

BAB II LANDASAN TEORI

2.1 *Data Mining*

Menurut (Hermawati, 2013) *Data mining* adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis induksi (*induction-based learning*) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari. *Knowledge Discovery in Databases* (KDD) adalah penerapan metode *saintifik* (untuk mengidentifikasi atau menemukan masalah) pada *data mining*. Dalam konteks ini *data mining* merupakan satu langkah dari proses KDD.

Secara garis besar tehnik-tehnik data mining adalah sebagai berikut :

a. Klasifikasi (*classification*)

Klasifikasi adalah menentukan sebuah *record* data baru ke salah satu dari beberapa kategori (atau klas) yang telah didefinisikan sebelumnya.

b. Regresi

Memprediksi nilai dari suatu variabel kontinyu yang diberikan berdasarkan nilai dari variabel yang lain, dengan mengasumsikan sebuah model ketergantungan linier atau nonlinier.

c. Klasterisasi (*clustering*)

Mempartisi data-set menjadi beberapa sub-set atau kelompok sedemikian rupa sehingga elemen-elemen dari suatu kelompok tertentu memiliki *set property* yang *dishare* bersama, dengan tingkat similaritas yang tinggi dalam satu kelompok dan tingkat similaritas antar kelompok yang rendah.

d. Kaidah asosiasi (*association rules*)

Mendeteksi kumpulan atribut-atribut yang muncul bersama (*co-occur*) dalam frekuensi yang sering, dan membentuk sejumlah kaidah dari kumpulan-kumpulan tersebut.

e. Pencarian pola sekuensial (*sequence mining*)

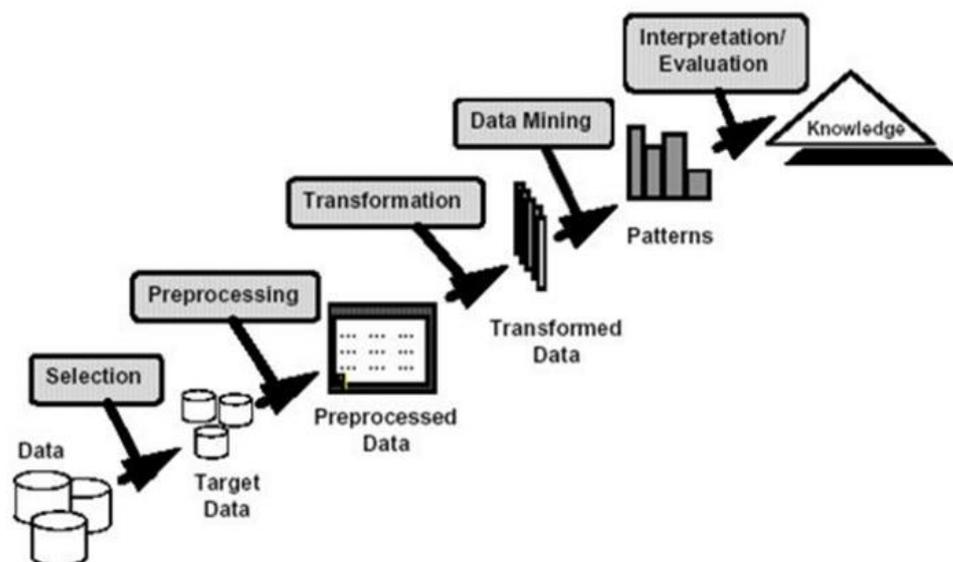
Mencari sejumlah *event* yang secara umum terjadi bersama-sama.

2.1.1 Operasi Data Mining

Operasi data mining menurut sifatnya dibedakan menjadi dua, yaitu:

- Prediksi (*prediction driven*) untuk menjawab pertanyaan apa dan sesuatu yang bersifat transparan. Operasi prediksi digunakan untuk validasi hipotesis, querying dan pelaporan.
- Penemuan (*discovery driven*) bersifat transparan dan untuk menjawab pertanyaan “mengapa?”. Operasi penemuan digunakan untuk analisis data eksplorasi, pemodelan prediktif, segmentasi database, analisis keterkaitan (*link analysis*) dan deteksi deviasi.

Adapun tahapan proses dalam penggunaan data mining dapat dilihat pada gambar 2.1 proses KDD.



Gambar 2.1 Proses KDD

Tahapan proses dalam penggunaan *data mining* yang merupakan proses *Knowledge Discovery in Database* (KDD) sseperti yang terlihat pada gambar 2.1 dapat diuraikan sebagai berikut :

- a. Memahami domain aplikasi untuk mengetahui dan menggali pengetahuan awal serta apa sasaran pengguna.
- b. Membuat target data-set yang meliputi pemilihan data dan fokus pada sub-set data.
- c. Pembersihan dan transformasi data meliputi eliminasi drau, *outliers*, *missing value* serta pemilihan fitur dan reduksi dimensi.
- d. Penggunaan algoritma *data mining* yang terdiri dari asosiasi, sekuensial, klasifikasi, dan lain-lain.
- e. Interpretasi, evaluasi dan visualisasi pola untuk melihat apakah ada sesuatu baru dan menarik dan dilakukan iterasi jika diperlukan.

