

## BAB II

### LANDASAN TEORI

#### 2.1 Sistem Informasi

Informasi adalah salah satu jenis utama sumber daya yang tersedia bagi manager. Informasi dapat dikelola seperti halnya sumber daya lain, dan perhatian pada topik ini bersumber dari dua pengaruh. Pertama, bisnis telah menjadi semakin rumit, dan kedua, komputer telah mencapai kemampuan yang semakin baik.

*Output* informasi dari komputer dapat digunakan oleh para menejer, non-menejer, serta orang-orang, dan organisasi-organisasi dalam lingkungan perusahaan. Menejer berada pada semua tingkat organisasional perusahaan, dan dalam semua bidang fungsional. Menejer melaksanakan fungsi dan peran, dan untuk berhasil, menejer memerlukan keahlian dalam komunikasi dan pemecahan masalah. Menejer perlu mengerti komputer (*computer litetate*), tetapi yang lebih penting mereka mengeti informasi (*informasi literate*).

Sangat bermanfaat jika menejer mampu melihat unitnya sebagai suatu sistem yang terdiri dari beberapa subsistem dan berada dalam supersistem yang lebih besar. Perusahaan adalah suatu sistem yang bersifat fisik, namun dikelola dengan suatu sistem konseptual. Sistem konseptual itu terdiri dari suatu pengolahan informasi yang mengubah data menjadi informasi dan menggambarkan sumber daya fisik.

Aplikasi utama komputer yang pertama adalah pengolahan data akuntansi. Aplikasi tersebut lalu diikuti empat aplikasi lain: sistem informasi menejemen (*menegement information sytem*), sistem pendukung keputusan (*decission support sytem*), kantor virtual (*virtual office*), dan sistem berbasis pengetahuan (*knowladge-based system*). Kelima aplikasi ini membentuk sistem informasi berbasis komputer(*computer-based information system*), atau CBIS.

Perusahaan-perusahaan membentuk suatu organisasi jasa informasi yang terdiri dari spesialis informasi untuk menyediakan keahlian dalam pengembangan sistem berbasis komputer. Para spesialis ini mencakup analisis sitem (*system*

analysts), pengelola *database* (*database administrator*), spesialis jaringan (*network specialists*), programmer dan operator. Dalam beberapa tahun terakhir para pemakai telah mulai melakukan sebagian besar pekerjaan para spesialis suatu fenomena yang disebut *end-user computing*.

Sangat sulit untuk membuktikan nilai ekonomis dari suatu aplikasi komputer, tetapi banyak analisis yang dilakukan untuk menjustifikasi tiap proyek potensial. Setelah berjalan proyek tersebut berkembang melalui siklus hidup sistem (*system life cycle*). Para spesialis informasi dapat berperan serta dalam tingkat yang beragam, tetapi keseluruhan siklus, termasuk pengembangan dan pemakaian, harus dikelola oleh menejer (Raymon McLeod, Jr.)

## 2.2 Definisi Penyakit

Penyakit adalah suatu keadaan abnormal dari tubuh atau pikiran yang menyebabkan ketidaknyamanan, disfungsi atau kesukaran terhadap orang yang di pengaruhinya, orang-orang biasanya berkonsultasi dengan seorang dokter.

Menurut Kathleen Meehan Arias, penyakit adalah suatu kesakitan yang biasanya memiliki sedikitnya dua sifat dari kriteria ini : agen etiologi telah diketahui, kelompok tanda serta gejala yang dapat diidentifikasi, atau perubahan anatomi yang konsisten.

## 2.3 Abortus

Abortus atau keguguran adalah berakhirnya kehamilan melalui cara apapun, spontan maupun buatan, sebelum janin mampu bertahan hidup. Batasan ini berdasar umur kehamilan dan berat badan. Dengan lain perkataan abortus adalah terminasi kehamilan sebelum 20 minggu atau dengan berat kurang dari 500 gr (Handono, 2009).

Ada beberapa macam istilah yang digunakan untuk membedakan abortus dalam ilmu kedokteran, yaitu :

1. *Spontaneous abortion*: gugur kandungan yang disebabkan oleh trauma kecelakaan atau sebab-sebab alami.
2. *Induced abortion* atau *procured abortion*: pengguguran kandungan yang disengaja. Termasuk di dalamnya adalah:

- a) *Therapeutic abortion*: pengguguran yang dilakukan karena kehamilan tersebut mengancam kesehatan jasmani atau rohani sang ibu, kadang-kadang dilakukan sesudah pemerkosaan
- b) *Eugenic abortion*: pengguguran yang dilakukan terhadap janin yang cacat.
- c) *Elective abortion*: pengguguran yang dilakukan untuk alasan-alasan lain.

