

**USULAN
PENELITIAN DOSEN PEMULA**



**PENERAPAN METODE KLASIFIKASI DATA MINING UNTUK PREDIKSI
KELULUSAN TEPAT WAKTU**

TIM PENGUSUL:

Ketua : Asep Saefulloh, Spd, M.Kom; NIDN 0407056701

Anggota : Sugeng Santoso, M.Kom; NIDN 0406047601

**SEKOLAH TINGGI MANAJEMEN DAN ILMU KOMPUTER RAHARJA
STMIK RAHARJA**

April 2013

HALAMAN PENGESAHAN

PENELITIAN DOSEN PEMULA

Judul Penelitian : Penerapan metode klasifikasi data mining untuk
Prediksi kelulusan tepat waktu

Kode>Nama Rumpun : 462/Teknologi Informasi

Ketua Peneliti :

- a. Nama Lengkap : Asep Saefulloh, SPd, M.Kom
- b. NIDN : 0407056701
- c. Jabatan Fungsional : Lektor
- d. Program Studi : Sistem Komputer
- e. Nomor HP : 085811342550
- f. Alamat email : asaefullah@gmail.com; asep.saefullah@raharja.info

Anggota Peneliti (1) :

- a. Nama lengkap : Sugeng Santoso., M.Kom
- b. NIDN : 0406047601
- c. Perguruan Tinggi : STMIK Raharja

Penelitian tahun ke : 1 (satu)

Biaya penelitian : - Diusulkan ke DIKTI Rp. 12.173.500
- Dana internal PT Rp. 2.000.000

Mengetahui
Ketua STMIK,

Tangerang, 15 Desember 2013
Ketua Peneliti,

(Ir. Untung Rahardja, MTI)
NIP.000594

(Asep Saefulloh, SPd, M.Kom)
NIP. 056007

Menyetujui,
Ketua Lembaga Penelitian dan Pengabdian Masyarakat,

(M. Yusup, S.Kom)
NIP. 026002

RINGKASAN

Perguruan Tinggi Raharja memiliki kekayaan data yang luar biasa, terdapat pada dua database yaitu database Absensi on Line (AO) dan database *Student Information System* (SIS). Database AO dijadikan pasokan data untuk mengelola Indeks Mutu Kumulatif Mahasiswa (IMK) sedangkan database SIS merupakan pemasok data untuk pengelolaan Indeks Prestasi Kumulatif (IPK). IMK dan IPK hanya berupa data belum memberikan informasi yang bermanfaat, selama ini untuk memperkirakan kelulusan tepat waktu mahasiswa dengan melihat pengaruh dari IMK dan IPK hanya berupa *forecasting*. Tujuan dari penelitian ini adalah untuk membantu Prodi dalam memprediksi jumlah kelulusan mahasiswanya setiap tahun angkatan, sehingga menjadi patokan dalam penentuan penerimaan mahasiswa baru tahun berikutnya. Adapun target khusus dari penelitian ini yaitu untuk mencari algoritma terbaik dari tiga algoritma yang diuji, kemudian hasil pemilihan algoritma tersebut dibuatkan interfacinya. Berangkat dari permasalahan di atas maka diteliti untuk memprediksi kelulusan tepat waktu menggunakan metode klasifikasi *data mining* dengan pemilihan algoritma C4.5, Naïve Bayes, dan Neural Network, sedangkan untuk desain penelitian menggunakan model CRISP-DM (*Cross Industry Standard Process for Data Mining*).

Dalam CRISP-DM penelitian melalui 6 tahapan yaitu :

1. *Business/Research Understanding Phase*
2. *Data Understanding Phase* (Fase Pemahaman Data)
3. *Data Preparation Phase* (Fase Pengolahan Data)
4. *Modeling Phase* (Fase Pemodelan)
5. *Evaluation Phase* (Fase Evaluasi)
6. *Deployment Phase* (Fase Penyebaran)

Dari ketiga algoritma yang akan diuji sebelum dipergunakan untuk prediksi kelulusan tepat waktu, maka yang disebut algoritma terbaik adalah algoritma yang paling tinggi tingkat *accuracy* pada model klasifikasi dan memiliki nilai AUC (*Area Under Curve*) antara 0.90-1.00 Hasil *data mining* dari algoritma terpilih dalam penelitian ini akan dirancang *interface* menggunakan *java engine* yang dapat menampilkan prediksi kelulusan tepat waktu beserta jumlah kelulusan tepat waktu setiap Program Studi. Manfaat dari penelitian ini adalah, data prediksi kelulusan tepat waktu dapat dipergunakan oleh Kepala Prodi atau Panitia Penerimaan Mahasiswa Baru dalam menetapkan jumlah calon mahasiswa yang akan diterima pada Perguruan Tinggi tersebut. Manfaat lainnya yaitu untuk melakukan pembinaan kepada mahasiswa secara dini agar terhindar dari ke tidak tepatan waktu penyelesaian studi.

BAB I PENDAHULUAN

1.1 Latar Belakang

Terkait dengan salah satu fungsi dari Perguruan Tinggi dalam pendidikan, pengajaran dan perihal ini menjadi salah satu butir akreditasi yaitu kelulusan tepat waktu bagi mahasiswa. Adanya informasi kelulusan tepat waktu tentu akan menjadikan suatu pengambilan keputusan yang tepat bagi manajemen Perguruan Tinggi dalam mengambil langkah berikutnya. Selama ini Perguruan Tinggi Raharja belum memiliki pola-pola prediksi kelulusan tepat waktu sebagai acuan untuk memprediksi jumlah lulus tepat waktu. Prediksi kelulusan tepat waktu yang dilakukan saat ini hanya berdasarkan *forecaster* dari data IPK dan IMK semester sebelumnya. Prediksi hampir sama dengan klasifikasi dan estimasi, hanya saja prediksi digunakan untuk menduga nilai-nilai tertentu yang akan terjadi di masa mendatang [Prab, 2012].

Sementara itu Perguruan Tinggi Raharja mempunyai *dataset* AO (Absensi Online) dan SIS (Student Information Services) yang selama ini belum dimanfaatkan secara maksimal. Adalah hal yang sangat disayangkan jika *dataset* yang begitu besar tidak dimanfaatkan untuk digali informasi apa yang terdapat didalamnya. Selain itu, selama ini ada anggapan dari para *forecaster* Perguruan Tinggi Raharja bahwa untuk memprediksi tingkat kelulusan tepat waktu cukup dengan melihat data IPK dan IMK sebelumnya. Berangkat dari permasalahan di atas maka dilakukanlah penelitian ini yaitu untuk melakukan *data mining* terhadap *dataset* AO dan SIS sehingga didapatkan informasi mengenai kelulusan tepat waktu dari mahasiswa Perguruan Tinggi Raharja.

