

**DIKLAT**  
**HMIF**

2018

# PEMBELAJARAN MESIN

---

AIK21356

SEMESTER 5



**DAFTAR ISI**

|                                      |    |
|--------------------------------------|----|
| DAFTAR ISI .....                     | 1  |
| UJIAN AKHIR SEMESTER 2011/2012.....  | 3  |
| SOAL LATIHAN.....                    | 4  |
| UJIAN AKHIR SEMESTER 2013/2014.....  | 6  |
| UJIAN TENGAH SEMESTER 2014/2015..... | 8  |
| UJIAN TENGAH SEMESTER 2015/2016..... | 9  |
| UJIAN AKHIR SEMESTER 2015/2016.....  | 10 |
| UJIAN TENGAH SEMESTER 2016/2017..... | 12 |
| UJIAN AKHIR SEMESTER 2016/2017.....  | 13 |
| UJIAN TENGAH SEMESTER 2016/2017..... | 14 |
| UJIAN TENGAH SEMESTER 2017/2018..... | 15 |
| LATIHAN.....                         | 16 |
| UJIAN AKHIR SEMESTER 2017/2018.....  | 17 |

**UJIAN AKHIR SEMESTER 2011/2012**

**Ujian Akhir Semester Gasal 2011/2012**  
**Program Studi Teknik Informatika**  
**Fakultas MIPA Universitas Diponegoro Semarang**

Mata Kuliah : Data Mining

Beban : 3 SKS

Semester : 5

Dosen : Sutikno, S.T., M.Cs.

Hari/ Tanggal : Kamis, 19 Januari 2012

Jam : 13.00 – 14.30 WIB

Waktu : 90 Menit

Sifat : Buku Terbuka

**Perhatian:** Segala bentuk **Kecurangan** (kerja sama, meminjam/memberi pinjaman catatan dan atau alat tulis, mencontek dan lain-lain) akan diberikan nilai **0 (nol)**.

1. Perhatikan data alumni, bobot atribut, dan kedekatan antar atribut pada tabel dibawah ini.

**Tabel Data Alumni**

| Mahasiswa | Asal SMA | Nilai Ujian Masuk | Nilai Ujian Nasional | Lama studi (semester) |
|-----------|----------|-------------------|----------------------|-----------------------|
| A         | Negeri   | A                 | A                    | 8                     |
| B         | Swasta   | B                 | B                    | 10                    |
| C         | Negeri   | B                 | B                    | 9                     |
| D         | Swasta   | A                 | A                    | 9                     |

**Tabel Bobot tiap-tiap atribut**

| Atribut             | Bobot |
|---------------------|-------|
| Asal SMA            | 0.25  |
| Nil. Ujian Masuk    | 1     |
| Nil. Ujian Nasional | 0.75  |

**Tabel Kedekatan nilai atribut Asal SMA**

| Asal SMA | Asal SMA | Kedekatan |
|----------|----------|-----------|
| Negeri   | Negeri   | 1         |
| Swasta   | Swasta   | 1         |
| Negeri   | Swasta   | 0.5       |

**Tabel Kedekatan nilai atribut Nilai Ujian Masuk**

| Pendidikan | Pendidikan | Kedekatan |
|------------|------------|-----------|
| A          | A          | 1         |
| B          | B          | 1         |
| A          | B          | 0.4       |

**Tabel Kedekatan nilai atribut Nilai Ujian Nasional**

| Pendidikan | Pendidikan | Kedekatan |
|------------|------------|-----------|
| A          | A          | 1         |
| B          | B          | 1         |
| A          | B          | 0.4       |

Jika di ketahui data baru :

- Asal SMA : Swasta
- nilai ujian masuk : B
- nilai ujian nasionalnya : B

Berapa perkiraan lama studi mahasiswa tersebut jika di cari dengan menggunakan Algoritma *Nearest Neighbor* berdasarkan beberapa tabel diatas?

**SOAL LATIHAN**

Sebuah perguruan tinggi jurusan Teknik Informatika, akan melakukan prediksi lama masa studi terhadap mahasiswa semester 2. Untuk melakukan hal tersebut dengan memanfaatkan data dalam database nilai mata kuliah (semester 1) dan masa studi yang telah tersimpan selama 5 tahun. Data- data tersebut diantaranya yaitu seperti pada tabel berikut.

| NO | NIM       | Nilai |   |    | Lama Studi (semester) |
|----|-----------|-------|---|----|-----------------------|
|    |           | PWI   | M | BI |                       |
| 1  | J2F004001 | A     | B | B  | 10                    |
| 2  | J2F004002 | A     | A | B  | 8                     |
| 3  | J2F004003 | B     | D | B  | 13                    |
| 4  | J2F004004 | B     | A | C  | 12                    |
| 5  | J2F004005 | C     | C | D  | 14                    |
| 6  | J2F004006 | A     | E | A  | 9                     |
| 7  | J2F004007 | D     | D | A  | 13                    |
| 8  | J2F004008 | A     | A | E  | 10                    |
| 9  | J2F004009 | B     | A | D  | 12                    |
| 10 | J2F004010 | C     | B | A  | 13                    |

| NO | NIM       | Nilai |   |    | Lama Studi (semester) |
|----|-----------|-------|---|----|-----------------------|
|    |           | PWI   | M | BI |                       |
| 11 | J2F004011 | D     | D | A  | 13                    |
| 12 | J2F004012 | E     | E | E  | 14                    |
| 13 | J2F004013 | A     | A | B  | 8                     |
| 14 | J2F004014 | B     | C | C  | 13                    |
| 15 | J2F004015 | C     | D | A  | 13                    |
| 16 | J2F004016 | D     | E | A  | 13                    |
| 17 | J2F004017 | E     | A | D  | 14                    |
| 18 | J2F004018 | A     | A | A  | 7                     |
| 19 | J2F004019 | B     | A | D  | 12                    |
| 20 | J2F004020 | A     | B | A  | 8                     |