### 2.3.1 Macam-Macam Penyakit Abortus

Menurut (Sarwono, 2008) membagi abortus menjadi beberapa klasifikasi yaitu :

1. Abortus spontan
 

Abortus yang terjadi tanpa tindakan mekanis atau medis untuk mengosongkan uterus.
2. Abortus imminens (keguguran mengancam)
 

Peristiwa terjadinya pendarahan dari uterus pada kehamilan sebelum 20 minggu.
3. Abortus incipiens (keguguran berlangsung)
 

Peristiwa pendarahan uterus pada kehamilan sebelum 20 minggu dengan adanya dilatasi serviks uteri yang meningkat, tetapi hasil konsepsi masih dalam uterus.
4. Abortus incomplet (keguguran tidak lengkap)
 

Pengeluaran sebagian hasil konsepsi pada kehamilan sebelum 20 minggu dengan masih ada sisa tertinggal dalam uterus.
5. Abortus complet (keguguran lengkap)
 

Pendarahan pada kehamilan muda di mana seluruh hasil konsepsi telah di keluarkan dari kavum uteri.
6. Abortus infeksiosa dan Abortus septik
 

Abortus infeksiosa adalah yang disertai infeksi pada genitalia, sedangkan abortus septik adalah abortus infeksiosa

berat dengan penyebaran kuman atau toksinnya ke dalam peredaran darah atau peritoeum.

### 2.3.2 Penyebab Abortus

Abortus pada wanita hamil dapat dipengaruhi oleh beberapa sebab diantaranya:

1. Kelainan pertumbuhan hasil konsepsi
2. Kelainan pada plasenta
3. Faktor ibu yang terserang penyakit
4. Kelainan yang terjadi pada organ kelamin ibu tersebut

## 2.4 Android

Android adalah sistem operasi untuk telepon seluler yang berbasis *linux*. Android menyediakan *platform* terbuka bagi para pengembang untuk menciptakan aplikasi mereka sendiri untuk digunakan oleh bermacam piranti. Selanjutnya aplikasi yang akan dibuat penulis menggunakan *Android Software Development Kit* dan *Eclipse Application Programming Interface* sebagai media penulisan *code* untuk pengembang aplikasi. ( Developer.android.com, 2012)

Android merupakan sistem operasi yang diperuntukan untuk smartphone dan perangkat bergerak lainnya. Sistem operasi Android memiliki kemampuan yang sangat baik dalam multitasking atau membuka banyak aplikasi dalam satu waktu. Sistem operasi Android bersifat free atau bebas untuk melakukan modifikasi terhadap aplikasi yang ada dalam Android dan juga dikembangkan sesuai dengan kebutuhan.

Android adalah aplikasi *opensource* yang banyak diterapkan pada *smartphone* dengan mangsa pasar menengah sehingga mudah dijangkau oleh konsumen. Android juga mengalami peningkatan penggunaan dari tahun ke tahun sehingga dimungkinkan akan menjadi tren untuk ke depan.

### 2.4.1 Sejarah dan Perkembangan Android

Pada mulanya terdapat berbagai macam sistem operasi pada perangkat selular, diantaranya sistem operasi Symbian, Microsoft Windows Mobile,

MobileLinux, iPhone, dan sistem operasi lainnya. Android adalah sistem operasi untuk telepon seluler yang berbasis Linux. Android menyediakan *platform* yang bersifat *open source* bagi para pengembang untuk menciptakan sebuah aplikasi. Awalnya, Google Inc. mengakuisi Android Inc. yang mengembangkan *software* untuk ponsel yang berada di Palo Alto, California Amerika Serikat. Kemudian untuk mengembangkan Android, dibentuklah *Open Handset Alliance*, yaitu konsorsium dari 34 perusahaan *hardware*, *software*, dan telekomunikasi, termasuk Google, HTC, Intel, Motorola, Qualcomm, T-Mobile, dan Nvidia (Safa'at, 2011)

Banyaknya sistem Operasi untuk perangkat mobile membuat calon pengguna kadang bingung menentukan pilihan tepat mana yang terbaik untuk dimilikinya, sementara itu Android menjadi OS yang paling cepat berkembang. Diperkenalkan pada 5 November 2007, kini OS Android sudah masuk generasi keempat yang diberi nama Ice Cream Sandwich. Kepopuleran Android mampu menggerus pangsa pasar iOS dari Apple, BlackBerry OS dari RIM, Windows Phone dari Microsoft, dan Symbian dari Nokia. Android dengan cepat menguasai pangsa pasar smartphone global, dengan perolehan 52% pada kuartal III tahun 2011, menurut data dari lembaga riset Gartner.

International Data Corporation (IDC) merilis data pengapalan smartphone dan market share di kuartal ketiga tahun 2012. Hasilnya seperti sudah banyak diketahui adalah sistem operasi Android Google makin menguasai pasar smartphone dunia dengan penguasaan pasar 75%. Ini artinya setiap empat smartphone yang dikapalkan, tiga diantaranya berbasis Android. Android juga membukukan rekor pengapalan terbesar dalam satu kuartal, yaitu mencapai 136 juta unit smartphone selama kuartal ketiga tahun 2012. Angka ini naik 91.5% ketimbang periode yang sama tahun lalu. iPhone (iOS) menduduki posisi kedua dengan total pengiriman 26.9 juta unit (14.9%), sedangkan penjualan BlackBerry yang terus menerus turun masih menduduki posisi ketiga dengan pengiriman 7.7 juta unit (4.3%), turun 34% ketimbang periode yang sama tahun sebelumnya (Developer.android.com, 2012)

Era handphone layar sentuh saat ini beredar dimana-mana, salah satu yang paling populer saat ini adalah yang menggunakan sistem operasi android. pada awalnya pengguna android adalah samsung. perlu diketahui bahwa android adalah salah satu produk google yang free atau gratis untuk kita pakai. selain itu juga kita dapat dengan mudah mempelajarinya. berikut adalah beberapa kelebihan dari android ( Retto, 2009 ) :

1. Open source

Ibarat sistem operasi untuk komputer desktop, Android adalah Linux untuk tablet. Android adalah sistem operasi yang berbasis kernel Linux. Android merupakan sistem operasi terbuka (open source). Artinya, Google merilis kode sumber (source code) Android untuk memudahkan pengembang aplikasi yang hendak menciptakan aplikasi untuk Android.

