

Contoh 1: *Business Intelligence* yang mendukung *Customer Relationship Management* dan *Enterprise System* dengan teknologi Web 2.0 dan *E-Commerce*.

I. LANDASAN TEORI

4.1. *Business Intelligence*

BI bukanlah sebuah produk atau sistem. BI merupakan arsitektur dan koleksi operasional serta aplikasi pendukung keputusan yang terintegrasi dengan database yang menyediakan akses mudah ke data operasional bisnis dan organisasi. BI dan prosesnya berdasar kepada sebuah roadmap sebagai panduan untuk mengembangkan aplikasi yang mendukung keputusan dengan menggunakan data yang terstruktur. BI membahas aplikasi untuk pendukung keputusan dan database. Aplikasi yang diimplementasikan berdasarkan roadmap BI disesuaikan dengan kebutuhan bisnis atau organisasi (Moss, 2003). Dalam pengembangan roadmap BI dibutuhkan informasi dan aktivitas sebagai berikut:

1. *Multidimensional analysis*, seperti *online analytical processing* (OLAP)
2. *Click-stream analysis*
3. *Data mining*
4. *Forecasting*
5. *Business analysis*
6. *Balanced scorecard preparation*
7. *Visualization*
8. *Querying, reporting, dan charting*
9. *Geospatial analysis*
10. *Knowledge management*
11. *Enterprise portal implementation*
12. *Mining for text, content, and voice*
13. *Digital dashboard access*
14. *Other cross-functional activities*

Beberapa contoh dari BI yang mendukung untuk pengambilan keputusan antara lain adalah sebagai berikut:

1. *Enterprise-wide data warehouses*
2. *Data marts (functional dan departmental)*

3. *Exploration warehouses (statistical)*
4. *Data mining*
5. *Web warehouses (click-stream data)*
6. *Operational data stores (ODSs)*
7. *Operational marts (oper marts)*
8. *Other cross-functional decision-support databases*

(Moss, 2003).

Penerapan sistem BI dipengaruhi oleh meningkatnya jumlah bisnis dan organisasi di berbagai sektor. Meskipun peningkatan jumlah solusi dan pengalaman BI selalu bertambah dalam banyak proyek implementasi di seluruh dunia, sistem BI masih harus terus ditingkatkan penggunaannya dikarenakan implementasi BI sering tidak berhasil atau tidak sepenuhnya mencapai target yang diharapkan. Analisis kasus dari data empiris yang memberikan contoh dan menilai sejauh mana berbagai atribut organisasi dan sifat pengguna mempengaruhi keberhasilan atau kegagalan pelaksanaan proyek BI sangat dibutuhkan untuk rencana bisnis dan *forecast*. Ketika perusahaan gagal untuk menyelesaikan implementasi sistem, diperlukan sebuah kerangka analisis di mana sistem BI didefinisikan dan diklasifikasikan secara khusus. Berdasarkan klasifikasi khusus, diharapkan mengurangi kegagalan pasca implementasi (Yeoh, 2010).

4.2. Enterprise System

Perkembangan sistem informasi yang berkembang pesat sangat berdampak pada strategi dan kebijakan dalam suatu organisasi (*enterprise*) baik yang berorientasi pada laba (*profit oriented*) ataupun lembaga yang selain profit oriented juga mempunyai misi sosial (*public services*) (Wheelen and Hunger, 2008).

Sifat global dan kompetitif di lingkungan bisnis telah memaksa usaha kecil dan menengah (UKM) agar lebih kompetitif, efisien dan responsif terhadap kebutuhan pelanggan mereka. Untuk mempertahankan atau meningkatkan daya saing dalam lingkungan bisnis yang dinamis saat ini ada baiknya UKM menerapkan sistem *enterprise resource planning* (ERP). Namun, keberhasilan pelaksanaan dan pemanfaatan sistem ERP memerlukan perhatian yang memadai untuk arsitektur *enterprise* (EA). Sistem ERP umumnya bersifat kompleks. Untuk sepenuhnya menyadari manfaat strategis implementasi ERP dan untuk merampingkan proses untuk efisiensi, EA harus digunakan sebagai kerangka kerja yang mengatur untuk

implementasi sistem ERP dan pemanfaatan. Penekanan pada EA ketika menerapkan ERP akan membutuhkan cara pendekatan yang sama sekali baru terlebih untuk akuisisi teknologi dan pemanfaatannya (Aarabi et al., 2011).

Metodologi pembangunan model arsitektur yang sering digunakan adalah Enterprise Architecture Planning (EAP), dimana tahapan pembangunannya adalah sebagai berikut (lihat gambar 4.1):

1. Inisiasi perencanaan/*Planning Initiation*

Tahapan ini bertujuan untuk mengidentifikasi dan mendokumentasi terhadap ruang lingkup, sasaran, visi, misi dan pemilihan metodologi dan penentuan alat-alat pemodelan yang akan digunakan, presentasi, dan rencana kerja.

2. Pemodelan Bisnis/*Business Modeling*

Tahapan ini bertujuan untuk mengidentifikasi struktur organisasi, model awal fungsi bisnis.

3. Survei Enterprise/*Enterprise Survey*

Tahapan ini berisi tentang model lengkap dari fungsi bisnis organisasi.

4. Sistem dan Teknologi Saat ini/*Current Systems and Technology*

Tahapan ini berisi tentang katalog sumber daya informasi (*Information Resource Catalog/IRC*) dan skema dari sistem dan teknologi yang telah ada.

5. Arsitektur Data/*Data Architecture*

Tahapan ini mengenai definisi entitas, model E-R diagram, matriks entitas fungsi dan dokumen arsitektur data.