2.2 Algoritma *K-Means*

Menurut (Santoso, 2007) dari beberapa tehnik klastering yang paling sederhana dan umum dikenal adalah klastering k-means. Dalam tehnik ini kita ingin mengelompokkan objek kedalam k atau kelompok klaster. Untuk melakukan klaster ini nilai k harus ditentukan terlebih dahulu. Algoritma yaitu :

- a. Pilih jumlah klaster k
- b. Inisialisasi k pusat klaster ini bisa dilakukan dengan berbagai cara yang paling sering dilakukan adalah dengan cara random. Pusat-pusat klaster diberi nilai awal dengan angka-angka random
- c. Tempatkan setiap data atau objek ke klaster terdekat kedekatan dua objek ditentukan berdasarkan jarak kedua onjek tersebut. Demikian juga kedekatan suatu data ke klaster tertentu ditentukan jarak antara data dengan pusat klaster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat klaster. Jarak paling dekat antara satu data dengan satu klaster tertentu akan menentukan suatu data masuk dalam klaster mana.

- d. Hitung kembali jarak pusat kluster dengan keanggotaan kluster yang sekarang pusat kluster adalah rata-rata dari semua data atau objek dalam kluster tertentu. Jika dikehendaki bisa juga memakai median dari kluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.
- e. Tugaskan lagi setiap objek dengan memakai pusat kluster yang baru. Jika pusat kluster sudah tidak berubah lagi, maka proses pengklasteran selesai. Atau kembali lagi ke langkah yang ketiga sampai pusat kluster tidak berubah lagi.

Rumus Jarak Eucclidean

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Dimana :

x : Titik data pertama

y : Titik data kedua

n: Jumlah karakteristik (atribut) dalam data mining

d(x,y) : Euclidean distance yaitu jarak data pada titik x dan titik y menggunakan kalkulasi matematika.

2.3 Bahasa Pemrograman Yang Digunakan

Komputer membutuhkan *software* untuk beroperasi dan membutuhkan sistem operasi atau program-program untuk membuat komponen-komponen komputer bekerja secara baik. Merupakan perangkat yang dapat dilihat oleh mata, tetapi tidak dapat diraba. *Software* juga sering digunakan untuk menunjukkan semua program yang dapat dipakai dalam sistem komputer. Dalam pengertian yang sempit, istilah ini menunjuk pada sebuah program yang dapat mempermudah pemakai dari berbagai jenis komputer untuk mendayagunakan *hardware* dengan baik. Untuk merancang dan membangun sistem ini pembuat membutuhkan *software-software* penunjang untuk memaksimalkannya antara lain :

2.3.1 HTML 5 (*Hypertext Markup Language 5*)

Menurut (Sianipar, 2015) HTML5 adalah sebuah markup untuk menstrukturkan dan menampilkan isi dari halaman web. HTML (yang pertama kali diciptakan pada tahun 1990 dan versi keempatnya, HTML4, pada tahun 1997) dan hingga bulan juni 2011 tetap dalam proses pengembangan. Tujuan utama pengembangan HTML5 adalah untuk memperbaiki teknologi HTML agar mendukung teknologi multimedia terbaru, mudah dibaca manusia dan juga mudah dimengerti oleh mesin.

HTML5 merupakan salah satu karya *World Wide Web Consortium*, W3C untuk mendefinisikan sebuah bahasa markup tunggal yang dapat ditulis dengan HTML ataupun XHTML. HTML5 merupakan jawaban atas pengembangan HTML 4.01 dan XHTML 1.1 yang selama ini berjalan terpisah. Dan di implementasikan secara berbeda-beda oleh banyak perangkat lunak pembuat web.

2.3.2 PHP (*Hypertext Preprocessor*)

Menurut (Ardhama, 2012) PHP *Hypertext Preprocessor* atau sering disebut php merupakan bahasa pemrograman berbasis server-side yang dapat melakukan parsing *script* php menjadi *script* web sehingga dari sisi *client* menghasilkan suatu tampilan yang menarik. Php merupakan pengembangan FI atau *Form Interface* yang dibuat oleh rasmus lerdoff pada tahun 1995.

Berbeda dengan HTML, kode PHP tidak diberikan secara langsung oleh server ketika ada permintaan atau request dari sisi client namun dengan cara pemrosesan dari server. Kode PHP disisipkan pada HTML. Perbedaan dari kode (*script*) HTML dan PHP yaitu setiap kode PHP ditulis selalu diberi tag pembuka yaitu `<?php` dan pada akhir kode PHP diberi tag penutup yaitu `?>`.

Kelebihan dan kekurangan PHP sebagai berikut :

- a. PHP menjadi populer karena kesederhanaannya dan kemampuannya dalam menghasilkan berbagai aplikasi web.
- b. Tidak detail untuk pengembangan skala besar.
- c. PHP adalah salah satu bahasa *server-side* yang di desain khusus untuk aplikasi web.
- d. Tidak memiliki sistem pemrograman berorientasi objek sesungguhnya.
- e. PHP termasuk dalam *open source product*.

2.3.3 Adobe Dreamweaver CS6

Menurut (Madcoms, 2013) menyatakan *Adobe Dreamweaver CS6* adalah perangkat lunak terkemuka untuk desain web yang menyediaka visual yang intuitif termasuk pada tingkat kode yang dapat digunakan untuk membuat dan mengedit website HTML, serta aplikasi mobile seperti smartphone, tablet dan perangkat lainnya.

