BAB 2
TINJAUAN PUSTAKA

2.1 Tinjauan Instansi

2.1.1 Sejarah Dinas Pariwisata dan Kebudayaan Provinsi Jawa Barat dan Badan Hukum Instansi

Dinas Pariwisata dan Kebudayaan Provinsi Jawa Barat (DISPARBUD JABAR) merupakan salah satu lembaga pemerintahan yang bergerak dalam bidang pelestarian serta pemberdayaan pariwisata dan kebudayaan yang ada di Provinsi Jawa Barat berlokasi di Jl.L.L.R.E Martadinata No.209 Bandung. Berdasarkan Peraturan Pemerintah Nomor : 25 Tahun 2000 tentang Kewenangan Pemerintah dan Kewenangan Propinsi sebagai Daerah Otonom, sebagaimana diamanatkan Undang-undang Nomor : 22 Tahun 1999 tentang Pemerintah Daerah dan Peraturan Pemerintah Nomor : 84 Tahun 2000 tentang Pedoman Organisasi Perangkat Daerah, maka Gubernur Kepala Daerah Tingkat I Jawa Barat mengeluarkan Peraturan Daerah Nomor : 15 Tahun 2000 tentang Pembentukan Dinas Daerah Propinsi Jawa Barat.

Dengan dikeluarkannya Peraturan Daerah Nomor : 15 Tahun 2000 tersebut, maka terbentuklah Dinas Kebudayaan dan Pariwisata Propinsi Jawa Barat yang merupakan gabungan 4 (empat) Instansi yaitu dari Dinas Pariwisata Propinsi Jawa Barat, Dinas Pendidikan Propinsi Jawa Barat Bidang Kebudayaan, Kantor Wilayah Departemen Pariwisata, Pos dan Telekomunikasi Propinsi Jawa Barat dan Kantor Wilayah Departemen Pendidikan Bidang Kebudayaan.

Pada tahun 2007 terbit Peraturan Pemerintah Nomor : 38 tentang Pembagian Urusan Pemerintahan antara Pemerintah, Pemerintah Daerah Provinsi dan Pemerintah Daerah Kabupaten/Kota yang pada akhirnya berimbas kepada Dinas Kebudayaan dan Pariwisata Provinsi Jawa Barat yaitu mengalami perubahan kembali. Tidak berupa gabungan tetapi tetap dengan posisi yang sama hanya mengalami perubahan nomenklatur saja. Berdasarkan Peraturan Daerah Provinsi Jawa Barat Nomor : 21 Tahun 2008 tentang Dinas Daerah Provinsi Jawa

Barat, nama Dinas Kebudayaan dan Pariwisata Propinsi Jawa Barat berubah nama menjadi Dinas Pariwisata dan Kebudayaan Provinsi Jawa Barat.

Menyimak hal tersebut diatas, Dinas Pariwisata dan Kebudayaan Provinsi Jawa Barat berupaya mengarahkan peningkatan sektor ekonomi melalui peningkatan sektor pariwisata yang dukung sektor kebudayaan menjadi salah satu faktor andalan yang mampu menggalakkan roda perekonomian sehingga mampu memberi kesempatan kepada masyarakat untuk membuka lapangan kerja dan berusaha secara mandiri yang pada gilirannya akan meningkatkan pendapatan daerah dan masyarakat juga pendapatan negara melalui devisa.

2.1.2 Logo Instansi



Gambar 2.1 Logo Disparbud

Logo Dinas Pariwisata dan Kebudayaan Jawa Barat ini mengacu pada Logo Provinsi Jawa Barat. Adapun arti dan makna dari logo ini yaitu:

1. Tulisan Gemah Ripah Repeh Rapih, melambangkan pepatah lama Sunda yang bermaksud menyatakan bahwa Jawa Barat adalah daerah yang kaya raya dan didiami oleh banyak penduduk yang rukun dan damai.
2. Gambar Bentuk Bulat Telur, pada lambang Jawa Barat ini berasal dari bentuk perisai sebagai penjagaan diri.
3. Kujang, melambangkan senjata suku bangsa Sunda yang merupakan penduduk asli Jawa Barat. Lima lubang pada kujang melambangkan lima sila dari Dasar Negara Indonesia yaitu Garuda Pancasila.
4. Padi, melambangkan bahan makanan pokok masyarakat Jawa Barat sekaligus juga melambangkan kesuburan pangan dan jumlah padi 17 melambangkan tanggal Proklamasi Republik Indonesia.

5. Kapas, melambangkan kesuburan sandang dan jumlah bunga kapas 8 buah melambangkan bulan Proklamasi Republik Indonesia.
6. Gunung yang terdapat dibawah padi dan kapas, melambangkan bahwa daerah Jawa Barat terdiri atas daerah pegunungan.
7. Sungai dan terusan yang terdapat dibawah gunung sebelah kiri, melambangkan di daerah Jawa Barat banyak terdapat sungai dan saluran air seperti persawahan dan perkebunan.
8. DAM yang terdapat ditengah-tengah bagian bawah antara gambar sungai dan petak, melambangkan kegiatan dibidang irigasi yang merupakan salah satu perhatian pokok mengingat Jawa Barat adalah daerah agraris. Hal ini juga melambangkan dam-dam yang berada Di Jawa Barat seperti Waduk Jatiluhur.

2.1.3 Visi dan Misi

A. Visi

" Terwujudnya Jawa Barat sebagai daerah budaya dan tujuan wisata andalan ".

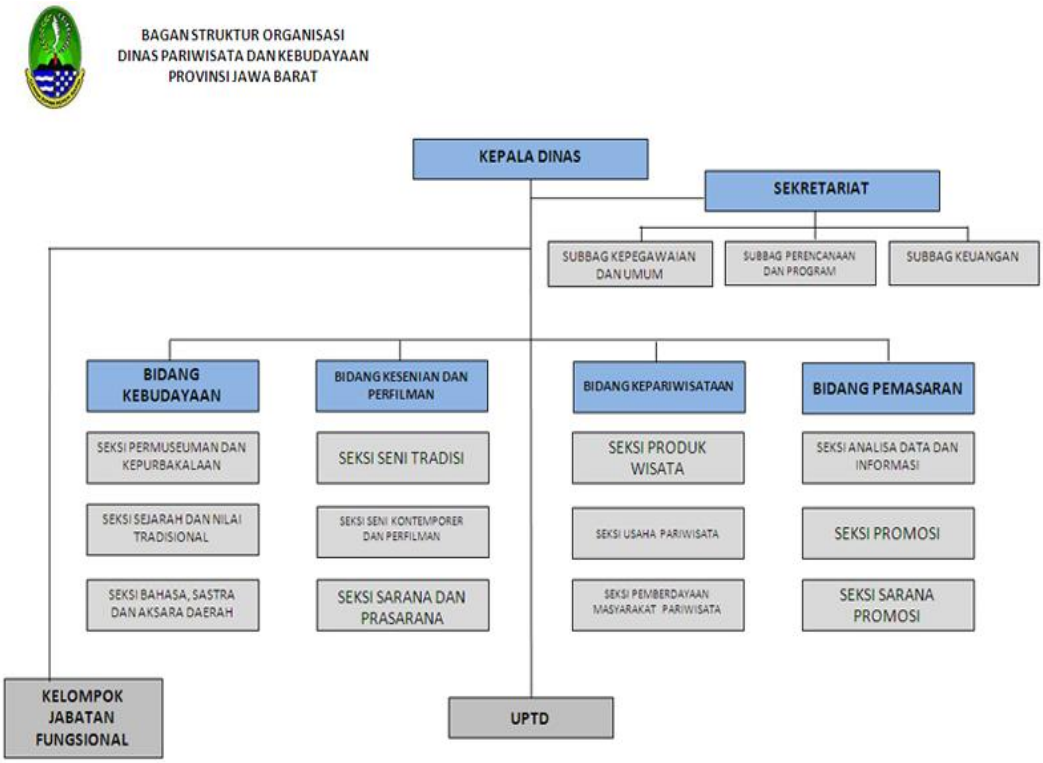
B. Misi

a) Pembinaan, pelestarian dan pengembangan asset budaya yang mendukung upaya pengembangan Pariwisata Jawa Barat.
b) Mengefektifkan kebudayaan sebagai asset daerah yang mendukung kepada pengembangan Usaha Jasa Pariwisata.
c) Mempromosikan Kepariwisataan Jawa Barat.
d) Meningkatkan sumber daya manusia Kebudayaan dan Kepariwisataan. Memuliakan nilai-nilai budaya yang terkandung dalam aspek jarahnitra dan tradisi Jawa Barat.

