

FORECASTING MOVING AVERAGE PADA DATA TIME SERIES KOTA BANDAR LAMPUNG

Rinaldi Revivo

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara
Jl. Letjen S. Parman No.1, Jakarta Barat, DKI Jakarta, Indonesia 11410
e-mail: rinaldi.535199101@stu.untar.ac.id

ABSTRAK

Perubahan iklim yang terjadi pada saat ini telah menimbulkan banyak permasalahan bagi lingkungan sekitar. Hal ini dapat dibuktikan dengan terjadinya kenaikan yang signifikan pada temperatur cuaca khususnya di Kota Bandar Lampung. Penelitian ini bertujuan untuk memprediksi data *time series* temperatur cuaca di Kota Bandar Lampung dengan metode *Forecasting*. Metode ini dikombinasikan dengan teknik *Moving Average* dengan 3 fitur pada data temperatur yang tersedia yaitu Temperatur Minimum (T_n), Temperatur Maksimum (T_x), dan Temperatur Rata-rata (T_{avg}). Kemudian, hasil yang diperoleh akan dievaluasi dengan teknik *Mean Square Error* (MSE) dan *Root Mean Square Error* (RMSE) lalu dibandingkan dengan data yang sudah ada. Nilai MSE dan RMSE menunjukkan perbedaan yang tergolong kecil yakni untuk T_n masing-masing dengan nilai 0.74 dan 0.86, T_x memiliki perbedaan cukup jauh sekitar 0.36 dan 0.60, sedangkan T_{avg} mencapai nilai persis 100. Penelitian ini membutuhkan eksperimen lebih lanjut karena adanya peluang bias pada data, sehingga disarankan untuk menggunakan metode lain maupun menambahkan teknik penanganan *missing value* atau teknik lainnya agar hasil yang diperoleh menjadi lebih akurat dan teruji.

Kata kunci: Peramalan, Temperatur, Iklim, Moving Average, Pembelajaran Mesin.

ABSTRACT

The current climate change has caused many problems for the surrounding environment. This can be proven by the significant increase in weather temperature, especially in Bandar Lampung City. This study aims to predict weather temperature time series data in Bandar Lampung City using the Forecasting method. This method is combined with the Moving Average technique with three features in the available temperature data, namely Minimum Temperature (T_n), Maximum Temperature (T_x), and Average Temperature (T_{avg}). Then, the results obtained will be evaluated using the Mean Square Error (MSE) and Root Mean Square Error (RMSE) techniques and compared with existing data. The MSE and RMSE values show relatively small differences, namely 0.74 and 0.86 for T_n , a significant difference of around 0.36 and 0.60 for T_x , while T_{avg} reached a value of exactly 100. This study requires further experimentation due to the possibility of bias in the data. Therefore, it is recommended to use other methods or add techniques for handling missing values or other techniques so that the results obtained are more accurate and reliable.

Keywords: Forecasting, Temperature, Climate, Moving Average, Machine Learning.

1. PENDAHULUAN

Kota Bandar Lampung adalah ibukota dari Provinsi Lampung yang berada di pulau Sumatra. Pulau Lampung merupakan pintu gerbang utama pulau Sumatera yang memiliki peran penting dalam hal-hal seperti perekonomian, pendidikan, dan kebudayaan. Sebagai gerbang utama, kegiatan-kegiatan yang dilakukan seperti distribusi barang melalui perairan Selat Sunda akan terpengaruhi oleh kondisi cuaca. Dengan demikian, prediksi cuaca di hari-hari berikutnya akan menjadi peran penting dalam kelancaran proses kegiatan.

Cuaca merupakan keadaan udara di suatu tempat pada waktu dan wilayah yang relatif sempit. Banyak unsur-unsur yang mempengaruhi cuaca yaitu suhu udara, tekanan udara, kelembapan udara, kecepatan angin, lama penyinaran matahari dan curah hujan. Cuaca yang sering terjadi perubahan akan mempengaruhi kegiatan manusia sehingga setiap harinya, manusia mencatat unsur-unsur tersebut yang nantinya dapat digunakan untuk memprediksi hasil cuaca keesokan harinya [1].