Keterangan Tabel:

PWI : Pemrograman Web dan Internet

M : Matematika

BI : Bahasa Inggris

Lama studi paling cepat 7 Semester dan paling lama 14 semester

- [1] Dari data-data tabel diatas, transformasikan data-data tersebut (3 baris awal) sebelum dilakukan pelatihan dengan menggunakan Jaringan Syaraf Tiruan *Backpropagation* dan Fungsi Aktivasi menggunakan *Sigmoid Biner*?
- [2] Gambarkan arsitektur jaringan syaraf tiruan *backpropagation* dari kasus diatas, jika digunakan 1 layer tersembunyi dan 2 *neuron* pada layer tersembunyi?
- [3] Hitunglah bobot-bobot jaringan baru setelah iterasi pertama dengan menggunakan data pola pertama dan di gunakan *alpha* (laju pelatihan) = 0.5 dan bobot-bobot awal sebagai berikut:  
Bobot-bobot jaringan dan bobot bias dari layer input ke hidden layer

|       | $Z_1$ | $Z_2$ |
|-------|-------|-------|
| $X_1$ | -0.1  | 0.2   |
| $X_2$ | 0.5   | -0.3  |
| $X_3$ | -0.1  | 0.8   |
| 1     | 0.4   | 0,1   |

Bobot-bobot jaringan dan bobot bias dari hidden layer ke layer output

|       | Y    |
|-------|------|
| $Z_1$ | 0.5  |
| $Z_2$ | -0.5 |

|          |     |
|----------|-----|
| <b>1</b> | 0.2 |
|----------|-----|

- [4] Hitung hasil prediksi lama studi suatu mahasiswa, dari arsitektur jaringan yang anda buat jika nilai mata kuliah semester 1 mahasiswa tersebut adalah Pemrograman Web dan Internet (PWI): B, Matematika (M): D, dan Bahasa Inggris (BI): B, serta bobot-bobot hasil pelatihan seperti pada tabel berikut:

Bobot-bobot jaringan dan bobot bias dari layer input ke hidden layer

|                      | <b>Z<sub>1</sub></b> | <b>Z<sub>2</sub></b> |
|----------------------|----------------------|----------------------|
| <b>X<sub>1</sub></b> | 0.1                  | 0.2                  |
| <b>X<sub>2</sub></b> | -0.5                 | 0.3                  |
| <b>X<sub>3</sub></b> | 0.2                  | -0.8                 |
| <b>1</b>             | 0.7                  | 0,1                  |

Bobot-bobot jaringan dan bobot bias dari hidden layer ke layer output

|                      | <b>Y</b> |
|----------------------|----------|
| <b>Z<sub>1</sub></b> | 0.1      |
| <b>Z<sub>2</sub></b> | -0.2     |
| <b>1</b>             | 0.3      |

**UJIAN AKHIR SEMESTER 2013/2014****Ujian Akhir Semester Gasal 2013/2014****Jurusan Ilmu Komputer/Informatika****Fakultas Sains dan Matematika Universitas Diponegoro Semarang****Mata Kuliah : Data Mining****Beban : 3 SKS****Dosen : - Nurdin Bahtiar, S.Si, MT****- Sutikno, S.T., M.Cs./****Hari/ Tanggal : Kamis, 9 Januari 2014****Jam : 13.00 – 14.30 WIB (90 Menit)****Sifat : Buku Terbuka**

**Perhatian:** Segala bentuk **KECURANGAN** (kerja sama, meminjam/memberi pinjaman catatan dan atau alat tulis, mencontek dan lain-lain) akan diberikan nilai **0 (NOL)**.

1. Perhatikan data alumni, bobot atribut, dan kedekatan antar atribut pada tabel dibawah ini.

**Tabel Data Alumni**

| Mahasiswa | Asal SMA | Nilai Ujian Masuk | Nilai Ujian Nasional | Predikat Kelulusan |
|-----------|----------|-------------------|----------------------|--------------------|
| Aji       | Negeri   | A                 | A                    | Cum laude          |
| Bima      | Swasta   | B                 | B                    | Memuaskan          |
| Candra    | Negeri   | B                 | A                    | Sangat Memuaskan   |
| Dimas     | Swasta   | A                 | A                    | Sangat Memuaskan   |

**Tabel Bobot tiap-tiap atribut**

| Atribut             | Bobot |
|---------------------|-------|
| Asal SMA            | 0.3   |
| Nil. Ujian Masuk    | 1     |
| Nil. Ujian Nasional | 0.7   |

**Tabel Kedekatan nilai atribut Asal SMA**

| Asal SMA | Asal SMA | Kedekatan |
|----------|----------|-----------|
| Negeri   | Negeri   | 1         |
| Swasta   | Swasta   | 1         |
| Negeri   | Swasta   | 0.5       |

**Tabel Kedekatan nilai atribut Nilai Ujian Masuk**

| Pendidikan | Pendidikan | Kedekatan |
|------------|------------|-----------|
| A          | A          | 1         |
| B          | B          | 1         |
| A          | B          | 0.5       |

**Tabel Kedekatan nilai atribut Nilai Ujian Nasional**

| Pendidikan | Pendidikan | Kedekatan |
|------------|------------|-----------|
| A          | A          | 1         |
| B          | B          | 1         |
| A          | B          | 0.5       |

Jika di ketahui data baru :

- Asal SMA : Swasta
- nilai ujian masuk : A
- nilai ujian nasionalnya : B

Apa perkiraan Predikat Kelulusan mahasiswa tersebut jika di cari dengan menggunakan Algoritma *Nearest Neighbor* berdasarkan beberapa tabel diatas?