2.1.4 Struktur Organisasi

Struktur Organisasi merupakan susunan wewenang kerangka kerja yang mewujudkan pola kerja tetap serta mengatur hubungan-hubungan di antara bidang-bidang kerja, maupun orang-orang yang mewujudkan kedudukan dan peranan masing-masing jabatan dalam mewujudkan kerjasama, struktur organisasi juga membuka adanya kesatuan arah dan langkah dalam melaksanakan kegiatan,

serta adanya kejelasan pembagian tugas, wewenang dan tanggung jawab dari orang-orang yang melaksanakan tugas tersebut. Berdasarkan Peraturan Daerah Jawa Barat No.21 Tahun 2008 Tentang Organisasi dan Tata Kerja Dinas Daerah Provinsi Jawa Barat, struktur organisasi Dinas Pariwisata dan Kebudayaan sebagai berikut:



Gambar 2.2 Struktur Organisasi Disparbud

2.1.5 Tugas Pokok dan Fungsi (Tupoksi) Disparbud Prov. Jabar

1. Deskripsi :

Berdasarkan Peraturan Gubernur Jawa Barat Nomor : 73 Tahun 2009 tentang Tugas Pokok, Fungsi, Rincian Tugas Unit dan Tata Kerja Dinas Pariwisata dan Kebudayaan Provinsi Jawa Barat. Dinas Pariwisata dan Kebudayaan Provinsi Jawa Barat mempunyai tugas pokok yaitu melaksanakan urusan pemerintahan

Daerah di bidang pariwisata dan kebudayaan berdasarkan asas otonomi, dekonsentrasi dan tugas pembantuan.

2. Fungsi :

a) Penyelenggaraan perumusan dan penetapan kebijakan teknis kepariwisataan, kebudayaan, kesenian dan perfilman serta pemasaran.
b) Penyelenggaraan pariwisata dan kebudayaan meliputi kepariwisataan, kebudayaan, kesenian dan perfilman serta pemasaran.
c) Penyelenggaraan pembinaan dan pelaksanaan tugas-tugas pariwisata dan kebudayaan meliputi kepariwisataan, kebudayaan, kesenian dan perfilman serta pemasaran.
d) Penyelenggaraan koordinasi dan pembinaan UPTD.
e) Penyelenggaraan tugas lain sesuai dengan tugas pokok dan fungsinya.

2.1.6 Deskripsi Pekerjaan

Pembagian tugas dalam suatu instansi lembaga sangat penting, supaya terjadi keharmonisan dalam mencapai suatu tujuan. Berikut ini terdapat deskripsi jabatan yang berlaku di Dinas Pariwisata Jawa Barat :

A. Tugas Kepala Dinas

Menyelenggarakan perumusan, penetapan, memimpin,mengkoordinasikan dan mengendalikan pelaksanaan kegiatan tugas pokok Dinas serta mengkoordinasikan dan membina UPTD.

B. Tugas Sekretaris

Menyelenggarakan koordinasi perencanaan dan program Dinas, pengkajian perencanaan dan program, pengelolaan keuangan, kepegawaian dan umum.

C. Tugas Bidang Kesenian dan Perfilman

Menyelenggarakan pengkajian bahan kebijakan teknis dan melestarikan, mengembangkan dan memanfaatkan kesenian meliputi seni tradisi, seni kontemporer dan perfilman, prasarana dan sarana kesenian.

D. Tugas Bidang Pemasaran :

Menyelenggarakan pengkajian bahan kebijakan teknis dan fasilitasi penyelenggaraan pembinaan dan pengembangan produk kepariwisataan, meliputi

produk pariwisata, usaha pariwisata, obyek dan daya tarik wisata serta pemberdayaan masyarakat pariwisata.

E. Tugas Bidang Kebudayaan :

Menyelenggarakan pengkajian bahan kebijakan teknis, fasilitasi dan penyelenggaraan pelestarian, pengembangan dan pemanfaatan kebudayaan meliputi permuseuman, peninggalan sejarah dan kepurbakalaan, kesejarahan, nilai tradisional dan kebahasaan daerah.

2.2 Pariwisata

Kepariwisataan itu sendiri merupakan pengertian jamak yang diartikan sebagai hal-hal yang berhubungan dengan pariwisata, yang dalam bahasa Inggris disebutkan *tourism*. Dalam kegiatan kepariwisataan ada yang disebut subyek wisata yaitu orang-orang yang melakukan perjalanan wisata dan obyek wisata yang merupakan tujuan wisatawan. Sebagai dasar untuk mengkaji dan memahami berbagai istilah kepariwisataan, berpedoman pada Bab 1 pasal 1 Undang-Undang Republik Indonesia Nomor 10 Tahun 2009 tentang kepariwisataan yang menjelaskan sebagai berikut : [3]

1. Wisata adalah kegiatan perjalanan yang dilakukan oleh sebagian atau kelompok orang dengan mengunjungi tempat tertentu untuk tujuan rekreasi, pengembangan pribadi, atau mempelajari keunikan daya tarik wisata yang dikunjungi dalam jangka waktu sementara.
2. Wisatawan adalah orang yang melakukan wisata.
3. Pariwisata adalah berbagai macam kegiatan wisata dan didukung berbagai fasilitas serta layanan yang disediakan oleh masyarakat, pengusaha, Pemerintah dan Pemerintah Daerah.

2.2.1 Objek Wisata

Pengertian Objek wisata Menurut Chafid Fandeli obyek wisata adalah perwujudan daripada ciptaan manusia, tata hidup, seni budaya serta sejarah bangsa dan tempat atau keadaan alam yang mempunyai daya tarik untuk dikunjungi wisatawan. Sedangkan obyek wisata alam adalah obyek wisata yang

daya tariknya bersumber pada keindahan sumber daya alam dan tata lingkungannya. Jenis Obyek wisata Penggolongan jenis objek wisata akan terlihat dari ciri-ciri khas yang ditonjolkan oleh tiap-tiap obyek wisata. Dalam UU No. 9 Tahun 1990 Tentang Kepariwisataan disebutkan bahwa obyek dan daya tarik wisata terdiri dari :

1. Objek dan daya tarik wisata ciptaan Tuhan Yang Maha Esa, yang berwujud keadaan alam, serta flora dan fauna.
2. Objek dan daya tarik wisata hasil karya manusia yang berwujud museum,peninggalan sejarah, wisata agro, wisata tirta, wisata buru, wisata petualangan alam, taman rekreasi dan tempat hiburan.