2. Google terus mengembangkan Android

Google nampak begitu serius mengembangkan ekosistem Android. Hingga 2012, Android telah dikembangkan sampai generasi keempat atau disebut Ice Cream Sandwich. Dari generasi ke generasi, Android terus mengalami kemajuan signifikan. Mulai dari performa teknis, hingga tampilan antarmukanya.

3. Ketersediaan aplikasi yang berkualitas

Karena Android bersifat open source, para programmer dan pengembang aplikasi pun antusias menciptakan aplikasi. Pengembang aplikasi tak perlu berhadapan dengan prosedur platform OS berbayar, karena aplikasi buatan mereka bebas didistribusikan dengan lisensi *open source*, *shareware* atau bahkan *freeware*.

4. Pengembang rajin meng-upgrade aplikasi

Pengembang aplikasi untuk Android rajin melakukan upgrade atas aplikasi yang telah mereka buat. Ini membuktikan keseriusan dan konsistensi pengembang pada aplikasi mereka agar tidak kehilangan konsumen.

5. Keberagaman pilihan perangkat

Android tidak hanya diproduksi oleh satu vendor saja. Google mempersilahkan vendor mana saja yang ingin mengadopsi Android. Google meyakini bahwa Android adalah tentang pilihan.

#### 6. Harga terjangkau

Keberagaman pilihan smartphone Android menciptakan persaingan dagang yang kuat di antara para vendor. Ini akan merangsang para vendor untuk menciptakan perangkat Android, yang ditujukan untuk segmen low end sampai high end. Sehingga semua kalangan dapat menjajal sensasi Android.

#### 7. Bebas mengkostumisasi ROM

ROM adalah singkatan dari Read Only Memory. Kostumisasi ROM pada Android biasa disebut rooting. Rooting bukanlah sesuatu yang ilegal dalam ekosistem Android. Fungsi rooting sejatinya digunakan pihak ketiga untuk mengembangkan aplikasi. Sedangkan untuk individu, mereka dapat mengetahui dan mengakses file sistem Android. Memang dibutuhkan kemampuan teknis untuk melakukan rooting, seperti teknik *flashing*, *restore*, *recovery*, dan sebagainya.

#### 8. Terkoneksi dengan seluruh layanan Google

Bagi Anda yang sudah sangat familiar dengan layanan Google, maka Android adalah pilihan yang tepat. Android didesain untuk mendukung layanan GMail, Google Docs, Google Maps, YouTube, Google Talk, Google+, dan tentu saja mesin pencari Google. Selain tujuh keunggulan Android di atas, ada satu hal penting yang sebenarnya menjadi kunci keberhasilan Android tak lain adalah dukungan vendor-vendor ternama terhadap Android.

Android didukung oleh sebuah aliansi bernama Open Handset Alliance (OHA), yang didirikan tepat saat Google memperkenalkan Android, 5 November 2007. OHA adalah sebuah konsorsium yang dipimpin oleh Google, dan sampai saat ini OHA sudah beranggotakan 84 perusahaan yang bergerak di bidang hardware, software, telekomunikasi, dan semikonduktor dari seluruh dunia (Developer.android.com, 2012 )

### 2.4.2 Dalvik Virtual Machine (DVM)

Salah satu elemen kunci dari android adalah Dalvik Virtual Machine (DVM). Android berjalan di Dalvik Virtual Machine (DVM) bukan di Java Virtual Machine (JVM), sebenarnya banyak persamaan dengan Java Virtual Machine (JVM) seperti Java ME (Java Mobile Edition), tetapi android menggunakan Virtual Machine sendiri yang diskostumisasi dan dirancang untuk memastikan bahwa beberapa fitur-fitur berjalan lebih efisien pada perangkat mobile. Pada dasarnya, konsep pemrograman yang dianut oleh Android adalah mengarah ke pola MVC (*Model View Controller*) dimana tampilan/layout disimpan di dalam file XML. Ada empat komponen sistem yang penting dan dibutuhkan untuk menjalankan aplikasi Android(Developer.android.com,2012 ).

#### 1. Activities

Aktivitas disini merupakan sebuah *user interface* yang dapat berinteraksi dengan pengguna. Contohnya yaitu daftar menu dari *item-item*, mengirim pesan teks, dan lain-lain.

#### 2. Service

Servis tidak memiliki *user interface*, namun bekerja di belakang layar misalnya mengambil data dari *server* untuk meng-*update* aplikasi android.

