

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Pita Lebar atau *Broadband***

Pita lebar (bahasa Inggris: *broadband*) merupakan sebuah istilah dalam Internet yang merupakan koneksi Internet transmisi data kecepatan tinggi. Internet (*Interconnected network*) adalah sebuah sistem jaringan komunikasi global yang menghubungkan komputer dan jaringan komputer di seluruh dunia secara global. Fasilitas menyediakan akses ke sejumlah layanan komunikasi termasuk halaman *world wide web* (www), surat elektronik (*email*), berita, hiburan dan data. Fasilitas akses internet tidak diasumsikan hanya melalui komputer, dimungkinkan juga menggunakan telepon Seluler, PDA, perangkat game elektronik, televisi digital, dan lain-lain [16]. Internet *broadband* adalah istilah generik yang digunakan untuk berbagai jenis koneksi internet dengan menggunakan teknologi *broadband*. Pengertian *broadband* dalam arti harfiah, berarti jangkauan frekuensi yang luas yang digunakan untuk mengirim dan menerima data. Sebelumnya, proses akses internet dial-up sangatlah lambat, kecepatan koneksi *dial-up* terlalu lamban karena saluran telepon tetap sibuk saat mengakses internet. Faktor-faktor inilah membuat metode koneksi *broadband* ini disukai untuk akses internet. Istilah, *broadband* mengacu pada koneksi *bandwidth* internet. Istilah *bandwidth* umumnya digunakan untuk merujuk pada kecepatan *transfer* data, dalam hal jaringan komputer dan koneksi internet. *transfer* data biasanya diukur dalam bit per detik (bps). Dalam koneksi internet *broadband*, kecepatan *transfer* sangat tinggi dibandingkan dengan koneksi *dial-up* internet [17].

#### **2.2 Faktor-Faktor yang Mempengaruhi Peramalan Pengguna Pitallebar**

##### **2.2.1 Jumlah Penduduk**

Tingkat pertumbuhan penduduk Indonesia dari tahun 2005 sampai dengan tahun 2013 rata-rata sebesar 1,39% dan 1,49% per tahun sesuai hasil Sensus Penduduk 2010. Pertumbuhan penduduk Indonesia secara nasional masih relatif cepat, hal ini dapat diketahui setiap tahun tingkat kelahiran semakin

meningkat. Disisi lain, jumlah penduduk yang besar merupakan pangsa pasar yang besar bagi operator telekomunikasi, terutama pada usia antara 5 tahun sampai dengan 60 tahun dalam menggunakan pitalebar. Penduduk dalam suatu negara menjadi faktor terpenting dalam pelaksanaan pembangunan karena menjadi subyek dan obyek pembangunan (Randy Suryoranu, 2014). Menurut proyeksi yang dilakukan oleh Perserikatan Bangsa-Bangsa (PBB) dengan menilik populasi absolut Indonesia di masa depan, maka negeri ini akan memiliki penduduk lebih dari 250 juta jiwa pada tahun 2015, lebih dari 270 juta jiwa pada tahun 2025, lebih dari 285 juta jiwa pada tahun 2035 dan 290 juta jiwa pada tahun 2045. Baru setelah 2050, populasi Indonesia akan berkurang. Menurut proyeksi PBB, pada tahun 2050 dua pertiga populasi Indonesia akan tinggal di wilayah perkotaan. Sejak 40 tahun yang lalu, Indonesia sedang mengalami sebuah proses urbanisasi yang pesat makanya sekarang sekitar separuh dari jumlah total penduduk Indonesia tinggal di wilayah perkotaan. Proses ini menunjukkan perkembangan positif bagi perekonomian Indonesia karena urbanisasi dan industrialisasi akan membuat pertumbuhan ekonomi lebih maju dan menjadikan Indonesia negeri dengan tingkat pendapatan menengah ke atas (Indonesia Investments, 2015) [8].

### **2.2.2 Produk Domestik Bruto (PDB)**

Pengertian Produk Nasional Bruto (PNB)/ *Gross National Product* (GNP) adalah jumlah barang dan jasa yang dihasilkan oleh faktor-faktor produksi milik warga negara baik yang tinggal di dalam negeri maupun di luar negeri, tetapi tidak termasuk warga negara asing yang tinggal di negara tersebut. Salah satu indikator penting untuk mengetahui kondisi ekonomi di suatu negara dalam suatu periode tertentu adalah data Produk Domestik Bruto (PDB), baik atas dasar harga berlaku maupun atas dasar harga konstan. PDB pada dasarnya merupakan jumlah nilai tambah yang dihasilkan oleh seluruh unit usaha dalam suatu negara tertentu, atau merupakan jumlah nilai barang dan jasa akhir yang dihasilkan oleh seluruh unit ekonomi [8].

PDB atas dasar harga berlaku menggambarkan nilai tambah barang dan

jasa yang dihitung menggunakan harga yang berlaku pada setiap tahun, sedangkan PDB atas dasar harga konstan menunjukkan nilai tambah barang dan jasa tersebut yang dihitung menggunakan harga yang berlaku pada satu tahun tertentu sebagai dasar. PDB atas dasar harga berlaku dapat digunakan untuk melihat pergeseran dan struktur ekonomi, sedang harga konstan digunakan untuk mengetahui pertumbuhan ekonomi dari tahun ke tahun. Perekonomian Indonesia yang diukur berdasarkan besaran Produk Domestik Bruto (PDB) atas dasar harga berlaku triwulan II-2015 mencapai Rp 2.866,9 triliun dan atas dasar harga konstan 2010 mencapai Rp 2.239,3 triliun (BPS, 2015) [8].

