

IMPLEMENTASI DATA MINING UNTUK MERAMALKAN PENJUALAN DI MINIMARKET IDOLA JL PATI-TAMBAKROMO KM 2 DESA KARANGMULYO RT 08 RW 1 DENGAN METODE TIME SERIES

Hamdani Widyatmoko, Anton Setiawan Honggowibowo, Nur Cahyani Dewi Retnowati
Teknik Informatika STTA Yogyakarta
informatika@stta.ac.id

Abstract

Minimarket idola on a daily basis there are many sales transactions, so that the data stored in the database is very large. The data can be used as much useful information for the owner of a minimarket in policy making. To explore the data that is used a lot of data mining technique. Data mining uses data analysis to discover patterns and relationships in data that may be used to make accurate predictions.

In this research, data mining is used to forecast the sales of goods in Minimarket Idola. Forecasting the future based on measuring the value of the patterns in the data collection. To perform sales forecasting in the future to use the method of time series. Forecasting time series data to predict what will happen based on past historical data.

Time series methods for forecasting sales in the calculation Minimarket Idola using exponential smoothing and moving average. Of the count sought the MAD (Mean Absolute Deviation) or forecasting errors. Where MAD is the smallest value of the calculation of exponential smoothing and moving average is the result of forecasting with a small error. Forecasting results will not always be appropriate because the market demand influenced by several factors. But it does not mean that the forecast is made useless.

Key Words : Data Mining, Time Series, Sales forecasting

Abstrak

Pada Minimarket Idola setiap harinya terjadi banyak transaksi penjualan, sehingga data yang disimpan di *database* sangat besar. Data yang banyak bisa dijadikan informasi yang bermanfaat bagi pemilik *minimarket* dalam pengambilan kebijakan. Untuk menggali data yang banyak tersebut digunakan teknik data mining. Data *mining* menggunakan analisis data untuk menemukan pola dan hubungan didalam data yang mungkin digunakan untuk membuat ramalan yang akurat.

Pada penelitian ini Data *mining* digunakan untuk meramalkan penjualan barang di Minimarket Idola. Peramalan yaitu mengestimasi nilai masa depan berdasarkan pola-pola didalam sekumpulan data. Untuk melakukan peramalan penjualan di waktu yang akan datang digunakan metode *time series*. Peramalan data *time series* memprediksi apa yang akan terjadi berdasarkan data historis masa lalu.

Metode *time series* untuk peramalan penjualan di Minimarket Idola menggunakan perhitungan *exponential smoothing* dan *moving average*. Dari perhitungan tersebut dicari nilai MAD (*Mean Absolute Deviation*) atau kesalahan peramalan. Dimana nilai MAD yang terkecil dari perhitungan *exponential smoothing* dan *moving average* merupakan hasil peramalan dengan

kesalahan yang kecil. Hasil peramalan tidak akan selalu tepat karena dipengaruhi beberapa faktor permintaan pasar. Namun tidak berarti bahwa ramalan yang dilakukan tidak berguna.

Kata Kunci : Data Mining, Time Series, Pola penjualan

1. Latar Belakang

Hampir semua *minimarket* menggunakan sistem komputerisasi dalam penyimpanan data penjualan, sehingga akan dihasilkan banyak data transaksi. Data transaksi yang ada dalam *database* penjualan barang menyimpan jumlah *record* transaksi penjualan yang memiliki volume yang sangat besar sehingga menyebabkan jumlah data terus menerus bertambah setiap harinya. Dari penumpukan data yang terjadi bisa digali untuk menemukan pola penjualan barang, sehingga data yang sedemikian banyak bisa dimanfaatkan untuk menganalisa pasar dan meramalkan penjualan pada waktu yang akan datang.

Dalam hal ini, sangat dibutuhkan aplikasi yang membantu *minimarket* untuk mendapatkan informasi dari jumlah data yang terkumpul. Salah satu cara untuk mendapatkan informasi tersebut adalah memanfaatkan teknik data *mining*. Data *mining* adalah suatu penerapan metode tertentu untuk melakukan penggalian data. Data yang menumpuk digali dengan cara menganalisa data transaksi penjualan barang yang telah terjadi di *minimarket* sehingga ditemukan pola penjualan untuk menganalisa pasar pada periode berikutnya.