2. Sebuah perguruan tinggi jurusan Teknik Informatika, akan melakukan prediksi Indeks Prestasi Kumulatif terhadap mahasiswa pada akhir semester 2. Untuk melakukan hal tersebut dengan memanfaatkan data dalam database nilai mata kuliah (semester 1) dan IPK semester 1 yang telah tersimpan selama 5 tahun. Data-data tersebut diantaranya yaitu seperti pada tabel berikut.

| NO | NIM       | Nilai |   |    | IPK |
|----|-----------|-------|---|----|-----|
|    |           | PWI   | M | BI |     |
| 1  | J2F004001 | A     | B | B  | 3,0 |
| 2  | J2F004002 | A     | A | B  | 3,3 |
| 3  | J2F004003 | B     | D | B  | 2,4 |
| 4  | J2F004004 | B     | A | C  | 2,8 |
| 5  | J2F004005 | C     | C | D  | 2,1 |
| 6  | J2F004006 | A     | E | A  | 2,6 |
| 7  | J2F004007 | D     | D | A  | 2,4 |
| 8  | J2F004008 | A     | A | E  | 2,7 |
| 9  | J2F004009 | B     | A | D  | 2,6 |
| 10 | J2F004010 | C     | B | A  | 2,9 |

| NO | NIM       | Nilai |   |    | IPK |
|----|-----------|-------|---|----|-----|
|    |           | PWI   | M | BI |     |
| 11 | J2F004011 | D     | D | A  | 2,1 |
| 12 | J2F004012 | E     | E | E  | 1,0 |
| 13 | J2F004013 | A     | A | B  | 3,2 |
| 14 | J2F004014 | B     | C | C  | 2,3 |
| 15 | J2F004015 | C     | D | A  | 2,1 |
| 16 | J2F004016 | D     | E | A  | 1,8 |
| 17 | J2F004017 | E     | A | D  | 2,0 |
| 18 | J2F004018 | A     | A | A  | 3,9 |
| 19 | J2F004019 | B     | A | D  | 2,3 |
| 20 | J2F004020 | A     | B | A  | 3,6 |

**Keterangan Tabel:**

PWI : Pemrograman Web dan Internet

M : Matematika

BI : Bahasa Inggris

- [A] Dari data-data tabel diatas, transformasikan data-data tersebut (3 baris awal) sebelum dilakukan pelatihan dengan menggunakan Jaringan Syaraf Tiruan *Backpropagation* dan Fungsi Aktivasi menggunakan *Sigmoid Biner*?
- [B] Gambarkan arsitektur jaringan syaraf tiruan *backpropagation* dari kasus diatas, jika digunakan 1 layer tersembunyi dan 2 *neuron* pada layer tersembunyi?
- [C] Hitunglah bobot-bobot jaringan baru setelah iterasi pertama dengan menggunakan data pola pertama dan di gunakan  $\alpha$  (laju pelatihan) = 0.2 dan bobot-bobot awal sebagai berikut:  
Bobot-bobot jaringan dan bobot bias dari layer input ke hidden layer

|       | $Z_1$ | $Z_2$ |
|-------|-------|-------|
| $X_1$ | 0.2   | 0.5   |
| $X_2$ | -0.3  | -0.1  |
| $X_3$ | 0.8   | 0.2   |
| 1     | 0.1   | -0.3  |

Bobot-bobot jaringan dan bobot bias dari hidden layer ke layer output

|       | Y    |
|-------|------|
| $Z_1$ | 0.5  |
| $Z_2$ | -0.5 |
| 1     | 0.2  |

- [D] Hitung hasil prediksi IPK mahasiswa, dari arsitektur jaringan yang anda buat jika nilai mata kuliah semester 1 mahasiswa tersebut adalah Pemrograman Web dan Internet (PWI): B, Matematika (M): D, dan Bahasa Inggris (BI): B, serta bobot-bobot hasil pelatihan seperti pada tabel berikut:

Bobot-bobot jaringan dan bobot bias dari layer input ke hidden layer

|       | $Z_1$ | $Z_2$ |
|-------|-------|-------|
| $X_1$ | 0.1   | 0.2   |
| $X_2$ | -0.5  | 0.3   |
| $X_3$ | 0.2   | -0.8  |
| 1     | 0.7   | 0.1   |

Bobot-bobot jaringan dan bobot bias dari hidden layer ke layer output

|       | Y    |
|-------|------|
| $Z_1$ | 0.1  |
| $Z_2$ | -0.2 |
| 1     | 0.3  |

# UJIAN TENGAH SEMESTER 2014/2015

## UJIAN TENGAH SEMESTER

JURUSAN : Ilmu Komputer / Informatika  
MATA KULIAH : Data Mining  
PENGAMPU : Nurdin Bahtiar, S.Si, MT

HARI, TANGGAL : Selasa, 11 November 2014  
WAKTU : 100 menit  
SIFAT : Buku tertutup, boleh menggunakan kalkulator

