

**PEMODELAN DATA MINING MENGGUNAKAN NEURAL
NETWORK UNTUK SELEKSI MAHASISWA PENERIMA
BANTUAN BEASISWA : STUDI KASUS STMIK PASIM SUKABUMI**



PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR
JAKARTA
2013



PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini,

Nama : Yogi Syarif Hidayat
N IM : 1111600407
Konsentrasi : Teknologi Sistem Informasi
Program Studi : Magister Ilmu Komputer

Menyatakan bahwa TESIS yang berjudul:

Pemodelan Data Mining Menggunakan Neural Network Untuk Seleksi
Mahasiswa Penerima Bantuan Beasiswa : Studi Kasus STMIK PASIM
Sukabumi

1. Merupakan hasil karya tulisan ilmiah sendiri dan bukan merupakan karya yang pernah diajukan untuk memperoleh gelar akademik oleh pihak lain,
2. Saya ijin untuk dikelola oleh Universitas Budi Luhur sesuai dengan norma hukum dan etika yang berlaku.

Pernyataan ini saya buat dengan penuh tanggung jawab dan saya bersedia menerima konsekuensi apapun sesuai aturan yang berlaku, apabila dikemudian hari pernyataan ini tidak benar.

Jakarta, September 2013



(Yogi Syarif Hidayat)

 PROGRAM STUDI MAGISTER ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

LEMBAR PENGESAHAN

Nama Mahasiswa : Yogi Syarif Hidayat
Nomor Induk Mahasiswa : 1111600407
Konsentrasi : Teknologi Sistem Informasi
Judul Tesis : Pemodelan Data Mining Menggunakan Neural Network Untuk Seleksi Mahasiswa Penerima Bantuan
Bersiswa : Studi Kasus STMIK PASIM Sukabumi.

Jakarta, 6 September 2013

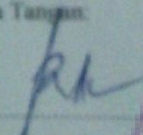
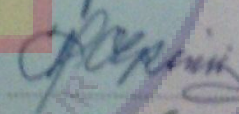
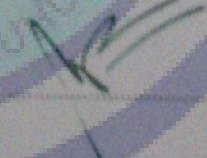
Tim Penguji:

Ketua,
Dr. Ir. Nazori AZ, MT

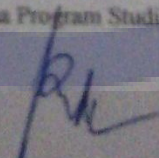
Anggota,
Dr. Moedjono, M.Sc

Pembimbing Utama,
Dr. Ir. Wendi Usino, M.Sc, MM

Tanda Tangan:

Ketua Program Studi


Dr. Ir. Nazori AZ, MT.

ii

ABSTRAK

Beasiswa merupakan *instrument* penting dalam proses berlangsungnya suatu kegiatan belajar pada suatu institusi perguruan tinggi, banyak kriteria untuk mendapatkan beasiswa, antara lain berdasarkan prestasi pendidikan, keaktifan mahasiswa, faktor ekonomi maupun kedisiplinan dari mahasiswa tersebut. Akurasi hasil penentuan penerima beasiswa digunakan untuk banyak alasan, utamanya adalah untuk kebutuhan bagi perguruan tinggi untuk memberikan kemudahan dan ketersediaan alokasi dana beasiswa pada masing-masing mahasiswa, dan kesempatan untuk membantu ekonomi mahasiswa yang kurang mampu untuk tetap bisa melanjutkan pendidikannya di perguruan tinggi tersebut. Tujuan pemberian beasiswa adalah merupakan kegiatan rutin yang harus dipenuhi oleh perguruan tinggi setiap tahun ajaran baru dan sebagai motivasi terhadap mahasiswa yang berprestasi untuk lebih meningkatkan lagi prestasinya dengan diberikan beasiswa, serta untuk membantu mahasiswa yang kurang mampu dalam segi finansial sehingga mereka tidak akan terputus ditengah jalan dalam menempuh ilmu dibangku perguruan tinggi dan khususnya untuk menyediakan sarana yang obyektif dan terpercaya bagi perguruan tinggi, mahasiswa dan orang tua mahasiswa lainnya dalam memonitor kegiatan belajar mengajar yang aktif sesuai dengan kurikulumnya. Untuk mengatasi permasalahan diatas, maka penulis menggunakan model data *mining* berbasis *neural network*. Model *neural network* digunakan untuk penentuan kelayakan seleksi penerimaan beasiswa bantuan untuk mahasiswa pada STMIK PASIM Sukabumi. Model ini dipilih karena proses *learning* dan klasifikasi pada algoritma *neural network* sederhana dan cepat. Secara umum, model algoritma *neural network* mempunyai tingkat akurasi yang tinggi.