#### 3. Broadcast Receivers

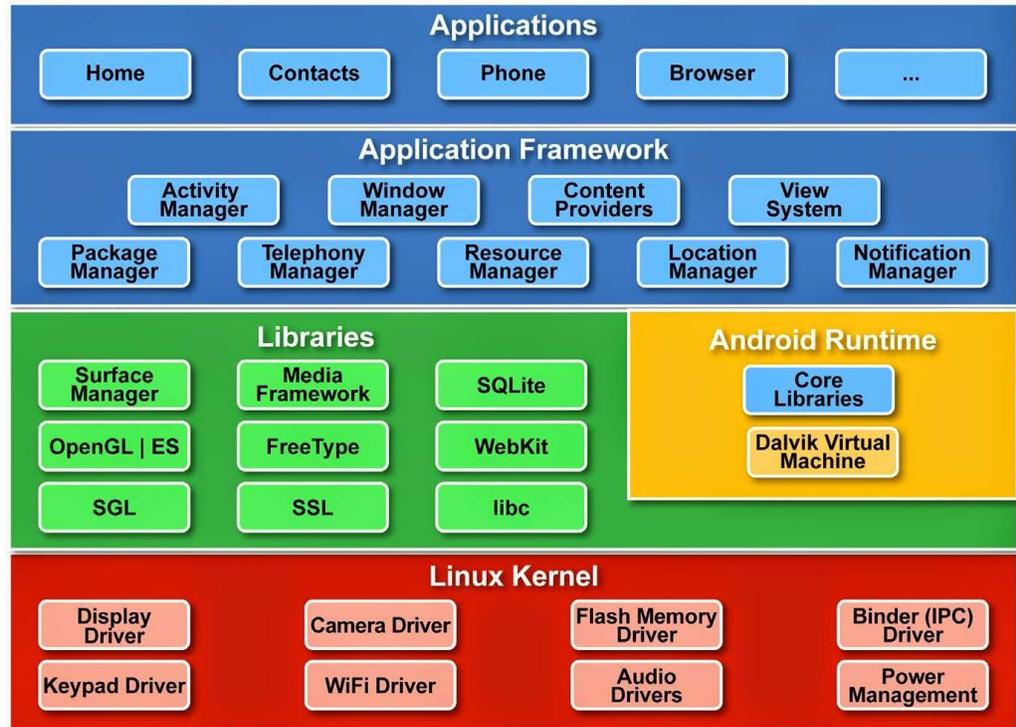
Merupakan komponen yang hanya menerima dan bereaksi atas pemberitahuan yang muncul, *Broadcast* bukan merupakan sebuah *user interface*, tetapi merupakan semacam *service* yang menjalankan aktivitas tertentu, sesuai dengan reaksi yang diterima

#### 4. Content providers

Merupakan seperangkat data dari aplikasi Android. Data disimpan di dalam file sistem, ataupun di SQLite *Database*.

### 2.4.3 Arsitektur Android

Sistem operasi Android dibangun berdasarkan kernel Linux dan memiliki arsitektur sesuai dengan Gambar 2.1



**Gambar 2.1** Arsitektur Android

Berikut ini merupakan penjelasan di setiap layer dari gambar arsitektur Android diatas (Hermawan, 2011 ) :

#### Layer Application

Android akan mengirim serangkaian aplikasi inti dari fungsionalitas handphone diantaranya *email client*, SMS, kalender, *map*, *browser*, *contact*, dan lain-lain. Semua aplikasi tersebut ditulis menggunakan bahasa pemrograman java.

#### Layer Application Framework

Dengan menyediakan *open development platform*, pengembangan mampu membuat aplikasi yang sangat beragam dan inovatif. Pengembang bebas untuk menambah fitur seperti akses lokasi informasi (*my location*),

menjalankan *background service*, mengatur alarm, menambah notifikasi ke status bar, dan masih banyak lagi.

### **Layer Libraries**

Android memasukkan seperangkat library C/C++ yang di gunakan oleh berbagai komponen dari sistem Android. Fitur ini terbuka untuk para pengembang melalui *Android Application Framework*. Beberapa contoh dari *Core Libraries* yaitu *System C Library*, *media Libraries*, *Surface manager*, *SQLite*, dan lain-lain.

### **Layer Android Runtime**

Android memasukkan seperangkat *Core Libraries* yang menyediakan sebagian besar fungsi yang tersedia di *Core Libraries* dari bahasa pemrograman java. Setiap aplikasi android berjalan di prosesnya sendiri, melalui *Dalvik Virtual Machine*.

### **Layer Linux Kernel**

Android menggunakan 2.3 untuk *Core System Service* seperti keamanan, manajemen memori, manajemen proses, *network stack*, dan driver model.

## **2.5 Eclipse**

Eclipse adalah sebuah IDE (Integrated Development Environment) untuk mengembangkan perangkat lunak dan dapat dijalankan di semua platform (platform-independent). Berikut ini adalah sifat dari Eclipse:

- a) Multi-platform: Target sistem operasi Eclipse adalah Microsoft Windows, Linux, Solaris, AIX, HP-UX dan Mac OS X.
- b) Multi-language: Eclipse dikembangkan dengan bahasa pemrograman Java, akan tetapi Eclipse mendukung pengembangan aplikasi berbasis bahasa pemrograman lainnya, seperti C/C++, Cobol, Python, Perl, PHP, dan lain sebagainya.
- c) Multi-role: Selain sebagai IDE untuk pengembangan aplikasi, Eclipse pun bisa digunakan untuk aktivitas dalam siklus pengembangan

perangkat lunak, seperti dokumentasi, test perangkat lunak, pengembangan web, dan lain sebagainya.