Namun kepariwisataan di jawa barat membagi objek wisata menjadi 8 karakter objek wisata, dimana karakter disini adalah proses pengelompokkan objek wisata sesuai kategorinya. dan 8 karakter adalah sebagai berikut :

a) Karakter Hutan
b) Karakter Pantai
c) Karakter Religius
d) Karakter Tradisi dan Komunikasi Adat.
e) Karakter Kepurbakalaan
f) Karakter Gelar Seni dan Ragam Festival
g) Karakter Wisata Belanja
h) Karakter Heritage

Dinas Pariwisata dan Kebudayaan Jawa barat (Disparbud Jabar) mengklasifikasian objek wisata menjadi 8 karakter objek wisata didasarkan atas kawasan Wisata Unggulan (KWU) yang ada di Jawa Barat. [2]

2.3 Text Mining

2.3.1 Pengertian *Text Mining*

Menurut Ronen Feldman dan James Sanger dalam buku *The Text Mining Handbook*, [4] *text mining* dapat didefinisikan secara luas sebagai suatu proses menggali informasi dimana seorang user berinteraksi dengan sekumpulan dokumen menggunakan tools analisis yang merupakan komponen-komponen dalam *data mining* yang salah satunya adalah kategorisasi. *Text mining* bisa dianggap subjek riset yang tergolong baru. *Text mining* dapat memberikan solusi

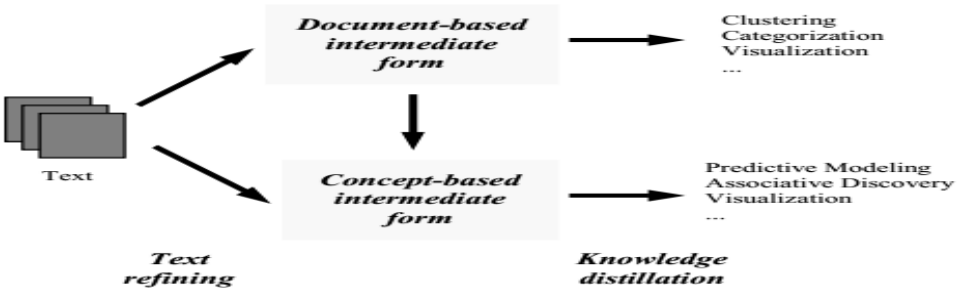
dari permasalahan seperti pemrosesan, pengorganisasian / pengelompokkan dan menganalisa *unstructured text* dalam jumlah besar.

Dalam memberikan solusi, text mining mengadopsi dan mengembangkan banyak teknik dari bidang lain, seperti *Data mining, Information Retrieval, Statistik* dan *Matematik, Machine Learning, Linguistic, Natural Languange Processing,* dan *Visualization*. Kegiatan riset untuk *text mining* antara lain ekstraksi dan penyimpanan text, preprocessing akan konten text, pengumpulan data statistik dan indexing dan analisa konten.

2.3.2 Tujuan *Text Mining*

Tujuan dari *Text Mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen tetapi Tujuan utama *text mining* adalah mendukung proses *knowledge discovery* pada koleksi dokumen yang besar. Jadi, sumber data yang digunakan pada *text mining* adalah kumpulan teks yang memiliki format yang tidak terstruktur atau minimal semi terstruktur. Adapun tugas khusus dari text mining antara lain yaitu pengkategorisasian teks(*text categorization*) dan pengelompokkan teks (*text clustering*)[4].

2.3.3 Kinerja *Text Mining*



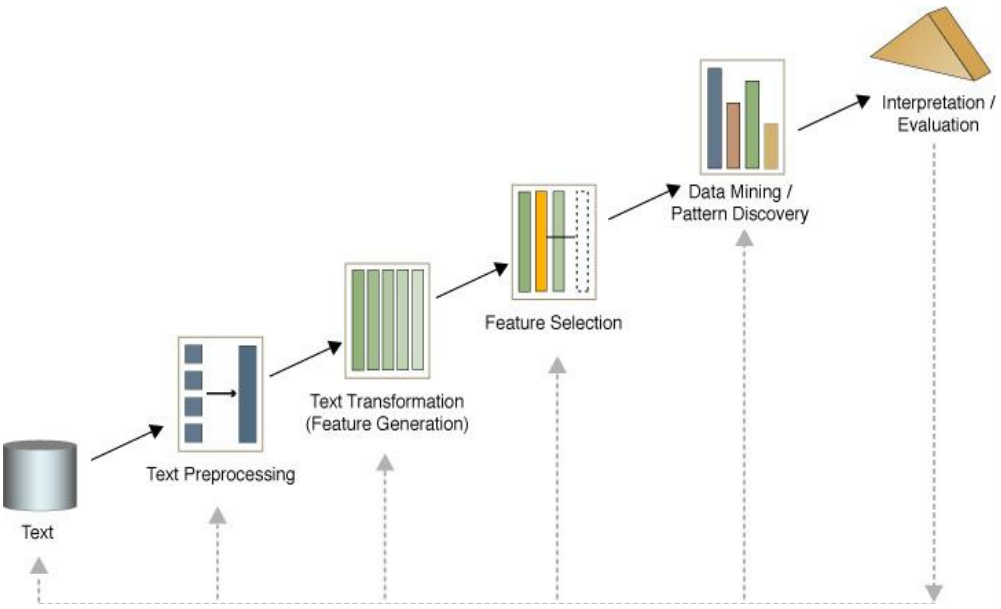
Tabel 2.3 Kinerja *Text Mining*[5]

Tan (1999) memberikan kerangka kerja dari *text mining* seperti pada Gambar 2.3 Pada tahap awal ditempuh langkah *text refining* yaitu pengubahan bentuk dari teks asli menjadi bentuk *intermediate (intermediate form),* yang dapat berbasis pada bentuk dokumen *(document-based intermediate form)* atau berbasis paad konsep *(concept-based intermediate form)*. Tahap berikutnya adalah tahap

knowledge distillation. Pada tahap ini jika bentuk intermediate berupa dokumen maka kegiatan distilasi pengetahuan dapat berupa kegiatan clustering dokumen, kategorisasi dokumen, visualisasi dan sebagainya. Untuk bentuk *intermediate* berupa konsep kegiatan distilasi dapat berupa *predictive modeling, assosiative discovery* dan visualisasi. Salah satu kegiatan penting dalam distilasi pengetahuan adalah klasifikasi atau kategorisasi teks dengan pendekatan supervised learning. Kategorisasi teks sendiri saat ini memiliki berbagai cara pendekatan antara lain berbasis numeris, misalnya pendekatan *probabilistic*, *support vector machine*, dan artificial *neural netw*ork, serta berbasis non numeris seperti *decision tree classification*.[5]

2.3.4 Proses *Text Mining*

Beberapa tahapan proses pokok dalam *text mining*, yaitu pemrosesan awal *text*, *(text preprocessing*), transformasi teks (*text transformation*)/ (*Feature Generation*), pemilihan fitur (*feature selection*), dan penemuan pola *text/data mining* (*pattern discovery*).[6]



Gambar 2.4 Proses *Text Mining* menurut even,zohar [6]

a. Text

Tahap pertama adalah permasalahan yang dihadapi pada *text mining* sama dengan permasalahan yang terdapat pada data mining, yaitu jumlah data yang

besar, dimensi yang tinggi, data dan struktur yang terus berubah, dan data *noise*. Perbedaan di antara keduanya adalah pada data yang digunakan. Pada *data mining*, data yang digunakan adalah *structured data*, sedangkan pada *text mining*, data yang digunakan text mining pada umumnya adalah *unstructured data*, atau minimal semistructured. Hal ini menyebabkan adanya tantangan tambahan pada *text mining* yaitu struktur *text* yang *complex* dan tidak lengkap, arti yang tidak jelas dan tidak standar, dan bahasa yang berbeda ditambah translasi yang tidak akurat.

b. Text Preprocessing

Tahap ini melakukan analisis semantik (kebenaran arti) dan sintaktik (kebenaran susunan) terhadap teks. Tujuan dari pemrosesan awal adalah untuk mempersiapkan teks menjadi data yang akan mengalami pengolahan lebih lanjut. Operasi yang dapat dilakukan pada tahap ini meliputi :[4]

1. *Text clean up.*

   Menghapus iklan dari halaman web, menormalkan teks dikonversi dari format biner.