6. Arsitektur Aplikasi/*Applications Architecture*

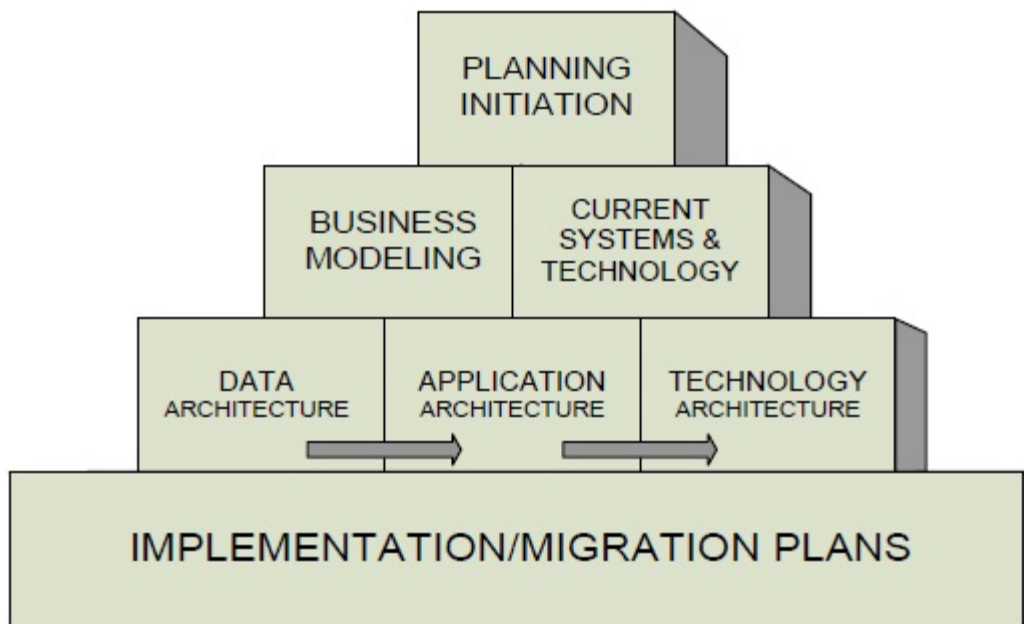
Tahapan ini mengenai definisi aplikasi, model proses bisnis, matriks aplikasi, analisis dampak, dokumen arsitektur aplikasi.

7. Arsitektur Teknologi/*Technology Architecture*

Tahapan ini mengenai distribusi data/aplikasi, dokumen arsitektur teknologi.

8. *Implementation Plan*

Tahapan ini berisi tentang urutan implementasi dan faktor-faktor sukses.
(Efraim and Linda, 2010)



Gambar 4.1 Komponen EAP

Untuk membantu dalam proses identifikasi dan dokumentasi setiap fungsi bisnis beserta lingkungan pendukungnya, *framework* atau *tools* yang biasanya digunakan adalah sebagai berikut:

1. Matriks Portofolio Aplikasi (*Applications Portofolio*) dari John Ward dan Joe Peppard, untuk melakukan estimasi semua sistem aplikasi yang ada di perusahaan atau organisasi.
2. Rantai Nilai (*Value Chain*) dari Michael E. Porter untuk membantu proses identifikasi dan definisi dari aktifitas bisnis utama di Universitas Respati Yogyakarta.
3. *Four-Stage Life Cycle* dari pendekatan Business System Planning (BSP) *IBM Corporation* untuk mendapatkan gambaran fungsi bisnis utama yang ada di Universitas Respati Yogyakarta.
4. *Work System Framework* dari Steven Alter untuk membantu mendokumentasikan sistem kerja di lingkungan organisasi Universitas Respati Yogyakarta.
5. *Business Process Management Notation* (BPMN) dari BPMI (*Business Process Management Initiative*), yang akan digunakan untuk memodelkan skema proses bisnis dari setiap bagian yang terlibat.

Tujuan utama perencanaan strategis informasi adalah mempersiapkan rencana bagi pengelolaan analisis, perancangan dan pengembangan sistem berbasis komputer. Dalam metodologi rekayasa informasi, tiap langkah dapat dilihat dari dua sisi,

yaitu data dan aktivitas. Untuk perencanaan strategi informasi di sisi data, arah tinjauan strategisnya adalah terhadap kebutuhan informasi yang dibutuhkan oleh enterprise. Sedangkan di sisi aktivitas, arah tinjauan strategisnya adalah dalam hal pemanfaatan teknologi untuk peningkatan kinerja enterprise (Yunis, 2007).

4.3. *E-Commerce, Web 2.0 dan Customer Relationship Management*

Hoffman (2008) yang dikutip oleh Jones (2011) mengatakan pada pertengahan tahun 1990-an para pengunjung website (*World Wide Web*) sangat bersemangat dikarenakan pada era ini sebuah situs sudah menyediakan *interface* sederhana dengan gambar yang interaktif dan berwarna. Sebelumnya, sebuah website hanya berbasis teks tanpa gambar. Pada pertengahan tahun 1990-an website yang mendukung *E-Commerce* sangat terbatas mulai dari fasilitas sampai dengan tampilan. Pada masa ini belum ada hal yang interaktif antara website dengan *user/customer*. Ini adalah era teknologi Web 1.0 dimana yang ada hanya teks dan gambar statis, dan ini adalah awal mula dari konten dinamis dan lebih interaktif (Jones, 2008).