Kata Kunci : Beasiswa, *data mining*, *neural network*.

ABSTRACT

Scholarship is an important instrument in the course of a learning activity at a higher education institution, many of the criteria for the scholarship, among others, based on educational achievement, student activity, economic factors and the discipline of the students. Accuracy of the determination of the scholarship recipients are used for many reasons, primarily to the need for universities to provide convenience and availability of scholarship funds allocated to each student, and the opportunity to help economically disadvantaged students to remain able to continue their education at the college. The purpose of scholarships is a routine activity that must be met by the college every new school year and as a motivation to the students who excel to further enhance his performance with given scholarships, as well as to assist underprivileged students in financial terms so that they will not be interrupted in the middle of the road in taking the bench science colleges and in particular to provide an objective and reliable means for universities, students and parents to monitor the activities of other students in active learning according to curriculum. For handling the problems above, the authors used data mining models based on neural network. Neural network model will be used for eligibility of students to receive scholarships at STMIK PASIM Sukabumi. This model was chosen because of the process of learning and neural network classification algorithm is simple and fast. In general, the neural network algorithm model has a high degree of accuracy.

Keywords: Scholarship, data mining, neural network

KATA PENGANTAR

Puji dan syukur kehadiran Allah SWT atas berkah dan rahmat-Nya, sehingga penulisan penelitian tesis yang berjudul ini “Pemodelan Data Mining Menggunakan Neural Network Untuk Seleksi Mahasiswa Penerima Bantuan Beasiswa : Studi Kasus STMIK PASIM Sukabumi” dapat terselesaikan dengan baik.

Tujuan pembuatan penelitian tesis ini adalah untuk memenuhi salah satu syarat dalam penyusunan tesis pada jenjang Strata Dua (S-2) pada Program Studi Magister Ilmu Komputer (MKOM) Program Pascasarjana Universitas Budi Luhur.

Ucapan Syukur dan terima kasih penulis persembahkan kepada semua pihak yang telah membantu dalam penyusunan penelitian tesis ini:

1. Bapak Dr. Ir. Wendi Usino, M.Sc, MM., sebagai dosen pembimbing yang telah memberikan waktu dan arahan selama penulisan tesis ini.
2. Bapak Dr. Ir. Nazori AZ, MT. sebagai Ketua Program Studi Magister Ilmu Komputer Universitas Budi Luhur.
3. Bapak Dr. Moedjiono, M.Sc. Direktur Program Pascasarjana Universitas Budi Luhur.
4. Segenap Staff Pengajar dan Sekretariat Tata Usaha Magister Ilmu Komputer Universitas Budi Luhur yang telah memberikan ilmunya dan bantuan akademik selama penulis mengikuti pendidikan di MKOM Universitas Budi Luhur.
5. Bapak Enur Zaenal Abidin, S.Kom, MM., Selaku Ketua STMIK PASIM Sukabumi yang telah memberikan kesempatan untuk melakukan penelitian.
6. Hj. Patonah, ibunda tercinta, H. A. Zarkoni (alm), ayahanda tercinta yang selalu memberikan dukungan, semangat, dan do'a bagi penulis.
7. Metha dan Asfha, yang selalu memberikan dukungan, semangat, dan do'a bagi penulis.
8. Teman-teman mahasiswa MKOM UBL, yang telah membantu dan saling memberi semangat dan motivasi.