Eclipse pada saat ini merupakan salah satu IDE favorit dikarenakan gratis dan open source, yang berarti setiap orang boleh melihat kode pemrograman perangkat lunak ini. Selain itu, kelebihan dari Eclipse yang membuatnya populer adalah kemampuannya untuk dapat dikembangkan oleh pengguna dengan komponen yang dinamakan *plug-in* ([id.wikipedia.org/wiki/eclipse](http://id.wikipedia.org/wiki/eclipse) 2016)

### 2.5.1 Arsitektur Eclipse

Sejak versi 3.0, Eclipse pada dasarnya merupakan sebuah *kernel*, yang mengangkat *plug-in*. Apa yang dapat digunakan di dalam Eclipse sebenarnya adalah fungsi dari *plug-in* yang sudah diinstal. Ini merupakan basis dari Eclipse yang dinamakan *Rich Client Platform* (RCP). Berikut ini adalah komponen yang membentuk RCP:

- a) Core Platform
- b) OSGi
- c) SWT (*Standard Widget Toolkit*)
- d) JFace
- e) Eclipse Workbench

Secara standar Eclipse selalu dilengkapi dengan JDT (*Java Development Tools*), *plug-in* yang membuat Eclipse kompatibel untuk mengembangkan program Java, dan PDE (*Plug-in Development Environment*) untuk mengembangkan *plug-in* baru. Eclipse beserta *plug-in*-nya diimplementasikan dalam bahasa pemrograman Java.

Konsep Eclipse adalah IDE yang terbuka (*open*), mudah diperluas (*extensible*) untuk apa saja, dan tidak untuk sesuatu yang spesifik. Jadi, Eclipse tidak saja untuk mengembangkan program Java, akan tetapi dapat digunakan untuk berbagai macam keperluan, cukup dengan menginstal *plug-in* yang dibutuhkan. Apabila ingin mengem-

bangkan program C/C++ terdapat *plug-in* CDT (*C/C++ Development Tools*). Selain itu, pengembangan secara visual bukan hal yang tidak mungkin oleh Eclipse, *plug-in* UML2 tersedia untuk membuat diagram UML. Dengan menggunakan PDE setiap orang bisa membuat *plug-in* sesuai dengan keinginannya. Salah satu situs yang menawarkan *plug-in* secara gratis seperti *Eclipse downloads by project*



**Gambar 2.2** Logo Eclipse

### 2.5.2 Bahasa Pemrograman Java

Java adalah bahasa pemrograman yang dapat dijalankan di berbagai literatur termasuk telepon genggam. Dikembangkan oleh *Sun Microsystems* dan diterbitkan tahun 1995. Java tidak boleh disalah pahami sebagai *JavaScript*. *JavaScript* adalah bahasa *scripting* yang digunakan oleh *web browser* ( Retto ,2009 )

Ada beberapa literat mengapa Java dipilih untuk aplikasi komunikasi bergerak :

1. Java merupakan bahasa pemrograman yang aman. Hal ini disebabkan karena kode Java harus selalu dieksekusi dengan *Java Virtual Machine (JVM)*, yang menyediakan lingkungan yang aman untuk eksekusi kode yang telah di *download*. Sebuah aplikasi biner dapat mengacaukan sebuah peralatan (misalnya tampilan *blue screen* pada *handphone*) namun aplikasi Java, dalam hal yang paling buruk, hanya dapat mengacaukan *virtual Machine* – nya saja.
2. Bahasa pemrograman Java adalah bahasa yang kokoh. Penggunaan *garbage collector* akan mengurangi waktu yang diperlukan oleh *programmer* untuk menemukan kebocoran dalam penggunaan memori. Demikian pula penggunaan

mekanisme *exception* pada Java mampu meningkatkan kemampuan *programmer* dalam membuat aplikasi yang kokoh.

Portabilitas Java sangat berguna dalam aplikasi wireless. Pertama, dengan menulis sebuah program saja, aplikasi tersebut dapat dijalankan di berbagai peralatan. Alasan kedua, adalah karena aplikasi tersebut dapat dikirimkan ke sebuah peralatan melalui jaringan wireless. Sedangkan untuk Macam-macam Java 2 *Software Developer Kit* (J2SDK) antara lain ( Mulyadi, 2010 ) :

1. J2SE (Java 2 Standard Edition)

Kategori ini digunakan untuk menjalankan dan mengembangkan aplikasi-aplikasi Java pada level PC (*Personal Computer*).

2. J2EE (Java 2 Enterprise Edition)

Kategori ini digunakan untuk menjalankan dan mengembangkan aplikasi-aplikasi Java pada lingkungan entriprise dengan menambahkan fungsionalitas-fungsionalitas Java semacam EJB, Java CORBA, Servlet dan JSP serta Java XML.

3. J2ME (Java 2 Micro Edition)

Kategori ini digunakan untuk menjalankan dan mengembangkan aplikasi-aplikasi Java pada *handled device* atau perangkat-perangkat semacam *handphone*, Palm, PDA, dan *Pocket PC*.