2. *Case Folding*

   adalah mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf ‘a’ sampai dengan ‘z’ yang diterima. Karakter selain huruf dihilangkan dan dianggap delimiter.

3. *Tokenization*

   Sebelum pengolahan yang lebih canggih, aliran karakter berkelanjutan harus dipecah menjadi konstituen bermakna. Hal ini dapat terjadi pada tingkat yang berbeda. Dokumen dapat dipecah menjadi bab-bab, bagian, paragraf, kalimat, kata, dan bahkan suku kata. Pendekatan yang paling sering ditemukan dalam sistem *text mining* melibatkan teks menjadi kalimat dan kata-kata, yang disebut *tokenization*.

4. *part-of-speech* (PoS) *tagging*,

   menghasilkan *parse tree* untuk tiap-tiap kalimat, dan pembersihan teks yang ambigu.

c. Text Transformation (Feature Generation)

Transformasi teks atau pembentukan atribut mengacu pada proses untuk mendapatkan representasi dokumen yang diharapkan. Pendekatan representasi dokumen yang lazim *bag of words*. Transformasi teks sekaligus juga melakukan pengubahan kata-kata ke bentuk dasarnya dan pengurangan dimensi kata di dalam dokumen.

d. *Feature Selection*

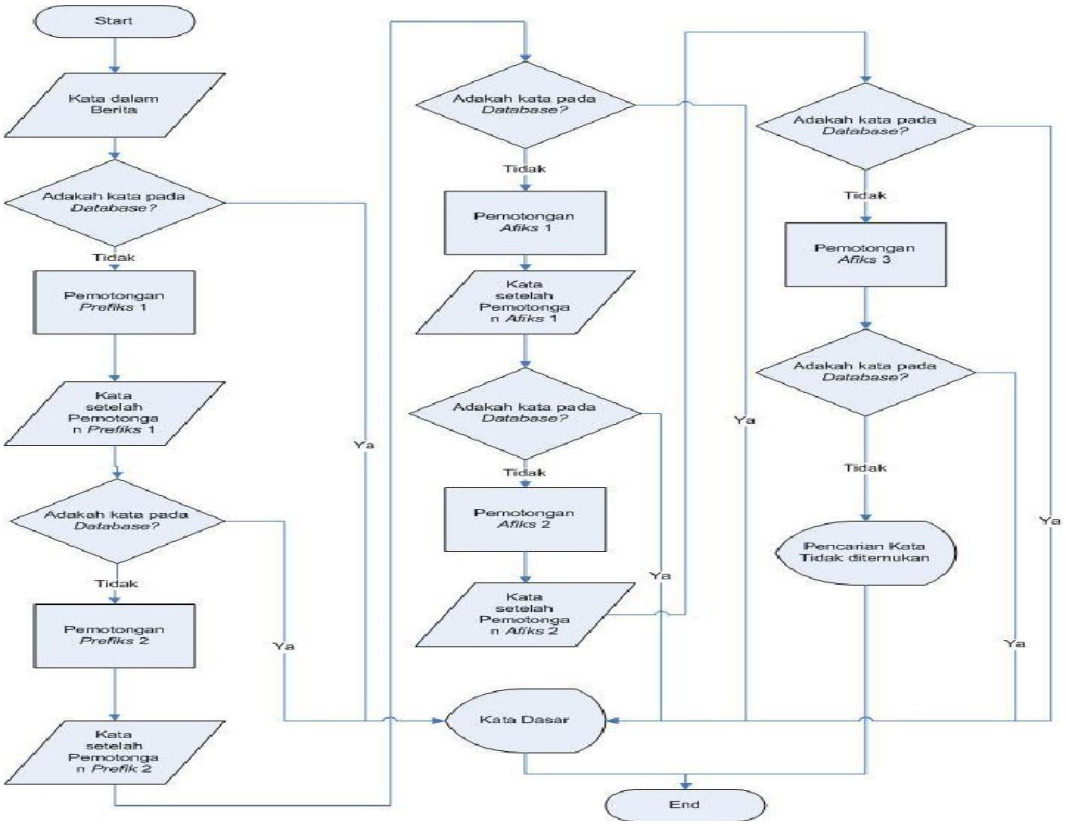
Pemilihan fitur (kata) merupakan tahap lanjut dari pengurangan dimensi pada proses transformasi teks. Operasi *feature selection* ini meliputi :

1. Stop words removal

Walaupun tahap sebelumnya sudah melakukan penghapusan kata-kata yang tidak deskriptif (*stopwords*), namun tidak semua kata-kata di dalam dokumen memiliki arti penting. Oleh karena itu, untuk mengurangi dimensi, pemilihan hanya dilakukan terhadap kata-kata yang relevan yang benar-benar merepresentasikan isi dari suatu dokumen. Langkah preprocessing yang menghilangkan atau menghapus kata-kata yang tidak penting atau tidak relevan disebut fitur seleksi. Banyak sistem, bagaimanapun, melakukan penyaringan jauh lebih agresif, menghilangkan 90 hingga 99 persen dari semua fitur[4].

2. Stemming

Stemming merupakan suatu proses yang mentransformasi kata-kata yang terdapat dalam suatu dokumen ke kata-kata akarnya (*root word*) dengan menggunakan aturan-aturan tertentu. Algoritma Nazief & Adriani sebagai algoritma *stemming* untuk teks berbahasa Indonesia yang memiliki kemampuan presentase keakuratan (presisi) lebih baik dari algoritma lainnya. Sebagai contoh, kata bersama, kebersamaan, menyamai, akan distem ke *root word*-nya yaitu "sama". Proses *stemming* pada teks ber Bahasa Indonesia berbeda dengan *stemming* pada teks berbahasa Inggris. Pada teks berbahasa Inggris, proses yang diperlukan hanya proses menghilangkan *sufiks*. Sedangkan pada teks berbahasa Indonesia, selain *sufiks, prefiks,* dan *konfiks* juga dihilangkan[7].



Gambar 2.5 *Flow Chart* Algoritma Nazief & Adriani [7]

e. Pattern Discovery

Text/Data mining/Pattern discovery merupakan tahap penting untuk menemukan pola atau pengetahuan (*knowledge*) dari keseluruhan teks. Tindakan yang lazim dilakukan pada tahap ini adalah operasi *text mining*, dan biasanya menggunakan teknik-teknik *data mining*. Dalam penemuan pola ini, proses *text mining* dikombinasikan dengan proses-proses *data mining*. [6]

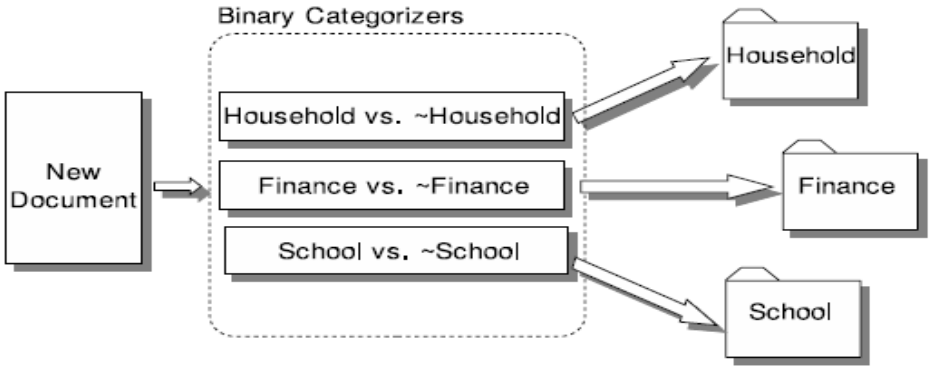
f. Interpretation/evaluation

Interpretation/*evaluation*,hasil dari proses mining akan diinterpretasikan ke dalam bentuk tertentu untuk kemudian dilakukan proses evaluasi. Masukan awal dari proses *text mining* adalah suatu data teks dan menghasilkan keluaran berupa pola sebagai hasil interpretasi atau evaluasi. Apabila hasil keluaran dari penemuan pola belum sesuai untuk aplikasi, dilanjutkan evaluasi dengan melakukan iterasi ke satu atau beberapa tahap sebelumnya. Sebaliknya, hasil interpretasi merupakan tahap akhir dari proses *text mining* dan akan disajikan ke pengguna dalam bentuk visual.[5]