Metode peramalan atau (*Forecasting*) adalah aktivitas/kegiatan untuk melakukan estimasi kondisi yang akan terjadi di masa depan dengan menggunakan data. *Forecasting* dapat dijadikan sebagai dasar suatu rencana jangka pendek, menengah, atau panjang. Metode *forecasting* dapat dibagi menjadi 2 yaitu metode peramalan kuantitatif dan metode peramalan kualitatif. *Time Series* merupakan himpunan observasi data terurut dalam waktu. Untuk data *time series* ini, metode peramalan yang digunakan adalah metode peramalan kuantitatif [2]. Sedangkan, metode *Moving Average* adalah teknik peramalan *time-series* yang menyederhanakan fluktuasi jangka pendek dengan menghitung rata-rata dari sejumlah observasi terakhir, lalu menggunakan rata-rata itu sebagai nilai ramalan untuk periode berikutnya [3].

2. TINJAUAN LITERATUR

Penelitian sebelumnya yang relevan dengan penelitian ini adalah terkait hubungan curah hujan dengan kasus demam berdarah dengue di Kota Bandar Lampung pada tahun 2016-2018 yang menyimpulkan bahwa dari hasil uji korelasi yang ditampilkan menunjukkan tingkat signifikansi antara curah hujan dan insiden DHF dengan kekuatan korelasi rata-rata yang positif [4]. Tujuan yang ingin dicapai dalam penelitian ini adalah untuk meramal data *time series* dari kota suhu terendah, suhu tertinggi, dan suhu rata-rata di Kota Bandar Lampung. Hasil prediksi dapat digunakan untuk dijadikan perbandingan dengan hasil pada metode lain selain metode *forecasting* dengan *moving average*.

Analisis *time series* mencakup proses autoregresif atau rata-rata bergerak (*moving average*), serta pembahasan mengenai pengaruh korelasi deret waktu terhadap jenis-jenis inferensi statistik lainnya, seperti estimasi rata-rata dan koefisien regresi. Tujuan utamanya adalah untuk mengembangkan model yang sesuai untuk menggambarkan struktur inheren dari deret tersebut yang kemudian dapat digunakan untuk menghasilkan nilai-nilai masa depan dari deret tersebut (yaitu untuk membuat perkiraan) [5].

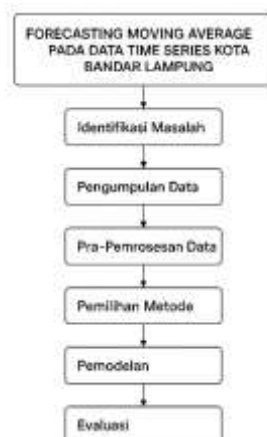
Untuk menganalisis pengaruh penggunaan data learning yang berbeda, maka perlu dilakukannya evaluasi terhadap hasil peramalan tersebut dari kedua perlakuan yang telah ditetapkan. Hal ini dilakukan untuk melihat perlakuan mana yang lebih memberikan selisih peramalan dan aktual yang paling kecil, sehingga dapat disimpulkan perilaku yang mana yang lebih tepat digunakan untuk melakukan peramalan cuaca. Adapun metode yang sering digunakan dalam melakukan pengujian metode *forecasting* salah satunya adalah metode *Mean Square Error* (MSE) dan *Root Mean Square Error* (RMSE) [5].