1. Diberikan suatu tabel penelitian tanaman mengenai kondisi tangkai, akar, daun, dan tanah sebagai berikut:

| No | Tangkai | Akar   | Daun      | Tanah  | Mati? |
|----|---------|--------|-----------|--------|-------|
| 1  | Lemas   | Kuat   | Menguning | Kering | Ya    |
| 2  | Lemas   | Kuat   | Menguning | Lembab | Ya    |
| 3  | Layu    | Kuat   | Menguning | Kering | Tidak |
| 4  | Kokoh   | Sedang | Menguning | Kering | Tidak |
| 5  | Kokoh   | Rapuh  | Hijau     | Kering | Tidak |
| 6  | Kokoh   | Rapuh  | Hijau     | Lembab | Ya    |
| 7  | Layu    | Rapuh  | Hijau     | Lembab | Tidak |
| 8  | Lemas   | Sedang | Menguning | Kering | Ya    |
| 9  | Lemas   | Rapuh  | Hijau     | Kering | Tidak |
| 10 | Kokoh   | Sedang | Hijau     | Kering | Tidak |

Menggunakan *naive bayes classifier*, jika suatu sampel tanaman dengan kondisi tangkainya LEMAS, akarnya RAPUH, daunnya MENGUNING, dan tanahnya LEMBAB, apakah kemungkinan yang terjadi pada tanaman tersebut (Mati / Tidak)?

2. Misalkan perhitungan nilai entropy pada pembuatan pohon keputusan melibatkan dua variabel  $X_1$  dan  $X_2$ , apakah maksud dari:
- Nilai entropy keduanya adalah 1
  - Nilai entropy keduanya adalah mendekati 0 (kecil sekali)
3. Jelaskan tiga alasan mengapa preprocessing data merupakan hal penting dalam data mining.
4. Tentukan BMI Result dari seseorang yang memiliki berat badan 64 kg dan tinggi badan 159 cm berdasarkan data table berikut menggunakan *Distance Weighted Nearest Neighbor*:

| No | Berat Badan (Kg) | Tinggi Badan (cm) | BMI Result  |
|----|------------------|-------------------|-------------|
| 1  | 70               | 160               | Overweight  |
| 2  | 58               | 179               | Underweight |
| 3  | 70               | 170               | Normal      |
| 4  | 65               | 170               | Normal      |
| 5  | 55               | 177               | Underweight |
| 6  | 59               | 175               | Normal      |
| 7  | 72               | 168               | Overweight  |
| 8  | 50               | 170               | Underweight |
| 9  | 67               | 173               | Normal      |
| 10 | 65               | 156               | Overweight  |

BMI: Body Mass Index

5. Gunakan K-Means untuk membagi empat objek berikut menjadi dua cluster!

| OBJEK  | BERAT | PH | CLUSTER |
|--------|-------|----|---------|
| Obat A | 1     | 1  |         |
| Obat B | 2     | 1  |         |
| Obat C | 4     | 3  |         |
| Obat D | 5     | 4  |         |

Tidak ada nilai bagi pelaku kecurangan.



# UJIAN TENGAH SEMESTER 2015/2016

## UJIAN TENGAH SEMESTER

PROGRAM STUDI : Teknik Informatika  
 MATA KULIAH : Data Mining  
 HARI, TANGGAL : Selasa, 3 November 2015  
 WAKTU / SIFAT : 90 menit / buku terbuka, boleh pakai kalkulator, Android Flight Mode  
 PENGAMPU : Nurdin Bahtiar, S.Si, MT & Sutikno, M.Cs

Jawablah setiap pertanyaan berikut dengan jelas dan singkat!

1. Gunakan tabel berikut untuk menentukan cluster menggunakan K-Means (2 cluster)!

| NO | X1  | X2  |
|----|-----|-----|
| 1  | 1   | 2   |
| 2  | 2.5 | 4.5 |
| 3  | 2   | 2   |
| 4  | 4   | 1.5 |
| 5  | 4   | 2.5 |

2. Diberikan data pemilihan tempat kos sebagai berikut:

| CONTOH | JARAK | HARGA  | FASILITAS | PILIH / TIDAK |
|--------|-------|--------|-----------|---------------|
| X1     | Dekat | \$     | Biasa     | Ya            |
| X2     | Jauh  | \$\$   | Biasa     | Tidak         |
| X3     | Jauh  | \$     | Mewah     | Ya            |
| X4     | Jauh  | \$\$\$ | Mewah     | Tidak         |
| X5     | Dekat | \$\$   | Mewah     | Ya            |

Buatlah pohon keputusan dari data di atas!

3. Berdasarkan tabel soal No.2 di atas, tentukan keputusan yang diambil jika ada tempat kos dengan spesifikasi (Jauh, \$\$, Mewah) menggunakan Naive Bayes Classifier!
4. Misalkan terdapat data laptop yang ingin dibeli / tidak berdasarkan pertimbangan sebagai berikut:

| NO | HARGA      | PROSESOR | HARD DISK | TOUCH SCREEN | BELI / TIDAK |
|----|------------|----------|-----------|--------------|--------------|
| 1  | < 6 juta   | Core i3  | 1 TB      | Ya           | Ya           |
| 2  | 6 – 8 juta | Core i3  | 500 GB    | Tidak        | Tidak        |
| 3  | 6 – 8 juta | Core i3  | 500 GB    | Tidak        | Ya           |
| 4  | < 6 juta   | Core i5  | 1 TB      | Tidak        | Tidak        |
| 5  | > 8 juta   | Core i5  | 500 GB    | Ya           | Tidak        |
| 6  | > 8 juta   | Core i5  | 1 TB      | Tidak        | Ya           |

Dengan bobot tiap atribut:

| ATRIBUT      | BOBOT |
|--------------|-------|
| HARGA        | 0.8   |
| PROSESOR     | 0.9   |
| HARD DISK    | 0.7   |
| TOUCH SCREEN | 0.5   |

Perbandingan nilai kedekatan:

| NILAI KEDEKATAN            | NILAI |
|----------------------------|-------|
| Untuk data yang sama       | 1     |
| Untuk data yang tidak sama | 0.5   |

Jika ada penawaran laptop dengan harga 7 juta, prosesor Core i5, Kapasitas harddisk 1 TB, dilengkapi fasilitas touchscreen, kira-kira dibeli tidak ya?