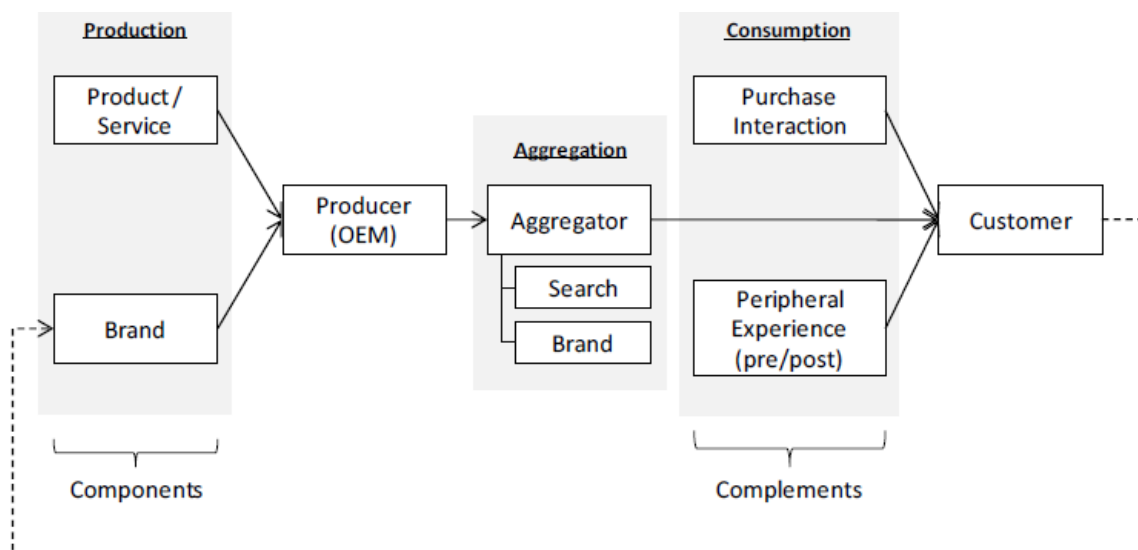
E-commerce dalam lebih dari 10 (sepuluh) tahun kemudian mulai melayani bisnis dengan komoditi besar dan antar negara serta antar benua dengan fasilitas yang lebih maju dalam dunia teknologi. Ini tidak hanya berbasis website namun sudah mendukung teknologi *mobile* dan perangkat lainnya dengan memanfaatkan teknologi *cloud* sebagai mediasi dan komunikasi data. Semua perkembangan dan inovasi teknologi ini seiring dengan perkembangan teknologi Web 2.0 dan disebut sebagai sebuah sistem operasi web, dimana "Web memfasilitasi kegiatan saling berbagi dan partisipasi dengan koneksi meskipun berbeda aplikasi" tanpa dibatasi oleh pemetaan geografis dan waktu (Allen, 2008).

Dalam teknologi Web 2.0, individu mengambil alih pengalaman Web mereka dengan melakukan kegiatan "mash-up" yang menggabungkan teks, grafik, audio, video dan animasi dari berbagai sumber. User memiliki peluang untuk memposting hasil dari YouTube, yang memungkinkan sebagian besar orang dengan koneksi internet dan kreativitas yang mereka miliki. Dengan latar belakang Web 2.0 ditambah dengan implementasi kecerdasan buatan akan menambah kemudahan bagi para pelaku bisnis untuk mengontrol dan melakukan aktivitas bisnis secara virtual.

Sumber daya yang didedikasikan pada teknologi *E-Commerce* diperlukan untuk mendukung Web 2.0 pada penelitian ini. Organisasi harus menyesuaikan diri untuk mendukung saluran *E-Commerce* di seluruh proses penjualan & pemasaran,

produksi dan dukungan pelanggan dalam rangka untuk menyebarkan dan mengambil keuntungan penuh dari lingkungan Web 2.0. Perubahan ini memerlukan adaptasi dan secara signifikan dalam hal bagaimana perusahaan mengembangkan dan mengkomunikasikan merek mereka. Bagi banyak bisnis dan organisasi yang menggunakan satu arah pemasaran dan strategi merek, memanfaatkan sifat komunikasi dua arah yang terdapat pada teknologi Web 2.0 dapat secara signifikan memberi perubahan yang positif terhadap kegiatan bisnis. Sumber daya, desain organisasi dan hambatan yang hadir dalam suatu perusahaan akan berdampak terhadap teknologi baru yang akan diimplementasikan.

Website *Agregator* menyediakan teknologi yang menyerap dan mengumpulkan isi sindikasi web yang dipasok dalam format “*web feed*” (RSS, Atom, dan format XML lainnya). Pasokannya berasal dari weblog, podcast, vlog, dan situs web media utama. Dengan teknologi aggregator, user bisa melihat posting dan konten dengan mudah di satu tempat saja. Perspektif *E-Commerce* dan teknologi Agreggator yang terdapat pada teknologi Web 2.0 dapat dilihat pada gambar 4.2.



Gambar 4.2 *Aggregators Interaction Map* (Lewis, 2011)

Ada beberapa prinsip dasar yang harus dipahami dalam melakukan implementasi CRM yang efektif, yaitu:

1. Sebuah produk atau jasa memiliki durasi waktu tertentu, baik dari segi pembuatan ataupun umur ekonomis. Maka ada baiknya jika perusahaan dapat

mengerti kebutuhan pelanggan di masa mendatang. Antisipasi kebutuhan lebih awal tersebut dapat “meniadakan” jeda waktu penciptaan produk atau jasa tersebut, sehingga di mata pelanggan, perusahaan “seolah-olah” dapat memenuhi permintaan mereka secara sangat cepat.

2. Perusahaan harus mengerti sungguh-sungguh bahwa informasi merupakan asset strategis yang tidak ternilai harganya. Jenis informasi yang paling berharga adalah yang menyangkut transaksi dan profil pelanggan sehingga perusahaan dapat mengerti perilaku dari masing-masing pelanggan yang ada.
3. Walaupun perusahaan memiliki ribuan bahkan jutaan pelanggan, setiap pelanggan harus diperhatikan sebagai seorang individu yang unik, dalam arti kata perusahaan harus dapat membangun hubungan atau relasi one-to-one dengan pelanggannya. Hanya dengan cara inilah pelanggan akan merasa diperhatikan dan diperlakukan spesial, sehingga yang bersangkutan akan menjadi loyal kepada perusahaan.
4. Banyak perusahaan yang menjual lebih dari satu produk atau jasa, namun perusahaan harus memiliki manajemen satu pintu (single point of contact) di mata pelanggan. Manajemen satu pintu ini bertujuan selain untuk tidak membuat bingung pelanggan, secara tidak langsung mewakili identitas perusahaan.
5. Pada akhirnya, dari pelanggan yang ada, volume dan frekuensi perdagangan yang diperoleh dari mereka harus selalu ditingkatkan. Cara perusahaan untuk dapat melakukan hal itu tanpa adanya unsur pemaksaan yang dirasakan pelanggan adalah dengan mempelajari seluruh rekaman transaksi pelanggan di masa lalu. Semakin lengkap dan detail rekaman transaksi yang dimiliki, akan semakin memperjelas kebutuhan dan perilaku pelanggan.
6. Perusahaan harus dapat menjangkau pelanggan di manapun dan kapanpun pelanggan berada, tanpa memperhatikan batas-batas waktu dan ruang. Dengan kata lain, perusahaan harus memiliki kanal distribusi (access channels) yang beragam, baik yang bersifat on-line maupun off-line.