1.2 Masalah Penelitian

1.2.1 Identifikasi Masalah

Dalam melakukan prediksi kelulusan tepat waktu mahasiswa Perguruan Tinggi Raharja terdapat berbagai macam masalah yaitu :

1. Metode prediksi masih menggunakan prinsip kekeluargaan sehingga dirasakan kurangnya tingkat profesionalisme dalam melakukan prediksi kelulusan tepat waktu.
2. Metode prediksi masih menggunakan prinsip apriori dalam

menentukan kelulusan tepat waktu.

3. Melakukan *datamining* terhadap database DMQ yang merupakan hasil dari pemilihan data dari dua database yaitu database AO dan database SIS untuk mendapatkan data kelulusan tepat waktu.

1.2.2 Batasan Masalah

Untuk lebih fokus, maka dalam penelitian ini hanya dibatasi masalah yang terkait dengan algoritma C4.5, Naïve Bayes, dan Neural Network menggunakan klasifikasi data mining dengan cara menganalisis sejumlah atribut yang menjadi parameter dalam prediksi kelulusan tepat waktu.

Penelitian ini menggunakan perangkat lunak *Weka (Weikato Environment Knowledge and Analysis)* versi 3.6.4 yang merupakan aplikasi data mining berbasis *open source* (GPL) dan berengine Java, dengan Graphical User Interface (GUI) menggunakan java.

1.2.3 Rumusan Masalah

Adapun rumusan masalah yang ada pada penelitian ini adalah

- a. Apakah algoritma C4.5, Naïve Bayes, dan Neural Network merupakan algoritma-algoritma yang dapat digunakan dalam menentukan prediksi kelulusan tepat waktu ?
- b. Diantara tiga algoritma yang dibahas dalam penelitian ini yaitu algoritma C4.5, Naïve Bayes, dan Neural Network, algoritma manakah yang terbaik dalam menentukan prediksi kelulusan tepat waktu ?
- c. Dari algoritma terpilih apakah dapat menampilkan data prediksi hasil *datamining* dengan menampilkan kelulusan tepat waktu ?

1.3 Tujuan dan Manfaat Penelitian

1.3.1 Tujuan

Penelitian ini bertujuan untuk:

- a. Membandingkan tingkat akurasi yang dihasilkan oleh teknik atau model data mining yaitu algoritma C4.5, Naïve Bayes, dan Neural Network dalam

memperkirakan kelulusan tepat waktu.

- b. Menjabarkan algoritma C4.5, Naïve Bayes, dan Neural Network kedalam rule.
- c. Menerapkan algoritma C4.5, Naïve Bayes, dan Neural Network dalam melakukan prediksi terhadap kelulusan tepat waktu.

1.3.2 Manfaat

Hasil penelitian ini diharapkan juga dapat memberikan manfaat praktis dan manfaat teoritis:

- a. Manfaat praktis dari hasil penelitian ini diharapkan agar dapat digunakan oleh Perguruan Tinggi Raharja dalam melakukan prediksi kelulusan tepat waktu.
- b. Manfaat berikutnya sebagai *early warning* bagi pengelola institusi, jika diketahui adanya mahasiswa yang akan tidak tepat lulus waktu untuk diberikan pembinaan.
- c. Data yang sudah *dimining* dan menghasilkan prediksi kelulusan tepat waktu, maka informasi tersebut dapat dipergunakan sebagai patokan dalam menentukan jumlah mahasiswa pada Penerimaan Mahasiswa Baru.

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

2.1.1 Definisi *Data Mining*

Penggunaan sistem komputerisasi dalam berbagai bidang baik itu dalam transaksi-transaksi bisnis, maupun untuk kalangan pemerintah dan sosial, telah menghasilkan data yang berukuran sangat besar. Data-data yang terkumpul ini merupakan suatu tambang emas yang dapat digunakan sebagai informasi dalam dunia bisnis. Perkembangan *data mining* yang pesat tidak dapat lepas dari perkembangan teknologi informasi yang memungkinkan data dalam jumlah besar terakumulasi.

Data Mining didefinisikan sebagai serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. ^([HanKam, 2001], 5) Patut diingat bahwa *data mining* sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Nama *data mining* berasal dari kemiripan antara pencarian informasi yang bernilai dari *database* yang besar dengan menambang sebuah gunung untuk sesuatu yang bernilai ^[Sum, 2006]. Keduanya memerlukan penyaringan melalui sejumlah besar material, atau menyelidiki dengan cerdas untuk mencari keberadaan sesuatu yang disebut bernilai tadi. *Data mining* ^[Mai, 2005] merupakan inti dari proses *Knowledge Discovery in Database* (KDD). KDD ^[Mai, 2005] adalah proses terorganisir untuk mengidentifikasi pola yang valid, baru, berguna, dan dapat dimengerti dari sebuah *data set* yang besar dan kompleks.

Data Mining sebenarnya memiliki akar yang panjang dari berbagai ilmu seperti kecerdasan buatan (*artificial intelligence*), *machine learning*, statistik dan *database*. Beberapa teknik yang sering disebut-sebut dalam literatur *data mining* antara lain: *clustering*, *classification*, *association rule mining* dan *sequential pattern*.

Perangkat lunak *data mining* adalah satu dari sekian banyak alat untuk menganalisa data. Dengan perangkat lunak tersebut, pengguna diberikan

kesempatan untuk menganalisa data dari berbagai sudut pandang, mengkategorikannya, dan menyimpulkan relasi yang teridentifikasi.

2.1.2 Gudang Data

Teknologi Data Warehouse adalah suatu set konsep dan perangkat yang memungkinkan penyediaan akses atas informasi pada sebuah Departemen Pemerintah. Data Warehouse memungkinkan suatu Departemen Pemerintah untuk melakukan analisa data dan membuat laporan-laporan yang dibutuhkan bagi para analis dan pengambil keputusan ^[Pram, 2006].

Peningkatan dramatis dalam *data capturing*, kekuatan pemrosesan, transmisi data dan kemampuan penyimpanan, memungkinkan organisasi untuk mengintegrasikan berbagai macam basis data mereka dalam gudang data (*data warehouse*). Gudang data, seperti juga *data mining*, adalah terminologi yang baru walaupun konsepnya telah ada selama bertahun-tahun. Gudang data merepresentasikan visi ideal dari pemeliharaan sebuah repositori terpusat dari seluruh organisasi. Pemusatan data dibutuhkan untuk memaksimalkan akses dan analisa pengguna. Perkembangan teknologi yang dramatis menyebabkan visi ini menjadi kenyataan bagi suatu organisasi. Perkembangan dramatis yang hampir sama juga terjadi pada perangkat lunak analisa data yang membuat pengguna dapat menganalisa data secara lebih mudah. Perangkat lunak analisa data inilah yang menyokong *data mining*.