5. Diberikan data aplikasi sosial media yang terinstall di HP Android yang dimiliki 10 pengguna yaitu:

| HP | Aplikasi yang terinstall               |
|----|--|
| 1  | WhatsApp, Skype, Line, BBM, Instagram, |
| 2  | WhatsApp, Skype, BBM, Instagram,       |
| 3  | WhatsApp, Skype, Instagram,            |
| 4  | WhatsApp, Line, Instagram,             |
| 5  | WhatsApp, Skype, Line, Path,           |

| HP | Aplikasi yang terinstall          |
|----|-----------------------------------|
| 6  | Line, BBM, Path,                  |
| 7  | WhatsApp, Skype, Line, Instagram, |
| 8  | WhatsApp, Skype, Line, BBM,       |
| 9  | WhatsApp, Instagram,              |
| 10 | Skype, Line, BBM, Instagram,      |

Tentukan nilai **support** dan **confidence** untuk aturan-aturan berikut (dalam %):

- a. {WhatsApp} → {BBM}  
 b. {WhatsApp, Skype} → {Instagram}  
 c. {Skype, Line} → {BBM}  
 d. {WhatsApp, Skype, Line} → {Instagram}  
 e. {Instagram, Line} → {WhatsApp}

# UJIAN AKHIR SEMESTER 2015/2016



## Ujian Akhir Semester Gasal 2015/2016 Jurusan Ilmu Komputer/Informatika Fakultas Sains dan Matematika Universitas Diponegoro Semarang

Mata Kuliah : Data Mining  
Beban : 3 SKS  
Dosen : - Nurdin Bahtiar, S.Si, MT  
- Sutikno, S.T., M.Cs./

Hari/ Tanggal : Selasa 12 Januari 2016  
Jam : 13.00 – 14.30 WIB (90 Menit)  
Sifat : Buku Terbuka

**Perhatian:** Segala bentuk **KECURANGAN** (kerja sama, meminjam/memberi pinjaman catatan dan atau alat tulis, mencontek dan lain-lain) akan diberikan nilai 0 (NOL).

1. Suatu perusahaan mempunyai data-data penerimaan pegawai seperti pada tabel dibawah. Dari tabel tersebut akan digunakan untuk memprediksi kinerja pegawai yang akan diterima.

| Nama | IPK | N. Test Wawancara | Kinerja Pegawai |
|------|-----|-------------------|-----------------|
| Amin | 3,3 | 70                | Baik            |
| Bima | 2,5 | 77                | Baik            |
| Cici | 3,7 | 40                | Buruk           |
| Dani | 2,5 | 50                | Baik            |
| Ema  | 3,8 | 70                | Buruk           |
| Fani | 2,1 | 60                | Buruk           |
| dst  |     |                   |                 |

- a) Dari data-data pelatihan diatas buatlah arsitektur jaringan **backpropagation**?  
b) Hitung bobot-bobot pelatihan jaringan dengan data pada pola pertama (epoch pertama) jika fungsi aktivasi yang digunakan pada hidden layer dan output layer adalah sigmoid biner, nilai laju pembelajaran=0,1; jumlah neuron pada hidden layer 2 buah dan bobot-bobot awal jaringan anda tentukan sendiri?  
c) Apa prediksi kinerja pegawai dari arsitektur jaringan yang anda buat jika diketahui seseorang mempunyai IPK 3,0 dan nilai wawancara 54 serta bobot-bobot hasil pelatihan seperti berikut:

- Bobot-bobot jaringan dan bobot bias dari layer input ke hidden layer

|       | $Z_1$ | $Z_2$ |
|-------|-------|-------|
| $X_1$ | 0.2   | -0.3  |
| $X_2$ | -0.1  | 0.3   |
| 1     | 0.7   | 0.1   |

- Bobot-bobot jaringan dan bobot bias dari hidden layer ke layer output

|       | $Y$  |
|-------|------|
| $Z_1$ | -0.1 |
| $Z_2$ | 0.5  |
| 1     | 0.3  |

2. Perhatikan 3 dokumen berikut:

Dokumen 1: "Teknologi Anti Maling di Ponsel Android."

Dokumen 2: "Smartphone Canggih, Teknologi Paling Tinggi."

Dokumen 3: "Teknologi Android Dimanfaatkan Militer Indonesia."

[A]. Dari 3 dokumen diatas jelaskan proses pre-processing, sebelum dilakukan proses penanganan query jika di gunakan model ruang vektor (VSM), skema pembobotannya Tf dan IDF, dan ukuran kemiripannya dengan vektor Cosine.

[B]. Jika diberikan query: "Sejarah Perkembangan Smartphone Versi Android", dokumen yang mana yang paling mirip dengan query tersebut?