2.3.4 MySQL

Menurut (Rudianto, 2011) MySQL dikembangkan oleh sebuah perusahaan Swedia bernama MySQL AB yang pada saat itu bernama TcX data consult AB sekitar tahun 1994-1995, namun cikal bakal kodenya sudah ada sejak 1979. Awalnya TcX membuat MySQL dengan tujuan mengembangkan aplikasi web untuk klien. Tcx merupakan perusahaan software dan konsultan *database*. Saat ini MySQL sudah diakui oleh Oracle Crop. MySQL adalah salah satu jenis *database server* yang sangat terkenal dan banyak digunakan untuk membangun aplikasi web yang mana database sebagai sumber dan pengolahan datanya.

Kepopuleran MySQL antara lain karena MySQL menggunakan SQL sebagai bahasa dasar untuk mengakses *database* sehingga mudah untuk digunakan. MySQL juga bersifat *open source* dan *free* pada berbagai *platform* kecuali pada *windows* bersifat *shareware*. MySQL didistribusikan

dengan *lisensi open source* GPL (*General Public License*) mulai versi 3.23, pada bulan juni 2000.

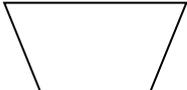
2.4 Diagram Flowchart

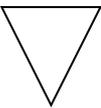
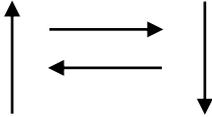
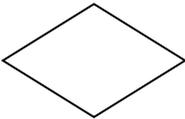
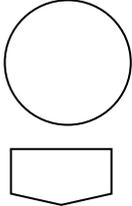
Menurut (Jogiyanto, 2005) menjelaskan bahwa bagan alir merupakan alat pengembangan sistem yang akan membantu menggambarkan sistem yang sedang berjalan pada perusahaan. Bagan alir (*flowchart*) adalah bagan (*chart*) yang menunjukkan alir (*flow*) di dalam program atau prosedur sistem secara logika. Bagan alir digunakan terutama untuk alat bantu komunikasi dan untuk dokumentasi. Terdapat beberapa jenis bagan alir yang biasa digunakan, yaitu sebagai berikut :

2.5 Bagan Alir Dokumen (*Document Flowchart*)

Bagan alir dokumen (*document flowchart*) atau disebut dengan bagan alir formulir (*form flowchart*) atau *paperwork flowchart* merupakan bagan alir yang menunjukkan arus dari laporan dan formulir termasuk tembusan-tembusannya. Adapun simbol bagan alir dapat dilihat pada tabel 2.1 sebagai berikut :

Tabel 2.1 Diagram Alir Dokumen (*document flowchart*)

SIMBOL	KETERANGAN
Dokumen 	Menunjukkan dokumen yang digunakan untuk input dan output baik secara manual maupun komputerisasi
Proses Manual 	Menunjukkan pekerjaan yang dilakukan secara manual.
Proses komputerisasi 	Menunjukkan proses dari operasi program komputer.

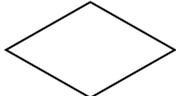
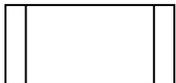
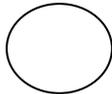
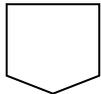
<p>Simpanan</p> 	Menunjukkan arsip atau Menunjukkan bahwa data dalam simbol ini akan disimpan ke suatu media tertentu
<p>Terminator</p> 	Menyatakan permulaan atau akhir suatu program
<p>Garis alir</p> 	Menyatakan jalannya arus suatu proses
<p>Decision</p> 	Digunakan untuk suatu penyeleksian kondisi di dalam program.
<p>Keyboard</p> 	Memasukkan data secara manual dengan menggunakan keyboard
<p>Hard disk</p> 	Media penyimpanan, menggunakan perangkat hard disk.
<p>Keterangan</p> 	Digunakan untuk memberikan keterangan yang lainnya.
<p>Penghubung</p> 	Menyatakan sambungan dari proses ke proses lainnya dalam halaman yang sama Menyatakan sambungan dari proses ke proses lainnya dalam halaman yang berbeda

2.6 Bagan Alir Program (*Program Flowchart*)

Bagan alir program (*program flowchart*) merupakan bagan yang menjelaskan secara rinci langkah-langkah dari proses program. Bagan alir program terdiri dari dua macam, yaitu bagan alir logika program (*program logic flowchart*) dan bagan alir program computer terinci (*detailed computer program flowchart*). Bagan alir logika program digunakan untuk menggambarkan tiap-tiap langkah didalam program komputer secara logika, sedangkan bagan alir program komputer terperinci digunakan untuk menggambarkan instruksi-instruksi program komputer secara terperinci. Bagan alir logika program dipersiapkan oleh analis sistem. Sedangkan bagan alir program komputer terperinci dipersiapkan pemrogram. Berdasarkan penjelasan diatas penulis dapat menarik kesimpulan bahwa bagan alir (*flowchart*) adalah suatu gambaran umum tentang sistem yang berjalan yang berfungsi sebagai alat bantu komunikasi dan untuk dokumentasi serta menyajikan kegiatan baik secara manual maupun komputerisasi. Adapun symbol bagan alir program dapat dilihat pada tabel 2.2 sebagai berikut

Tabel 2.2 Bagan Alir Program (*Program Flowchart*)

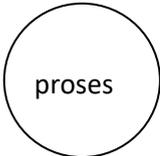
Simbol	Keterangan
Terminator 	Digunakan untuk memberikan awal dan akhir suatu proses.
Proses 	Menunjukkan proses dari operasi program komputer.
<i>Preparation</i> 	Proses inisialisasi/pemberian harga awal.
Input/Output Data 	Proses input/output data, parameter, informasi.