2.1.3 Proses Penemuan Pengetahuan

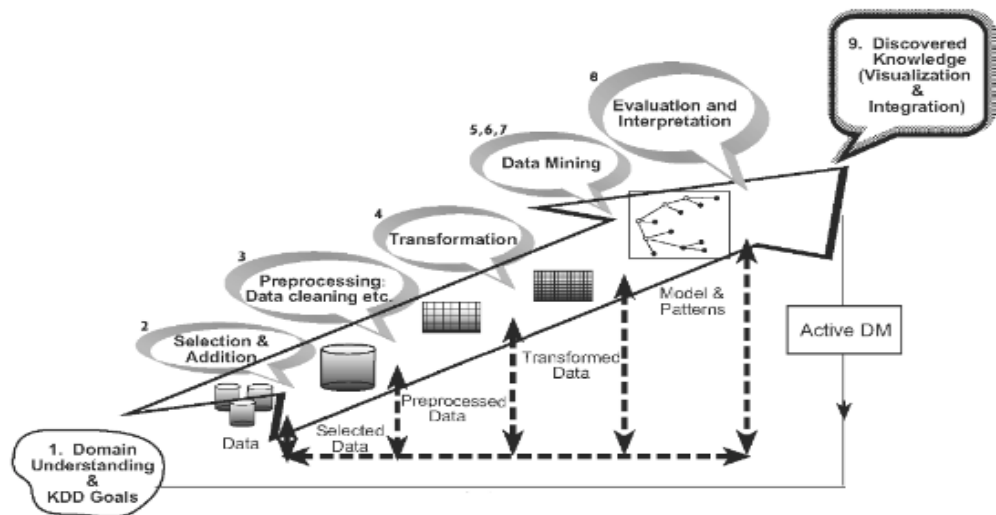
Permasalahan yang berhubungan dengan *data mining* biasanya berupa statistik. Perbedaan antara statistik dan teknik *data mining* adalah dalam jumlah data. Data pada saat ini luar biasa besar, sangat diperlukan untuk mengotomasi beberapa aspek dari pengumpulan data, persiapan dan pemrosesan. Walaupun demikian, untuk analisa kumpulan data tersebut pada akhirnya masih memerlukan pekerjaan manusia.

Pada saat ini teknologi informasi berskala besar berevolusi memisahkan antara transaksi dan sistem analisa, *data mining* menyediakan hubungan antara keduanya. Perangkat lunak *data mining* menganalisa relasi dan pola yang tersimpan dalam data transaksi. Data mining merupakan inti

dari proses *Knowledge Discovery in Database* (KDD) [Mai, 2005]. KDD adalah proses terorganisir untuk mengidentifikasi pola yang valid, baru, berguna, dan dapat dimengerti dari sebuah data set yang besar dan kompleks [Mai, 2005].

Langkah-langkah dalam KDD [Mai, 2005].

1. Pembentukan pemahaman domain aplikasi. Pada tahap ini menentukan tujuan dari *end-user* dan bagian terkait dimana KDD dilakukan. Mengembangkan pemahaman tentang domain aplikasi ini adalah awal langkah persiapan. Ini mempersiapkan adegan untuk memahami apa yang harus dilakukan dengan banyak keputusan (tentang transformasi, algoritma, representasi, dll). Orang-orang yang bertanggung jawab atas proyek KDD perlu memahami dan menentukan tujuan dari pengguna akhir dan lingkungan di mana pengetahuan penemuan proses akan berlangsung (termasuk pengetahuan awal yang relevan). Sebagai KDD hasil proses, mungkin ada revisi dan perbaikan dari langkah ini. Memiliki memahami tujuan KDD, *preprocessing* data dimulai, sebagaimana didefinisikan dalam berikutnya tiga langkah (perhatikan bahwa beberapa metode di sini mirip dengan algoritma Data Mining, tetapi digunakan dalam konteks *preprocessing*).



Gambar 2.1. langkah-langkah KDD

2. Memilih dan menciptakan satu data set untuk mendukung proses penemuan *knowledge* akan dilakukan. Penentuan data yang akan

digunakan untuk proses KDD dilakukan pada tahap ini. Mencari data yang tersedia, memperoleh data tambahan yang dibutuhkan, mengintegrasikan semua data untuk KDD ke dalam sebuah data set, termasuk atribut yang diperlukan dalam proses KDD. Terdapat interaktif dan iteratif dari KDD tersebut. Dimulai dengan data yang tersedia baik mengatur dan kemudian mengembang dan mengamati efeknya dalam KDD.

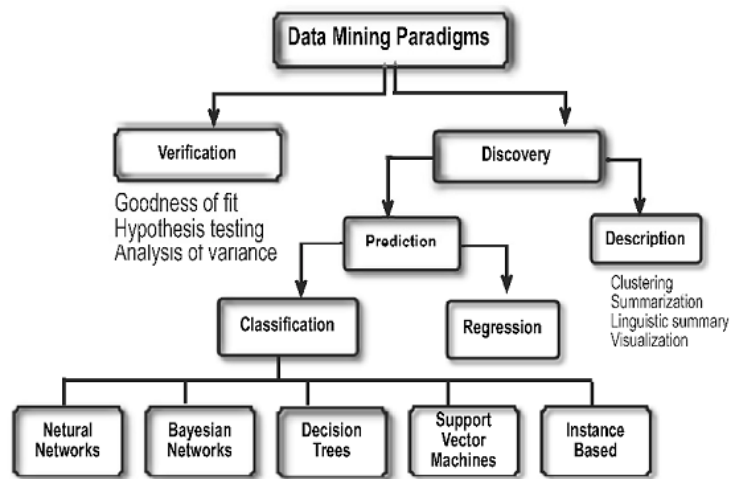
3. *Preprocessing* dan *cleansing*. Dalam tahap ini kehandalan data ditingkatkan. Termasuk data *clearing*, seperti menangani data yang tidak lengkap, menghilangkan gangguan atau *outlier*. Termasuk menggunakan metode statistik yang kompleks, atau melakukan penambangan spesifik data dengan algoritma dalam KDD.
4. Transformasi data. Pada tahap ini, generasi data yang lebih baik untuk data mining dipersiapkan dan dikembangkan, membuat data menjadi lebih baik menggunakan metode reduksi dimensi dan transformasi atribut. Sebagai contoh, dalam pemeriksaan medis, hasil bagi atribut mungkin sering menjadi faktor yang paling penting, dan tidak satu persatu. Di pemasaran, kita mungkin perlu mempertimbangkan efek di luar kendali kita serta upaya dan temporal isu (seperti mempelajari pengaruh akumulasi iklan). Namun, bahkan jika kita tidak menggunakan transformasi yang tepat di awal, kita dapat memperoleh efek mengejutkan bahwa petunjuk kepada kita tentang transformasi diperlukan (Pada iterasi berikutnya). Dengan demikian proses KDD mencerminkan kepada dirinya sendiri dan menyebabkan pemahaman tentang transformasi yang dibutuhkan (seperti pengetahuan ringkas dari sebuah ahli dalam bidang tertentu mengenai indikator terkemuka kunci).
5. Memilih tugas data mining yang cocok. Pada tahap ini ditentukan tipe data mining yang akan digunakan, apakah klasifikasi, regresi, atau clustering, tergantung pada tujuan KDD dan tahap sebelumnya.
6. Memilih algoritma data mining. Pemilihan algoritma yang paling tepat untuk menemukan pola dilakukan pada tahap ini. Ada dua tujuan utama dalam Data Mining: prediksi dan deskripsi. Prediksi sering

disebut sebagai *Supervised Data Mining*, sementara deskriptif Data Mining meliputi aspek-aspek *unSupervised* dan visualisasi Data Mining. Sebagian besar data teknik pertambangan didasarkan pada pembelajaran induktif, di mana model yang dibangun secara eksplisit maupun implisit oleh generalisasi dari jumlah yang memadai pelatihan data training. Asumsi yang mendasari pendekatan induktif adalah bahwa model terlatih ini berlaku untuk kasus

masa depan. Strategi ini juga memperhitungkan tingkat meta-learning untuk set tertentu dari data yang tersedia.