3. Diketahui data kelulusan mahasiswa yang terdiri dari 2 parameter nilai pada semester 1 yaitu Nilai matematika dan nilai bahasa inggris seperti pada tabel berikut.

| No | Nama   | N. Matematika | N. B. Inggris | LULUS TEPAT WAKTU |
|----|--------|---------------|---------------|-------------------|
| 1  | Adi    | 35            | 78            | Tidak             |
| 2  | Budi   | 80            | 98            | Ya                |
| 3  | Candra | 50            | 54            | Tidak             |
| 4  | Dani   | 60            | 78            | Tidak             |
| 5  | Fani   | 90            | 34            | Tidak             |
| 6  | Galuh  | 34            | 56            | Tidak             |
| 7  | Hani   | 78            | 81            | Ya                |

Jika data-data diatas akan di dibagi menjadi 2 kelas yaitu lulus tepat waktu dan tidak tepat waktu dengan metode **Support Vektor Machine**:

- Data yang mana yang merupakan support vektor dan jelaskan alasannya?
- Tentukan Hyperplanenya?
- Jika diketahui data uji/ data testing suatu mahasiswa semester 1 dengan nilai matematika 82 dan nilai bahasa inggris 75, apakah mahasiswa tersebut lulus tepat waktu berdasarkan hyperplane yang telah ditemukan?

# UJIAN TENGAH SEMESTER 2016/2017

## UJIAN TENGAH SEMESTER

PROGRAM STUDI : Teknik Informatika  
 MATA KULIAH : Data Mining  
 HARI, TANGGAL : Kamis, 13 Oktober 2016  
 WAKTU / SIFAT : 90 menit / buku terbuka, boleh pakai kalkulator, Android Flight Mode  
 PENGAMPU : Nurdin Bahtiar, S.Si, M.T & Sutikno, M.Cs

**Jawablah setiap pertanyaan berikut dengan jelas dan singkat!**

1. Gunakan tabel berikut untuk menentukan clusternya menggunakan K-Means (2 cluster)!

| NO | X1 | X2 |
|----|----|----|
| 1  | 1  | 2  |
| 2  | 2  | 3  |
| 3  | 2  | 2  |
| 4  | 4  | 1  |
| 5  | 5  | 2  |
| 6  | 5  | 1  |

2. Mukidi ingin membeli sepeda motor. Tetapi ia tidak tahu sepeda motor apa yang cocok buat dirinya. Ia mencoba menanyakan kepada teman-temannya. Mukidi bertekad akan membeli sepeda motor berdasarkan rekomendasi dari teman-temannya. Berikut data pemilihan sepeda motor yang dibeli / tidak oleh teman-teman Mukidi:

| NAMA    | CC  | HARGA      | TRANSMISI | BELI / TIDAK |
|---------|-----|------------|-----------|--------------|
| Andi    | 125 | \$         | Manual    | Ya           |
| Beny    | 110 | \$\$\$     | Manual    | Tidak        |
| Cinta   | 110 | \$         | Matic     | Ya           |
| Diandra | 110 | \$\$\$\$\$ | Matic     | Tidak        |
| Endah   | 125 | \$\$\$     | Matic     | Ya           |

**Perintah:**

Tentukan keputusan yang diambil Mukidi jika ada tawaran sepeda motor dengan spesifikasi (**110 CC**, Harga **\$\$\$**, Transmisi **Matic**) menggunakan Naïve Bayes Classifier!

3. Pada *decision tree learning*, terdapat istilah **entropy** dan **gain**. Jelaskan arti kedua istilah tersebut disertai ilustrasi contoh!
4. Misalkan terdapat data penerimaan beasiswa yang lolos / tidak sebagai berikut:

| NO | PENGHASILAN ORANG TUA | IPK   | ADA REKOMENDASI | AKTIF ORGANISASI | LOLOS BEASISWA / TIDAK |
|----|-----------------------|-------|-----------------|------------------|------------------------|
| 1  | < 2 juta              | < 3.0 | Ada             | Ya               | Ya                     |
| 2  | 2 – 5 juta            | < 3.0 | Tidak ada       | Tidak            | Tidak                  |
| 3  | 2 – 5 juta            | < 3.0 | Tidak ada       | Tidak            | Ya                     |
| 4  | < 2 juta              | > 3.0 | Ada             | Tidak            | Tidak                  |
| 5  | > 5 juta              | > 3.0 | Tidak ada       | Ya               | Tidak                  |
| 6  | > 5 juta              | > 3.0 | Ada             | Ya               | Ya                     |

Dengan bobot tiap atribut:

| ATRIBUT               | BOBOT |
|-----------------------|-------|
| PENGHASILAN ORANG TUA | 0.8   |
| IPK                   | 0.9   |
| ADA REKOMENDASI       | 0.7   |
| AKTIF ORGANISASI      | 0.6   |

Perbandingan nilai kedekatan:

| NILAI KEDEKATAN            | NILAI |
|----------------------------|-------|
| Untuk data yang sama       | 1     |
| Untuk data yang tidak sama | 0.6   |

Jika adaajuan penerima beasiswa dengan karakteristik: **penghasilan orang tua 4 juta**, **IPK > 3.0**, **ada rekomendasi**, dan **aktif organisasi**, kira-kira akan lolos didanani beasiswa tidak ya?

5. Diberikan data belanja bahan minuman dari 5 pelanggan sebagai berikut:

| ID | Bahan yang dibelanjakan       |
|----|-------------------------------|
| 1  | Gula, Teh, Kopi, Susu, Krimer |
| 2  | Gula, Teh, Susu, Krimer       |
| 3  | Gula, Teh, Krimer             |
| 4  | Gula, Kopi, Krimer            |
| 5  | Gula, Teh, Kopi, Susu         |

Tentukan **support** dan **confidence** (dalam %) untuk:

- a. {Gula} → {Susu}  
 b. {Gula, Teh} → {Krimer}  
 c. {Teh, Kopi} → {Susu}  
 d. {Gula, Teh, Kopi} → {Krimer}  
 e. {Krimer, Kopi} → {Gula}

-ooOoo-

**UJIAN AKHIR SEMESTER 2016/2017**

**Ujian Akhir Semester Gasal 2016/2017**  
**Departemen Ilmu Komputer/Informatika**  
**Fakultas Sains dan Matematika Universitas Diponegoro Semarang**

**Mata Kuliah** : Data Mining  
**Beban** : 3 SKS  
**Dosen** : - Nurdin Bahtiar, S.Si, MT  
           - Sutikno, S.T., M.Cs./

**Hari/ Tanggal** : Jum'at, 16 Desember 2016  
**Jam** : 10.00 – 11.30 WIB (90 Menit)  
**Sifat** : Buku Terbuka

**Perhatian:** Segala bentuk **KECURANGAN** (kerja sama, meminjam/memberi pinjaman catatan dan atau alat tulis, mencontek dan lain-lain) akan diberikan nilai **0 (NOL)**.