Maka dari permasalahan yang ada, penulis pada tugas akhir ini akan mengembangkan sistem *data mining* (penggalian data) dengan metode *time series*. *Time Series* merupakan data yang terdiri atas satu objek tetapi meliputi beberapa periode waktu misalnya harian, mingguan, bulanan, tahunan dan lain-lain. Data masa lalu dikumpulkan, dianalisis dengan metode-metode tertentu untuk dapat mempolakan dan meramalkan hasil di waktu yang akan datang. Ramalan yang baik adalah yang mendekati kenyataan. Oleh karena itu metode perhitungan yang diterapkan untuk mengimplementasikan data mining ini adalah menggunakan metode *times series* dengan perhitungan *exponential smoothing* dan *moving average*.

2. Landasan Teori

Data Mining

Data *mining* adalah suatu proses yang menggunakan teknik statistik, matematika, kecerdasan tiruan dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari *database* yang besar (Turban 2005). Data *mining* menggunakan analisis data untuk menemukan pola dan hubungan didalam data yang mungkin digunakan untuk membuat *forecasting* atau ramalan yang akurat.

Peramalan Data Time Series

Forecasting atau peramalan yaitu mengestimasi nilai masa depan berdasarkan pola-pola didalam sekumpulan data (Turban 2005). Peramalan Data *Time Series* memprediksi apa yang akan terjadi berdasarkan data historis masa lalu. *Time series* adalah serangkaian nilai pengamatan (observasi) yang diambil selama kurun waktu tertentu pada umumnya dalam

interval yang panjang (Turban 2005). Peramalan dapat diterapkan bila terdapat kondisi seperti tersedianya informasi data historis, informasi tersebut dapat dikuantitatifkan dalam bentuk numerik dan dapat diasumsikan bahwa beberapa aspek pola masa lalu akan terus berlanjut di waktu mendatang.

Perhitungan Peramalan *Time Series*

Rumus umum metode *Single Exponential Smoothing* adalah sebagai berikut :

$$(S_{t+1} = \alpha X_t + (1 - \alpha) S_t)$$

S_{t+1} : periode peramalan (waktu hendak dilakukan peramalan)

X_t : data asli pada periode ke- t

S_t : nilai peramalan pada periode ke- t

α : nilai perkiraan fluktuasi (diisi nilai antara 0 s/d 1)

Secara matematis, *moving average* dinyatakan dalam persamaan sebagai berikut:

$$F_{t+1} = \frac{X_t + X_{t-1} + \dots + X_{t-n+1}}{n}$$

F_{t+1} : nilai ramalan untuk periode waktu ke - $t+1$,

X_t : nilai aktual periode ke - t ,

n : banyak data (ordo).

Sedangkan, untuk mengetahui tingkat kesalahan peramalan penjualan atau *Mean Absolute Deviation* (MAD) menggunakan rumus sebagai berikut :

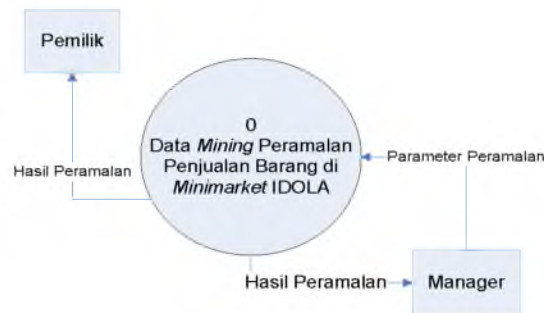
$$MAD = \frac{\sum | \text{Nilai Forecast} - \text{Nilai Nyata} |}{n}$$

n : jumlah data untuk keperluan peramalan
 $\sum$: jumlah data untuk keperluan peramalan

