BAB II

TINJAUAN PUSTAKA

2.1 Peramalan

Peramalan adalah kegiatan untuk memperkirakan apa yang akan terjadi pada masa yang akan datang [3]. Peramalan (*forecasting*) merupakan bagian vital bagi setiap organisasi bisnis dan untuk setiap pengambilan keputusan manajemen yang sangat signifikan.

Peramalan Deret Waktu atau *Time Series Forecasting* adalah serangkaian pengamatan terhadap suatu variabel yang diambil dari waktu ke waktu dan dicatat secara berurutan menurut urutan waktu [4].

2.2 Pola Data Musiman (*Seasonal)*

Pola data musiman adalah salah satu dari 4 bentuk pola data deret waktu yang menunjukkan perulangan indeks pada kurun waktu tertentu seperti hari, minggu, bulan atau kuartal [5]. Variasi musiman adalah fluktuasi yang muncul secara reguler setiap tahun yang biasanya disebabkan oleh iklim, kebiasaan (mempunyai pola tetap dari waktu ke waktu). Contoh yang menunjukan variasi musiman seperti penjualan pakaian akan meningkat pada saat hari raya, penjualan buku dan tas sekolah akan meningkat pada saat awal sekolah.

2.3 *Time Series* Model

Peramalan merupakan alat bantu yang penting dalam perencanaan yang efektif dan efisien. Menurut [6], teknik peramalan terbagi menjadi dua bagian, yang pertama metode peramalan subjektif dan metode peramalan objektif. Metode peramalan subjektif mempunyai model kualitatif dan metode peramalan objektif mempunyai dua model, yaitu model *time series* dan model kausal. Model kualitatif berupaya memasukkan faktor-faktor subyektif dalam model peramalan, model ini akan sangat bermanfaat jika data kuantitatif yang akurat sulit diperoleh. Contoh dari metode ini ialah metode delphi, opini juri eksekutif, komposit kekuatan dan survei pasar konsumen.

Model *time series* merupakan model yang digunakan untuk memprediksi masa depan dengan menggunakan data historis. Dengan kata lain, model *time series* mencoba melihat apa yang terjadi pada suatu kurun waktu tertentu dan menggunakan data masa lalu untuk memprediksi [6]. Contoh dari model *time series* ini antara lain *Moving average*, *Exponential Smoothing* dan proyeksi trend.

2.4 Penelitian Sebelumnya

Beberapa jurnal yang telah dipublikasikan terkait untuk peramalan jumlah penderita atau jumlah kasus penyakit menggunakan metode *sarima* antara lain sebagai berikut :

1. Pada tahun 2008 Ika Milasari melakukan penelitian dengan judul Peramalan Jumlah Penderita Demam Berdarah Menggunakan Model Arima Musiman (Studi Kasus RSU Sidoarjo). Metode arima musiman

atau *sarima* menghasilkan peramalan untuk jumlah penderita demam berdarah 2 periode kedepan [2].

2. Pada tahun 2017 Slamet Widodo melakukan penelitian dengan judul Peramalan Jumlah Kasus Tuberkulosis di Kabupaten Malang Menggunakan Metode *Autoregressive Integrated Moving Average*. Pada penelitian ini menghasilkan kesimpulan arima (0,1,3) adalah model terbaik untuk peramalan jumlah kasus tuberkulosis [1].

Dengan melihat kedua penelitian ini maka metode *sarima* digunakan untuk peramalan jumlah kasus penyakit.

2. 5 Metode *Seasonal* *Autoregressive Integrated Moving Average*

***SARIMA* adalah metode peramalan** *Time Series* untuk data model stokastik dengan pola data musiman [4]. Metode *SARIMA* dapat dipakai untuk data yang bersifat musiman yang dapat memberikan hasil peramalan yang tidak jauh beda dengan data aktual.

Langkah-langkah pemodelan *Seasonal Autoregressive Integrated Moving Average* (*SARIMA*) adalah [7]:

1. Proses identifkasi model.
2. Pendugaan parameter model.
3. Pemeriksaan residual(sisaan).
4. Mengunakan model untuk peramalan jika model memenuhi syarat.

Secara umum, notasi *SARIMA* adalah :

ARIMA $(p,d,q)\ (P,D,Q)^s$ (1)

Dimana :

p,d,q : Bagian tidak musiman dari model

$(P,D,Q)^s$: Bagian musiman dari model

S : Jumlah periode per musim

Rumus umum dari ARIMA $(p,d,q)(P,D,Q)^s$ adalah sebagai berikut [7]:

$$\Phi_P(B^S)\phi_p(B)(1-B)^d(1-B^S)^D Z_t = \theta_q(B)\Theta_q(B^S)a_t \quad (2)$$

dengan :

$\phi_p(B)$: AR *Non Seasonal*

$\Phi_P(B^S)$: AR *Seasonal*

$(1-B)^d$: *differencing non seasonal*

$(1-B^S)^D$: *differencing seasonal*

$\theta_q(B)$: MA *non seasonal*

$\Theta_q(B^S)$: MA *seasonal*

Tahap differencing adalah olah data dengan menghitung nilai perbedaan antara Zt dengan Zt-1. Data dihitung selisihnya dengan data periode sebelumnya. Hal ini dilakukan untun menghilangkan unsur tren pada data.