Teknologi CRM yang dipergunakan lebih beragam dan bervariasi. Aspek-aspek pemilihan teknologi yang harus diperhatikan sehubungan dengan proses membangun relasi dengan pelanggan. Beberapa kriteria dapat kita simpulkan sebagai berikut (Lewis, 2011):

1. Kriteria pelanggan dapat bervariasi, seperti individu, keluarga, institusi, sampai dengan sebuah komunitas.
2. Bisnis atau organisasi dapat melakukan hubungan dengan pelanggan dari berbagai tempat yang diinginkan, apakah itu di rumah, tempat kerja, kendaraan, tempat publik, dan lain sebagainya.
3. Berbagai jenis peralatan yang dapat dipergunakan untuk melakukan hubungan antara bisnis dan pelanggan, seperti telepon, faksimili, komputer, televisi, PDA (Personal Digital Assistant), kios, dan lain sebagainya.

Contoh 1: Pengembangan Model *Forecasting* dan *Classification* sebagai Penentu Tingkat Kerawanan Wilayah Terhadap Penyakit Demam Berdarah *Dengue* (DBD) dengan Visualisasi Model Spasial

LANDASAN TEORI

II.1. Demam Berdarah Dengue (DBD)

DBD merupakan penyakit infeksi yang disebabkan oleh virus *dengue* dengan gejala sakit demam, nyeri otot, dan atau nyeri sendi yang disertai penurunan dari sel darah putih, adanya bercak kemerahan di kulit, pembesaran kelenjar getah bening, penurunan jumlah trombosit dan kondisi terberat adalah perdarahan dari hampir seluruh jaringan tubuh.

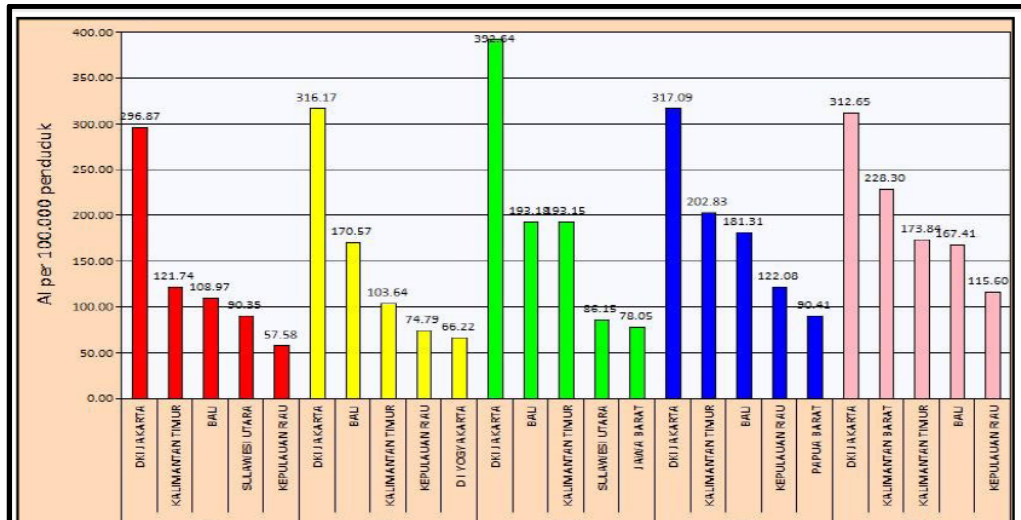
Beberapa faktor diketahui berkaitan dengan peningkatan transmisi virus *dengue* yaitu :

1. Vektor (nyamuk), terutama berhubungan dengan sanitasi lingkungan.
2. Penjamu (manusia) penderita di lingkungan.
3. Lingkungan : curah hujan, suhu, sanitasi dan kepadatan penduduk (Anonim, 2009).

Persebaran Kasus

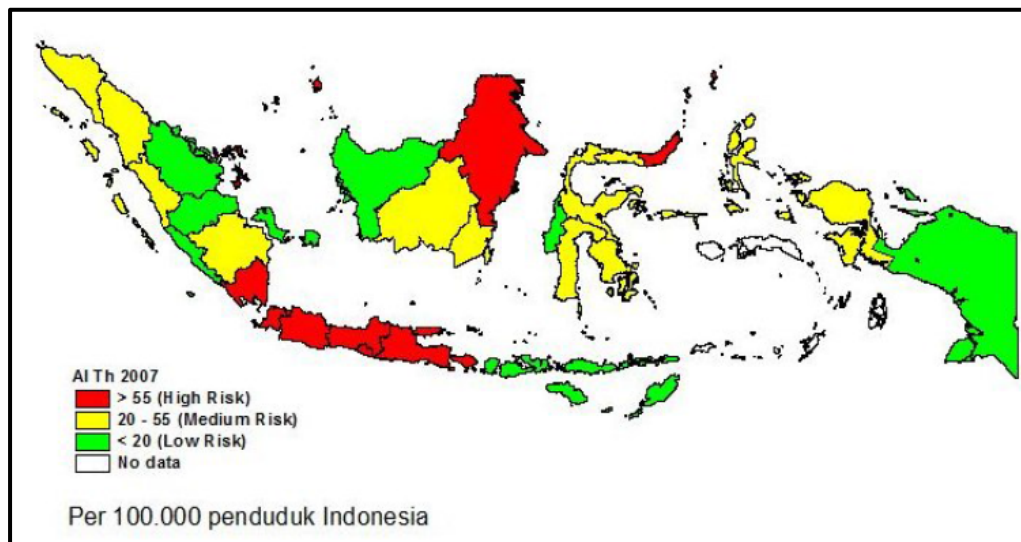
Di Indonesia DBD telah menjadi masalah kesehatan masyarakat selama 41 tahun terakhir. Sejak tahun 1968 telah terjadi peningkatan persebaran jumlah propinsi dan kabupaten/kota yang endemis DBD, dari 2 provinsi dan 2 kota, menjadi 32 (97%) dan 382 (77%) kabupaten/kota pada tahun 2009. Provinsi Maluku, dari tahun 2002 sampai tahun 2009 tidak ada laporan kasus DBD. Selain itu terjadi juga peningkatan jumlah kasus DBD, pada tahun 1968 hanya 58 kasus menjadi 158.912 kasus pada tahun 2009. (Buletin Epidemiologi, 2010).