2.3.5 Kategori Teks

Kategorisasi 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. Pada kategorisasi teks, diberikan sekumpulan kategori (label) dan koleksi dokumen yang berfungsi sebagai data latih, yaitu data yang digunakan untuk membangun model, dan kemudian dilakukan proses untuk menemukan kategori yang tepat untuk dokumen test, yaitu dokumen yang digunakan untuk menentukan akurasi dari model. Misalkan ada sebuah dokumen x sebagai inputan, maka output yang dihasilkan oleh model tersebuat adalah kelas atau kategori y dari beberapa kategori tertentu yang telah didefinisikan sebelumnya (y1,…,yk). Adapun contoh dari pemanfaatan kategorisasi teks adalah pengkategorisasian berita ke dalam beberapa kategori seperti bisnis, teknologi, kesehatan dan lain sebagainya; pengkategorisasian email sebagai spam atau bukan; pengkategorisasian kilasan film sebagai film favorit, netral atau tidak favorit; pengkategorisasian paper yang menarik dan tidak menarik; dan penggunaan dari kategorisasi teks yang paling umum adalah kategorisasi otomatis dari web pages yang dimanfaatkan oleh portal Internet seperti Yahoo. Kategorisasi otomatis ini memudahkan proses *browsing* artikel berdasarkan topik tertentu yang dilakukan oleh user. Salah satu algoritma kategorisasi yang sering digunakan adalah algoritma *Naive bayes*. Algoritma ini merupakan algoritma yang menerapkan metode probabilistic learning method.[4]

Kategorisasi teks adalah banyak digunakan, namun lamban, nama untuk klasifikasi dokumen. Ini adalah perwujudan paling murni dari model *spreadsheet* dengan jawaban berlabel. Setelah data ditransformasikan ke format *spreadsheet* biasa numerik, standar Data-metode *text mining* yang berlaku. Gambar 1.4 mengilustrasikan aplikasi klasifikasi dokumen. Dokumen tersebut akan disusun dalam folder, satu folder untuk setiap topik. Sebuah dokumen baru disajikan, dan tujuannya adalah untuk menempatkan dokumen ini di sesuai folder.[8]

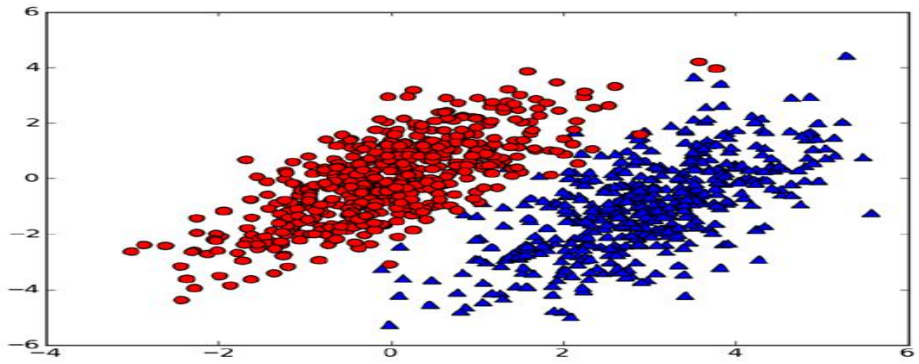


Gambar 2.6 Kategori Teks[8]

2.4 Naïve Bayes Classifier

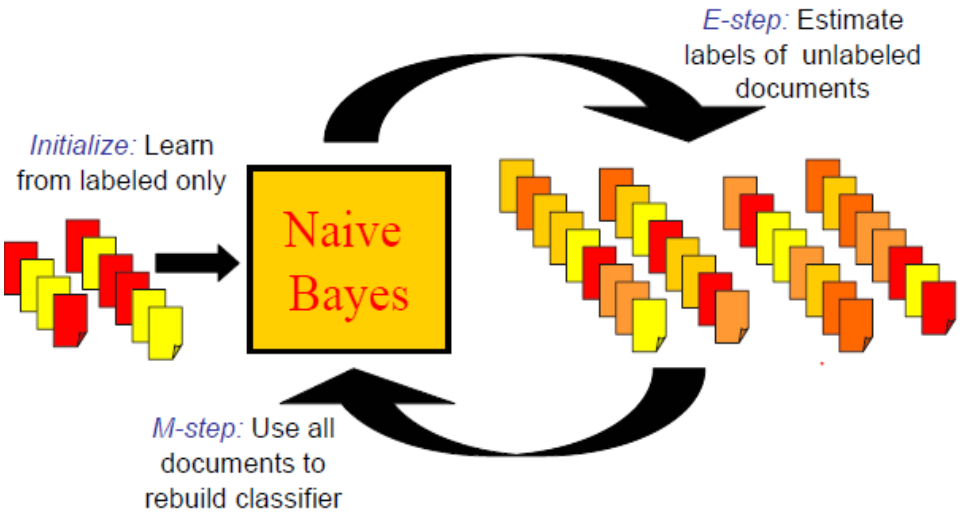
Bayesian Classifier adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. *Bayesian Classifier* didasarkan pada teorema bayes yang memiliki kemampuan klasifikasi serupa dengan *decision tree* dan *neural network*. *Bayesian Classifier* terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar.[10]

Naïve Bayes Classifier menggunakan teori probabilitas sebagai dasar teori, Metode ini adalah metode yang dipergunakan untuk proses klasifikasi teks. Terdapat 2 tahap pada proses klasifikasi teks. Tahap pertama adalah pelatihan terhadap himpunan artikel contoh (*training example*). Contoh dari asumsi bahwa kita memiliki dataset dengan dua kelas data di dalam. data ini ditunjukkan pada Gambar 27.



Gambar 2.7 Dua probabilitas distribusi dengan diketahui parameter menggambarkan distribusi[11]

Sedangkan tahap kedua adalah proses klasifikasi dokumen yang belum dikategorikan. Gambar ini menunjukkan proses klasifikasi dokumen dengan *naïve bayes classifier*.