3. Perancangan Sistem

Diagram Konteks

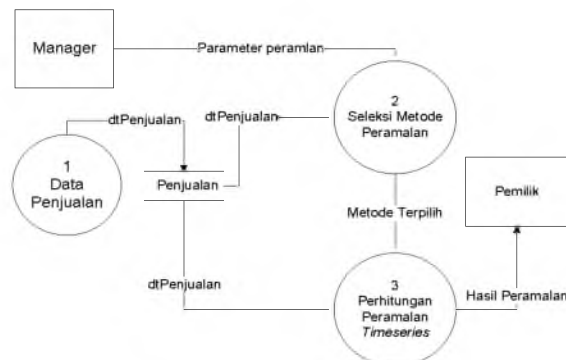
Diagram konteks memperlihatkan bahwa entitas yang terlibat dalam sistem adalah manager dan pemilik. Manager akan menginputkan parameter peramalan ke sistem, yang berhubungan dengan kategori barang dan periode peramalan. Sedangkan pemilik akan menerima hasil peramalan penjualan dari sistem yang digunakan sebagai dasar pengambilan kebijakan. Diagram konteks ada pada gambar 1.



Gambar 1 Diagram Konteks.

DFD Level 1

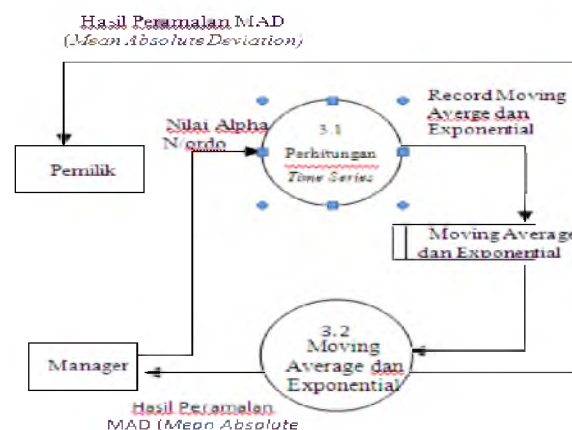
Hasil proses penjabaran DFD level 1 adalah terdapat tiga buah sub proses yaitu , sub proses pencatatan penjualan, sub proses seleksi metode peramalan, dan sub proses peramalan. DFD Level 1 ada di gambar 2.



Gambar 2 DFD Level 1.

DFD Level 2 Proses Perhitungan Time Series

Proses perhitungan *time series* mengolah parameter nilai alpha dan ordo yang diinputkan oleh *manager*. Nilai alpha untuk perhitungan *exponential smoothing* dan ordo untuk *moving average*. Selanjutnya parameter diproses sesuai perhitungan *exponential smoothing* dan *moving average*. Dimana nilai MAD antara *exponential smoothing* dan *moving average* yang paling kecil merupakan hasil peramalan yang memiliki kesalahan terkecil. DFD Level 2 proses perhitungan *time series* ada pada Gambar 3.



Gambar 3 DFD Level 2 Proses perhitungan Time series

Flowchart Data Mining Peramalan Penjualan

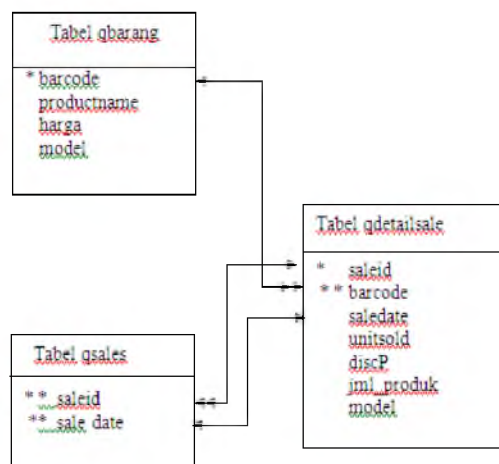
Pada *flowchart* data mining untuk peramalan penjualan ini, inputan parameter peramalan seperti input nilai alpha dan ordo serta periode peramalan berdasar harian, mingguan dan bulanan hanya bisa dilakukan oleh *user* yang memiliki hak akses sebagai *manager*. *Flowchart* data mining peramalan penjualan ditunjukkan pada Gambar 4.



Gambar 4 Flowchart Sistem Data Mining Peramalan Penjualan

Relasi Antar Tabel

Relasi *database* pada penerapan *time series* pada data mining untuk peramalan transaksi penjualan barang dapat dilihat pada Gambar 5



Gambar 5 Relasi Antar Tabel.