<p style="text-align: center;">Garis alir</p> 	<p>Digunakan untuk menunjukkan arus dari proses.</p>
<p style="text-align: center;"><i>Decision</i></p> 	<p>Digunakan untuk suatu penyeleksian kondisi di dalam program.</p>
<p style="text-align: center;">Proses terdefinisi</p> 	<p>Simbol yang digunakan untuk menunjukkan suatu operasi yang rinciannya ditunjukkan ditempat lain.</p>
<p style="text-align: center;">Penghubung</p>  	<p>Menyatakan sambungan dari proses ke proses lainnya dalam halaman yang sama</p> <p>Menyatakan sambungan dari proses ke proses lainnya dalam halaman yang berbeda</p>

2.7 Data Flow Diagram

Menurut (Salahuddin, 2015) menjelaskan bahwa data flow diagram adalah representasi grafik yang menggambarkan aliran informasi dan transformasi informasi yang diaplikasikan sebagai data yang mengalir dari masukan (*input*) dan keluaran (*output*). DFD dapat digunakan untuk merepresentasikan sebuah sistem atau perangkat lunak pada beberapa level abstraksi. DFD dapat dibagi menjadi beberapa level; yang lebih detail untuk merepresentasikan aliran informasi atau fungsi yang lebih detail. DFD menyediakan mekanisme untuk pemodelan fungsional ataupun pemodelan aliran informasi. Notasi-notasi pada DFD adalah seperti tabel 2.3 dibawah ini :

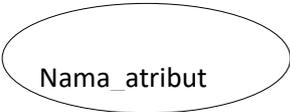
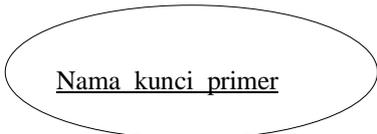
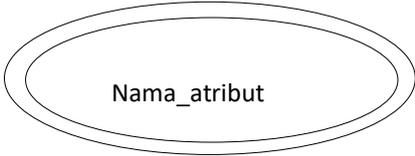
Tabel 2.3 simbol *Data Flow Diagram*

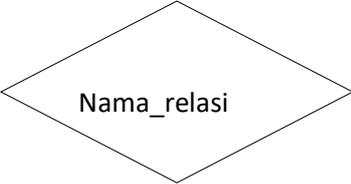
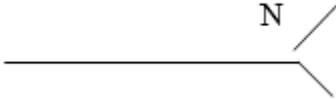
Notasi	Keterangan
	<p>Proses atau fungsi atau prosedur; pada pemodelan perangkat lunak yang akan diimplementasikan dengan pemrograman terstruktur, maka pemodelan notasi inilah yang ahrus menjadi fungsi atau prosedur didalam kode program</p>
	<p><i>File</i> atau basis data atau penyimpanan (<i>storage</i>) pada pemodelan perangkat lunak yang akan diimplementasikan dengan pemrograman terstruktur, maka pemodelan notasi inilah yang harusnya dibuat mejadi table-tabel basis data yang dibutuhkan, table-tabel ini juga harus sesuai dengan perancangan table-tabel pada basis data (ERD)</p>
	<p>Entitas luar (<i>external entity</i>) atau masukan atau keluaran atau orang-orang yang memakai atau berinteraksi dengan perangkat lunak yang dmodelkan atau system lain yang terkait dengan aliran data dari system yang dimodelkan</p>
	<p>Aliran data; merupakan data yang dikirim antar proses, dari penyimpanan keproses, atau dari proses kemasukan (<i>input</i>) atau keluaran (<i>ouput</i>)</p>

2.8 ERD (*Entity Relational Diagram*)

Menurut (Salahuddin, 2015) pemodelan awal basis data yang paling banyak digunakan adalah Entity Relatioanal Diagram (ERD). ERD dikembangkan berdasarkan teori himpunan dalam bidang matematika. ERD digunakan untuk pemodelan basis data relasional. Notasi-notasi pada ERD adalah seperti pada tabel 2.4 di bawah ini :

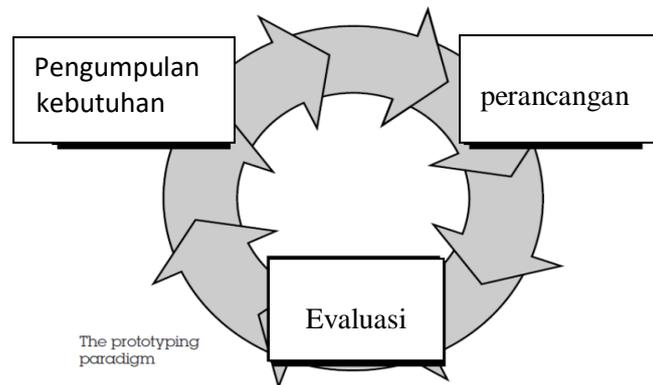
Tabel 2.4 Notasi pada model ERD

Simbol	Keterangan
Entitas /entity 	Entitas/entity merupakan data inti yang akan disimpan; bakal table pada basis data
Atribut 	Atribut/field;atau kolom data yang butuh disimpan dalam suatu entitas
Atribut Kunci Primer 	<i>Field</i> atau kolom yang butuh disimpan dalam suatu entitas dan digunakan sebagai kunci akses recod yang diinginkan; biasanya berupa id
Atribut Multi nilai 	<i>Field</i> atau kolom data yang butuh disimpan dalam suatau entitas yang dapat memiliki nilai lebih dari satu.