Daerah Istimewa Yogyakarta merupakan salah satu propinsi yang memiliki angka kejadian tertinggi (*inciden rate*) terhadap penyakit DBD. Hal ini dibuktikan dengan perbandingan data sebagai berikut :



(Sumber : Ditjen PP & PL Tahun 2013)

Gambar 2.1. Lima propinsi tertinggi Angka Insiden DBD per 100.000 Penduduk di Indonesia Tahun 2009-2013



Gambar 2.2. Persebaran Insiden DBD per 100.000 Penduduk Menurut Provinsi di Indonesia Tahun 2009-2013

Dengan melihat jumlah kejadian penyakit DBD dari periode waktu tertentu, tampak telah terjadi perubahan pola penyakit DBD, dimana dahulu DBD adalah penyakit pada anak-anak dibawah 15 tahun, saat ini telah menyerang

seluruh kelompok umur, bahkan lebih banyak pada usia produktif. Perlu penelitian lebih lanjut hal yang mempengaruhinya, apakah karena virus yang semakin virulen (ganas) atau karena pengaruh lain.

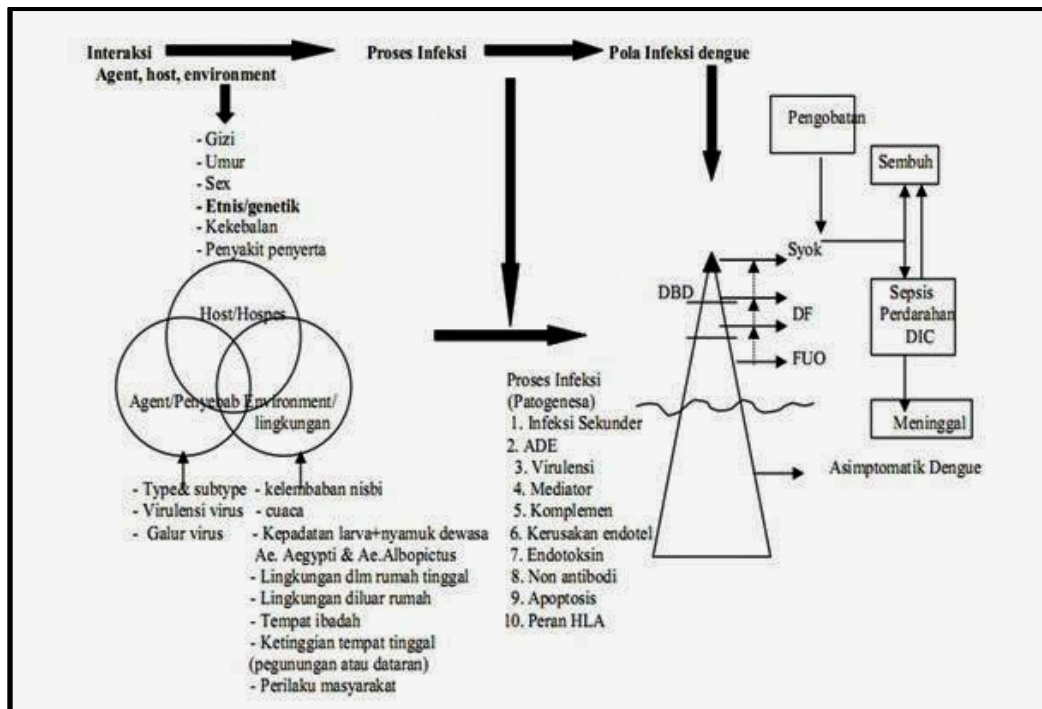
II.2. Epidemiologi

Epidemiologi merupakan salah satu bagian dari pengetahuan Ilmu Kesehatan Masyarakat (*public health*) yang menekankan perhatiannya terhadap keberadaan penyakit dan masalah kesehatan lainnya dalam masyarakat. Keberadaan penyakit masyarakat itu didekati oleh epidemiologi secara kuantitatif. Karena itu epidemiologi akan mewujudkan dirinya sebagai suatu metode pendekatan yang banyak memberikan perlakuan kuantitatif dalam menjelaskan masalah kesehatan. (Suhardiono, 2005)

Epidemiologi mempelajari tentang distribusi, frekuensi dan determinan penyakit pada populasi. Berikut definisi dari distribusi, frekuensi dan determinan faktor resiko :

1. Distribusi : orang, tempat atau waktu.
2. Frekuensi, ukuran frekuensi : Insiden dan atau prevalen.
3. Determinan faktor resiko : faktor yang mempengaruhi atau faktor yang memberi resiko atas terjadinya penyakit atau masalah kesehatan.

Epidemiologi mengukur kejadian dan distribusi kejadian tersebut menurut variabel orang, tempat, waktu dan berupaya menentukan faktor yang menyebabkan terjadinya kejadian itu di kelompok populasi. Ilmu ini dikembangkan dari pengalaman mempelajari beberapa wabah penyakit pada waktu-waktu tertentu dengan angka kematian yang tinggi.



Gambar 3.1. Penilaian Resiko *Epidemiology* dan Faktor yang Mempengaruhi Penyebaran Penyakit DBD

Keterkaitan antar faktor yang berpengaruh terhadap penyebaran penyakit DBD dikaji dari disiplin ilmu kesehatan masyarakat. Epidemiologi merupakan bagian dari disiplin ilmu kesehatan masyarakat yang mempunyai peran penting terhadap pola penyebaran penyakit DBD. Kasus kejadian penyakit DBD di suatu daerah dibandingkan antara *inciden rate* (IR) dan tingkat populasi penduduk di daerah tersebut.