1. Berikut data pelatihan rekam medis prediksi penyakit diabetes millitus yang terdiri dari 3 gejala yaitu usia, keturunan dan lelah.

| Data ke- | Usia | Keturunan | Lelah | Status Diabete Millitus |
|----------|------|-----------|-------|-------------------------|
| 1        | 43   | ada       | ya    | Ya                      |
| 2        | 52   | Tidak ada | tidak | Tidak                   |
| 3        | 40   | Tidak ada | ya    | Tidak                   |
| 4        | 45   | Tidak ada | ya    | Ya                      |
| 5        | 50   | Tidak ada | ya    | Tidak                   |
| 6        | 63   | ada       | ya    | Ya                      |
| dst      |      |           |       |                         |

- a. Dari data-data pelatihan diatas buatlah arsitektur jaringan backpropagation?  
 b. Hitung bobot-bobot pelatihan jaringan dengan data pada pola pertama (epoch pertama) jika fungsi aktifasi yang digunakan pada hidden layer dan output layer adalah sigmoid biner, nilai laju pembelajaran=0,5; jumlah neuron pada hidden layer 3 buah dan bobot-bobot awal jaringan anda tentukan sendiri?
2. Di ketahui data penerbangan yang memiliki 2 input parameter yaitu Kecepatan Angin dan Suhu, seperti pada tabel berikut:

| No | Kec. Angin (knot) | Suhu (°C) | Delay (ya/tidak) |
|----|-------------------|-----------|------------------|
| 1  | 10                | 28        | Tidak            |
| 2  | 16                | 27        | Ya               |
| 3  | 19                | 26        | Ya               |
| 4  | 12                | 30        | Tidak            |
| 5  | 17                | 25        | Ya               |
| 6  | 20                | 29        | Tidak            |

Jika data-data diatas akan di dibagi menjadi 2 kelas yaitu penerbangan *delay* dan penerbangan tidak *delay* dengan metode **Support Vektor Machine**:

- a. Data yang mana yang merupakan *support vektor* dan jelaskan alasannya?  
 b. Tentukan Hyperplanenya?  
 c. Jika diketahui data uji/ data testing dengan kecepatan angin 18 knot dan suhu 26 °C, apakah penerbangan *delay* atau tidak berdasarkan hyperplane yang telah ditemukan?
3. Perhatikan 3 dokumen berikut:  
 Dokumen 1 : "Teknologi masa depan: mobil terbang sampai robot."  
 Dokumen 2 : "Gambaran robot canggih."  
 Dokumen 3 : "Teknologi masa depan sudah ada sekarang ."
- a. Dari 3 dokumen diatas jelaskan proses pre-processing, sebelum dilakukan proses penanganan query jika di gunakan model ruang vektor (VSM), skema pembobotannya Tf dan IDF, dan ukuran kemiripannya dengan vektor Cosine.  
 b. Jika diberikan query: "Teknologi robot masa depan", dokumen yang mana yang paling mirip dengan query tersebut?

~ tik051216 ~

**UJIAN TENGAH SEMESTER 2017/2018**

Ujian Akhir Semester Gasal 2017/2018  
Departemen Ilmu Komputer/ Informatika  
Fakultas Sains dan Matematika Universitas Diponegoro

|             |                    |               |                               |
|-------------|--------------------|---------------|-------------------------------|
| Mata Kuliah | : Penambangan Data | Dosen         | : - Nurdin Bahtiar, S.Si, M.T |
| Beban       | : 3 SKS            |               | - Indra Waspada, ST, MTI      |
| Semester    | : 3                | Hari/ Tanggal | : Jumat/ 8 Desember 2017      |
| Sifat       | : buku tertutup    | Waktu         | : 75 menit                    |

1. {Bobot 60} Suatu situs media online, misalnya detik.com memiliki halaman kategori sebagai berikut: Berita, Politik, Keuangan, Hiburan, Olahraga, dan Seni. Sesi yang didapat pengguna dipandang sebagai satu kunjungan ke situs web. Dalam satu sesi, Pengguna yang sama dapat mengakses konten dari kategori yang berbeda. Apabila selama 30 menit pengguna tidak aktif maka akan berganti ke sesi baru. Aktivitas penelusuran web ini dapat dianalogikan dengan aktivitas belanja di toko, yaitu sesi dikaitkan dengan transaksi atau keranjang belanja, sedangkan halaman yang diakses dianalogikan dengan item yang dibeli.  
Pada tabel 1 diperoleh suatu daftar data sesi dan kategori media yang diakses pada sesi tertentu. **Kebutuhan klien** (pengelola media online) terhadap data ini adalah **menemukan asosiasi antar kategori media**.
  - a. {50} Jelaskan secara detil tiap tahap untuk menghasilkan solusi bagi klien dengan mengacu **kerangka kerja CRISP-DM** menggunakan algoritma **Apriori** dengan nilai support minimum 0,25 dan batas nilai confident 0,75.
  - b. {10} Jika anda pada posisi sebagai klien, berikan salah satu contoh kebijakan yang akan anda lakukan setelah mendapat informasi asosiasi antar kategori media dari hasil soal 1.a.