7. Penggunaan algoritma data mining. Pada tahap ini dilakukan implementasi dari algoritma data mining yang telah ditentukan pada tahap sebelumnya. Misalnya dengan menseting parameter kontrol algoritma, seperti jumlah minimum kasus dalam daun tunggal dari pohon keputusan.
8. Evaluasi, pada tahap ini dilakukan evaluasi dan penerjemahan dari pola yang diperoleh, sehubungan dengan tujuan yang ditetapkan pada langkah pertama. Langkah ini berfokus pada komprehensibilitas dan kegunaan dari model induksi. Pada langkah ini pengetahuan ditemukan juga terdokumentasi untuk penggunaan lebih lanjut. Langkah terakhir adalah penggunaan dan umpan balik secara keseluruhan pada pola dan hasil penemuan diperoleh dengan *Data Mining*.
9. Penggunaan pengetahuan yang ditemukan yakni memasukkan pengetahuan ke dalam sistem lain untuk ditindaklanjuti. Pengetahuan menjadi aktif dalam arti bahwa kita dapat membuat perubahan ke sistem dan mengukur dampak. Sebenarnya keberhasilan langkah ini menentukan efektivitas proses KDD secara keseluruhan. Ada banyak tantangan dalam langkah ini, kehilangan "kondisi laboratorium". Misalnya, pengetahuan itu ditemukan dari sebuah snapshot statis tertentu (biasanya sampel) dari data, tapi sekarang data menjadi dinamis.

Berikut adalah gambar klasifikasi data mining ^[Mai, 2005]:



Gambar 2.2. Klasifikasi Data Mining

Ada banyak metode data mining digunakan untuk tujuan yang berbeda dan tujuan. Pengklasifikasian tersebut untuk untuk membantu dalam memahami berbagai metode, keterkaitan dan pengelompokan. Hal ini berguna untuk membedakan antara dua jenis Data Mining: verifikasi berorientasi (sistem memverifikasi hipotesis pengguna) dan penemuan berorientasi (sistem menemukan aturan baru dan pola mandiri).

Pada gambar II.2 di atas menunjukkan langkah-langkah dalam proses data mining, proses dalam tahapan data mining terdiri dari tiga langkah utama [Mai, 2005], yaitu:

1. Data Preparation

Pada langkah ini, data dipilih, dibersihkan, dan dilakukan *preprocessed* mengikuti pedoman dan *knowledge* dari ahli domain yang menangkap dan mengintegrasikan data internal dan eksternal ke dalam tinjauan organisasi secara menyeluruh.

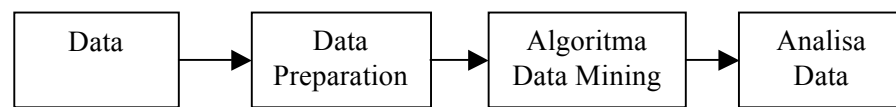
2. Penggunaan Algoritma Data Mining

Pada langkah ini digunakan algoritma data mining untuk memudahkan dalam melakukan identifikasi data dan mengintegrasikan keseluruhan data yang sudah ditemukan.

3. Tahap Analisa

Keluaran dari data mining dievaluasi untuk melihat apakah *knowledge*

domain ditemukan dalam bentuk *rule* yang telah diekstrak dari jaringan.



Gambar 2.3. Langkah-langkah dalam Proses *Data Mining* [Mai & Rok, 2010]

2.1.4 Teknik-Teknik *Data Mining*

Dengan definisi *data mining* yang luas, ada banyak jenis teknik analisa yang dapat digolongkan dalam *data mining*. Banyak praktisi *data mining* menyetujui tentang suatu area aplikasi tertentu. Berbagai teknik *data mining* dapat digunakan untuk setiap tipe fungsi. Teknik-teknik ini terdiri atas algoritma spesifik yang dapat digunakan untuk setiap fungsi. Memahami bagaimana teknik-teknik ini bekerja dapat membantu dalam memilih teknik yang sesuai untuk memecahkan suatu problema. Beberapa teknik *data mining* antara lain:

2.1.4.1 *Association Rule Mining*

Association rule mining adalah teknik *mining* untuk menemukan aturan assosiatif antara suatu kombinasi item. Contoh dari aturan assosiatif dari analisa pembelian di suatu pasar swalayan adalah bisa diketahui berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu. Dengan pengetahuan tersebut pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu. Penting tidaknya suatu aturan assosiatif dapat diketahui dengan dua parameter, *support* yaitu persentase kombinasi item tersebut dalam *database* dan *confidence* yaitu kuatnya hubungan antar item dalam aturan assosiatif.

Algoritma yang paling populer untuk masalah ini adalah Apriori [Agra dan Sri, 2007] dengan paradigma *candidate generate and test*, yaitu pembuatan kandidat kombinasi item yang mungkin berdasar aturan tertentu lalu diuji apakah kombinasi item tersebut memenuhi syarat *support* minimum. Kombinasi item yang memenuhi syarat tersebut disebut *frequent itemset*,

yang nantinya dipakai untuk membuat aturan-aturan yang memenuhi syarat *minimum confidence*. Algoritma baru yang lebih efisien bernama FP-Growth [Han, 2007]

2.1.4.2 Classification

Classification adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Model itu sendiri bisa berupa aturan “jika-maka”, berupa *decision tree*, formula matematis atau *neural network*. *Decision tree* adalah salah satu metode *classification* yang paling populer karena mudah untuk diinterpretasi oleh manusia.

Algoritma *decision tree* yang paling terkenal adalah C4.5 ([Quin 1993], 102), tetapi akhir-akhir ini telah dikembangkan algoritma yang mampu menangani data skala besar yang tidak dapat ditampung di memori seperti RainForest [J. Geh 2007]. Metode-metode *classification* yang lain adalah *Bayesian*, *neural network*, *genetic algorithm*, *fuzzy*, *case-based reasoning*, dan *k-nearest neighbor*. Proses *classification* biasanya dibagi menjadi dua fase: *learning* dan *test*. Pada fase *learning*, sebagian data yang telah diketahui kelas datanya diumpungkan untuk membentuk model perkiraan. Kemudian pada fase *test*, model yang sudah terbentuk diuji dengan sebagian data lainnya untuk mengetahui akurasi dari model tersebut. Bila akurasinya mencukupi model ini dapat dipakai untuk memprediksi kelas data yang belum diketahui.