4. Implementasi Dan Analisa Sistem

Implementasi sistem merupakan tahap mendeskripsikan suatu sistem agar sistem tersebut siap untuk diaplikasikan. Berdasarkan tahapan implementasi tersebut dapat diketahui sistem yang dibuat benar-benar mencapai tujuan yang diinginkan atau tidak. Berdasarkan analisis dari desain sistem yang telah dilakukan, maka telah diimplementasikan

sebuah sistem Data Mining Menggunakan Metode *Time Series*. Implementasi Hasil Peramalan dapat dilihat pada Gambar 6.

Laporan Peramalan Penjualan Bulanan						
Periode	Nama	Populasi	Exponential Smoothing	Moving Average	Metode	Hasil
Januari 2010 s/d Desember 2010	BARANG A	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG B	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG C	11	11.11111111 0.10	11.11111111 0.10	Berdasarkan Exponential Smoothing	0.00
Januari 2010 s/d Desember 2010	BARANG D	11	11.11111111 0.30	11.11111111 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG E	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Exponential Smoothing	0.00
Januari 2010 s/d Desember 2010	BARANG F	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG G	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG H	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG I	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG J	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG K	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG L	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG M	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG N	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG O	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG P	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG Q	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG R	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG S	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG T	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG U	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG V	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG W	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG X	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG Y	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00
Januari 2010 s/d Desember 2010	BARANG Z	15	15.55314000 0.30	15.55314000 0.30	Berdasarkan Average	0.00

Gambar 6 Tampilan Tabel Hasil Peramalan

Analisa Hasil

Sistem yang telah dibuat merupakan penerapan dari metode yang ada, yaitu metode *timeseries* dibuat dalam bentuk sebuah program dengan perhitungan *exponential smoothing* dan *moving average*. Sehingga hasil yang didapatkan tidak jauh berbeda dengan perhitungan manualnya.

Tahap perhitungan yang diperlukan :

1. Tahap Pertama yaitu menentukan periode waktu peramalan.

Periode waktu peramalan bisa berdasar harian, minnguan, bulanan. Dari permasalahan diatas diketahui periode peramalan mulai dari bulan ke-1 (Januari 2010) s/d bulan ke-10 (Oktober 2010) untuk peramalan penjualan bulan ke-11 (November 2010).

2. Tahap Kedua yaitu memberikan nilai α (alfa) pada perhitungan *exponential smoothing* (α bernilai 0 hingga 1).

Nilai α berpengaruh pada hasil *exponential smoothing*. Besar kecilnya nilai konstanta α yang digunakan oleh *manager* atau pemilik merupakan nilai *confidence* atau kepercayaan diri untuk melakukan peramalan diwaktu yang akan datang. Penggunaan nilai α bisa dipengaruhi oleh kebiasaan *manager* dalam mengetahui pola data penjualan. Dapat dijelaskan bahwa semakin besar nilai α penghalusan atau *smoothing* yang dilakukan kecil, nilai α kecil penghalusan atau *smoothing* semakin besar, α optimum akan meminimumkan nilai MAD.