### **2.2.3 Pendapatan Per Kapita**

Pendapatan per kapita adalah besarnya pendapatan rata-rata penduduk di suatu negara yang diperoleh dari hasil pembagian pendapatan nasional suatu negara dengan jumlah penduduk negara tersebut. Biasanya, pendapatan per kapita sering disebut dengan Produk Domestik Bruto (PDB) per kapita (BPS, 2014). Pendapatan per kapita sering digunakan untuk mengukur kemakmuran sebuah negara. Semakin besar pendapatan per kapita, negara tersebut akan dinilai semakin makmur. Pendapatan rata-rata per kapita penduduk Indonesia pada tahun 2013 sebesar Rp 32.463.700 per tahun atau Rp 2.705.308 per bulan. Walaupun pertumbuhan ekonomi melambat dari tahun ke tahun, Pemerintah pada tahun 2015 telah menaikkan upah minimum regional (UMR) menjadi Rp 3,1 juta. Pendapatan per kapita menunjukkan daya beli masyarakat dalam untuk memenuhi kebutuhan hidup [8].

### **2.2.4 Laju Pertumbuhan Ekonomi**

Ekonomi Indonesia triwulan III-2015 terhadap triwulan III-2014 (y-on-y) tumbuh 4,73 persen meningkat dibanding triwulan II-2015 yang tumbuh 4,67 persen. Pertumbuhan didorong oleh hampir semua lapangan usaha, dimana pertumbuhan tertinggi dicapai di Bidang Informasi dan Komunikasi yang tumbuh 10,83 persen (BPS, 2015). Pertumbuhan ekonomi Indonesia 5 (lima) tahun terakhir dari tahun 2011 sampai dengan tahun 2015 menunjukkan penurunan atau

perlambatan, yaitu: 2011 sebesar 6,49%, 2012 sebesar 6,23%, tahun sebesar 2013 sebesar 5,78%, 2014 sebesar 5,03 dan 2015 sebesar 4,67%.

Struktur ekonomi Indonesia secara spasial pada triwulan II-2015 didominasi oleh kelompok provinsi di pulau Jawa dan pulau Sumatera. Kelompok provinsi di pulau Jawa memberikan kontribusi terbesar terhadap Produk Domestik Bruto, yakni sebesar 58,35%, diikuti oleh pulau Sumatera sebesar 22,31%, dan pulau Kalimantan 8,22%. Semakin baik laju pertumbuhan ekonomi suatu negara, semakin meningkat daya beli masyarakat. Akibatnya, meningkat pula kesejahteraan masyarakatnya. Seiring dengan meningkatnya perekonomian, pendapatan masyarakat Indonesia pun ikut meningkat. Jumlah masyarakat miskin menurun dan jumlah masyarakat menengah meningkat [8].

### **2.2.5 Inflasi**

Inflasi merupakan kenaikan harga barang dan jasa secara umum, dimana barang dan jasa tersebut merupakan kebutuhan pokok masyarakat atau turunnya daya jual mata uang suatu negara. Bila harga barang secara umum naik terus-menerus, maka masyarakat akan panik sehingga perekonomian tidak berjalan normal, karena di satu sisi ada masyarakat yang berlebihan uang memborong barang, sementara yang kekurangan uang tidak bisa membeli barang. Di sisi lain, bila inflasi berkepanjangan, maka produsen banyak yang bangkrut karena produknya relatif akan semakin mahal sehingga tidak ada yang mampu membeli. Termasuk produk-produk teknologi informasi komunikasi (TIK), termasuk produk pitalebar. Bila diperhatikan laporan dari BPS, sumbangsih TIK terhadap inflasi rendah tetapi masyarakat daya beli menurun akan mempengaruhi juga produk-produk TIK. Akibat inflasi secara umum adalah menurunnya daya beli masyarakat karena secara riil tingkat pendapatannya juga menurun. Jadi, misalkan besarnya inflasi pada tahun yang bersangkutan naik sebesar 5% sementara pendapatan tetap, maka berarti secara riil pendapatan mengalami penurunan sebesar 5% yang akibatnya relatif akan menurunkan daya beli sebesar 5% juga [8].

### 2.3 ***Cross-Industry Standard Process For Data Mining (CRISP-DM)***

CRISP-DM atau *Cross-Industry Standard Process for Data Mining* adalah salah satu model proses data mining (*data mining framework*) yang awalnya (1996) dibangun oleh 5 perusahaan yaitu Integral Solutions Ltd (ISL), Teradata, Daimler AG, NCR Corporation dan OHRA. Framework ini kemudian dikembangkan oleh ratusan organisasi dan perusahaan di Eropa untuk dijadikan *methodology standard non-proprietary* bagi *data mining*. Versi pertama dari metodologi ini dipresentasikan pada 4th CRISP-DM SIG Workshop di Brussels pada bulan Maret 1999 dan langkah langkah proses *data mining* berdasarkan model ini di publikasikan pada tahun berikutnya. Sekitar tahun 2006 dan 2008 terbentuklah grup CRISP-DM 2.0 SIG yang berkeinginan untuk mengupdate CRISP-DM *process model*. Namun, produk akhir dari inisiatip ini tidak diketahui. Banyak hasil penelitian yang mengungkapkan bahwa CRISP-DM adalah *data mining* model yang masih digunakan secara luas di kalangan *industry*, sebahagian dikarenakan keunggulannya dalam menyelesaikan banyak persoalan dalam proyek-proyek *data mining*. Mariscal, Marba dan Fernandez menyatakan CRISP-DM sebagai *defacto* menjadi standar untuk pengembangan proyek *data mining* dan *knowledge discovery* karena paling banyak digunakan dalam pengembangan *data mining*. Hasil survei “Penggunaan Metodologi dalam Proyek *Data Mining*”, memperlihatkan pengguna CRISP-DM di tahun 2002 mencapai 51%, kemudian menurun menuju 41% di tahun 2004. Meskipun persentasi penggunaan CRISP-DM menurun 10%, jumlah pengguna metodologi ini masih terbilang lebih banyak daripada pengguna metodologi lain. Model proses CRISP-DM memberikan gambaran tentang siklus hidup proyek *data mining*. CRISP-DM memiliki 6 tahapan yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment* [18].