<p>Relasi</p> 	<p>Relasi yang menghubungkan antar entitas; biasanya diawali dengan kata kerja.</p>
<p>Asosiasi / Association</p> 	<p>Penghubung adalah penghubung antar relasi dan entitas dimana di kedua ujungnya memiliki <i>multiplicity</i> kemungkinan jumlah pemakaian, Kemungkinan jumlah maksimum keterhubungan antara entitas satu dengan entitas yang lain disebut dengan kardinalitas.</p>

2.9 Model *Prototype*

(S, Rosa & M, 2016) Model *prototype* (*prototyping model*) dimulai dari mengumpulkan kebutuhan pelanggan terhadap perangkat lunak yang akan dibuat. Lalu dibuatlah program *prototype* agar pelanggan lebih terbayang dengan apa yang sebenarnya diinginkan. Program *prototype* biasanya merupakan program yang belum jadi. Program ini biasanya menyediakan tampilan dengan simulasi alur perangkat lunak sehingga tampak seperti perangkat lunak yang sudah jadi. Program *prototype* ini dievaluasi oleh pelanggan atau *user* sampai ditemukan spesifikasi yang sesuai dengan keinginan pelanggan atau *user*. (S, Rosa & M, 2016) Berikut adalah gambar dari model *prototype* :



Gambar 2.2 Model *Prototype*

(S, Rosa & M, 2016) *Mock-up* adalah sesuatu yang digunakan sebagai model desain yang digunakan untuk mengajar, demonstrasi, evaluasi desain, promosi, atau keperluan lain. Sebuah *mock-up* disebut sebagai *prototype* perangkat lunak jika menyediakan atau mampu mendemonstrasikan sebagian besar fungsi sistem perangkat lunak dan memungkinkan pengujian desain sistem perangkat lunak. Iterasi terjadi pada pembuatan *prototype* sampai sesuai dengan keinginan pelanggan (*customer*) atau *user*. Proses pada model *prototyping* dapat dijelaskan sebagai berikut :

- a. Pengumpulan kebutuhan : *Developer* dan klien bertemu dan menentukan tujuan umum, kebutuhan yang diketahui dan gambaran bagian-bagian yang akan dibutuhkan berikutnya.
- b. Perancangan : Perancangan dilakukan cepat dan rancangan mewakili semua aspek perangkat lunak yang diketahui, dan rancangan ini menjadi dasar pembuatan *prototype*.
- c. Evaluasi : Klien mengevaluasi *prototype* yang dibuat dan digunakan untuk memperjelas kebutuhan perangkat lunak. Perulangan ketiga proses ini terus berlangsung hingga semua kebutuhan terpenuhi. *Prototype-prototype* dibuat untuk memuaskan kebutuhan klien dan untuk membangun perangkat lunak lebih cepat, namun tidak semua *prototype* bisa dimanfaatkan.

2.10 Penelitian Terdahulu

Pada bagian ini akan membahas mengenai tabel penelitian terdahulu yang berhubungan dengan karya ilmiah penulis, yang terdapat pada tabel 2.5 berikut:

Tabel 2.5 Penelitian terdahulu

No	Nama	Judul	Tahun Terbit	Keterangan
1	Nurjoko Dan Hendra Kurniawan	Aplikasi datamining untuk memprediksi tingkat kelulusan mahasiswa menggunakan algoritma apriori di ibi darmajaya Bandar lampung	2016	Pertumbuhan yang pesat dari akumulasi data telah menciptakan kondisi kaya akan data tapi minim informasi. Data mining merupakan penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar yang diharapkan dapat mengatasi kondisi tersebut. Dengan memanfaatkan data induk mahasiswa dan data kelulusan mahasiswa, diharapkan dapat menghasilkan informasi tentang tingkat kelulusan dengan data induk mahasiswa melalui teknik data mining. Kategori tingkat kelulusan di ukur dari lama studi dan IPK. Algoritma yang digunakan adalah algoritma apriori, informasi yang ditampilkan berupa nilai <i>support</i> dan <i>confidence</i> dari masing-masing kategori tingkat kelulusan.
2	windha mega pradnya dhuhita	<i>Clustering</i> menggunakan metode k-means untuk menentukan status gizi balita	2015	Kekurangan gizi atau yang biasa disebut malnutrisi merupakan salah satu masalah kesehatan yang cukup sering menimpa balita-balita di Indonesia. Merunut data dari WHO, jumlah balita yang meninggal dunia akibat kekurangan gizi di Indonesia tahun 2012 adalah 29 dari 1000 kelahiran. Kepedulian orang tua dan aparat desa (dalam hal ini petugas Pusat Pelayanan Kesehatan Masyarakat – PUSKESMAS) untuk memantau gizi balita sangat diperlukan. Penelitian yang dilakukan mencoba untuk melakukan pengelompokan 50 balita di desa Karang Songo kedalam 5 cluster status gizi. Pengelompokan status gizi balita di Desa Kembang Songo menggunakan metode K-Means dilakukan melalui beberapa tahapan yaitu : penentuan tujuan bisnis, pengumpulan data 50 balita di Desa karang Songo, pengelompokan status gizi balitake dalam 5 cluster yaitu cluster 1 - gizi buruk; cluster 2 - gizi kurang; cluster 3 - gizi baik; cluster 4 - gizi lebih; cluster 5 - obesitas, perhitungan cluster menggunakan software SPSS, analisa hasil data output, pengelompokan status gizi balita menggunakan tabel Growth Chart, dan yang terakhir melakukan pengujian dengan membandingkan hasil pengelompokan algoritma K-means dan tabel Growth Chart
3	suhada dan yulmaini	Pemanfaatan teknologi internet untuk pemesanan	2015	Usaha Kecil dan Menengah (UKM) Keripik Singkong terletak di Desa Sungai Langka Kabupaten Pesawaran tepatnya di Dusun 9. UKM ini bergerak dibidang kuliner yaitu