Penulis menyadari bahwa dalam penulisan tugas akhir ini masih terdapat banyak kekurangan. Namun demikian semoga laporan tesis ini dapat memberikan manfaat bagi semua pihak yang membacanya.

Sukabumi, Juni 2013

Yogi Syarif Hidayat

DAFTAR ISI

Lembar Judul	i
Lembar Persetujuan Tesis	ii
Abstrak	iii
Kata Pengantar	v
Daftar Isi	vii
Daftar Gambar	x
Daftar Tabel	xi
Daftar Lampiran	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Masalah Penelitian	2
1.2.1 Identifikasi Masalah	2
1.2.2 Batasan Masalah	2
1.2.3 Rumusan Masalah	3
1.3 Tujuan dan Manfaat Penelitian	3
1.3.1 Tujuan Penelitian	3
1.3.2 Manfaat Penelitian	3
1.4 Sistematika Penulisan	4
1.5 Daftar Pengertian	4
BAB II LANDASAN TEORI DAN KERANGKA PEMIKIRAN	6
2.1 Tinjauan Pustaka	6
2.1.1 Pengertian Database	6
2.1.2 Pengertian Data Warehouse	6
2.1.3 Pengertian Data Mining	7
2.1.4 Algoritma Klasifikasi Data Mining	12

2.1.5	Support Vector Machine	15
2.1.6	Neural Network	24
2.1.6.1	Konsep Neural Network	24
2.1.6.2	Studi Kasus Neural Network	27
2.1.7	Evaluasi dan Validasi Metode Klasifikasi Data Mining..	29
2.1.8	Rapid Miner 5	31
2.1.9	Evaluasi Model	32
2.2	Tinjauan Studi	32
2.3	Tinjauan Objek Penelitian	36
2.3.1	Struktur Organisasi STMIK PASIM Sukabumi	37
2.3.2	Dukungan Sistem Informasi	37
2.4	Kerangka Pemikiran	38
2.5	Hipotesis Penelitian	40
BAB III	METODOLOGI DAN RANCANGAN PENELITIAN	41
3.1	Metode Penelitian	41
3.2	Metode Pemilihan Sampel	41
3.3	Metode Pengumpulan Data	42
3.4	Instrumentasi Penelitian	42
3.5	Teknik Analisis dan Pengumpulan Data	42
3.6	Langkah-langkah Penelitian	37
3.7	Jadwal Penelitian	46
BAB IV	PEMBAHASAN HASIL PENELITIAN	48
4.1	Pengelompokkan Data	48
4.2	Hasil Penelitian	49
4.2.1	Perhitungan Data Training	49
4.2.2	Model Neural Network	52
4.2.3	Model SVM	55

4.3	Evaluasi dan Validasi	56
4.3.1	Pengujian Model	56
4.3.2	Analisis Hasil	60
4.4	Implikasi Penelitian	61
4.4.1	Aspek Sistem	61
4.4.2	Aspek Manajerial	61
4.4.3	Aspek Penelitian Lanjut	62
4.5	Rencana Implementasi Sistem	62
4.5.1	Tahapan Implementasi Sistem	62
4.5.2	Kegiatan Implementasi Sistem	63
BAB V	PENUTUP	65
5.1	Kesimpulan	65
5.2	Saran	65
Daftar Pustaka		
Lampiran		