### 2.4 **Peramalan atau *Forecasting***

Peramalan merupakan alat bantu yang penting dalam perencanaan yang efektif dan efisien. Peramalan (*forecasting*) adalah kegiatan memperkirakan atau memprediksikan apa yang akan terjadi pada masa yang akan datang dengan waktu

yang relative lama. Sedangkan ramalan adalah suatu situasi atau kondisi yang akan diperkirakan akan terjadi pada masa yang akan datang. Untuk memprediksi hal tersebut diperlukan data yang akurat di masa lalu, sehingga dapat dilihat prospek situasi dan kondisi di masa yang akan datang. Pada umumnya kegunaan peramalan adalah sebagai berikut :

1. Sebagai alat bantu dalam perencanaan yang efektif dan efisien.
2. Untuk menentukan kebutuhan sumber daya di masa mendatang.
3. Untuk membuat keputusan yang tepat.

Kegunaan peramalan terlihat pada suatu pengambilan keputusan. Keputusan yang baik adalah keputusan yang didasarkan atas pertimbangan apa yang akan terjadi pada waktu keputusan dalam berbagai kegiatan perusahaan. Walaupun demikian perlu diketahui bahwa ramalan selalu ada unsur kesalahannya, sehingga yang perlu diperhatikan adalah usaha untuk memperkecil kesalahan dari ramalan tersebut.

Berikut ini adalah beberapa definisi atau pengertian Peramalan atau Pengertian Forecasting menurut para ahli :

- a. Pengertian Peramalan (*Forecasting*) menurut Lalu Sumayang, Peramalan adalah perhitungan yang objektif dan dengan menggunakan data-data masa lalu, untuk menentukan sesuatu dimasa yang akan datang.
- b. Pengertian Peramalan (*Forecasting*) menurut Sudjana, Peramalan adalah proses perkiraan (penguakan) besarnya atau jumlah sesuatu pada waktu yang akan datang berdasarkan data pada masa lampau yang dianalisis secara ilmiah khususnya menggunakan metode statistika.
- c. Pengertian *Forecasting* atau Peramalan menurut Jay Heizer & Barry Render, Peramalan adalah seni dan ilmu untuk memprediksi kejadian di masa depan dengan melibatkan pengambilan data historis dan memproyeksikannya ke masa mendatang dengan model pendekatan sistematis.
- d. Pengertian *Forecasting* atau Peramalan menurut William J. Stevenson , Peramalan adalah input dasar dalam proses pengambilan keputusan manajemen operasi dalam memberikan informasi tentang permintaan di masa mendatang dengan tujuan untuk menentukan berapa kapasitas atau persediaan yang

diperlukan untuk membuat keputusan staffing, budget yang harus disiapkan, pemesanan barang dari supplier dan partner dari rantai pasok yang dibutuhkan dalam membuat suatu perencanaan.

- e. Pengertian Peramalan (*Forecasting*) menurut Arman Hakim Nasution, Peramalan adalah proses untuk memperkirakan beberapa kebutuhan dimasa datang yang meliputi kebutuhan dalam ukuran kuantitas, waktu, dan lokasi yang dibutuhkan dalam rangka memenuhi permintaan barang ataupun jasa. [19].

## 2.5 Jenis-jenis Peramalan

### 1. Peramalan Kualitatif

Peramalan Kualitatif adalah peramalan yang didasarkan atas data kualitatif pada masa lalu. Hasil peramalan yang dibuat sangat bergantung pada orang yang menyusunnya. Hal ini penting karena hasil peramalan tersebut ditentukan berdasarkan pendapat dan pengetahuan serta pengalaman penyusunnya.

### 2. Peramalan Kuantitatif

Peramalan kuantitatif adalah peramalan yang didasarkan atas data kuantitatif masa lalu. Hasil peramalan yang dibuat sangat tergantung pada metode yang dipergunakan dalam peramalan tersebut. Semakin kecil penyimpangan antara hasil ramalan dengan kenyataan yang terjadi maka semakin baik pula metode yang digunakan. Peramalan Kuantitatif dapat diterapkan bila terdapat kondisi berikut :

- a. Tersedia informasi (data) tentang masa lalu.
- b. Informasi (data) tersebut dapat dikuantitatifkan dalam bentuk data numerik.
- c. Dapat diasumsikan bahwa beberapa aspek pola masa lalu akan terus berlanjut pada masa yang akan datang [20].

## 2.6 Metode *Exponential Smoothing*

Metode *Exponential Smoothing* merupakan prosedur perbaikan terus-menerus pada peramalan terhadap objek pengamatan terbaru. Metode peramalan ini menitik-beratkan pada penurunan prioritas secara eksponensial pada objek

pengamatan yang lebih tua. Dalam pemulusan eksponensial atau exponential smoothing terdapat satu atau lebih parameter pemulusan yang ditentukan secara eksplisit, dan hasil ini menentukan bobot yang dikenakan pada nilai observasi. Dengan kata lain, observasi terbaru akan diberikan prioritas lebih tinggi bagi peramalan daripada observasi yang lebih lama. Metode *exponential smoothing* dibagi lagi berdasarkan menjadi beberapa metode.