Gambar 2.8 klasifikasi *naïve bayes classifier*.[9]

selanjutnya yaitu melakukan klasifikasi dokumen dengan mencari nilai maksimum dari :[12]

$$\arg\max_{vj=v} P\,(\,vj \mid a_1, a_2, \ldots\ldots, a_n) \tag{2.1}$$

konsep dasar yang digunakan oleh naïve bayes adalah Teorema Bayes, dengan perhitungan nilai probabilitas bersyarat :

$$P\,(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \tag{2.2}$$

yaitu, Peluang kejadian A sebagai B ditentukan dari peluang B saat A, peluang A, peluang A, dan peluang B. pada saat pengaplikasiannya nanti rumus ini berubah menjadi :

$$P\,(Ci|D) = \frac{P(D|Ci)\,P(Ci)}{P(D)} \tag{2.3}$$

dengan menerapkan teorema bayes persamaan (2.1) maka dapat ditulis :

$$V_{MAP} = \arg\max_{vj=v} \frac{P\,(a_1, a_2, \ldots\ldots, a_n\,|vj)P(vj)}{P(a_1, a_2, \ldots\ldots, a_n)} \tag{2.4}$$

karena nilai $P(a_1, a_2, \ldots\ldots, a_n)$ konstan, sehingga untuk semua *vj* besarnya sama maka nilainya dapat diabaikan, menjadi persamaan :

$$V_{MAP} = \underset{vj=v}{\arg\max} P(a_1, a_2, \ldots\ldots, a_n \mid vj)\, P(vj) \qquad (2.5)$$

dengan mengasumsikan bahwa setiap kata dalam $< a_1, a_2, \ldots\ldots, a_n >$ sulit untuk dihitung, maka akan diasumsikan $P(a_1, a_2, \ldots\ldots, a_n \mid vj)$ bahwa setiap kata pada dokumen tidak mempunyai keterkaitan, dan ditulis persamaan menjadi :

$$V_{MAP} = \underset{vj=v}{\arg\max} P(vj) \prod_i P(ai \mid vj) \qquad (2.6)$$

nilai $P(vj)$ ditentukan pada saat pelatihan, yang nilainya didekati dengan persamaan :

$$P(vj) = \frac{|\,Docs\ j\,|}{|Contoh\,|} \qquad (2.7)$$

dimana :

$P(vj)$: probabilitas setiap dokumen terhadap sekumpulan dokumen.

$|\,Docs\ j\,|$: banyaknya dokumen yang memiliki kategori j dalam pelatihan,

$|Contoh\,|$: banyaknya dokumen dalam contoh yang digunakan saat pelatihan.

untuk nilai $P(wk \mid vj)$ ditentukan dengan persamaan :

$$P(wk \mid vj) = \frac{nk+1}{n+|Kosakata\,|} \qquad (2.8)$$

Dimana :

$P(wk \mid vj)$: Probabilitas kemunculan kata *wk* pada suatu dokumen dengan kategori *vj*.

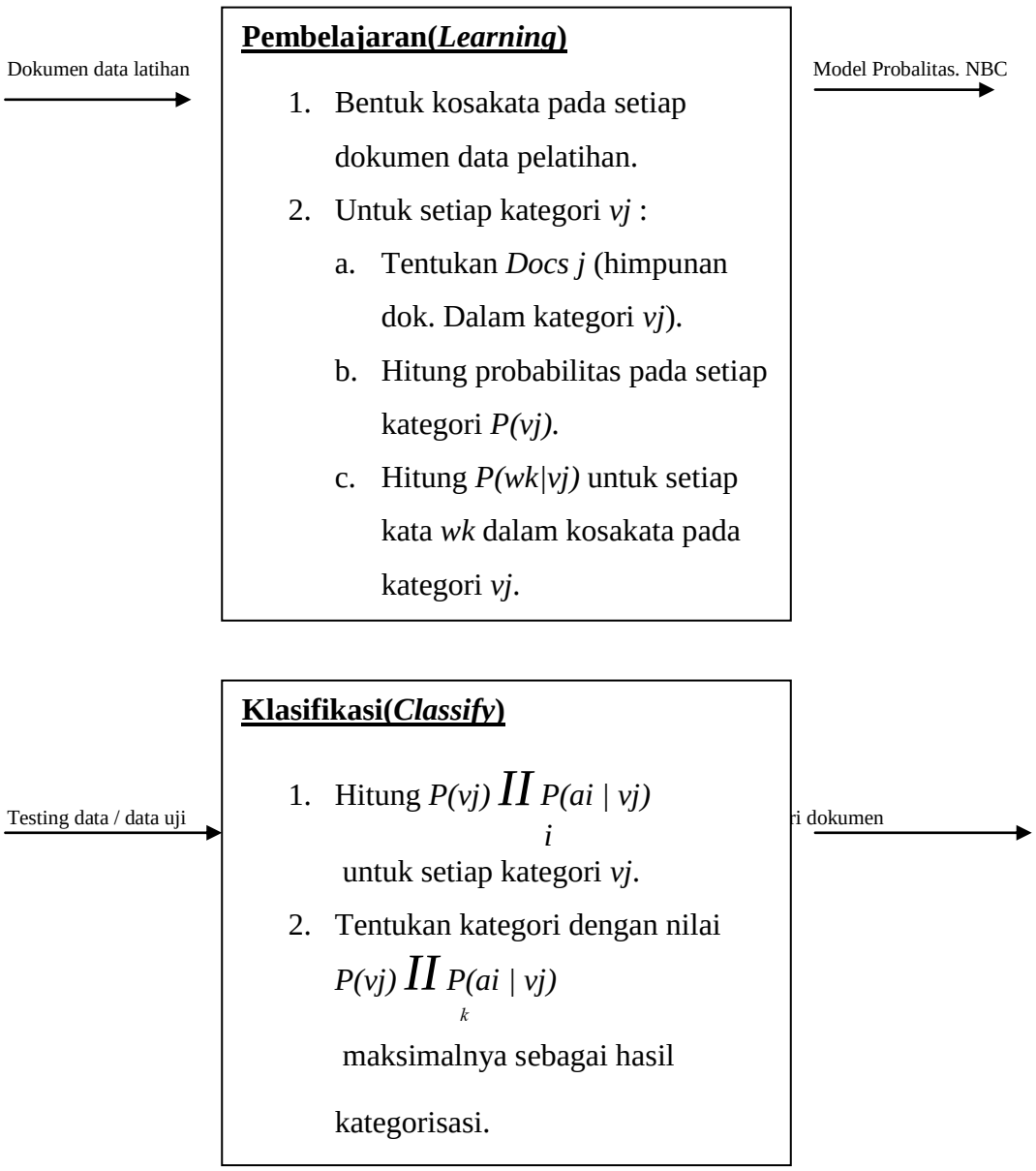
nk : frekuensi munculnya kata *wk* dalam dokumen yang berkategori *vj*.

sedangkan *n* : banyaknya seluruh kata dalam dokumen berkategori *vj*.

| Kosakata | : banyaknya kata dalam contoh pelatihan.

Pada persamaan 2.8 terdapat suatu penambahan 1 pada pembilang, hal ini dilakukan untuk mengantisipasi jika terdapat suatu kata pada dokumen uji yang tidak ada pada setiap dokumen data pelatihan.

Selanjutnya, pembelajaran algoritma *Naïve Bayes Classifier* dapat dilihat pada gambar 2.9 sebagai berikut :



Gambar 2.9 Tahapan Proses Klasifikasi Dokumen Dengan Algoritma Naïve Bayes Classifier [12]

2.5 Konsep Dasar Pemrograman Objek

2.5.1 Pemrograman Berorientasi Objek (PBO)

Pemrograman berorientasi objek berarti sebuah teknik pemrograman yang dalam proses pengembangannya menggunakan terminologi objek, di mana setiap objek memiliki atribut beserta dengan fungsi yang dapat saling berinteraksi satu dengan yang lain. Contohnya objek makhluk hidup – baik manusia ataupun binatang maka kita akan mengenalinya dari bentuk, ukuran, beratnya kemudian dari perilakunya kita dapat melihat bahwa makhluk hidup dapat melihat, makan, berjalan, berlari dan selain itu terdapat fungsi lainnya seperti fungsi peredaran darah yang berhubungan dengan fungsi pernapasan atau fungsi pencernaan.