DAFTAR GAMBAR

	halaman
II.1 Proses KDD	8
II.2 Langkah-langkah Dalam Proses Data Mining	9
II.3 Decision Boundary	16
II.4 Paradigma Kernel Trick	18
II.5 Linear Separation	19
II.6 Arsitektur SVM	20
II.7 Ilustrasi data	23
II.8 Multi-Layer Perceptron	27
II.9 Rapidminer	31
II.10 Struktur Organisasi STMIK PASIM Sukabumi	37
II.11 Sistem Informasi Akademik STMIK PASIM Sukabumi	37
II.12 Kerangka Pemikiran	39
III.1 Langkah-langkah Penelitian	45
IV.1 Multilayer Perceptron	49
IV.2 Desain Model Neural Network	52
IV.3 Neural Network yang Dihasilkan	53
IV.4 Nilai Bobot Akhir untuk Hidden Layer Data Training	54
IV.5 Nilai Bobot Akhir untuk Output Layer Data Training	54
IV.6 Desain model SVM	55
IV.7 Model algoritma SVM	55
IV.8 Desain Model Validasi	56
IV.7 Kurva ROC neural network	59
IV.8 Kurva ROC SVM	59

DAFTAR TABEL

	halaman
II.1 AND Problem	20
II.2 XOR	21
II.3 Nilai Awal untuk Input, Bobot dan Bias	27
II.4 Perhitungan Untuk Bobot dan Bias Terbaru	29
II.5 Model Confusion Matrix	29
II.6 Studi Terdahulu Yang Relevan	34
III.1 Sampel	41
III.2 Atribut dan Nilai Kategori	43
III.3 Data Training	35
III.4 Jadwal Penelitian	46
IV.1 Atribut dan Nilai Kategori	48
IV.2 Hidden Layer	51
IV.3 Output Layer	52
IV.4 Model Confusion Matrix neural network	57
IV.5 Model Confusion Matrix SVM	57
IV.6 Nilai Accuracy, Precision dan Recall	58
IV.7 Perhitungan Nilai Accuracy, Precision dan Recall Data Training	58
IV.8 Perhitungan Nilai Accuracy, Precision dan Recall Data Uji Validasi	49
IV.9 Perbandingan Nilai AUC	60
IV.10 Nilai Accuracy dan AUC	60
IV.11 Rencana implementasi sistem	62

DAFTAR PUSTAKA

- [Bramer. 2007] Bramer, Max. , *Principles of Data Mining*, Springer, London, 2007.
- [Gorunescu 2011] Gorunescu, *Data Mining: Concepts, Models, and Techniques*, Verlag Berlin Heidelberg: Springer, 2011.
- [Han,J 2007] Han, Jiawei and Kamber, M. *Data Mining: Concepts and Techniques*, University of Illinois at Urbana-Champaign, 2007.
- [Marimin 2008] Marimin, *Teknik dan Aplikasi Pengambilan Keputusan Kriteria Majemuk*, PT. Grasindo, Jakarta, 2008.
- [Kusrini 2008] Kusrini dan Luthfi, E.T, *Algoritma Data Mining*, Andi Publishing, Yogyakarta, 2008.
- [Larose, D. T. 2005] Larose, D. T., *Discovering Knowledge in Data*, New Jersey: John Willey & Sons, Inc, 2005.
- [Larose, D. T. 2007] Larose, D. T., *Data Mining Methods and Models - 2007*, : John Willey & Sons, Inc, 2007.
- [Larose, D. T. 2007] Larose, D. T and Markov, Zdravkov, *Uncovering Patterns in Web Content, Structure and Usage*, : John Willey & Sons, Inc, 2007.
- [Maimon 2005] Maimon, Oded and Rokach, Lior, *Data Mining and Knowledge Discovery Handbook*, Springer, York, 2005.
- [Myatt, Glenn J.. 2007] Myatt, Glenn J *Making Sense of Data: A Practical Guide to Exploratory Data Analysis and Data Mining*, John Wiley & Sons, Inc., New Jersey, 2007.
- [Santoso 2007] Santoso, Budi, *Data Mining Terapan dengan Matlab*, PT. Graha Ilmu, Surabaya, 2007.
- [Vercellis. 2009] Vercellis , *The Business Intelligence*, Politecnico di milano, Italy, 2009.
- [Wu, Xindong. 2009] Wu, Xindong & Kumar, Vipin. , *The Top Ten Algorithms in Data Mining*, CRC Press, Boca Raton, 2009.