Forecasting Method	Data Pattern	Quantity of Historical Data	Forecast Horizon
Naïve	Stationary	1 or 2	Very short
Moving averages	Stationary	Number equals to the periods in the moving average	Very short
Exponential Smoothing:			
Simple	Stationary	5 to 10	Short
Adaptive Response	Stationary	10 to 15	Short
Holts	Linear Trend	10 to 15	Short to medium
Winters	Trend & Stationary	At least 4 or 5 per season	Short to medium
Regressions Based:			
Trend	Linear and non linear trend with or without seasonality	Minimum of 10 with 4 or 5 per season if seasonality is included	Short to medium
Causal	Can handle nearly all data patterns	Minimum of 10 per independent variable	Short, medium, and long
Time series decomposition	Can handle trend, seasonal, and cyclical patterns	Enough to see two peaks and troughs in the cycle	Short, medium, and long
ARIMA	Stationary or transformed to stationary	Minimum of 50	Short, medium, and long

Gambar 2.1 Metode Forecasting

(Sumber : Wilson-Keating 2001:51)[21]

### 2.6.1 *Single Exponential Smoothing*

Juga dikenal sebagai *simple exponential smoothing* yang digunakan pada peramalan jangka pendek, biasanya hanya 1 bulan ke depan. Model mengasumsikan bahwa data berfluktuasi di sekitar nilai mean yang tetap, tanpa trend atau pola pertumbuhan konsisten [21]. Rumus untuk *Simple exponential smoothing* adalah sebagai berikut:

$$F_{t+1} = \alpha * X_t + (1 - \alpha) * F_t \quad (1)$$

Dimana :

$F_t$  = peramalan untuk periode t.

$X_t + (1-\alpha)$  = Nilai aktual time series

$F_{t+1}$  = peramalan pada waktu t + 1

$\alpha$  = konstanta perataan antara 0 dan 1

### 2.6.2 Double Exponential Smoothing

Metode ini digunakan ketika data menunjukkan adanya trend. *Exponential smoothing* dengan adanya *trend* seperti pemulusan sederhana kecuali bahwa dua komponen harus diupdate setiap periode – level dan trendnya. Level adalah estimasi yang dimuluskan dari nilai data pada akhir masing-masing periode. *Trend* adalah estimasi yang dihaluskan dari pertumbuhan rata-rata pada akhir masing-masing periode [21].

Rumus *double exponential smoothing* adalah:

$$S_t = \alpha * Y_t + (1 - \alpha) * (S_{t-1} + b_{t-1}) \quad (2)$$

$$b_t = \gamma * (S_t - S_{t-1}) + (1 - \gamma) * b_{t-1} \quad (3)$$

$$F_{t+m} = S_t + b_t m \quad (4)$$

dimana:

$S_t$  = peramalan untuk periode t.

$Y_t + (1-\alpha)$  = Nilai aktual time series

$b_t$  = trend pada periode ke - t

$\alpha$  = parameter pertama perataan antara nol dan

1 = untuk pemulusan nilai observasi

$\gamma$  = parameter kedua, untuk pemulusan trend

$F_{t+m}$  = hasil peramalan ke - m

$m$  = jumlah periode ke muka yang akan diramalkan

### 2.6.3 Triple Exponential Smoothing

Metode ini digunakan ketika data menunjukkan adanya trend dan perilaku musiman Untuk menangani musiman, telah dikembangkan parameter persamaan ketiga yang disebut metode “HoltWinters” sesuai dengan nama penemunya.

Terdapat dua model Holt-Winters tergantung pada tipe musimannya yaitu Multiplicative seasonal model dan Additive seasonal model. Metode *exponential smoothing* yang telah dibahas sebelumnya dapat digunakan untuk hampir segala jenis data stasioner atau non – stasioner sepanjang data tersebut tidak mengandung faktor musiman. Tetapi bilamana terdapat musiman, metode ini dijadikan cara untuk meramalkan data yang mengandung faktor musiman, namun metode ini sendiri tidak dapat mengatasi masalah tersebut dengan baik. Meskipun demikian, metode ini dapat menangani faktor musiman secara langsung [21]. Rumus yang digunakan untuk *triple exponential smoothing* adalah:

Pemulusan trend:

$$B_t = g (S_t - S_{t-1}) + (1 - g) b_{t-1} \quad (5)$$

Pemulusan Musiman:

$$I = b \ t \ X \\ t \ S + (1-b) \ t -L +m \quad (6)$$

Ramalan:

$$F_{t+m} = (S_t + b \ t \ m) I_{t-L+m} \quad (7)$$

Dimana  $L$  adalah panjang musiman (misal, jumlah kuartal dalam suatu tahun),  $b$  adalah komponen trend,  $I$  adalah faktor penyesuaian musiman, dan  $F_{t+m}$  adalah ramalan untuk  $m$  periode ke muka.

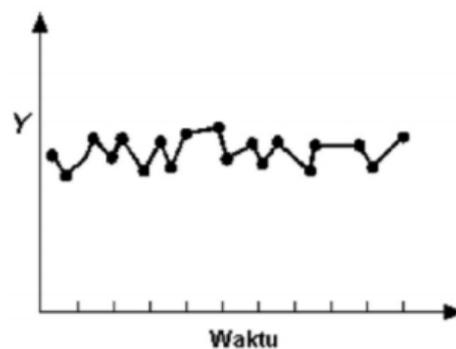
## 2.7 Jenis Pola Data

Model *time series* seringkali dapat digunakan dengan mudah untuk meramal, sedangkan model kausal dapat digunakan dengan keberhasilan yang lebih besar untuk pengambilan keputusan dan kebijaksanaan. Bilamana data yang diperlukan tersedia, suatu hubungan peramalan dapat dihipotesiskan baik sebagai fungsi dari waktu atau sebagai fungsi dari variabel bebas, kemudian diuji. Langkah penting dalam memilih model *time series* yang tepat adalah dengan mempertimbangkan jenis pola data, sehingga metode yang paling tepat dengan pola tersebut dapat diuji. Pola data dapat dibedakan menjadi empat jenis siklis dan trend.

*Time series* merupakan data yang dikumpulkan, dicatat atau diobservasi sepanjang waktu secara berurutan dengan beberapa periode waktu dapat tahun, kuartal, bulan, minggu dan pada beberapa kasus hari atau jam. Data *time series* di analisis untuk menemukan pola variasi masa lalu yang dapat dipergunakan untuk memperkirakan nilai untuk masa depan (*forecast*) karena dengan mengamati data runtut waktu akan terlihat empat komponen yang akan mempengaruhi pola data masa lalu dan sekarang yang benderung berulang di masa mendatang (Mukhyi, 2008). Klasifikasi model *time series* berdasarkan bentuk atau fungsi antara lain linier dan nonlinier, contoh dari model *time series linier* yaitu *moving average*, *Exponential Smoothing*.