Ciri-ciri inilah yang menjadi ide dasar bagaimana mengembangkan sebuah perangkat lunak yang kompleks dengan menggunakan model objek tersebut. Dengan demikian, proses pengembangan perangkat lunak akan lebih mudah karena hanya akan menyesuaikan model pemrograman dengan objek yang kita buat. PBO memiliki tujuan untuk memberikan pemahaman sistem kepada user atau client. Hal ini karena user/client dapat lebih mudah memahami alur pemrograman dengan kasus yang dihadapi, hal ini berbeda dengan pemrograman terstruktur karena lebih berorientasi kepada programmer untuk menyelesaikan sebuah kasus.[13]

2.6 Pengertian UML (*Unified Modeling Language)*

UML singkatan dari Unified Modeling Language yang berarti bahasa pemodelan standar. (Chonoles, 2003 : Bab1) mengatakan sebagai bahasa, berarti UML memiliki sintaks dan semantik. Ketika akan membuat model menggunakan konsep UML ada aturan-aturan yang harus diikuti. Bagaimana elemen pada model-model yang dibuat berhubungan satu dengan lainnya dan harus mengikuti standar yang ada. UML bukan hanya sekedar diagram, tetapi juga menceritakan konteksnya.[14]

UML adalah himpunan struktur danteknik untuk pemodelan desain program berorientasi objek (OOP) serta aplikasinya. UML adalah metodologi untuk mengembangkan sistem OOP dan sekelompok perangkat tool untuk

mendukung pengembangan sistem tersebut UML mulai diperkenalkan oleh Object Management Group, sebuah organisasi yang telah mengembangkan model, teknologi, dan standar OOP sejak tahun 1980-an. Sekarang UML sudah mulai banyak digunakan oleh para praktisi OOP. UML merupakan dasar bagi perangkat (tool) desain berorientasi objek dari IBM.

UML adalah suatu bahasa yang digunakan untuk menentukan, memvisualisasikan, membangun, dan mendokumentasikan suatu sistem informasi. UML dikembangkan sebagai suatu alat untuk analisis dan desain berorientasi objek oleh Grady Booch, Jim Rumbaugh, dan Ivar Jacobson.Namun demikian UML dapat digunakan untuk memahami dan mendokumentasikan setiap sistem informasi.Penggunaan UML dalam industri terus meningkat. Ini merupakan standar terbuka yang menjadikannya sebagai bahasa pemodelan yang umum dalam industri peranti lunak dan pengembangan sistem.

UML diaplikasikan untuk maksud tertentu, biasanya antara lain untuk :

1. Merancang perangkat Lunak
2. Sarana Komunikasi antara perangkat lunak dengan proses bisnis.
3. Menjabarkan sistem secara rinci untuk analisa dan mencari apa yang diperlukan sistem.
4. Mendokumentasi sistem yang ada, proses-proses dan organisasinya.

2.6.1 Diagram-Diagram UML

Beberapa literature menyebutkan bahwa UML menyediakan Sembilan jenis diagram, yang lain menyebutkan sepuluh karena ada beberapa diagram yang digabung, misalnya diagram komunikasi, diagram urutan dan diagram perwaktuan digabung menjadi diagram interaksi. Namun demikian model-model itu dapat dikelompokkan menjadi 10 macam diagram untuk memodelkan aplikasi berorientasi objek, yaitu :

UML menyediakan 10 macam diagram untuk memodelkan aplikasi berorientasi objek, yaitu:

1. *Use Case Diagram* untuk memodelkan proses bisnis.

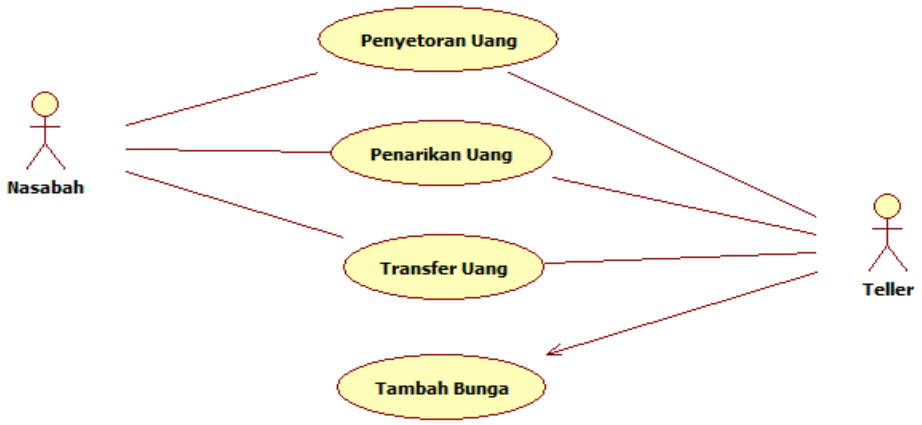
2. *Conceptual Diagram* untuk memodelkan konsep-konsep yang ada di dalam aplikasi.
3. *Sequence Diagram* untuk memodelkan pengiriman pesan (message) antar objek.
4. *Collaboration Diagram* untuk memodelkan interaksi antar objek.
5. *State Diagram* untuk memodelkan perilaku objek di dalam sistem.
6. *Activity Diagram* untuk memodelkan perilaku userdan objek di dalam sistem.
7. *Class Diagram* untuk memodelkan struktur kelas.
8. *Objek Diagram* untuk memodelkan struktur objek.
9. *Component Diagram* untuk memodelkan komponen objek.
10. *Deployment Diagram* untuk memodelkan distribusi aplikasi.

Berikut akan dijelaskan 4 macam diagram yang paling sering digunakan dalam pembangunan aplikasi berorientasi objek, yaitu use case diagram, sequence diagram, collaboration diagram, dan class diagram.

a. *Use Case Diagram*

Use case diagram digunakan untuk memodelkan bisnis proses berdasarkan perspektif pengguna sistem. *Use case* diagram terdiri atas diagram untuk *use case* dan *actor*. *Actor* merepresentasikan orang yang akan mengoperasikan atau orang yang berinteraksi dengan sistem aplikasi. *Use case* merepresentasikan operasi-operasi yang dilakukan oleh *actor*. *Use case* digambarkan berbentuk elips dengan nama operasi dituliskan didalamnya. Actor yang melakukan operasi dihubungkan dengan garis lurus ke *use case*.

Gambar di bawah ini merupakan salah satu contoh bentuk diagram *use case.*

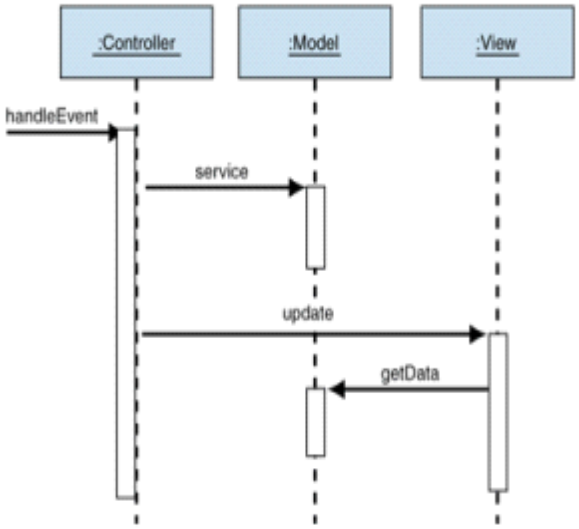


Gambar 2.10 Diagram *Use Case*

b. *Sequence Diagram*

Sequence diagram menjelaskan secara detil urutan proses yang dilakukan dalam sistem untuk mencapai tujuan dari *use case*. Interaksi yang terjadi antar *class*, operasi apa saja yang terlibat, urutan antar operasi, dan informasi yang diperlukan oleh masing-masing operasi.

Gambar di bawah ini merupakan salah satu contoh bentuk diagram *sequence* diagram.