2.1.4.3 Clustering

Berbeda dengan *association rule mining* dan *classification* dimana kelas data telah ditentukan sebelumnya, *clustering* melakukan pengelompokan data tanpa berdasarkan kelas data tertentu. Bahkan *clustering* dapat dipakai untuk memberikan label pada kelas data yang belum diketahui itu. Karena itu *clustering* sering digolongkan sebagai metode *unsupervised learning*. Prinsip dari *clustering* adalah memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/*cluster*. *Clustering* dapat dilakukan pada data yang memiliki beberapa atribut yang dipetakan sebagai ruang

multidimensi. Banyak algoritma *clustering* memerlukan fungsi jarak untuk mengukur kemiripan antar data, diperlukan juga metode untuk normalisasi bermacam atribut yang dimiliki data. Beberapa kategori algoritma *clustering* yang banyak dikenal adalah metode partisi dimana pemakai harus menentukan jumlah k partisi yang diinginkan lalu setiap data diuji untuk dimasukkan pada salah satu partisi, metode lain yang telah lama dikenal adalah metode hierarki yang terbagi dua lagi: *bottom-up* yang menggabungkan *cluster* kecil menjadi *cluster* lebih besar dan *top-down* yang memecah *cluster* besar menjadi *cluster* yang lebih kecil. Kelemahan metode ini adalah bila salah satu penggabungan/pemecahan dilakukan pada tempat yang salah, tidak dapat didapatkan *cluster* yang optimal. Pendekatan yang banyak diambil adalah menggabungkan metode hierarki dengan metode *clustering* lainnya seperti yang dilakukan oleh Chameleon ^[G. Kar 2007].

2.1.5 Alat *Data Mining*

Saat ini banyak sekali *data mining tool* dikembangkan oleh lembaga riset suatu universitas atau perusahaan yang bergerak di bidang teknologi informasi. Untuk melakukan prediksi kelulusan tepat waktu, beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan untuk prediksi (dalam keadaan yang tepat). Beberapa teknik klasifikasi yang dapat digunakan untuk prediksi seperti Neural Network, Decision Tree, Naïve Bayes dan K-Nearest Neighbor ^[Pra, 2012].

Pada penelitian ini menggunakan *Waikato Environment Knowledge and Analysis* (WEKA), dengan algoritma C4.5, Naïve Bayes, dan Neural Network.

2.1.5.1 *Waikato Environment Knowledge and Analysis* (WEKA)

WEKA adalah sebuah perangkat lunak *data mining/machine learning* yang dibangun oleh *Department of Computer Science University of Waikato* di New Zealand. Weka dikembangkan mulai tahun 1999 sampai sekarang. WEKA ditulis dengan bahasa pemrograman Java. Saat ini WEKA didistribusikan dibawah lisensi

publik GNU. *Software* ini banyak digunakan untuk riset, edukasi dan aplikasi pelengkap (komplemen) *data mining* yang dibuat oleh Witten dan Frank.

Karena WEKA didistribusikan dibawah lisensi GNU, maka WEKA dapat digunakan secara penuh tanpa adanya biaya untuk membayar lisensi. Universitas Waikato menyediakan *source* WEKA di situs *web* mereka. Pengguna dapat langsung *men-download* dari <http://www.cs.waikato.ac.nz/~ml/weka>. Mereka menyediakan dalam berbagai versi, baik versi untuk linux maupun untuk windows. Saat ini versi yang terbaru adalah WEKA 3.4.6

Secara umum WEKA menyediakan 3 fasilitas untuk *data mining*. Ketiga fasilitas tersebut adalah *data processing*, *data mining*, dan *visualization tools*. *Data processing* adalah proses untuk membuang data yang tidak diperlukan. WEKA menyebut istilah *data processing* dengan *filter*. Beberapa *filter* yang disediakan seperti, *discretize*, *nominaltobinary* dan *add cluster*.

Data mining adalah fasilitas untuk *me-mining* data sehingga ditemukan pola atau asosiasi dari sekumpulan data. WEKA menyediakan tiga fungsi *data mining*. Ketiga fungsi tersebut adalah *classification*, *clustering* dan *association rule*. Masing–masing fungsi tersebut terdiri dari beberapa algoritma. Misalnya *classification* terdiri dari *Adtree*, *ID3*, *J48* dan sebagainya. *Clustering* terdiri atas *CobWeb*, *EM*, *FarhtestFirst*, *MakeDensityBaseClusterer* dan *SimpleKMeans*. Sedangkan *Association Rules* terdiri atas *Apriori*, *PredictiveApriori* dan *Tertius*.

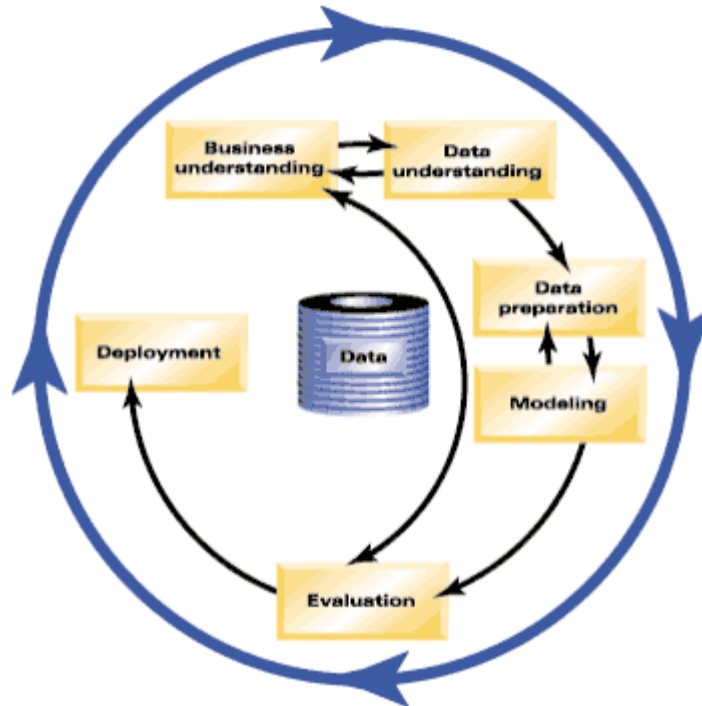
BAB III

METODE PENELITIAN

Metodologi atau tahapan penelitian diperlukan sebagai kerangka dan panduan proses penelitian, sehingga rangkaian proses penelitian dapat dilakukan secara terarah, teratur dan sistematis.

3.1. Langkah-Langkah Penelitian

Penelitian ini didesain dengan menggunakan model CRISP-DM (*Cross Industry Standard Process for Data Mining*), dalam metode ini terdapat 6 tahapan [Larose, 2005].