Pada Tugas Akhir ini menggunakan bahasa java J2ME yang dirancang untuk dapat menjalankan program Java pada perangkat-perangkat semacam handphone dan PDA, yang memiliki karakteristik yang berbeda dengan sebuah komputer biasa.

## 2.6 Android SDK

Untuk dapat mengembangkan suatu produk aplikasi dengan berbasis sistem operasi Android. maka dibutuhkan Android SDK sebagai tools (alat-alat) yang dibutuhkan untuk menunjang Eclipse yang digunakan. Android SDK tersedia

secara free (bebas di gunakan dan gratis) dan menunjang berbagai platform (jenis sistem operasi) Android dari versi pertama kali diperkenalkan hingga sekarang. Dengan adanya Android SDK dalam pembuatan projek android dapat dipermudah. *Programer developer* dapat mengkompile hasil dari projek nya langsung dengan virtual emulator android, yang sudah disediakan dalam tools Android SDK tersebut ( Developer.android.com, 2012 ) Dengan cara tersebut maka developer dapat dengan cepat mengetahui hasil dari program yang dibuatnya berjalan lancar pada platform. Sebagai platform aplikasi netral, Android dapat membuat aplikasi yang kita butuhkan yang bukan merupakan aplikasi bawaan *Handphone/Smartphone*. Beberapa fitur-fitur Android yang paling penting adalah:

- *Framework* Aplikasi yang mendukung penggantian komponen dan *reusable*.
- Mesin *Virtual Dalvik* dioptimalkan untuk perangkat *mobile*.
- *Integrated browser* berdasarkan *engine open sourceWebKit*.

## 2.7 Data Mining

Tan (2006) mendefinisikan *data mining* sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Istilah *data mining* kadang disebut juga *knowledge discovery*. Selanjutnya perbedaan *data mining* dan *data warehouse* adalah *data mining* merupakan bidang yang sepenuhnya menggunakan apa yang sepenuhnya digunakan oleh *data warehouse*, bersama dengan bidang yang menangani masalah pelaporan dan manajemen data. Sementara *data warehouse* sendiri bertugas untuk menarik/meng-*query* data dari basis data mentah untuk memberikan hasil data yang nantinya digunakan oleh bidang yang menangani manajemen, pelaporan, dan *data mining*. Dengan *data mining* inilah, penggalian informasi baru dapat dilakukan dengan bekal data mentah yang diberikan oleh *data warehouse*. Hasil yang diberikan oleh ketiga bidang tersebut berguna untuk mendukung aktivitas bisnis cerdas (*business intellegent*).

Pekerjaan yang berkaitan dengan *data mining* dibagi menjadi 4 (empat) kelompok, yaitu (Prasetyo, E, 2012) :

1. *Prediction Modelling* (model prediksi), berkaitan dengan sebuah model yang dapat melakukan pemetaan dari setiap himpunan variabel ke setiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan baru yang didapat.
2. *Cluster Analysis* (analisis kelompok), pengelompokan data-data kedalam sejumlah kelompok (*cluster*) berdasarkan kesamaan karakteristik masing-masing data pada kelompok-kelompok yang ada.
3. *Association Analysis* (analisis asosiasi), digunakan untuk menemukan pola yang menggambarkan kekuatan fitur dalam data dengan tujuan untuk menemukan pola yang menarik dengan cara yang efisien.
4. *Anomaly Detection* (deteksi anomali), berkaitan dengan pengamatan sebuah data dari sejumlah data yang secara signifikan mempunyai karakteristik yang berbeda dari sisa data yang lain.

### 2.7.1 Tugas Utama Data Mining

Secara umum *data mining* mempunyai tugas utama :

#### I. Klasifikasi (*Classification*)

Klasifikasi bertujuan untuk mengklasifikasikan *item* data menjadi satu dari beberapa kelas standart. Sebagai contoh, suatu program email dapat mengklasifikasikan email yang sah dengan email spam. Beberapa algoritma klasifikasi antara lain *pohon keputusan*, *nearest neighbor*, *naïve bayes*, *neural networks*, dan *support vector machines*

#### II. Regresi (*Regression*)

Regresi merupakan permodelan dan investigasi hubungan dua atau lebih variabel. Dalam analisis regresi ada satu lebih variabel independent / prediktor yang biasa diwakili dengan notasi dengan notasi  $x$  dan satu variabel respon yang biasa diwakili dengan notasi  $y$  (Santosa, 2007).

#### III. Pengelompokan (*Clustering*)

*Clustering* merupakan metode pengelompokan sejumlah data ke dalam kluster(*group*) sehingga dalam setiap kluster berisi data yang semirip mungkin.

#### IV. Pembelajaran Aturan Asosiasi (*Association Rule Learning*)

Pembelajaran aturan asosiasi mencari hubungan antara variabel. Sebagai contoh suatu toko mengumpulkan data kebiasaan pelanggan dalam berbelanja. Dengan menggunakan pembelajaran atur asosiasi, toko tersebut dapat menentukan produk yang sering dibeli bersamaan dan menggunakan informasi ini untuk tujuan pemasaran.