3. METODE PENELITIAN

Jenis algoritma *forecasting* yang digunakan adalah algoritma *Moving average*. Teknik ini memanfaatkan kesalahan *forecast* sebelumnya dengan pendekatan regresi untuk meramalkan pengamatan di masa yang akan mendatang. Teknik evaluasi untuk algoritma kali ini adalah RME dan RMSE. *Root Mean Square Error* (RMSE) dan *Mean Square Error* (MSE) teknik evaluasi yang menilai tingkat kesalahan dari hasil prediksi dimana semakin besarnya nilai, semakin tinggi tingkat kesalahan. Alur dari penelitian ini dapat di lihat pada Gambar 1.

3.1 Dataset

Dalam penelitian Dataset yang digunakan merupakan data iklim yang diambil dari situs web BMKG yaitu <https://www.bmkg.go.id> dengan atribut arah angin saat kecepatan maksimum (*ddd_x*), arah angin terbanyak (*ddd_car*), curah hujan (*RR*), kecepatan angin maksimum (*ff_x*), kecepatan angin rata-rata (*ff_avg*), kelembapan rata-rata (*RH_avg*), lamanya penyinaran matahari (*ss*), temperatur maksimum (*Tx*), temperatur minimum (*Tn*), temperatur rata-rata (*Tavg*). Atribut yang digunakan adalah temperatur minimum, temperatur rata-rata dan temperatur maksimum.



Gambar 1. Alur Penelitian

3.2 Pra-pemrosesan Data

Data iklim yang diambil dari situs web BMKG memiliki *missing values* pada tanggal dan atribut tertentu. Data yang memiliki missing value diproses dengan cara jika total *missing values* kurang dari 20%, maka data numerik diisi dengan nilai rata-rata 7 hari terakhir. Sedangkan data yang nilainya bukan numerik diisi dengan nilai modulus dari 7 hari terakhir. Namun, jika total *missing values* lebih dari 20%, data numerik akan diisi dengan nilai rata-rata dibulan yang sama dari tahun-tahun sebelumnya dan data yang nilainya bukan numerik diisi dengan nilai modulus dari data di bulan yang sama dari tahun-tahun sebelumnya.

Setelah menangani data *missing values*, dilakukan normalisasi. Normalisasi digunakan untuk menyamakan jangkauan data setiap atribut yang mempunyai jangkauan yang berbeda. Normalisasi yang digunakan dalam proses ini adalah dengan *Standard Scaler*. Pada eksperimen kali ini, dataset yang didapat dari BMKG tidak berurutan sehingga data harus disortir terlebih dahulu menurut tanggal. Berikut Gambar 2 merupakan contoh datanya.

	Tanggal	Tn	Ts	Tavg	md_avg	rh	ss	ff_s	ddd_s	ff_avg	ddd_car
0	2010-08-01	20.0	31.7	29.2	76.0	0.3	0.0	4.0	312.0	2.0	SW
1	2010-08-02	21.0	31.5	28.1	76.0	17.4	7.4	6.0	315.0	1.0	NDW
2	2010-08-03	20.0	31.4	29.1	76.0	0.0	9.0	7.0	130.0	3.0	SE
3	2010-08-04	21.0	31.4	28.0	79.0	3.6	1.0	3.0	90.0	2.0	N
4	2010-08-05	21.0	30.6	27.4	87.0	0.0	1.0	3.0	300.0	1.0	N
5	2010-08-06	20.0	31.2	28.0	84.0	46.0	1.0	6.0	90.0	3.0	E
6	2010-08-07	20.0	29.4	26.4	90.0	0.0	0.0	3.0	180.0	3.0	S
7	2010-08-08	20.0	31.2	27.6	83.0	1.0	0.0	3.0	220.0	0.0	S
8	2010-08-09	20.0	31.4	28.3	94.0	0.0	0.0	6.0	180.0	2.0	SE
9	2010-08-10	19.0	32.2	28.5	84.0	0.0	4.0	3.0	180.0	3.0	SE

Gambar 2. Bentuk Dataset

Data yang akan diproses adalah data temperatur minimum, temperatur maksimum, dan temperatur rata-rata sehingga data yang tidak digunakan akan dihilangkan terlebih dahulu sehingga tersisa hanya data temperatur. Berikut Gambar 3 adalah bentuk final dari dataset sebelum diprediksi.