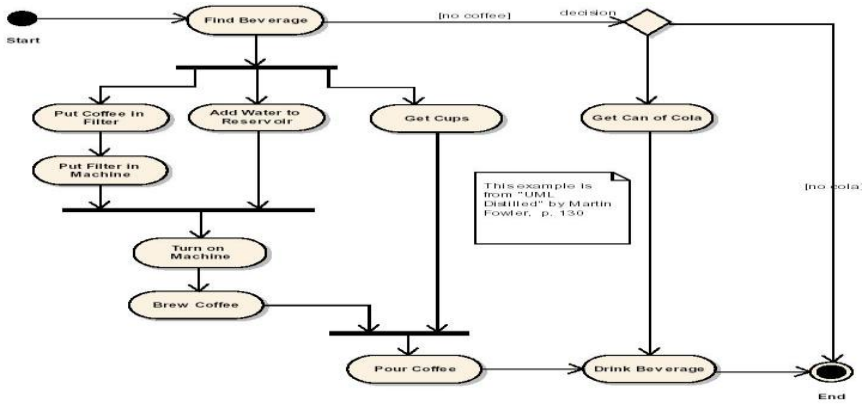


Gambar 2.11 Diagram *Sequence*

b. *Activity Diagram*

Activity diagram adalah representasi grafis dari alur kerja tahapan aktifitas. Diagram ini mendukung pilihan tindakan, iterasi dan *concurrency*. Pada pemodelan UML, *activity diagram* dapat digunakan untuk menjelaskan bisnis dan alur kerja operasional secara *step-by-step* dari komponen suatu sistem. *Activity diagram* menunjukkan keseluruhan dari aliran kontrol.

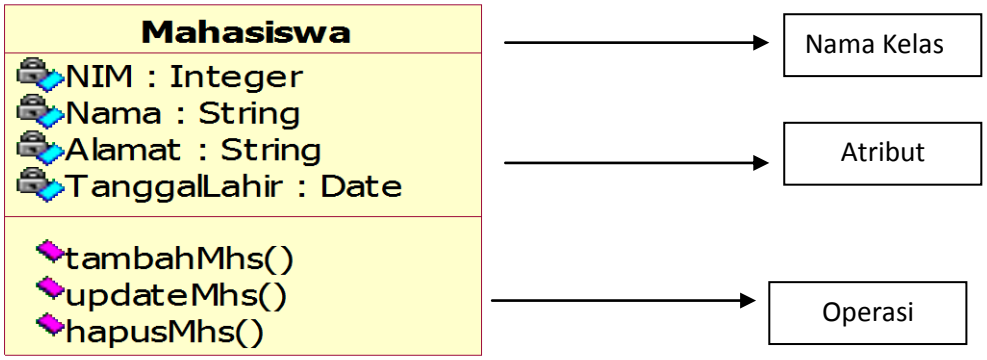
Gambar di bawah ini merupakan salah satu contoh bentuk diagram *activity* diagram.



Gambar 2.12 Diagram *Activity*

c. *Class Diagram*

Class diagram merupakan diagram yang selalu ada di permodelan sistem berorientasi objek. *Class* diagram menunjukkan hubungan antar class dalam sistem yang sedang dibangun dan bagaimana mereka saling berkolaborasi untuk mencapai suatu tujuan.



Gambar 2.13 *Class* Diagram

2.7 MYSQL

MySQL adalah sebuah perangkat lunak sistem manajemen basis data SQL atau DBMS yang multithread, multi-user, dengan sekitar 6 juta instalasi di seluruh dunia. MySQL AB membuat MySQL tersedia sebagai perangkat lunak gratis dibawah lisensi GNU General Public License (GPL), tetapi mereka juga menjual dibawah lisensi komersial untuk kasus-kasus dimana penggunaannya tidak cocok dengan penggunaan GPL.

Tidak sama dengan proyek-proyek seperti Apache, dimana perangkat lunak dikembangkan oleh komunitas umum, dan hak cipta untuk kode sumber dimiliki oleh penulisnya masing-masing, MySQL dimiliki dan disponsori oleh sebuah perusahaan komersial Swedia MySQL AB, dimana memegang hak cipta hampir atas semua kode sumbernya. Kedua orang Swedia dan satu orang Finlandia yang mendirikan MySQL AB adalah David Axmark, Allan Larsson, dan Michael "Monty" Widenius.

Ada beberapa kelebihan yang dimiliki MySQL sehingga dapat menarik banyak pengguna. Kelebihan tersebut yaitu:

1. Fleksibilitas. Saat ini, MySQL telah dioptimasi untuk duabelas platform seperti HP-UX, Linux, Mac OS X, Novell Netware, OpenBSD, Solaris, Microsoft Windows dan lain-lain. MySQL juga menyediakan source code yang dapat diunduh secara gratis, sehingga pengguna dapat mengkompilasi sendiri sesuai platform yang digunakan. Selain itu, MySQL juga dapat dikustomisasi sesuai keinginan penggunanya, misalnya mengganti bahasa yang digunakan pada antarmukanya.
2. Performansi. Sejak rilis pertama, pengembang MySQL fokus kepada performa. Hal ini masih tetap dipertahankan hingga sekarang dengan terus meningkatkan fiturnya.
3. Lisensi. MySQL menawarkan berbagai pilihan lisensi kepada penggunanya. Lisensi open source yang ditawarkan yaitu lisensi GNU General Public License dan Free/Libre and Open Source Software (FLOSS) License Exception. Selain itu ditawarkan juga lisensi komersil berbayar yang memiliki fasilitas dukungan teknis.

2.8 Macromedia Dreamweaver Adobe Cs 5

Dreamweaver merupakan sebuah aplikasi untuk merancang pembuatan website. Dreamweaver dibuat oleh perusahaan Macromedia sehingga dinamakan Macromedia Dreamweaver. Sejak Macromedia diakuisisi Adobe Inc., namanya berubah menjadi Adobe Dreamweaver. Versi Pertama Dreamweaver dibawah adobe adalah Cs5, mengikuti versi rilisnya yang dipaketkan dalam Adobe Creative Suite 5.

Untuk pengguna tingkat lanjut, Dreamweaver menyediakan tampilan Code sehingga pengguna dapat merancang tampilan yang lebih lengkap menggunakan kode. Pengguna juga dimudahkan dengan berbagai fasilitas yang dimiliki Dreamweaver seperti *tag auto-completion* untuk penulisan kode HTML. Format yang didukung Dreamweaver juga cukup lengkap, mulai dari HTML, JavaScript, CSS, sampai XML.

2.9 PHP

PHP adalah salah satu bahasa Sever-side yang didesain khusus untuk aplikasi web.PHP dapat disisipkan diantara bahasa HTML dan karena bahasa Server side, maka PHP akan dieksekusi di server,sehingga yang dikirimkan ke browse adalah “hasil jadi” dalam bentuk HTML, dan kode PHP anda tidak akan terlihat.PHP dahulunya merupakan proyek pribadi dari Rasmus Lerdof (dengan dikeluarkannya PHP versi 1) yang digunakan untuk membuat home pribadinya.Versi pertama ini berupa kumpulan script PERL.Untuk versi keduanya, Rasmus menulis ulang script-script PERL tersebut menggunakan bahasa C.kemudian menambahkan fasilitas untuk Form HTML dan koneksi Mysql. Adapun PHP didapat dari singkatan Personal Home Pages.

2.10 WAMP *(Windows Apache Mysql PHP)*

Wamp adalah suatu web server yang mengkoneksikan antara apache, mysql, dan php dengan sistem operasi windows. Dengan demikian kita tidak perlu menkonfigurasi koneksi antara apache, mysql, dan php karena sudah tersedia dan terkonfigurasi di dalam WAMP.