### 2.8 Klasifikasi (*Classification*)

Salah satu tugas utama *data mining* adalah klasifikasi. Klasifikasi digunakan untuk menempatkan bagian yang tidak diketahui pada data ke dalam kelompok yang sudah diketahui. Klasifikasi menggunakan variabel target dengan nilai nominal. Dalam satu set pelatihan, variabel target sudah diketahui. Dengan pembelajaran dapat ditemukan hubungan antara fitur dengan variabel target. Ada dua langkah dalam proses klasifikasi (Han and Kember, 2006):

1. Pembelajaran (*learning*): pelatihan data dianalisis oleh algoritma klasifikasi.
2. Klasifikasi: data yang diujikan digunakan untuk mengakulasi akurasi dari aturan klasifikasi. Jika akurasi dianggap dapat diterima, aturan dapat diterapkan pada klasifikasi data tuple yang baru.

### 2.9 Teori Bayes

Menurut(Prasetyo,E.2012). Bayes merupakan teknik prediksi probalistik sederhana yang berdasar pada penerapan teorema Bayes (atau aturan bayes)

dengan asumsi independensi (ketidak tergantungan) yang kuat (naif). Dengan kata lain, dalam *Naïve Bayes*, model yang di gunakan adalah "model fitur independen"

Dalam Bayes (terutama *Naïve Bayes*), maksud independensi yang kuat pada fitur adalah bahwa sebuah fitur pada sebuah data tidak berkaitan dengan ada atau tidak adanya fitur lain dalam data yang sama. Contohnya, pada kasus klasifikasi hewan dengan fitur penutup kulit, melahirkan, berat dan menyusui. Disini ada ketergantungan pada fitur menyusui karena hewan yang menyusui biasanya melahirkan, atau hewan bertelur tidak menyusui. Dalam Bayes, hal tersebut tidak dipandang sehingga masing-masing fitur seolah tidak memiliki hubungan apapun.

Prediksi Bayes didasarkan pada teorema Bayes dengan rumus pada 2.1 :

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)} \dots\dots\dots(2.1)$$

Keterangan :

- P(H | E) = Probabilitas akhir bersyarat (conditional probability) suatu hipotesis H terjadi jika diberikan bukti (evidence) E terjadi.
- P(E| H) = Probabilitas sebuah bukti E terjadi akan mempengaruhi hipotesis H .
- P(H) = Probabilitas awal hipotesis H terjadi tanpa memandang bukti apapun.
- P(E) = Probabilitas awal bukti E terjadi tanpa memandang hipotesis atau bukti yang lain.

Ide dasar dari aturan Bayes adalah bahwa hasil dari hipotesis atau peristiwa (H) dapat diperkirakan pada beberapa bukti (E) yang diamati. Ada beberapa hal penting dari aturan Bayes yaitu :

1. Sebuah probabilitas awal/priori H atau PH adalah probabilitas dari suatu hipotesis sebelum bukti diamati.
2. Sebuah probabilitas akhir H atau (PH) adalah probabilitas dari suatu hipotesis setelah bukti diamati.

Teorema Bayes juga bisa menangani beberapa bukti, misalnya ada  $E_1$ ,  $E_2$ , dan  $E_3$  sehingga akhir untuk hipotesis (H) dapat dihitung dengan rumus pada 2.2 :

$$P(H | E_1, E_2, E_3) = \frac{P(E_1, E_2, E_3 | H) * P(H)}{P(E_1, E_2, E_3)} \dots \dots \dots (2.2)$$

Karena yang digunakan untuk bukti adalah independen, bentuk di atas dapat diubah dengan rumus pada 2.3 :

$$P(H | E_1, E_2, E_3) = \frac{P(E_1 | H) * P(E_2 | H) * P(E_3 | H) * P(H)}{P(E_1) * P(E_2) * P(E_3)} \dots \dots \dots (2.3)$$

## 2.10 Metode Naive Bayes untuk klasifikasi

(Prasetyo, E. 2012). Kaitan antara Naïve bayes dengan klasifikasi, korelasi hipotesis, dan bukti dengan klasifikasi adalah bahwa hipotesis dalam teorema Bayes merupakan label kelas yang menjadi target pertama dalam klasifikasi, sehingga bukti- bukti merupakan fitur-fitur yang menjadi masukan dalam model klasifikasi. Formulasi Naïve Bayes untuk klasifikasi dengan rumus pada 2.4 :

$$P(Y | X) = \frac{P(Y) \prod_{i=1}^q P(X_i | Y)}{P(X)} \dots \dots \dots (2.4)$$

Keterangan :

$P(Y | X)$  = probabilitas data dengan vektor X pada kelas Y

$P(Y)$  = probabilitas awal kelas Y

$\prod_{i=1}^q P(X_i | Y)$  = probabilitas independen kelas Y dari semua fitur dalam vector X

Karena  $P(X)$  selalu tetap, sehingga dalam perhitungan prediksi nantinya cukup hanya dengan menghitung  $P(Y) \prod_{i=1}^q P(X_i | Y)$  dengan memilih yang terbesar sebagai kelas yang di pilih sebagai hasil prediksi. Sementara probabilitas

independen  $\prod_{i=1}^q P(X_i|Y)$  tersebut merupakan pengaruh semua fitur dari data terhadap setiap kelas  $Y$ , yang dinotasikan dengan rumus pada 2.5 :

$$P(X|Y=y) = \prod_{i=1}^q P(X_i|Y=y) \dots \dots \dots (2.5)$$

Setiap fitur  $X = \{X_1, X_2, X_3, \dots, X_q\}$  terdiri atas  $q$  atribut ( $q$  dimensi).