Untuk mengetahui faktor resiko perilaku masyarakat terhadap kejadian DBD dilakukan penelitian yang bersifat survei dengan rancangan sekat lintang (*cross-sectional*) yang bertujuan untuk memperoleh gambaran mengenai hubungan tingkat pengetahuan, sikap dan perilaku masyarakat terhadap penyakit DBD. (Suhardiono, 2005)

II.3. Basis Pengetahuan

Perencanaan pengembangan basis pengetahuan dilakukan dengan pendekatan 2 buah model, yaitu : (Hong Ren, Jian Li , Zheng-An Yuan , Jia-Yu Hu , Yan Yu and Yi-Han Lu., 2013)

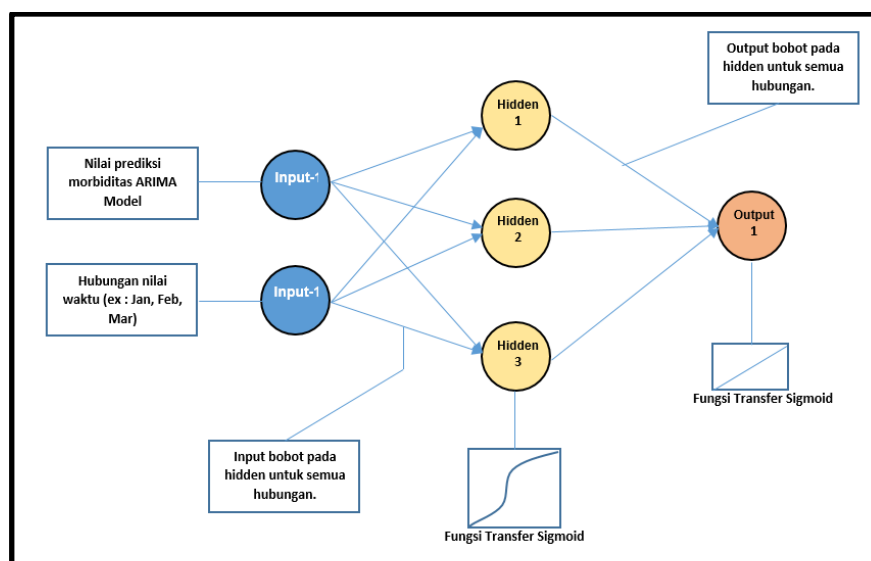
a. *Autoregressive Integrated Moving Average Model (ARIMA)*

ARIMA-BPNN terdiri dari *Autoregressive Integrated Moving Average Model* (ARIMA) dan *Back Propagation Artificial Neural Network* (BPNN). Pengembangan model dilakukan untuk meramalkan kejadian kasus penyakit DBD di wilayah Yogyakarta. Pengembangan model direncanakan menggunakan *data training* dalam kurun waktu selama 120 bulan dari angka kejadian kasus penyakit DBD dari bulan Januari 2003 sampai dengan Desember 2013, kemudian dilakukan validasi dalam kurun waktu 12 bulan dari angka kejadian kasus penyakit DBD dimulai dari bulan Januari 2014 sampai dengan Desember 2014. Kemudian pengujian model peramalan (*forecast*) kejadian DBD dilakukan pada bulan Januari 2016 sampai dengan Desember 2016 di Yogyakarta. Penggunaan pendekatan model proses ini dibagi menjadi 3 tahapan yaitu :

1. Langkah pertama yaitu menetapkan model ARIMA sebagai model yang terbaik dan sesuai serta untuk melakukan prediksi nilai (*value*) dari setiap rangkaian waktu. Model proses dibangun dan dirancang untuk mendapatkan keuntungan asosiasi dari hubungan seasonally dan sekuensial yang biasanya berada pada pengumpulan data secara periodik. Parameter model ditentukan dengan menggunakan metode *Least Squares* bersyarat. Residual analysis, Root Mean Square Error (RMSE), normalized Bayesian Information Criterion (BIC), dan

stationary R square dilakukan untuk mendapatkan perbandingan yang sesuai antara ARIMA models.

- Langkah kedua yaitu melakukan *training* data menggunakan BPNN (*Back Propagation Neural Network*). Pada perencanaan pengembangan model arsitektur BPNN terdiri dari 3 layer. Dua neuron yang dikumpulkan untuk melakukan prediksi nilai morbiditas dari model ARIMA dan nilai (angka) korespondensi waktu sebagai layer input, 3 neuron untuk melakukan estimasi nilai morbiditas secara aktual dan mensimulasikannya pada *hidden layer*, dan 1 neuron melakukan tranfer untuk meramalkan insiden pada *output layer*. Neuron di hidden layer memiliki fungsi tangen hiperbolik dan pengalihan *sigmoid neuron*, sedangkan pada lapisan output memiliki fungsi *transfer linier*. Algoritma *Bayesian Back Propagation* digunakan untuk melatih jaringan (*train network*) dan memberikan pendekatan pemersatu untuk menangani masalah model kompleksitas dan model yang sesuai. Ilustrasi dapat dilihat pada gambar berikut :



Gambar 3.1. Kombinasi dari model ARIMA dan model BPNN.

3. Langkah ketiga adalah melakukan validasi terhadap model yang dikombinasikan selama 12 bulan dari data morbiditas bulan Januari 2015 sampai dengan Desember 2015 untuk meramalkan kejadian penyakit DBD pada tahun 2016.

Tingkat rata-rata *error (mean error rate)* (MER) digunakan untuk menjelaskan perbandingan prediksi dan nilai aktual antara model ARIMA dan kombinasi model ARIMA-BPNN pada tahun 2016.

b. Model Kombinasi ARIMA dan *Back Propagation Neural Network* (BPNN)

Untuk membangun model kombinasi ARIMA-BPNN diperlukan nilai morbiditas yang diambil dari nilai waktu terbaik model ARIMA dan digunakan sebagai input (2×131 matrix), sedangkan nilai morbiditas yang sebenarnya digunakan sebagai data target (1×131 matrik) (Gambar 3.1). Hasil nilai model nilai akan diterapkan pada tahun penelitian yaitu pada tahun 2016 dengan melihat fluktuasi per 100.000 penduduk.

