00
0.00
2
1
1
0
0.0
tdk tmt

Keterangan Tabel 2:

HPL : Hasil Pelatihan
KSB : Kondisi Sebenarnya

Lampiran 2.

Tabel 3. Hasil analisa prediksi terhadap data yang akan di uji

NPM
Convert Data Hadir
Convert Data

Nilai
Convert Data

Bayar
T
HPU
KSB
KJST

Sem

I
Sem

II
Sem III
Sem. I
Sem. II

Sem. III
Sem. I
Sem. II
Sem. III

10441029
3
1
1
2.70
2.74
0.33
2
2
1
0
0.0
tdk tmt
Benar

10441030
4
2
1
1.95
1.53
0.67
2
2
1
0
0.0
tdk tmt
Benar

10441007
4
4
4
2.95
2.89
2.94
2
2
2
1
1.0

tamat
Benar

10451039

3
2
4
2.85
1.20
2.37
2
1
1
1
0.6
tamat
Benar

10451022

3
3
3
2.70
2.55
2.68
1
2
2
1
0.7
tamat
Benar

10331018

1
1
1
1.45
1.60
0.68
1
2
1
0
0.0
tdk tmt
Benar

10331009

3

2

2

3.45

2.10

2.42

1

2

1

0

0.2

tdk tmt

Benar

10311069

4

4

3

3.05

2.84

1.78

2

2

2

1

1.0

tamat

Benar

10311001

4

4

4

3.00

2.89

2.50

1

2

1

1

0.9

tamat

Benar

10311052

2

2

4

0.65

2.47

2.56

2

2

2

0

0.1

tdk tmt

Benar

10311056

4

4

4

2.70

2.74

2.56

1

1

2

1

1.0

tamat

Benar

09441059

4

3

3

2.75

2.63

2.50

2

2

1

1

0.6

tamat

Benar

09441055

4

4
2
3.35
2.84
2.61
2
2
2
1
1.0
tamat
Benar

08451005
3
3
2
2.60
0.00
2.21
1
2
2
1
0.8
tamat
Benar

09451047
4
3
3
2.80
2.90
2.68
1
2
2
1
0.7
tamat
Benar

09451042
3
1
1

3.10
2.15
1.21
2
2
2
0
0.0
tdk tmt
Benar

09451016
3
2
2
3.10
3.00
2.63
2
2
1
0
0.0
tdk tmt
Benar

09321015
3
2
4
2.72
2.50
2.70
2
2
1
1
0.5
tamat
Salah

09321014
2
2
4
2.67
2.50

2.60
2
2
2
0
0.3
tdk tmt
Benar

09311039
3
3
3
2.45
2.79
2.72
2
2
2
1
0.8
tamat
Benar

08451035
3
4
4
2.70
2.65
2.42
2
2
2
1
1.0
tamat
Benar

08457040
3
3
3
3.05
2.90
2.21
2

2
2
1
0.7
tamat
Benar

11457101
4
4
4
3.05
2.55
2.74
2
2
2
1
1.0
tamat
Benar

0742017
1
3
3
1.35
0.00
2.05
2
2
2
1
0.1
tamat
Salah

0742021
1
1
3
2.05
1.40
1.00
2
2
2

0
0.0
tdk tmt
Benar

Keterangan Tabel 3:

T : Target
HPU : Hasil Pengujian
KSB : Kondisi Sebenarnya
KJST : Ketepatan Jaringan Syaraf Tiruan