$$S_{t+1} = \alpha \cdot X_t + (1 - \alpha) S_t$$

No	Bulan	Tahun	Penjualan (X _t) bungkus	Exponential Smoothing ($S_{t+1} = \alpha X_t + (1-\alpha) S_t$)	St
1	Januari	2010	75	0.00	
2	Februari	2010	60	$S_{2+1} = (0.1(60) + (1-0.1) \times 75)$ $S_2 = 6.0 + (0.9 \times 75)$ $S_2 = 6.0 + 67.5 = 73.5$	75
3	Marat	2010	85	$S_{3+1} = (0.1(85) + (1-0.1) \times 73.5)$ $S_3 = 8.5 + (0.9 \times 73.5)$ $S_3 = 8.5 + 66.15 = 74.65$	73.5
4	April	2010	60	$S_{4+1} = (0.1(60) + (1-0.1) \times 74.65)$ $S_4 = 6.0 + (0.9 \times 74.65)$ $S_4 = 6.0 + 67.19 = 73.19$	74.65
5	Mei	2010	76	$S_{5+1} = (0.1(76) + (1-0.1) \times 73.19)$ $S_5 = 7.6 + (0.9 \times 73.19)$ $S_5 = 7.6 + 65.87 = 73.47$	73.19
6	Juni	2010	64	$S_{6+1} = (0.1(64) + (1-0.1) \times 73.47)$ $S_6 = 6.4 + (0.9 \times 73.47)$ $S_6 = 6.4 + 66.12 = 72.52$	73.47
7	Juli	2010	52	$S_{7+1} = (0.1(52) + (1-0.1) \times 72.52)$ $S_7 = 5.2 + (0.9 \times 72.52)$ $S_7 = 5.2 + 65.27 = 70.47$	72.52
8	Agustus	2010	89	$S_{8+1} = (0.1(89) + (1-0.1) \times 70.47)$ $S_8 = 8.9 + (0.9 \times 70.47)$ $S_8 = 8.9 + 63.42 = 72.32$	70.47
9	September	2010	120	$S_{9+1} = (0.1(120) + (1-0.1) \times 72.32)$ $S_{10} = 12 + (0.9 \times 72.32)$ $S_{10} = 12 + 65.09 = 77.09$	72.32
10	Oktober	2010	82	$S_{10+1} = (0.1(82) + (1-0.1) \times 77.09)$ $S_{11} = 8.2 + (0.9 \times 77.09)$ $S_{11} = 8.2 + 69.38 = 77.58$	77.09
			JUMLAH =	763	

Jadi hasil perhitungan *exponential smoothing* untuk peramalan penjualan bulan ke 11 atau bulan November adalah 77.58 atau dibulatkan menjadi 78 bungkus.

3. Tahap Ketiga menentukan jumlah n (ordo) yang digunakan untuk perhitungan *moving average*.

$$F_{t+1} = \frac{X_t + X_{t-1} + \dots + X_{t-n+1}}{n}$$

Dengan nilai n = 3 bulanan

No	Bulan	Tahun	Penjualan (X _t) bungkus	Nilai Peramalan dengan n=3
1	Januari	2010	75	
2	Februari	2010	60	
3	Marat	2010	85	$= (75 + 60 + 85) / 3$ $= 73.33$
4	April	2010	60	$= (60 + 85 + 60) / 3$ $= 68.33$
5	Mei	2010	76	$= (85 + 60 + 76) / 3$ $= 73.67$
6	Juni	2010	64	$= (60 + 76 + 64) / 3$ $= 66.67$
7	Juli	2010	52	$= (76 + 64 + 52) / 3$ $= 64.00$
8	Agustus	2010	89	$= (82 + 52 + 89) / 3$ $= 74.33$
9	September	2010	120	$= (52 + 89 + 120) / 3$ $= 87.00$
10	Oktober	2010	82	$= (89 + 120 + 82) / 3$ $= 97.00$
			JUMLAH =	763

Jadi hasil perhitungan *Moving average* dengan n= 3, untuk peramalan penjualan bulan ke 11 atau bulan November adalah 97.00 bungkus.

4. Tahap Keempat menghitung nilai *Mean Absolut Deviation* (MAD) atau kesalahan peramalan pada hasil perhitungan *exponential smoothing* dan *moving average*.

MAD = (Nilai Forecast - Nilai Nyata)

Menghitung nilai MAD *exponential smoothing* $\frac{n}{\Sigma X}$ Diketahui hasil St atau peramalan untuk bulan ke -11 adalah 77.58 dan nilai nyatanya X_t adalah 82, $n=10$ dan $\Sigma X=763$. maka nilai MAD sebagai berikut ;

$$\begin{aligned} MAD_{es} &= \frac{10}{763} (82 - 77.58) \\ &= 0.013 \times 4.42 \\ &= \mathbf{0.05} \end{aligned}$$

Jadi nilai MAD untuk *exponential smoothing* adalah **0.05**.

Menghitung nilai MAD *moving average*. Diketahui hasil St atau peramalan untuk bulan ke -11 adalah 97 dan nilai nyatanya adalah 82, $n=10$ dan $\Sigma X=763$ maka nilai MAD sebagai berikut ;

$$\begin{aligned} MAD_{ma} &= \frac{10}{763} \times (97 - 82) \\ &= 0.013 \times 15 \\ &= \mathbf{0.18} \end{aligned}$$

Jadi nilai MAD untuk *moving average* adalah **0.18**.