Gambar 3.1. Tahap CRISP-DM (*Cross Industry Standard Process for Data Mining*)

1. *Business/Research Understanding Phase*
2. *Data Understanding Phase* (Fase Pemahaman Data)
3. *Data Preparation Phase* (Fase Pengolahan Data)
4. *Modeling Phase* (Fase Pemodelan)
5. *Evaluation Phase* (Fase Evaluasi)
6. *Deployment Phase* (Fase Penyebaran)

3.1.1. *Business/Research Understanding Phase*

Data yang diperoleh dari database DMQ pada Perguruan Tinggi Raharja, ternyata selama ini belum pernah dilakukan penggalian kekayaan terhadap data tersebut. Data tersebut sudah cukup banyak dan jika dilakukan *data mining* untuk menghasilkan kelulusan tepat waktu tentunya akan sangat

bermanfaat.

Diketahui bahwa Perguruan Tinggi Raharja belum memanfaatkan database tersebut, dan dalam menentukan prediksi kelulusan masih menggunakan metode apriori dengan tingkat subyektifitas yang tinggi. Sampai saat ini, masih belum ditemukan algoritma yang paling akurat dalam prediksi kelulusan tepat waktu. Untuk itu, dalam penelitian ini akan mengkaji dan membuat model hasil komparasi *algoritma C4.5*, *Naïve Bayes* dan *Neural Network* dalam menentukan algoritma yang paling akurat dan menghasilkan rule prediksi kelulusan tepat waktu. Dari algoritma terpilih untuk memudahkan dalam melihat prediksi kelulusan tepat waktu, dibuatkan GUI (*Graphical User Interface*) menggunakan *engine Java*.

3.1.2. *Data Understanding Phase (Fase Pemahaman Data)*

Data yang diperoleh dari database DMQ Perguruan Tinggi Raharja pada tahun 2012 yaitu sebanyak 5842 mahasiswa. Attribute atau variable yang digunakan sebanyak 9 atribut (Nim, Nama Mahasiswa, Kujur, Jenjang, Status, Jenis Kelamin, IPK, IMK, R.IPK Jurusan). Dilakukan pemrosesan terhadap data tersebut sehingga digunakan sebanyak 6 atribut atau variable yang digunakan dalam prediksi kelulusan tepat waktu adalah: Nim, Nama Mahasiswa, Jenjang Pendidikan, Jurusan, IPK dan IMK. Dari 6 atribut 2 adalah *predictor* yaitu IPK dan IMK dan 1 atribut tujuan yaitu kelulusan tepat waktu. Attribute yang menjadi parameter dalam prediksi kelulusan tepat waktu dapat dilihat sebagai berikut:

Table 3.1. Atribut dan Nilai Kategori IPK, IMK dan Prediksi

No.	Atribut	Nilai
1.	IPK	Sangat Baik Baik Cukup Kurang
2.	IMK	

		Sangat Baik Baik Cukup Kurang
3.	Prediksi	
		Lulus Tepat Waktu Lulus Tidak Tepat Waktu

3.1.3. *Data Preparation Phase (Fase Pengolahan Data)*

Dari 5842 data mahasiswa diambil data dari angkatan 2009 dan 2010 dengan pertimbangan sudah melewati semester II (tingkat stabilitas dalam menghadiri perkuliahan sudah tinggi) dan masih ada semester yang mereka akan tempuh (untuk memprediksi kelulusan tepat waktu). Setelah melakukan query terhadap database DMQ maka diperoleh 1033 record mahasiswa untuk angkatan 2009 dan 2010. Dalam penelitian ini diambil 10% sampel random atau sebanyak 103 record data dari database DMQ yang berisi data IMK dan IPK yang akan diolah.

Untuk selanjutnya dilakukan teknik *preprocessing* agar kualitas data yang diperoleh lebih baik dengan cara, (Vecellis, 2009) :

1. *Data validation*, untuk mengidentifikasi dan menghapus data yang ganjil (*outlier/noise*), data yang tidak konsisten, dan data yang tidak lengkap (*missing value*)
2. *Data integration and transformation*, untuk meningkatkan akurasi dan efisiensi algoritma. Data yang digunakan dalam penulisan ini bernilai kategorikal. untuk model *neural network*, data ditransformasi ke dalam angka menggunakan *software Weka*.
3. *Data size reduction and discretization*, untuk memperoleh data set dengan jumlah atribut dan *record* yang lebih sedikit tetapi bersifat informatif. Di dalam data *training* yang digunakan dalam penelitian ini, dilakukan seleksi atribut dan penghapusan data duplikasi menggunakan *software Weka*.

3.1.4. *Modeling Phase* (Fase Pemodelan)

Pada tahapan ini merupakan tahapan pemrosesan data training yang diklasifikasikan oleh model dan kemudian menghasilkan sejumlah aturan. Pada penelitian ini menggunakan tiga algoritma yaitu *algoritma C4.5*, *Naïve Bayes* dan *Neural Network*.

3.1.5. *Evaluation Phase* (Fase Evaluasi)

Pada fase ini dilakukan pengujian terhadap model-model yang bertujuan untuk mendapatkan model yang paling akurat. Evaluasi dan validasi dilakukan dengan menggunakan metode *Confusion Matrix* dan kurva ROC (*Receiver Operating Characteristic*).

3.1.6. *Deployment Phase* (Fase Penyebaran)

Setelah pembentukan model selanjutnya dilakukan analisa dan pengukuran pada tahap sebelumnya, pada tahap ini diterapkan model atau rule yang paling akurat dalam prediksi kelulusan tepat waktu dan selanjutnya dapat digunakan untuk mengevaluasi data baru.

BAB IV

BIAYA DAN JADWAL PENELITIAN

4.1 Anggaran Biaya

No	Jenis Pengeluaran	Biaya yang diusulkan
1	Gaji dan upah	Rp. 2.510.000
2	Bahan habis pakai dan peralatan	Rp. 6.275.000
3	Perjalanan	Rp. 1.506.000
4	Publikasi, seminar, laporan	Rp. 3.882.500

		Rp. 14.173.500
--	--	-----------------------

4.2 Jadwal Penelitian

Berikut ini adalah jadwal penelitian yang dilakukan:

Jadwal kegiatan	Bulan Ke / 2013-2014					
	1	2-3	4-6	7-9	10-11	12
Studi objek penelitian						
Observasi data						
Pengumpulan bahan literatur						
Penyusunan desain penelitian						
Pengolahan data						
Pemodelan						
Evaluasi sistem						
Penyebaran						
Pembuatan paper/artikel						
Pengiriman naskah paper/artikel						

DAFTAR PUSTAKA

- [Bor 2006] Borgelt, Cristian. 2006. *An Implementation FP-Growth Algorithm*. Diambil November 2012, dari <http://www.borgelt.net/papers/fpgrowth.pdf>.
- [Cas 2007] Castro, Jose. 2007. *C4.5 Decision Tree Creation and Pruning*. Diambil November 2012 dari http://mlrg.cecs.ucf.edu/MLRG_documents/c4.5.pdf.