II.4. Model Peramalan (*Forecast*)

Dalam penelitian ini data yang digunakan adalah data asli (*real data*) kejadian kasus penyakit DBD. (Joaquín P.O, Fátima M.H, Gerardo R.S, René S.S, Rodolfo A. P. R., Adriana M.S., 2010).

a. *Source Data*

Data kasus kematian dan kejadian karena DBD. Data diambil dari Dinas Kesehatan terkait tempat kejadian. Dari semua *record* mulai dari tahun 2003-2013

akan dilakukan klasifikasi berdasarkan parameter penyebab kematian akibat DBD dan daerah tempat kejadian.

b. Data Preprocessing

Untuk setiap daerah tempat kejadian, tingkat mortalitas per 100.000 mulai tahun 2004-2014 yang akan dihitung. Rumus yang digunakan adalah sebagai berikut :

$$Rate = \frac{\text{angka kematian}}{\text{populasi}} * 100.000 \dots\dots\dots(3.7)$$

dimana :

1. Angka kematian = angka kematian disebabkan DBD di tahun kejadian.
2. Populasi = jumlah populasi di daerah pada tahun kejadian.

c. Forecast to The Future

Banyak tantangan penting di masa depan yaitu mengenai *big data* manajemen dan analisis yang timbul dari sifat data antara lain besar, beragam, dan berkembang. Ini adalah tantangan bagi peneliti dan praktisi untuk menangani peramalan (*forecast*) selama beberapa tahun ke depan :
(Wei Fan, Albert Bifet, 2014)

1. *Analytics Architecture.*
2. *Statistical significance.*
3. *Distributed mining.*
4. *Time evolving data.*
5. *Compression.*
6. *Visualization.*

II.5. Model Classification

Klasifikasi bertujuan untuk mengelompokkan kumpulan data secara otomatis oleh mesin (komputer), hal tersebut agar dapat mengenali objek data setelah melalui pemodelan *forecast*. Klasifikasi dapat dilakukan dengan beberapa metode, antara lain : (Ahmed A.B.E.D & Elabary I.S., 2014).

1. *Support Vector Machine (SVM)*
2. *Artificial Neural Network (ANN)*
3. *Artificial Neural Fuzzy Inference System (ANFIS)*
4. *K-Nearest Neighbour*
5. *Decision Tree*

II.6. Pemodelan Data *Spatial & Autocorelation*

Autokorelasi spasial adalah korelasi antara variabel dengan dirinya sendiri berdasarkan ruang atau dapat juga diartikan suatu ukuran kemiripan dari objek di dalam suatu ruang (jarak, waktu dan wilayah). (Lembo, 2006)

Jika terdapat pola sistematis di dalam penyebaran sebuah variabel, maka terdapat autokorelasi spasial. Adanya 2 autokorelasi spasial mengindikasikan bahwa nilai atribut pada daerah tertentu terkait oleh nilai atribut tersebut pada daerah lain yang letaknya berdekatan (bertetangga).

Adapun langkah-langkah dalam memodelkan data spasial dan auto correlation adalah sebagai berikut.

- a. Memetakan kejadian penyakit DBD per bulan antar tahun dalam suatu visualisasi model spasial.

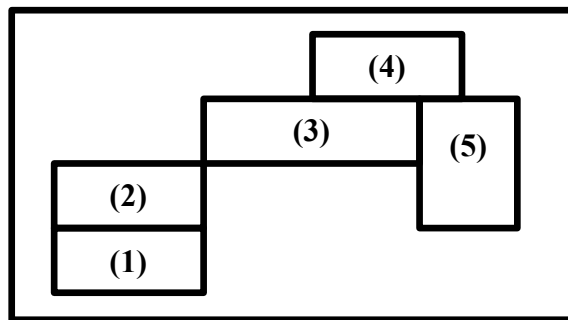
- b. Membandingkan pola yang terbentuk dari bulan ke bulan antar tahun.
- c. Menghitung ukuran dependensi spasial menggunakan indeks Moran's I.
- d. Mengidentifikasi pola sebaran kejadian berdasarkan indeks Moran's I.
- e. Menguji dependensi spasial indeks Moran's I.
- f. Menghitung ukuran dependensi spasial menggunakan indeks Geary's C.
- g. Mengidentifikasi pola sebaran kejadian berdasarkan indeks Geary's C.
- h. Menguji dependensi spasial indeks Geary's C.
- i. Membuat dan menganalisis Moran's Scatterplot.
- j. Menghitung dan menguji dependensi spasial indeks LISA.

a. Matrik Pembobot Spasial

Matriks penimbang spasial (w) adalah unsur penting dalam menggambarkan kedekatan antara suatu lokasi dengan lokasi lain dan ditentukan berdasarkan informasi atau kedekatan antara suatu lokasi dengan lokasi lain (*neighborhood*). LeSage (1999) menyajikan *queen contiguity* (persinggungan sisi-sudut) untuk mengukur kedekatan lokasi menggunakan asas persinggungan (*contiguity*) lokasi.

Matrik pembobot spasial dapat ditentukan dengan beragam metode. Salah satu metode penentuan matrik pembobot spasial yang digunakan dalam penelitian ini adalah *Queen contiguity* (persinggungan sisi-sudut). Matrik pembobot (w_{ij}) berukuran $n \times n$, dimana setiap elemen matrik menggambarkan ukuran kedekatan antara pengamatan i dan j . Gambar 1 diberikan ilustrasi mengenai perhitungan matrik pembobot menggunakan *Queen contiguity*. Ilustrasi tersebut menggunakan lima daerah sebagai pengamatannya. Elemen

matrik didefinisikan 1 untuk wilayah yang bersisian (*common side*) atau titik sudutnya (*common vertex*) bertemu dengan daerah yang menjadi perhatian, sedangkan daerah lainnya didefinisikan elemen matrik pembobot sebesar nol. Untuk daerah 3, didapatkan $w_{32}=1$, $w_{34}=1$, $w_{35}=1$ dan yang lain sama dengan nol. Matrik w_{ij} ini memiliki ukuran matrik 5x5. Ilustrasi lebih lengkapnya disajikan pada gambar 2.1.