### 2.7.1 Data Stationer

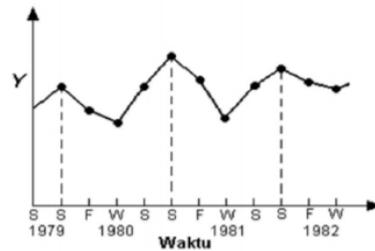
Pola data ini terjadi jika terdapat data yang berfluktuasi disekitar nilai rata-rata yang konstan. Suatu produk yang penjualannya tidak meningkat atau menurun selama waktu tertentu termasuk jenis pola ini. Pola khas dari data horizontal atau stasioner seperti ini dapat dilihat dalam Gambar 2.2.



**Gambar 2.2** Pola Data Stationer / Horizontal [11]

### 2.7.2 Data Musiman

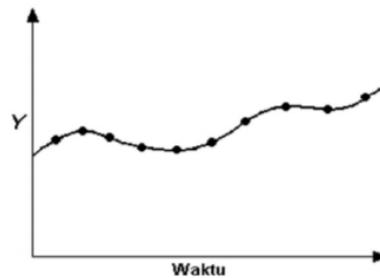
Pola data ini terjadi jika terdapat suatu deret data yang dipengaruhi oleh faktor musiman (misalnya kuartal tahun tertentu, bulanan, atau hari-hari pada minggu tertentu). Penjualan dari produk seperti minuman ringan, es krim, dan bahan bakar pemanas ruang semuanya menunjukkan jenis pola ini. Untuk pola musiman kuartalan dapat dilihat Gambar 2.3.



**Gambar 2.3** Pola Data Musiman [11]

### 2.7.3 Data Siklis

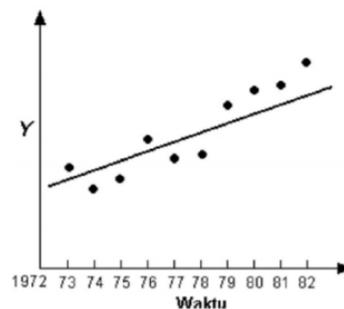
Pola data ini terjadi jika terdapat data yang dipengaruhi oleh fluktuasi ekonomi jangka panjang seperti yang berhubungan dengan siklus bisnis. Contoh: Penjualan produk seperti mobil, baja, dan peralatan utama lainnya. Jenis pola ini dapat dilihat pada Gambar 2.4.



**Gambar 2.4** Pola Data Siklis [11]

### 2.7.4 Data Trend

Pola data ini terjadi jika terdapat kenaikan atau penurunan sekuler jangka panjang dalam data. Contoh: Penjualan banyak perusahaan, GNP dan berbagai indikator bisnis atau ekonomi lainnya. Jenis pola ini bisa dilihat pada Gambar 2.5



**Gambar 2.5** Pola Data Trend [11]

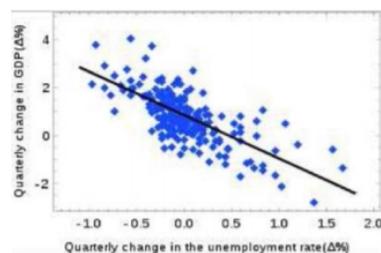
## 2.8 Ordinary Least Square

Dalam statistik dan ekonometrik, kuadrat terkecil biasa (*Ordinary Least Square*) atau kuadrat terkecil linier (*Linear Least Square*) merupakan metode untuk memperkirakan parameter yang tidak diketahui dalam model regresi linier. Metode ini meminimalkan jumlah jarak kuadrat antara data yang diamati langsung pada dataset, dan data yang telah diprediksi dengan menggunakan metode pendekatan linear. Dengan mendapatkan nilai parameter yang optimal maka error yang didapatkan dari hasil peramalan yang dilakukan akan kecil dan ketepatan peramalan semakin endekati terhadap data yang real [11].

### 2.8.1 Regresi Linear Sederhana (*Simple Linear Regression*)

Dalam statistik, regresi linier sederhana adalah penaksir kuadrat terkecil dari model regresi linier dengan variable prediktor tunggal. Dengan kata lain, regresi linier sederhana mencocokkan dengan garis lurus melalui serangkaian titiktitik n sedemikian rupa yang membuat jumlah kuadrat residual dari model. Pada gambar dibawah ini merupakan contoh penyebaran data yang ada pada persamaan linier untuk dicarikan parameter dengan rumus dibawah ini:

$$\hat{\beta} = \frac{\sum x_i y_i - \frac{1}{n} \sum x_i \sum y_i}{\sum x_i^2 - \frac{1}{n} (\sum x_i)^2} \quad (8)$$



**Gambar 2.6** Contoh *Linier Regression* [11]

## 2.9 Evaluasi Hasil Peramalan

Evaluasi hasil peramalan digunakan untuk mengetahui keakuratan hasil peramalan yang telah dilakukan terhadap data yang sebenarnya. Terdapat banyak metode untuk melakukan perhitungan kesalahan peramalan. Beberapa metode yang digunakan adalah [11]:

### 2.9.1 *Mean Square Error*

Cara yang cukup sering digunakan dalam mengevaluasi hasil peramalan yaitu dengan menggunakan metode *Mean Squared Error* (MSE). Dengan menggunakan MSE, error yang ada menunjukkan seberapa besar perbedaan hasil estimasi dengan hasil yang akan diestimasi. Hal yang membuat berbeda karena adanya keacakan pada data atau karena tidak mengandung estimasi yang lebih akurat [11].