- [Con 2004] Connolly, Thomas. Begg, Carolyn. 2004. *Database System A Practical Approach, Implementation, and Management* Fourth Edition: Addison Wesley.
- [Gar 2012] Garner, R Stephen. WEKA: *The Waikato Environment for Knowledge Analysis*. Diambil November 2012 dari <http://www.cs.waikato.ac.nz/~ml/publications/1995/Garner95-WEKA.pdf>
- [Geh 1998] Gehrke, Johannes. Ramakrishnan, Raghu. Ganti, Venkatesh. 1998. *RainForest – A Framework for Fast Decision Tree Construction of Large Datasets*. Diambil November 2012 dari <http://www.cs.cornell.edu/johannes/papers/1998/vldb1998-rainforest.pdf>
- [Gor 2011] Gorunescu Florin. 2011. *Data Mining: Concept, Models and Techniques*. Springer-Verleg Berlin Heidelberg
- [Han 2007] Han, Jiawei. Pei, Jian. 2007. *Mining Frequent Pattern by Pattern-Growth: Metodology and Implication*. Diambil November 2012 dari <http://www.acm.org/sigs/sigkdd/explorations/issues/2-2-2000-12/han.pdf>
- [Hut 2005] Hutabarat, Bernaridho I. 2005. *Data Warehousing dengan SQL Server 2005*: Alex Media Komputindo.
- [Kar 2007] Karypis, George. Han, Eui-Hong(Sam). Kumar, Vipin. 2007. *CHAMELEON: A hierarchical clustering algorithm using dynamic modeling*. Diambil November 2012 dari <http://www-users.cs.umn.edu/~kumar/papers/chameleon.ps>
- [Kam 2008] KAMII (Komunitas Mahasiswa Sistem Informasi Yogyakarta. 2008. *Sistem Informasi* (http://kamii_yogyakarta.tripod.com). [Accessed 12 Desember 2012].
- [Liu 2007] Liu, Bing. 2007. *Integrating Classification and Association Rule Mining*. Diambil November 2012 dari http://www.comp.nus.edu.sg/~dm2/publications/kdd98_1.ps
- [Lar 2006] Larose T. Daniel. 2006. *Data Mining Methods and Models*. John Wiley & Sons, Inc Publication

- [Mai 2005] Maimon, Rockah. 2005. *Data Mining and Knowledge Discovery Handbook*. Springer
- [Prab 2012] Prabowo. 2012. *Aneka Teknik, Piranti dan Penerapan Data Mining: Studi Kasus Peramalan Harga Saham Industri Telekomunikasi Berbasis Jaringan Saraf Tiruan*. Modul Perkuliahan Universitas Budi Luhur
- [Rah 2009] Rahardja, Hidayati, Badar. 2009. *Educational Measurement dengan Menerapkan Indeks Mutu Kumulatif*. Jurnal Forum Ilmiah INDONUSA. Vol.6 No. 3:213-221
- [Sum 2006] Sumathi, Sivanandam. 2006. *Introduction to Data Mining and Its Applications*. Springer Verlag Berlin Heidelberg
- [Sof 2008] Sofa, H, S.IP, M.Pd. 2008. *Nilai Informasi bagi Pengambilan Keputusan*. (<http://massofa.wordpress.com>). [Accessed 19 Desember 2012]
- [Wu 2009] Wu Xindom, Kumar Vivin. 2009. *The Top Ten Algorithms in Data Mining*. Chapman & Hill Book
- [Wit 2011] Witten H. Ian, Frank Eibe, Hall A. Mark. 2011. *Data Mining: Practical Machine Learning Tolls and Techniques*. Elseiver Inc

Lampiran-Lampiran :

Lampiran 1. Justifikasi anggaran penelitian

1. Honor						
Pelaksana	Honor/jam (Rp)	Waktu (jam/minggu)	Minggu	Honor per tahun (Rp)		
				Thn I	Thn ...	Thn n
Ketua	15.687.5	8		1506000		
Anggota 1	10.458.33	8		1004000		
Sub total (Rp)				2510000		
2. Peralatan penunjang						
Material	Justifikasi Pemakaian	Kuantitas	Harga Satuan (Rp)	Harga Peralatan Penunjang (Rp)		
				Thn I	Thn ...	Thn n
laptop	Pengolahan data	1	4500000	4500000		
printer	pencetakan	1	1425000	1425000		
Sub total (Rp)				5925000		
3. Bahan Habis Pakai						
Material	Justifikasi Pemakaian	Kuantitas	Harga Satuan (Rp)	Biaya per Tahun (Rp)		
				Thn I	Thn ...	Thn n
Tinta printer	Penggantian tinta printer	1 bh	318000	318000		
Kertas A4	Cetak dokumen	1 rim	32000	32000		
Sub total (Rp)				350000		
4. Perjalanan						
Perjalanan	Justifikasi Perjalanan	Kuantitas	Harga Satuan (Rp)	Biaya per Tahun (Rp)		
				Thn I	Thn ...	Thn n
Perjalanan ke kota UI depok	Pencarian referensi	4 x PP	188250	1506000		
Sub total (Rp)				1506000		
5. Lain-lain						
Kegiatan	Justifikasi	Kuantitas	Harga Satuan (Rp)	Biaya per Tahun (Rp)		
				Thn I	Thn ...	Thn n
Laporan semesteran	Pembuatan laporan	2	250000	500000		
Desiminasi hasil penelitian	seminar	1	1000000	1382500		
Call paper	Mengikuti seminar	1	2000000	2000000		
Sub total (Rp)				3882500		
TOTAL ANGGARAN YANG DIPERLUKAN SETIAP TAHUN (Rp)				Thn I	Thn ...	Thn n
				14.173.500		

TOTAL ANGGARAN YANG DIPERLUKAN SELURUH TAHUN (Rp)	14.173.500
---------------------------------------------------	------------

Lampiran 2. Susunan organisasi tim peneliti dan pembagian tugas

Format Susunan Organisasi Tim Peneliti/Pelaksana dan Pembagian Tugas

No	Nama / NIDN	Instansi Asal	Bidang Ilmu	Alokasi Waktu (jam/minggu)	Uraian Tugas
1.	Asep Saefulloh/0407056701	STMIK Raharja	Tek. Sistem Informasi	8/minggu	Mengorganisir seluruh kegiatan, melakukan kerjasama dg instansi terkait, pengolahan data base, membuat laporan, diseminasi, paper
2.	Himawan/0313098502	STMIK Raharja	Teknik Informatika	8/minggu	Melakukan programming sistem, laporan, diseminasi, paper

Lampiran 3. Biodata ketua dan anggota

Format Biodata Ketua/Anggota Tim Peneliti/Pelaksana

A. Identitas Diri

1	Nama Lengkap (dengan gelar)	Asep Saefulloh, S.Pd, M.Kom
2	Jenis Kelamin	L
3	Jabatan Fungsional	Lektor
4	NIP/NIK/Identitas lainnya	056007
5	NIDN	0407056701
6	Tempat, Tanggal Lahir	Garut, 07 Mei 1967
7	E-mail	asaefullah@gmail.com ; asep.saefullah@raharja.co
8	Nomor Telepon/HP	085811342550
9	Alamat Kantor	Jl. Jenderal Sudirman no. 40 Cikokol Tangerang
10	Nomor Faks	021-5529742
11	Lulusan yang Telah Dihasilkan	S-1 = 45
12	Nomor Telepon	021-5529692
13	Mata Kuliah yang Diampu	1. Elektronika 2. Sistem digital 3. Pengantar Teknologi Informasi 4. Sistem basis data