		keripik singkong pada usaha kecil menengah (ukm) di desa sungai langka		mengolah makanan ringan keripik singkong dengan 5 aneka rasa diantaranya adalah rasa jagung bakar, rasa keju, rasa balado, rasa original, dan rasa barbque. Pemesanan keripik singkong yang ada di UKM ini hanya dilakukan melalui komunikasi dari mulut ke mulut, sehingga pemesanan masih dalam lingkup kecil. Hal ini menjadi kendala dari UKM ini karena hanya di daerah sekitar Kabupaten Pesawaran saja. Metode yang dilakukan dalam kegiatan pengabdian ini adalah pengumpulan data kegiatan dengan menggunakan metode Observasi dan wawancara. Metode pengembangan sistem aplikasi web menggunakan System Development Life Cycle (SDLC) dalam model Waterfall. Hasil dari kegiatan pengabdian ini adalah aplikasi Web yang dapat memperluas jangkauan pasar dan efektif dalam pemesanan produk, sehingga tidak hanya di wilayah Kabupaten Pesawaran saja tetapi dapat juga di luar wilayah Kabupaten Pesawaran.
4	sri karnila	Penerapan aplikasi crm untuk career centre Pada perguruan tinggi ibi darmajaya	2012	Perguruan tinggi IBI Darmajaya juga dapat menerapkan CRM untuk meningkatkan pengelolaan pelanggan seperti layaknya perusahaan pada lembaga pusat pelatihan. Darmajaya training centre IBI Darmajaya memberikan informasi terkait pelatihan, dapat menjadi 'career centre' yang menjembatani perguruan tinggi dengan perusahaan dan alumni. Alumni dan perusahaan mitra atau perusahaan yang menyerap alumni sebagai tenaga kerja, baik yang bekerja sama ataupun tidak, dapat merupakan pelanggan IBI Darmajaya yang mendapat pelayanan informasi pelatihan dan lowongan pekerjaan. Career centre saat ini memberikan informasi pelatihan, lowongan pekerjaan dilakukan saat proses yaitu saat alumni masih menempuh pendidikan atau berstatus mahasiswa dan pada perusahaan yang bekerja sama, hal ini memungkinkan pelayanan serta keterhubungan antara alumni, perusahaan dan IBI Darmajaya kurang efektif, karena informasi yang diberikan oleh career centre belum tentu diketahui oleh alumni bahkan perusahaan rekanaan. Untuk itu dibutuhkan sebuah aplikasi CRM pada career centre, sehingga IBI Darmajaya memiliki media komunikasi yang menghasilkan informasi sesuai kebutuhan alumni dan perusahaan sehingga dapat membina hubungan secara berkelanjutan antara alumni dan perusahaan – perusahaan dengan perguruan tinggi.
5	Nurjoko Dan Abdi Kurniawan	Penerapan data mining menggunakan association rules untuk mendukung strategi pemasaran calon	2015	Pertumbuhan yang pesat dari akumulasi data telah menciptakan kondisi kaya akan data tapi minim informasi. Data mining merupakan penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar yang diharapkan dapat mengatasi kondisi tersebut. Dengan memanfaatkan data calon mahasiswa dan data mahasiswa, diharapkan dapat

		mahasiswa baru (studi kasus ibi darmajaya)		menghasilkan informasi tentang sumber informasi yang digunakan oleh calon mahasiswa yang dapat dimanfaatkan oleh manajemen dalam melakukan usaha-usaha yang terkait dengan peningkatan jumlah mahasiswa dengan melakukan pemilihan media promosi yang tepat melalui teknik data mining. Algoritma yang digunakan adalah algoritma apriori, informasi yang ditampilkan berupa nilai support dan confidence dari masing-masing pola yang terbentuk. Dari uji coba terhadap data calon mahasiswa dan data mahasiswa dengan model Association rules dihasilkan beberapa aturan dengan tingkat nilai maksimum confidence sebesar 81,1% Untuk sumber informasi yang berasal dari media Koran dan tingkat nilai maksimum confidence sebesar 77,8% untuk sumber informasi melalui media brosur. Saat ini media promosi yang tepat untuk digunakan oleh tim pemasaran dalam menjangkau calon mahasiswa baru adalah media Koran dan media brosur dengan tidak mengabaikan media promosi yang lain
6	Anindya khrisna wardhani	K-MEANS ALGORITHM IMPLEMENTATION FOR CLUSTERING OF PATIENTS DISEASE IN KAJEN CLINIC OF PEKALONGAN	2016	<i>In determining the consistency of health data, can use data mining techniques that can dig the hidden information from multidimensional data sets that have been obtained. In addition, data which connected with other data can also be done by these data mining techniques. One of the data mining techniques is quite well known namely clustering. The methods are quite popular in data mining techniques that called k-means method. It is used to facilitate medical recorder for analyzing the general health situation of population groups in archiving health care data. The results of this analysis, the clustering of the disease based on age, sex, duration of disease and disease diagnosis. This research used tool Rapid Miner 5.3. Based on the data from clinic centers Kajen Pekalongan, the result of clustering is 376 items of acute and 624 unacute diseases from 1000 total of data.</i>
7	Fiqih satria dan RZ. Abdul Aziz	DALAM MENENTUKAN CLUSTER DATA MAHASISWA PEMOHON BEASISWA (STUDI KASUS : STMIK PRINGSEWU)	2016	Penelitian ini bertujuan untuk mengetahui langkah-langkah analisis cluster dengan metode metode Ward dan metode K-Means, serta membandingkan hasil analisis kedua metode tersebut untuk mengclusterkan data mahasiswa terkait pengambilan keputusan untuk menentukan mahasiswa yang berhak menerima beasiswa Peningkatan Prestasi Akademik (PPA) dan beasiswa Biaya Bantuan Akademik (BBA) di STMIK PRINGSEWU. Analisis cluster dilakukan menggunakan Aplikasi IBM SPSS Versi 23. Hasil pengclusteran kedua metode tersebut dibandingkan dengan menggunakan rasio simpangan baku dalam kelompok (Sw) dan simpangan baku antar kelompok (Sb). Pada penelitian ini Berdasarkan nilai rasio simpangan baku dalam cluster dan simpangan baku antar cluster menunjukkan bahwa metode Ward dan