$$MSE = \frac{1}{N} \sum_{t=h}^N (y_t - \hat{y}_t)^2 \quad (9)$$

Dimana:

MSE = *Mean Square Error*

N = Jumlah Sampel

$t y$  = Nilai Aktual Indeks

$t y^{\wedge}$  = Nilai Prediksi Indeks

RMSE merupakan mengakarkan nilai dari MSE yang sudah dicari sebelumnya. RMSE digunakan untuk mencari keakuratan hasil peramalan dengan data history dengan menggunakan rumus (Makridakis, 1999). Semakin kecil nilai yang dihasilkan semakin bagus pula hasil peramalan yang dilakukan [11].

$$RMSE = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n}} \quad (10)$$

### 2.9.2 *Mean Absolute Percentage Error*

Metode ini melakukan perhitungan perbedaan antara data asli dan data hasil peramalan. Perbedaan tersebut diabsolutkan, kemudian dihitung ke dalam bentuk persentase terhadap data asli. Hasil persentase tersebut kemudian didapatkan nilai *mean*-nya. Suatu model mempunyai kinerja sangat bagus jika nilai MAPE berada di bawah 10%, dan mempunyai kinerja bagus jika nilai MAPE berada di antara 10% dan 20%.

Dalam fase peramalan, menggunakan MSE sebagai suatu ukuran ketepatan juga dapat menimbulkan masalah (Makridakis, 1999). Ukuran ini tidak memudahkan perbandingan antar deret berkala yang berebeda dan untuk selang

waktu yang berlainan, karena MSE merupakan ukuran absolut. Lagi pula, interpertasinya tidak bersifat intuitif bahkan untuk para spesialis sekalipun, karena ukuran ini menyangkut penguadratan sederetan nilai.

Alasan yang telah disebutkan di atas dalam hubungan dengan keterbatasan MSE sebagai suatu ukuran ketepatan peramalan, Maka diusulkan ukuran – ukuran alternatif, yang diantaranya menyangkut galat persentase. Tiga ukuran berikut sering digunakan [11]:

Galat Persentase (*Percentage Error*)

$$PE_t = \left( \frac{X_t - F_t}{X_t} \right) (100) \quad (11)$$

Nilai Tengah Galat Persentase (*Mean Percentage Error*)

$$MPE = \sum_{i=1}^n \frac{PE_i}{n} \quad (12)$$

Nilai Tengah Galat Persentase Absolut (*Mean Absolut Percentage Error*)

$$MAPE = \sum_{i=1}^n \frac{|PE_i|}{n} \quad (13)$$

Dimana :

$X_t$  = Data history atau Data aktual pada periode ke - t

$F_t$  = Data hasil ramalan pada periode ke - t

n = jumlah data yang digunakan

t = periode ke – t

## 2.10 *Decission Tree*

*Decision Tree* adalah struktur pohon, dimana setiap node pohon merepresentasikan atribut yang telah diuji, setiap cabang merupakan suatu pembagian hasil uji, dan node daun (*leaf*) merepresentasikan kelompok kelas tertentu. Level node teratas dari sebuah *decision tree* adalah node akar (*root*) yang biasanya berupa atribut yang paling memiliki pengaruh terbesar pada suatu kelas tertentu. Pada umumnya *decision tree* melakukan strategi pencarian secara *top-down* untuk solusinya [22].

Pada proses mengklasifikasi data yang tidak diketahui, nilai atribut akan diuji dengan cara melacak jalur dari node akar (*root*) sampai node akhir (*leaf*) dan

kemudian akan diprediksi kelas yang dimiliki oleh suatu data baru tertentu. Menurut Siswanto, pohon keputusan (*decision tree*) adalah model visual untuk menyederhanakan proses pembuatan keputusan secara rasional. Dengan adanya visualisasi memungkinkan untuk memahami proses pembuatan keputusan yang berstruktur, bertahap, dan rasional. Pembuatan keputusan sendiri berarti memilih alternatif-alternatif keputusan yang tersedia, karena unsur ketidakpastian maka berbagai kemungkinan keadaan akan dihadapi oleh masing-masing alternatif keputusan itu.

Oleh karena itu, diagram keputusan mempunyai noda keputusan dan noda cabang. Dari beberapa pengertian diatas, disimpulkan bahwa pohon keputusan adalah metode yang digunakan sebagai pengambilan keputusan dari berbagai alternatif keputusan yang tersedia melalui proses pembuatan keputusan yang berstruktur, bertahap, dan rasional dengan diagram keputusan yang mempunyai noda keputusan dan noda cabang.

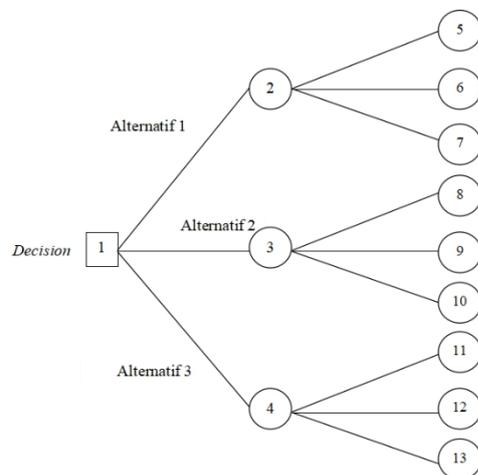
### **Analisis Pohon Keputusan (*Decision Tree*)**

Diagram pohon sering kali membantu dalam memahami dan menyelesaikan persoalan probabilitas. Diagram pohon biasanya digambarkan dengan lambang yang baku. Dimulai dengan suatu nokhta kemudian dibuat cabang-cabang sebanyak peristiwa yang mungkin dapat dihasilkan dari percobaan. Pada masing-masing cabang dituliskan probabilitas terjadinya peristiwa yang bersangkutan. Jika percobaan dilakukan lagi, maka langkah-langkah itu diulang. Setiap cabang berakhir pada nokhta yang kemudian diisi dengan probabilitas peristiwa bersama. Pada nokhta yang paling awal dituliskan angka 1 yang artinya jumlah probabilitas dari seluruh peristiwa yang mungkin.