Tabel 1. Daftar Sesi dan Kategori yang diakses

| ID Sesi | Daftar Kategori Media yang diakses |
|---------|------------------------------------|
| 1       | { Berita, Keuangan }               |
| 2       | { Berita, Keuangan }               |
| 3       | { Olahraga, Keuangan, Berita }     |
| 4       | { Seni }                           |
| 5       | { Olahraga, Berita, Keuangan }     |
| 6       | { Berita, Seni, Hiburan }          |

2. {Bobot 40} Pada suatu corpus yang mengandung tiga buah dokumen, berikan analisis anda untuk menentukan kata yang **paling penting** diantara 3 kata berikut: "data", "mining", dan "text" **pada tiap dokumen**. Gunakan pembobotan yang mengacu pada TF-IDF.

| ID Dokumen | Isi Dokumen   |
|------------|---|
| 1          | This is a data mining course  |
| 2          | We are studying text mining. Text mining is a subfield of data mining |
| 3          | Mining text is interesting and I am interested in it                  |

== Selamat Mengerjakan ==



**LATIHAN 2017/2018****UJIAN TENGAH SEMESTER**

PROGRAM STUDI : Teknik Informatika  
MATA KULIAH : Data Mining  
HARI, TANGGAL : Jumat, 6 Okt 2017  
WAKTU : 75 menit  
SIFAT : Open rumus selebar kertas HVS  
PENGAMPU : Nurdin Bahtiar, S.Si, M.T  
Indra Waspada, S.T, M.T

**Jawablah setiap pertanyaan berikut dengan jelas!**

1. Buatlah sebuah flowchart untuk menghitung jarak Euclidean dari dua buah titik di koordinat kartesius! (20 poin)
2. Jelaskan dan beri contoh 3 (tiga) karakteristik umum dari data yang harus diselesaikan sebelum dianalisis menggunakan teknik data mining! (30 poin)
3. Buatlah sebuah algoritma lengkap (dalam bentuk pseudocode) untuk memberikan prediksi menggunakan Naïve Bayes Classifier! (50 poin)

-ooOoo-

**UJIAN AKHIR SEMESTER 2017/2018**

Ujian Akhir Semester Gasal 2017/2018  
Departemen Ilmu Komputer/ Informatika  
Fakultas Sains dan Matematika Universitas Diponegoro

|             |                    |               |                               |
|-------------|--------------------|---------------|-------------------------------|
| Mata Kuliah | : Penambangan Data | Dosen         | : - Nurdin Bahtiar, S.Si, M.T |
| Beban       | : 3 SKS            |               | - Indra Waspada, ST, MTI      |
| Semester    | : 3                | Hari/ Tanggal | : Jumat/ 8 Desember 2017      |
| Sifat       | : buku tertutup    | Waktu         | : 75 menit                    |

1. {Bobot 60} Suatu situs media online, misalnya detik.com memiliki halaman kategori sebagai berikut: Berita, Politik, Keuangan, Hiburan, Olahraga, dan Seni. Sesi yang didapat pengguna dipandang sebagai satu kunjungan ke situs web. Dalam satu sesi, Pengguna yang sama dapat mengakses konten dari kategori yang berbeda. Apabila selama 30 menit pengguna tidak aktif maka akan berganti ke sesi baru. Aktivitas penelusuran web ini dapat dianalogikan dengan aktivitas belanja di toko, yaitu sesi dikaitkan dengan transaksi atau keranjang belanja, sedangkan halaman yang diakses dianalogikan dengan item yang dibeli.
- Pada tabel 1 diperoleh suatu daftar data sesi dan kategori media yang diakses pada sesi tertentu. Kebutuhan klien (pengelola media online) terhadap data ini adalah **menemukan asosiasi antar kategori media**.
- {50} Jelaskan secara detil tiap tahap untuk menghasilkan solusi bagi klien dengan mengacu **kerangka kerja CRISP-DM** menggunakan algoritma **Apriori** dengan nilai support minimum 0,25 dan batas nilai confident 0,75.
  - {10} Jika anda pada posisi sebagai klien, berikan salah satu contoh kebijakan yang akan anda lakukan setelah mendapat informasi asosiasi antar kategori media dari hasil soal 1.a.

Tabel 1. Daftar Sesi dan Kategori yang diakses

| ID Sesi | Daftar Kategori Media yang diakses |
|---------|------------------------------------|
| 1       | { Berita, Keuangan }               |
| 2       | { Berita, Keuangan }.              |
| 3       | { Olahraga, Keuangan, Berita }     |
| 4       | { Seni }                           |
| 5       | { Olahraga, Berita, Keuangan } .   |
| 6       | { Berita, Seni, Hiburan }          |

$$\text{Support} = \frac{A \cap B}{\text{Semua}}$$

$$\text{Conf} = \frac{A \cap B}{A}$$

2. {Bobot 40} Pada suatu corpus yang mengandung tiga buah dokumen, berikan analisis anda untuk menentukan kata yang **paling penting** diantara 3 kata berikut: "data", "mining", dan "text" pada **tiap dokumen**. Gunakan pembobotan yang mengacu pada TF-IDF.

| ID Dokumen | Isi Dokumen   |
|------------|---|
| 1          | This is a data mining course  |
| 2          | We are studying text mining. Text mining is a subfield of data mining |
| 3          | Mining text is interesting and I am interested in it                  |

== Selamat Mengerjakan ==