Dalam menetapkan nilai p dan q dapat dibantu dengan mengamati pola *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF) dengan acuan pada tabel 2.1 [8].

Tabel 2.1 Pola ACF dan PACF untuk identifikasi parameter tentatif

Model	**Pola ACF**	**Pola PACF**
AR (p)	Menurun secara cepat (*dies down*)	Muncul *spike* yang signifikan hingga *lag* ke-p dan *cut off* setelah *lag* ke-p
MA (q)	muncul *spike* yang signifikan hingga *lag* ke-q dan *cut off* setelah *lag* ke-q	menurun secara cepat (*dies down*)
ARMA (p,q)	menurun secara cepat (*dies down*)	menurun secara cepat (*dies down*)
AR (p) atau MA (q)	muncul *spike* yang signifikan hingga *lag* ke-q dan *cut off* setelah *lag* ke-q	muncul *spike* yang signifikan hingga *lag* ke-p dan *cut off* setelah *lag* ke-p

Tahapan pemeriksaan residual terbagi menjadi pemeriksaan L-jung Box (*white noise),* dan kenormalan residual. Pengujian untuk memenuhi asumsi *white noise* dilakukan pengujian L-jung Box dengan hipotesa [9] :

$H_0 : \rho_1 = \rho_2 = \rho_{...} = \rho_k = 0$

H_1 : minimal ada satu nilai $\rho_k \neq 0$, dimana k = 1,2,...,K

dengan statistik uji :

$$Q = n(n+2)\sum_{k=1}^{K}(n-k)^{-1}\,\hat{\rho}_k \quad (3)$$

Dimana n adalah banyak pengamatan dan $\hat{\rho}_k$ adalah sampel ACF residual pada lag ke-k, dan K adalah lag maksimum. Daerah kritis = Q > $x^2_{(\alpha,K-m)}$ atau *P-value* < *α =5%.*

Pengujian kenormalan distribusi dapat dilakukan uji Kolmogorov-Smirnov dengan hipotesa [10]:

H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

Dengan statistik uji :

$$D = \overset{Sup}{x}[|F_n(x) - F_0(x)|] \quad (4)$$

Dimana D adalah nilai deviasi absolut maksimum antara F (x) n dan ( ) 0 F x , Sup masing-masing merupakan fungsi *Kolmogorov* peluang kumulatif yang dihitung dari data sampel, fungsi peluang kumulatif distribusi normal, dan nilai supremum untuk semua α_t . Daerah Kritis: Tolak H0 jika $D \geq D_{(1-\alpha,n)}$ atau *P-value* < α , dengan α = 5% [9].

Model terbaik ditentukan berdasarkan nilai galat, semakin kecil nilai galat yang dihasilkan maka semakin baik model tersebut sekaligus menentukan model tersebut layak untuk digunakan pada peramalan [11].

2.6 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 :

2.6.1 Mean Square Error (MSE)

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 data aktual. MSE merupakan rata-rata selisih kuadrat antara nilai yang diramalkan dan yang diamati. Kekurangan penggunaan MSE adalah bahwa metode evaluasi ini cenderung menonjolkan nilai deviasi yang besar karena adanya pengkuadratan [12].

$$MSE = \frac{\sum(kesalahan\ peramalan)^2}{n}$$

dimana :

n = jumlah periode data

2.6.2 Mean Absolute Deviation (MAD)

Mean Absolute Deviation adalah ukuran kesalahan peramalan keseluruhan untuk sebuah model. Nilai MAD dihitung dengan mengambil jumlah nilai absolut dari kesalahan peramalan dibagi dengan jumlah periode data (n) [12].

$$MAD = \frac{\sum|data\ aktual - peramalan|}{n}$$

dimana :

n = jumlah periode data

2.6.3 Mean Absolute Percentage Error (MAPE)

Masalah yang sering muncul pada perhitungan akurasi dengan metode perhitungan MSE dan MAD adalah nilainya bergantung pada besarnya nilai dari *item* yang diramalkan. Apabila satuan dari item yang diramalkan adalah ribuan atau jutaan, nilai dari MAD dan atau MSE bisa jadi sangat besar. Untuk menghindari hal ini, dapat digunakan metode penghitungan akurasi dengan menghitung *meanabsolute percentage error* atau MAPE. MAPE adalah nilai rata-rata dari selisih absolut antara nilai hasil peramalan dan nilai aktual, nilai ditunjukkan dalam persentase [13].

$$\text{MAPE} = \sum_{t=1}^{T} \frac{|\text{peramalan} - \text{aktual}|}{n} * 100$$