				<p>K-Means memiliki kinerja yang sama ketika digunakan pada clustering beasiswa PPA, kedua metode memiliki nilai rasio (Sw/Sb) yang sama yaitu 0,749959584 %. Dengan demikian maka hasil cluster yang akan digunakan sebagai acuan pada penentuan penerima beasiswa PPA adalah hasil analisis cluster menggunakan metode Ward ataupun K-Means karena keduanya memiliki hasil cluster dan kinerja yang sama. Sedangkan pada clustering beasiswa BBA, metode Ward memiliki kinerja yang lebih baik daripada metode KMeans karena metode Ward memiliki nilai rasio (Sw/Sb) 0.5346668 % lebih kecil dibandingkan dengan nilai rasio (Sw/Sb) metode K-Means yaitu 0,831525302 %. Dengan demikian maka hasil cluster yang dapat digunakan sebagai acuan pada penentuan penerima beasiswa BBA adalah hasil analisis cluster menggunakan metode Ward.</p>
8	ASRONI, Ronald Adrian	Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang	2015	<p>The selection process among outstanding students in a department has a big problem. This process is not fair because only involve one criteria and ignore the other criteria. We need the best student to participate in a competition held by the Indonesia Security Incident Response Team on Internet Infrastructure (ID SIRTII) of the Ministry of Communication and Information. This process uses Weka software to calculate the best student. It provides the various method to explore the data. One of them is clustering method. There are many algorithms in clustering method. In this research, we will investigate widely about one of that algorithms. Its name is K-Means. This algorithm (K-Means) will give the recommendations about the best student based on the cluster. It will represent the many clusters of a student group. The best cluster can be calculated more to get the names of the best students group. They are eligible to enter the competition. K-means involve the GPA (Grade Point Average) and related course to support the academic skill in order to get the best student. This research helps the teacher select the best student to enter the competition. Many similar cases can use this algorithm in order to get the best student.</p>
9	Nurul Rohmawati W, Sofi Defiyanti, Mohamad Jajuli	IMPLEMENTASI ALGORITMA K-MEANS DALAM PENGLASTERAN MAHASISWA PELAMAR BEASISWA	2015	<p>Pengelompokan data pelamar beasiswa Bantuan Belajar Mahasiswa (BBM) dikelompokkan menjadi 3 kelompok yaitu berhak menerima, dipertimbangkan dan tidak berhak menerima beasiswa. Pengelompokan menjadi 3 kelompok ini berguna untuk memudahkan dalam menentukan penerima beasiswa BBM. Algoritma k-means merupakan algoritma dari teknik clustering yang berbasis partisi. Teknik ini dapat mengelompokkan data mahasiswa pelamar beasiswa. Tujuan dari penelitian ini adalah untuk pengukuran kinerja algoritma, Pengukuran ini dilihat dari hasil cluster dengan menghitung nilai</p>