B. Riwayat Pendidikan

	S-1	S-2	S-3
Nama Perguruan Tinggi	UPI/IKIP Bandung	Universitas Budi Luhur	
Bidang Ilmu	Pend. Tek. Elektronika Komunikasi	Tek. Sistem Informasi	
Tahun Masuk-Lulus	1987-1993	2011-2013	
Judul Skripsi/Tesis/Disertasi	FM Dolby stereo	Metode klasifikasi data mining untuk memprediksi kelulusan tepat waktu	
Nama Pembimbing/Promotor	Drs. Mukhidin, M.Pd	Dr. Moedjiono, M.Sc	

C. Pengalaman Penelitian Dalam 5 Tahun Terakhir
(Bukan Skripsi, Tesis, maupun Disertasi)

No.	Tahun	Judul Penelitian	Pendanaan	
			Sumber*	Jml (Juta Rp)
1	2013	Aplikasi pembayaran sewa studio music 99 berbasis WEB	Studio music 99	1.500.000
2	2012	Pendeteksi kebocoran gas LPG menggunakan mikrokontroler AT89S2051 melalui handphone sebagai media informasi	institusi	6.250.000
3	2011	Kendali motor industry dengan metode client server berbasis WEB	Institusi, industri	9.000.000
4	2010	Robot IP Network (Robin) yang dikendalikan melalui website	Institusi	5.450.000

* Tuliskan sumber pendanaan baik dari skema penelitian DIKTI maupun dari sumber lainnya.

D. Pengalaman Pengabdian Kepada Masyarakat dalam 5 Tahun Terakhir

No.	Tahun	Judul Pengabdian Kepada Masyarakat	Pendanaan	
			Sumber*	Jml (Juta Rp)
1	2013	Memberikan materi seminar guru se Kota Tangerang	-	-
2	2012	Memberikan materi workshop computer kepada siswa SMA	-	-
3	2011	Membuat jaringan RT/RW net di Perumnas II Kota Tangerang	Swadaya masyarakat	4.500.000

* Tuliskan sumber pendanaan baik dari skema pengabdian kepada masyarakat DIKTI maupun dari sumber lainnya.

E. Publikasi Artikel Ilmiah Dalam Jurnal dalam 5 Tahun Terakhir

No.	Judul Artikel Ilmiah	Nama Jurnal	Volume/Nomor/Tahun
1	Aplikasi pembayaran sewa studio music 99 berbasis	CSRID	Vol 5 NO 1/2013

	web		
2	Pendeteksi kebocoran gas LPG menggunakan mikrokontroler AT89S2051 melalui handphone sebagai media informasi	CSRID	Vol 4 No 2/2012
3	Ilusi tampilan CCIT dengan LED menggunakan teknologi POV	CSRID	Vol 2 No 1/2010
4	Robot IP Network yang dikendalikan melalui website	CSRID	Vol 2 No 2/2010
5	Kendali motor industry dengan metode client server berbasis web	CCIT	2011
6	Alrm fire detection with GUI	CCIT	2010
7	Simulasi kendali kecepatan mobil secara otomatis	CCIT	Vol 2 No 2/2009

F. Pemakalah Seminar Ilmiah (*Oral Presentation*) dalam 5 Tahun Terakhir

No.	Nama Pertemuan Ilmiah / Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1	SenmiBL	Analisis WEB e-commerce pada situs jakartanotebook.com dengan metode SWOT	Desember 2012/Jakarta
2	SNIF	Aplikasi face recognition menggunakan metode eigen face dan principal component analysis	Oktober 2012/Medan
3	SEMANTIK	Aplikasi steganografi untuk menyembunyikan teks dalam media image dengan metode LSB	Juni 2012/Semarang

G. Karya Buku dalam 5 Tahun Terakhir

No.	Judul Buku	Tahun	Jumlah Halaman	Penerbit
1	Pengendalian robot berbasis internet	2013	198	Pemprov Banten
2	Statistik deskriptif for IT	2012	238	Andi-Yogyakarta
3	Kewirausahaan	2011	343	Andi-Yogyakarta
4	Graphical User Interface	2011	186	Pemprov Banten
5	Mudah membuat Website dengan CSS-HTML	2009	125	Andi-Yogyakarta

H. Perolehan HKI dalam 5–10 Tahun Terakhir

No.	Judul/Tema HKI	Tahun	Jenis	Nomor P/ID
1	Peralatan belajar creative communication and innovative technology	2008	patent	C00200800523

I. Pengalaman Merumuskan Kebijakan Publik/Rekayasa Sosial Lainnya dalam 5 Tahun Terakhir

No.	Judul/Tema/Jenis Rekayasa Sosial Lainnya yang Telah Diterapkan	Tahun	Tempat Penerapan	Respon Masyarakat
1				

2				
3				
dst				

J. Penghargaan dalam 10 tahun Terakhir (dari pemerintah, asosiasi atau institusi lainnya)

No.	Jenis Penghargaan	Institusi Pemberi Penghargaan	Tahun
1	Sertifikat semi finalis Indonesia ICT Award	INAICTA-Depkominfo	2010

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan Hibah penelitian dosen pemula.

Tangerang, 10 Mei 2013
 Pengusul,

(Asep Saefulloh, SPd, M.Kom)

Lampiran 4. Surat pernyataan ketua peneliti



SEKOLAH TINGGI MANAJEMEN DAN ILMU KOMPUTER

STMIK RAHARJA

SK.MENDIKNAS. RI: 74/D/O/2001

Jl. Jenderal Sudirman No. 40 Modern Cikokol-Tangerang 15117 Telp.(02)5529692

SURAT PERNYATAAN KETUA PELITI/PELAKSANA

Yang bertanda tangan di bawah ini:

Nama : Asep Saefulloh, SPd, M.Kom
NIDN : 0407056701
Pangkat / Golongan : IIIC
Jabatan Fungsional : Lektor

Dengan ini menyatakan bahwa proposal penelitian saya dengan judul:

**PENERAPAN METODE KLASIFIKASI DATA MINING UNTUK PREDIKSI
KELULUSAN TEPAT WAKTU**

yang diusulkan dalam skema penelitian dosen pemula untuk tahun anggaran 2014 **bersifat original dan belum pernah dibiayai oleh lembaga/sumber dana lain.**

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku dan mengembalikan seluruh biaya penelitian yang sudah diterima ke kas negara.

Demikian pernyataan ini dibuat dengan sesungguhnya dan dengan sebenar-benarnya.

Mengetahui
Ketua Lembaga,

Cap dan tanda tangan

(M. Yusup, S.Kom)
026002

Tangerang, 10 Mei 2013
Yang menyatakan,

meterai Rp6000
tanda tangan

(Asep Saefulloh, S.Pd, MKom)
056007