Sumber: Lesage (1999)

Gambar 2.1. Ilustrasi *Contiguity*

Matriks pembobot yang dapat terbentuk dari Gambar 2.1 adalah sebagai berikut.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

b. Moran's I

Moran's I mengukur korelasi satu variabel misal x (x_i dan x_j) dimana $i \neq j$, $i=1,2,\dots,n$, $j=1,2,\dots,n$ dengan banyak data sebesar n , maka formula dari Moran's I adalah pada persamaan (1) (Paradis, 2010).

$$L = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{ij} (X_i - \bar{X})(X_j - \bar{X})}{S_0 \sum_{i=1}^n (X_i - \bar{X})^2} \dots\dots\dots(1)$$

x pada persamaan (1) merupakan rata-rata dari variabel x, w_{ij} merupakan elemen dari matrik pembobot, and S_0 adalah jumlahan dari elemen matrik pembobot, dimana $S_0 = \sum_i \sum_j w_{ij}$. Nilai dari indeks I ini berkisar antara -1 dan 1. Identifikasi pola menggunakan kriteria nilai indeks I , jika $I > I_0$, maka mempunyai pola mengelompok (*cluster*), jika $I = I_0$, maka berpola menyebar tidak merata (tidak ada autokorelasi), dan $I < I_0$, memiliki pola menyebar. I_0 merupakan nilai ekspektasi dari I yang dirumuskan $E(I) = I_0 = -1/(n - 1)$ (Lee dan Wong, 2001).

Pengujian hipotesis terhadap parameter I dapat dilakukan sebagai berikut.

H_0 : tidak ada autokorelasi spasial.

H_1 : terdapat autokorelasi positif (indeks Moran's I bernilai positif).

H_2 : terdapat autokorelasi negatif (indeks Moran's I bernilai negatif).

Menurut Lee dan Wong (2001) dalam Kartika (2007) statistik uji dari indeks Moran's I diturunkan dalam bentuk statistik peubah acak normal baku. Hal ini didasarkan pada teori Dalil Limit Pusat dimana untuk n yang besar dan ragam diketahui maka $Z(I)$ akan menyebar normal baku sebagai berikut.

$$Z_{hitung} = \frac{I - E(I)}{\sqrt{Var(I)}} \dots\dots\dots(2)$$

dengan I adalah indeks Moran's I , Z_{hitung} adalah nilai statistik uji indeks Moran's I , $E(I)$ adalah nilai ekspektasi indeks Moran's I , dan $Var(I)$ adalah nilai varians dari indeks Moran's I .

$$Var(I) = \frac{n\{(n^2 - 3n + 3)S_1 - nS_2 + 3S_0^2\}}{(n-1)(n-2)(n-3)S_0^2} - \frac{k\{n(n-1)S_1 - 2S_2 + 6S_0^2\}}{(n-1)(n-2)(n-3)S_0^2} - \frac{1}{(n-1)^2} \dots\dots\dots(3)$$

$$S1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (W_{ij} + W_{ji})^2, k = \sum_{i=1}^n (x_i - \bar{x})^4 / ((\sum_{i=1}^n x_i - \bar{x})^2)^2$$

$$S2 = \sum_{i=1}^n (w_i + w_i)^2, w_i = \sum_{j=1}^n W_{ij} \text{ dan } W_i = \sum_{j=1}^n W_{ji}$$

Pengujian ini akan menolak hipotesis awal jika nilai $Z_{hitung} > Z_{(\alpha)}$ (autokorelasi positif) atau $Z_{hitung} < -Z_{(\alpha)}$ (autokorelasi negatif). Positif autokorelasi spasial mengindikasikan bahwa antar lokasi pengamatan memiliki keeratan hubungan.

c. Geary's C

Indeks ini dirumuskan sebagai berikut (Lee dan Wong, 2010)

$$C = \frac{(n-1) \sum_i \sum_j w_{ij} (x_i - x_j)^2}{2S_0 \sum_i (x_i - \bar{x})^2} \quad (4)$$

Nilai w_{ij} , x_i , x_j , n , dan S_0 yang digunakan dalam persamaan (4) sama dengan nilai pada persamaan (1). Identifikasi pola sebaran menggunakan indeks C, jika nilai C terletak diantara 0 (nol) dan 1, maka pola sebarannya adalah mengelompok (*cluster*), jika nilai C mendekati 1, polanya menyebar tidak merata (tidak ada autokorelasi), dan pada nilai C terletak di antara 1 dan 2, pola sebarannya merata (Lee dan Wong, 2001).

Pengujian terhadap parameter C dapat dilakukan sebagai berikut.

H₀: tidak ada autokorelasi spasial

H₁: terdapat autokorelasi positif (indeks Geary's C bernilai kurang dari 1)

H₁: terdapat autokorelasi negatif (indeks Geary's C bernilai lebih dari 1).

Statistik Uji:

$$Z_{hitung} = \frac{C - E(C)}{\sqrt{Var(C)}} \quad \dots\dots\dots(5)$$

dengan C merupakan indeks Geary's C, Z_{hitung} merupakan nilai statistik uji indeks Geary's C, $E(C)$ merupakan nilai ekspektasi indeks Geary's C (bernilai 1), dan $Var(C)$ merupakan nilai varians dari indeks Geary's C.

$$Var(I) = \frac{(n-1)S_1[n^2-3n+3]-(n-1)b^2}{n(n-2)(n-3) S_0^2} - \frac{(n-1)S_2[n^2+3n-6-(n^2-n+2)b^2]}{4n(n-2)(n-3) S_0^2} + \frac{n^2-3-(n-1)^2b_2}{n(n-2)(n-3)} \dots\dots\dots(6)$$

S_0 , S_1 , dan S_2 pada persamaan (6) sama dengan Moran's I.

Pengujian ini akan menolak hipotesis awal jika nilai $Z_{hitung} > Z(\alpha)$ (autokorelasi positif) atau $Z_{hitung} < -Z(\alpha)$ (autokorelasi negatif). Positif autokorelasi spasial mengindikasikan bahwa antar lokasi pengamatan memiliki keeratan hubungan.