5. Tahap Kelima menentukan nilai MAD paling kecil.

Dari perhitungan diatas dihasilkan nilai MAD untuk *exponential smoothing* adalah 0.05 dan nilai MAD *moving average* adalah 0.18. Jadi nilai Mean Absolut Deviation (MAD) atau kesalahan peramalan yang paling kecil adalah MAD *exponential smoothing* = 0.05. Dengan demikian peramalan untuk penjualan bulan berikutnya atau bulan ke-11 (November) lebih baik menggunakan hasil perhitungan *exponential smoothing*. Hasil peramalan penerapan data mining dengan metode *time series* itu dapat dilihat pada gambar 7.

HASIL PERAMALAN PENJUALAN					
Laporan Peramalan Penjualan bulanan Dari : Januari-2010 Sampai : Oktober-2010					
Nama Barang : SARIMI BASO SAPI RES					
No	Bulan	Tahun	Penjualan	Exponential Smoothing	Moving Average
1	1	2010	75		
2	2	2010	60	75.00	
3	3	2010	65	73.50	
4	4	2010	60	74.65	66.33
5	5	2010	75	73.18	73.67
6	6	2010	64	73.47	65.67
7	7	2010	52	72.52	64.00
8	8	2010	55	70.47	66.33
9	9	2010	120	72.52	67.00
10	10	2010	62	77.00	97.00

Peramalan Exponential Smoothing Periode Berikutnya : 77.58 Unit, Nilai MAD : 0.05
 Peramalan Moving Average Periode Berikutnya : 97.00 Unit, Nilai MAD : 0.18
 Lebih baik menggunakan Exponential Smoothing

Gambar 7 Hasil Peramalan Sarimi Baso Sapi

5 Kesimpulan

1. Penerapan Data Mining dengan metode *time series* memberikan informasi peramalan penjualan diwaktu mendatang kepada pemilik *minimarket* Idola yang dihasilkan dari penggalan *database* penjualan.

2. Perhitungan *exponential smoothing* dan *moving average* menghasilkan peramalan penjualan yang tidak jauh berbeda dengan nilai penjualan yang terjadi. Tingkat ketepatannya ditentukan dengan nilai MAD yang paling kecil dari dua perhitungan tersebut.
3. *Mean Absolute Deviation* (MAD) digunakan untuk mengukur ketepatan nilai atau kesalahan peramalan. Semakin kecil nilai MAD semakin kecil nilai kesalahannya.

Sebagai saran, sistem ini dapat dikembangkan dengan menambahkan beberapa metode perhitungan pada *time series*, selain *exponential smoothing* dan *moving average* agar hasil peramalan penjualan memiliki pilihan perhitungan yang lebih akurat dan mengetahui jenis pola penjualan setiap barang.

6. Referensi

- [1] Arief, M.Rudiyanto, Pemrograman Web Dinamis Menggunakan PHP dan MySQL, CV. Andi Offset, Yogyakarta, 2011.
- [2] Hadi, Sutrisno, MA, *Statistik Jilid 3*, CV. Andi Offset. Yogyakarta, 2002.
- [3] HM.,Jogiyanto, Analisis dan Desain Sistem Informasi: Pendekatan Teori dan Praktek Aplikasi Bisnis, CV. Andi Offset, Yogyakarta, 2005.
- [4] R.Spiegel, Murray dan Stephens, Larry, Schaums Outlines Teori dan Soal- Soal Statistik Edisi ke-3, Erlangga, Jakarta, 2007.
- [5] Sihombing, Erika, Implementasi Data Mining Menggunakan Metode Apriori Pada Transaksi Di Chorus Minimarket, Compiler, Volume 1 Nomor 1. Mei 2012.
- [6] Turban, Efram, Aronson, Jay E, dan Peng-Liang, Ting, *Decision Support Systems and Intelligent Systems* (Sistem Pendukung Keputusan dan Sistem Cerdas), CV. Andi Offset, Yogyakarta, 2005.