				<p>kemurnian (purity measure) dari masing – masing cluster yang di hasilkan. Data yang digunakan dalam penelitian ini adalah data mahasiswa yang mengajukan beasiswa kepada Fakultas Ilmu Komputer UNSIKA sebanyak 36 mahasiswa. Data akan diubah menjadi 3 dataset dengan format yang berbeda-beda, yakni data atribut kodifikasi sebagian, atribut kodifikasi keseluruhan dan atribut data asli. Nilai purity pada dataset data kodifikasi sebagian untuk hasil cluster algoritma k-means sebesar 61.11%. Pada dataset kodifikasi keseluruhan nilai purity hasil cluster algoritma k-means sebesar 80.56%. Dan untuk dataset data asli nilai purity hasil cluster algoritma k-means sebesar 75%. Maka dapat di simpulkan bahwa algoritma k-means lebih cocok digunakan pada dataset dengan format atribut yang dikodifikasi keseluruhan.</p>
10	Triowali Rosandi	PERBANDIN GAN METODE NAÏVE BAYES CLASSIFIER DENGAN DECISION TREE(C4.5) UNTUK MENGANALI SA KELANCARA N PEMBIAYAA N (STUDI KASUS: KSSPS/BMT AL-FADHILA	2016	<p>Salah satu algoritma klasifikasi yang sering digunakan dan mendapat banyak perhatian para peneliti dalam memprediksi pembiayaan bermasalah pada perbankan adalah <i>Naive Bayes Clasifier</i> dan <i>Decision Tree (C4.5)</i>. Tetapi kelemahan yang dihadapi pada kedua algoritma tersebut adalah lamanya waktu dan tingkat akurasi prediksi yang digunakan untuk melakukan prediksi. Masalah ini juga menjadi perhatian banyak peneliti untuk memperbaikinya agar kinerja waktu dan akurasi prediksi menjadi lebih singkat tetapi kinerja akurasi tetap baik. Karena hal tersebut maka kedua algoritma tersebut akan kita badingkan dari tingkat keakurasiannya maupun kinerja waktu untuk meningkatkan kinerja prediksi. Langkah-langkah analisis pembiayaan dengan metode <i>Naive Bayes Clasifier</i> adalah dengan menghitung jumlah class/label, menghitung jumlah kasus yang sama dengan class yang sama, mengalikan semua hasil variable, membandingkan hasil class dengan hasil terbesar akan dijadikan sebagai keputusan. Metode <i>Decision Tree (C4.5)</i> merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau segmentasi atau pengelompokan dan bersifat prediktif. Langkah-langkah analisis pembiayaan dengan metode ini adalah dengan memilih atribut sebagai akar lalu membuat cabang untuk masing-masing nilai lanjutkan dengan membagi kasus dalam cabang dan mengulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama. Hasil analisa kedua metode tersebut dibandingkan dengan melihat kemudahan penggunaan dan ketepatan hasil perhitungan masing-masing algoritma. Dari perhitungan yang telah di ujicoba hal tersebut menunjukkan bahwa metode <i>Decision Tree (C4.5)</i> memiliki tingkat akurasi lebih tinggi dan efesiensi waktu yang lebih cepat daripada metode <i>Naive Bayes Classifier</i>.</p>

11	muhardi dan nisar	penentuan penerimaan beasiswa dengan algoritma fuzzy c-means di universitas megow pak tulang bawang	2015	<p>Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan kepada perorangan yang bertujuan untuk digunakan demi keberlangsungan pendidikan yang ditempuh. Penentuan beasiswa dapat dikelompokkan berdasarkan kriteria Indeks Prestasi Kumulatif (IPK), tingkat kemiskinan, tanggungan orang tua dan prestasi mahasiswa dalam proses rekrutmen beasiswa. Algoritma <i>Fuzzy C-Means</i> merupakan satu algoritma yang mudah dan sering digunakan dalam pengelompokan data karena membuat suatu perkiraan yang efisien dan tidak memerlukan banyak parameter. Beberapa penelitian telah menghasilkan kesimpulan bahwa metode <i>Fuzzy C-Means</i> dapat digunakan untuk mengelompokkan data berdasarkan atribut-atribut tertentu. Penerapan algoritma <i>Fuzzy C-Means</i> dalam penentuan beasiswa di kelompokkan menjadi tiga <i>cluster</i> yaitu menerima, dipertimbangkan dan tidak berhak menerima beasiswa, sampel data sebanyak 75 data mahasiswa diperoleh tiga <i>cluster</i> berdasarkan nilai rata-rata penentuan beasiswa kemudian setiap <i>cluster</i> diklasifikasikan berdasarkan kriteria mana yang lebih diprioritaskan dengan nilai terbesar pada jarak akhir merupakan <i>cluster</i> yang menerima beasiswa, sedangkan <i>cluster</i> dengan nilai terkecil merupakan <i>cluster</i> yang tidak berhak menerima beasiswa.</p>
12	Petrus Dwi Ananto Pamungkas	MENENTUKAN KEMUNGKINAN MASUKNYA CALON MAHASISWA BARU PADA SEBUAH PERGURUAN TINGGI SWASTA MANGGUNAKAN TEKNIK KLASIFIKASI POHON KEPUTUSAN DENGAN APLIKASI RAPIDMINER 5.1	2015	<p>Teknik Klasifikasi dengan pembentukan pohon keputusan merupakan salah satu teknik pengolahan data yang sangat besar yang mampu mempresentasikan data menjadi informasi yang mudah untuk dipahami oleh hampir semua orang dibandingkan dalam bentuk tabel. Biasanya, hasil pengumpulan data adalah dalam bentuk tabel. Tapi tidak semua orang mampu membaca tabel dengan cepat dan tepat sesuai keinginan pembuatnya, apalagi jika datanya sangat besar. Data dalam penelitian ini ada ratusan data yang disajikan dalam bentuk tabel. Semua informasi calon mahasiswa dikumpulkan dalam satu tabel dengan banyak item, mulai dari data pribadi, orang tua hingga asal sekolah. Tabel yang dihasilkan tersebut akan sulit dipahami sehingga penyerapan informasinya menjadi lama dan tidak akurat. Penyajian informasi dalam bentuk pohon keputusan menjadi alternative tepat untuk data yang sangat banyak tersebut. Dengan menggunakan aplikasi RapidMiner 5.1 maka pembuatan pohon keputusan dari data calon mahasiswa menjadi lebih mudah dan cepat sehingga pihak manajemen perguruan tinggi dapat melihat dan mengetahui faktor-faktor apa saja yang membuat calon mahasiswa untuk tetap melanjutkan proses pendaftaran hingga menjadi mahasiswa.</p>