Umumnya, metode Naïve Bayes ini mudah dihitung untuk fitur bertipe kategoris. Namun untuk tipe numerik (kontinu), ada perlakuan khusus sebelum dimasukkan dalam Naïve Bayes, yaitu :

1. Melakukan diskretisasi pada setiap fitur kontinu dan mengganti nilai fitur kontinu tersebut dengan nilai interval diskret. Pendekatan ini dilakukan dengan mentransformasi fitur kontinu kedalam fitur ordinal.
2. Dari distribusi probabilitas diasumsikan bentuk tertentu untuk fitur kontinu dan memperkirakan parameter distribusi dengan data peralihan. Distribusi Gaussian biasanya dipilih untuk mempresentasikan probabilitas bersyarat dari fitur kontinu pada sebuah kelas  $P(X_i|Y)$ , sedangkan distribusi Gaussian dikarakteristikan dengan dua parameter. mean  $\mu$ , dan varian,  $\sigma^2$ . Untuk setiap kelas  $y_j$ , probabilitas bersyarat kelas  $y_j$  untuk fitur  $X_i$  dengan rumus pada 2.6 :

$$P(X_i = x_i|Y = y_j) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp \left\{ -\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2} \right\} \dots \dots \dots (2.6)$$

Keterangan :

$\mu_{ij}$  : mean sampel  $X_i$  ( $\bar{X}$ ) dari semua data latih yang menjadi milik kelas  $y_j$

$\sigma_{ij}^2$  : varian sampel ( $S^2$ ) dari data latih.

## 2.11 Penelitian Sebelumnya

Naïve Bayes merupakan metode populer yang banyak digunakan untuk klasifikasi. Berbagai riset yang telah dilakukan berkaitan dengan kasus prediksi yang menggunakan metode Naïve Bayes, antara lain :

Penelitian yang berjudul “*Sistem Prediksi Prestasi (IPK) Berdasarkan Latar belakang Sekolah Asal dan Atribut Mahasiswa Ketika Masuk Kuliah Menggunakan Naïve Bayes*” oleh Meinggian Vilian Sari. Adapun Data yang diolah

dalam penelitian ini adalah data mahasiswa Teknik Informatika Universitas Muhammadiyah Gresik angkatan 2010 semester 6 sejumlah 103 mahasiswa. Atribut yang digunakan adalah IPK dan dibagi menjadi 2 kelas, yaitu IPK Tinggi dan Rendah. Hasil penelitian menunjukkan bahwa akurasi tertinggi yang didapatkan pada pengujian sistem ini adalah 84.62%.

Selain itu, penelitian yang terkait mengenai klasifikasi adalah penelitian yang dilakukan oleh Ervina Rizka Anandita dengan judul "*Klasifikasi Tebu dengan Menggunakan Algoritma Naïve Bayes Classification pada Dinas Kehutanan dan Perkebunan Pati*". Data yang digunakan pada penelitian ini adalah data dari Dinas Kehutanan dan Perkebunan Kabupaten Pati, dengan atribut hasil panen berdasarkan jenis tebu yang dimiliki oleh perkebunantebu yang meliputi jenis tebu, hasil produksi, umurpanen, tinggi tanaman, diameter batang, daerah tanam, bobot batang, rendeman dan macam got yang digunakan. Berdasarkan atribut tersebut, klasifikasi tebu dibagi menjadi 2 kelas yaitu jenis tebu yang produktif atau tidak produktif dengan akurasi tertinggi sebesar 73,3%.

Penelitian dilakukan oleh Mohammad Zoqi Sarwani, Magister Ilmu Komputer, Universitas Brawijaya. Dalam acara Seminar Nasional Sistem Informasi Indonesia (SESINDO) di Institut Teknologi Sepuluh Nopember (ITS) Surabaya Tahun 2015. Penelitiannya berjudul : "*Analisis Twitter untuk Mengetahui Karakter Seseorang Menggunakan Algoritma Naïve Bayes Classifier*". Dalam penelitian ini dilakukan analisis terhadap *tweets* atau *posting* kata yang terdapat pada Twitter yang dijadikan sebagai atribut. Kelas kepribadian yang digunakan berdasarkan konsep MBTI (Myer-Briggs Type Indicator), dari 16 indikator yang kemudian dikelompokkan menjadi 4 kelas tipe kepribadian yaitu : "Guardian", "Artisan", "Idealis" dan "Rasional". Proses klasifikasi menggunakan algoritma Naïve Bayes Classifier. Pada uji coba yang dilakukan dalam penelitian ini, peneliti menggunakan 10 pengguna Twitter sebagai data latih dan 10 pengguna Twitter sebagai data uji. Hasil yang diperoleh tersebut kemudian dibandingkan dengan data dari pakar, sehingga didapat keakurasian data mencapai 100%.