Kelebihan dari metode pohon keputusan [22] adalah :

1. Daerah pengambilan keputusan yang sebelumnya kompleks dan sangat global, dapat diubah menjadi lebih simpel dan spesifik.
2. Eliminasi perhitungan-perhitungan yang tidak diperlukan, karena ketika menggunakan metode pohon keputusan maka *sample* diuji hanya berdasarkan kriteria atau kelas tertentu.

3. Fleksibel untuk memilih fitur dari internal node yang berbeda, fitur yang terpilih akan membedakan suatu kriteria dibandingkan kriteria yang lain dalam node yang sama. Kefleksibelan metode pohon keputusan ini meningkatkan kualitas keputusan yang dihasilkan jika dibandingkan ketika menggunakan metode perhitungan satu tahap yang lebih konvensional.
4. Dalam analisis multivariat, dengan kriteria dan kelas yang jumlahnya sangat banyak. Metode pohon keputusan dapat menghindari munculnya permasalahan ini dengan menggunakan kriteria yang jumlahnya lebih sedikit pada setiap node internal tanpa banyak mengurangi kualitas keputusan yang dihasilkan.

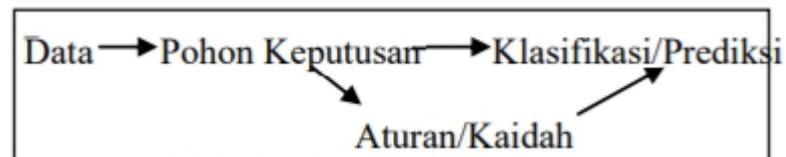


**Gambar 2.7** *Decision Tree* (Pohon Keputusan) [22]

### 2.11 *Iterative Dichotomiser Three (ID3)*

ID3 merupakan sebuah metode yang digunakan untuk membangkitkan pohon keputusan. Input dari algoritma ini adalah sebuah database dengan beberapa variabel yang juga dikenal dengan atribut. Setiap masukan dalam *database* menyajikan sebuah objek dari domain yang disebut dengan variabel bebas (*independent variable*). Sebuah atribut didesain untuk mengklasifikasikan objek yang disebut dengan variabel tidak bebas (*dependent variable*). Proses klasifikasi dilakukan dari node paling atas yaitu akar pohon (root). Dilanjutkan ke bawah melalui cabang-cabang sampai dihasilkan node daun (*leaves*) dimana node daun ini menunjukkan hasil akhir klasifikasi. Sebuah objek yang diklasifikasikan

dalam pohon harus dites nilai entropinya. Entropi adalah ukuran dari teori informasi yang dapat mengetahui karakteristik impurity dan homogeneity dari kumpulan data. Dari nilai entropi tersebut kemudian dihitung nilai information gain (IG) masing-masing atribut independent terhadap atribut dependent-nya. IG merupakan nilai rata-rata entropi pada semua atribut.



**Gambar 2.8** Alur Diagram Pohon Keputusan [23]

Langkah menghitung information gain adalah sebagai berikut:

(1) Hitung info kelas dengan probabilitas dari sample yang mempunyai class  $C_i$ .

$$I(s_1, s_2, \dots, s_m) = - \sum_{i=1}^m P_i \cdot \log_2(P_i) \qquad P_i = \frac{s_i}{s}$$

(2) Hitung Entropy attribute:

$$E(A) = \sum_{j=1}^v \frac{s_{1j} + \dots + s_{mj}}{s} I(s_1, s_2, \dots, s_m)$$

(3) Hitung gain:

$$\text{Gain}(A) = I(s_1, s_2, \dots, s_m) - E(A)$$

Adapun sample data yang digunakan oleh ID3 memiliki beberapa syarat sebagai berikut: (1) Deskripsi atribut-nilai. Atribut yang sama harus mendeskripsikan tiap contoh dan memiliki jumlah nilai yang sudah ditentukan. (2) Kelas sudah didefinisikan sebelumnya, dan digambarkan dengan jelas. Untuk kelas yang kontinu dipecah-pecah menjadi kategori-kategori yang relative, misalnya saja metal dikategorikan sebagai "hard, quite hard, flexible, soft quite soft". (3) Jumlah contoh yang cukup. Karena dibutuhkan test case yang cukup membedakan pola yang valid dari peluang suatu kejadian.

Dengan pemodelan *decision tree*, persoalan klasifikasi yang cukup sederhana dapat dilakukan dengan akurat karena dalam mengklasifikasikan suatu objek, tentunya terdapat aturan-aturan yang menjadi penentu termasuk kategori manakah objek tersebut. Dengan algoritma ID3, pohon keputusan dibangun

berdasarkan aturan-aturan tersebut, sehingga proses learning akan memiliki akurasi yang tinggi karena klasifikasinya dilakukan dengan melakukan inferensi terhadap aturan-aturan tersebut [23].

## 2.12 Penelitian Terdahulu

Tabel 2.1 Penelitian Terdahulu

Penulis (Tahun)	Algoritma/ Teknik Pemodelan	Tipe Klasifikasi	Kelebihan	Kekurangan	Tujuan
Azwar Aziz (2016) [8]	Teknik <i>Time Series</i>	<i>Exponential Smoothing Method</i> dan Regresi berganda	Metode <i>double exponential smoothing</i> dipilih sebagai metode terbaik untuk mendapatkan jumlah pengguna pitalebar di masa yang akan datang.	Metode Single dan Triple Exponential Smoothing metode yang nilai error sangat besar.	Untuk mendapatkan Forecasting permintaan pengguna broadband di Indonesia.
Chintia Vairra Hudiyandi, dkk. (2019) [9]	Algoritma <i>Double Moving Average</i> , Algoritma <i>Double Exponential Smoothing</i>	Metode <i>Double Moving Average</i> , Metode <i>Double Exponential Smoothing</i>	Metode <i>Double Exponential Smoothing</i> dapat melakukan peramalan dengan lebih akurat dibandingkan <i>Double Moving Average</i> dalam peramalan jumlah kedatangan wisatawan mancanegara di bandara Ngurah Rai.	Nilai MAPE sebesar 3,355	Untuk Peramalan Jumlah Kedatangan Wisatawan Mancanegara di Bandara Ngurah Rai dengan menggunakan Perbandingan Double Moving Averagedan Double Exponential Smoothing.

Hegar Winda Tresnani, dkk. (2018) [10]	<i>Algoritma grey holt winter exponential smoothing</i>	Metode <i>grey holt winter exponential smoothing</i> dengan <i>golden section</i>	MAPE minimum yang didapat berdasarkan metode <i>grey holt winter exponential smoothing</i> adalah sebesar 16.06%, sedangkan MAPE minimum yang dihasilkan oleh metode <i>grey holt winter exponential smoothing</i> dengan <i>golden section</i> adalah sebesar 13.92% yang dilakukan pada pengujian sebanyak 96 data	MAPE yang didapat melalui optimasi parameter menggunakan <i>golden section</i> , pada pengujian sebanyak 96 data dengan season per bulan memiliki nilai kesalahan minimum sebesar 13.92%	Untuk menemukan optimasi parameter <i>smoothing</i> terbaik sehingga didapatkan kesalahan peramalan minimum.
Alda Raharja, dkk. (2011) [11]	Algoritma <i>Time Series</i>	Metode <i>exponential smoothing</i> dan <i>moving average</i>	Peramalan metode <i>double exponential smoothing</i> didapatkan hasil yang lebih baik dibandingkan dengan metode <i>moving average</i> , dengan selisih error sebesar 0.29% dan selisih error RMSE sebesar 74.15.	Nilai evaluasi kesalahan peramalan tetap berada pada interval tertentu, yaitu 2% - 3% untuk nilai MAPE.	Untuk menentukan metode yang lebih baik dalam peramalan.

Tias dkk. [12]	Safitri, (2017) Algoritma <i>Time Series</i>	Metode <i>Exponential Smoothing Holt-Winters</i> dan ARIMA	- Peramalan jumlah kedatangan wisatawan mancanegara ke Bali Ngurah Rai menurut pintu masuk dengan metode <i>exponential smoothing Holt-Winters</i> menghasilkan nilai <i>Mean Square Error</i> (MSE) yaitu 1436553590 dan nilai <i>Mean Absolute Percentage Error</i> (MAPE) yaitu 8,86198%. Sedangkan metode ARIMA menghasilkan nilai <i>Mean Square Error</i> (MSE) yaitu 1353169319 dan nilai <i>Mean Absolute Percentage Error</i> (MAPE) yaitu 9,40981%. - Lebih efektif menggunakan metode <i>exponential smoothing Holt-Winters</i> dibandingkan metode ARIMA karena nilai MAPE yang lebih kecil daripada nilai MAPE yang dihasilkan metode ARIMA.	Metode <i>exponential smoothing Holt-Winters</i> yaitu membutuhkan banyak waktu dalam menentukan paramater alpha, beta dan gamma dengan cara trial and error.	Untuk mengetahui perbandingan model peramalan menggunakan metode <i>Exponential Smoothing Holt-Winters</i> dan ARIMA, serta menentukan metode yang terbaik.
Sofiyatun dan Haryadi Surjono. (2019) [13]	<i>Algoritma</i> Pengambilan keputusan dengan metode <i>Decision Tree</i>	Metode <i>Exponential Smoothing, Moving Average, Trend Projection, Addictive Decomposition Average All,</i>	Melalui perhitungan dan penyusunan pohon keputusan ( <i>Decision tree</i> ) secara manual dapat diketahui bahwa kota Bekasi adalah keputusan yang paling tepat diantara alternatif yang lain	Membutuhkan <i>expected Value</i> dari perusahaan agar bisa di akurasi yang lebih baik lagi.	Untuk mengetahui metode peramalan terbaik untuk penjualan tiket pesawat pada PT. HIS Tour & Travel

		<i>Addictive Decomposition Centered Moving Average, Multiplicative Average All, Multiplicative Decomposition Centered.</i>	dengan hasil Rp 280.560.000		Cabang MM 2100 Cibitung.
Hikmatullah dkk (2019). [14]	Algoritma <i>Iterative Dichotomiser Three</i> (ID3)	Metode <i>Decission Tree</i>	Penerapan metode <i>Iterative Dichotomiser Three</i> (ID3) dalam klasifikasi diagnosa kesehatan kehamilan berdasarkan pada hasil pengujian klasifikasi dengan menggunakan algoritma <i>Iterative Dichotomiser Three</i> mempunyai nilai optimal akurasi sebesar 80,33 %.	<i>Gain</i> sebesar 0,126447	Untuk mengetahui penerapan metode <i>Iterative Dichotomiser Three</i> (ID3) dalam klasifikasi diagnosa kesehatan kehamilan
J. Dairy Sci. 94 :1873–1892 [24]	Algoritma <i>Decission tree</i>	Metode <i>Decission tree</i>	Decision tree adalah metode yang paling efektif untuk memastikan strategi perawatan yang paling ekonomis untuk <i>commercial dairy herds</i> (ternak sapi perah komersial) dan juga sebagai alat instruksional yang berguna untuk memahami interaksi kompleks yang mempengaruhi ekonomi dari perawatan CM ( <i>Clinical Mastitis</i> ).	Menghasilkan lebih sedikit EMV ( <i>Expected Monetary Value</i> ).	Untuk mengetahui strategi terbaik mengobati mastitis yang disebabkan oleh <i>gram-positive pathogens</i>