	Tanggal	Tn	Ts	Tavg
0	2010-01-01	20.0	31.7	29.2
1	2010-01-02	21.0	31.5	28.1
2	2010-01-03	20.0	31.4	29.1
3	2010-01-04	21.0	31.4	28.0
4	2010-01-05	21.0	30.6	27.4
5	2010-01-06	20.0	31.2	28.0
6	2010-01-07	20.0	29.4	26.4
7	2010-01-08	20.0	31.2	27.6
8	2010-01-09	20.0	31.4	28.3
9	2010-01-10	19.0	32.2	28.5

Gambar 3. Bentuk Final Dataset

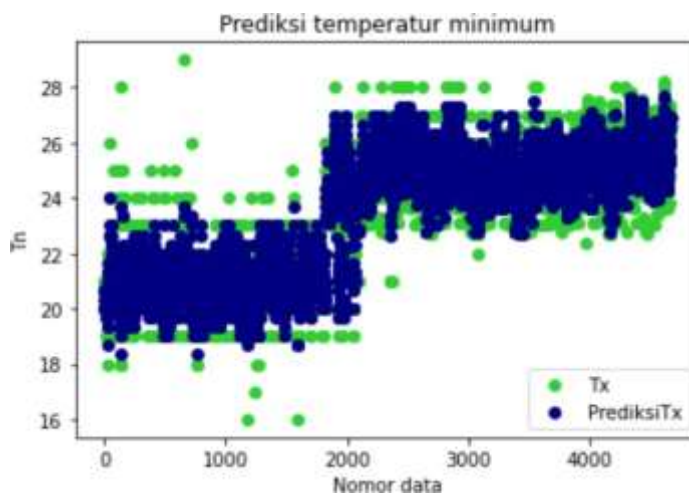
4. HASIL DAN PEMBAHASAN

Dari data di atas, prediksi dihasilkan dengan moving average. Contoh hasil prediksi temperatur minimum, temperatur rata-rata, dan temperatur maksimum dapat dilihat pada Gambar 4 berikut.

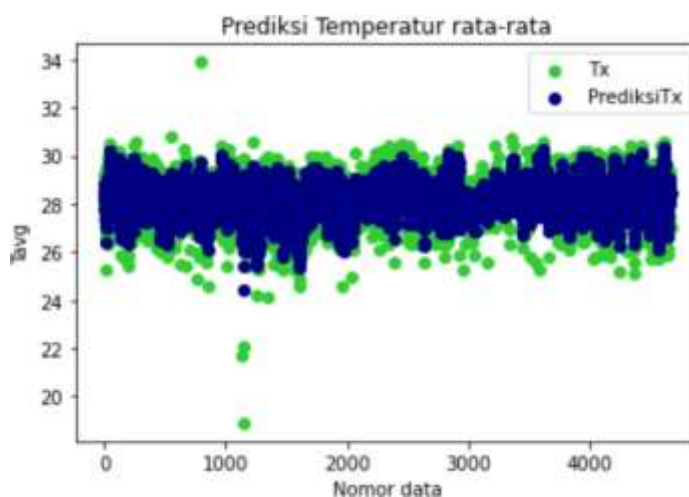
PrediksiTn	PrediksiTavg	PrediksiTx
20.000000	28.500000	31.500000
20.000000	28.500000	31.500000
20.333333	28.800000	31.533333
20.666667	28.400000	31.433333
20.666667	28.166667	31.133333
...
26.066667	28.500000	31.933333
26.466667	28.600000	32.266667
26.866667	28.900000	32.533333
27.000000	29.066667	32.766667
26.933333	28.466667	31.766667

Gambar 4. Hasil Prediksi Temperatur dengan *Moving Average*

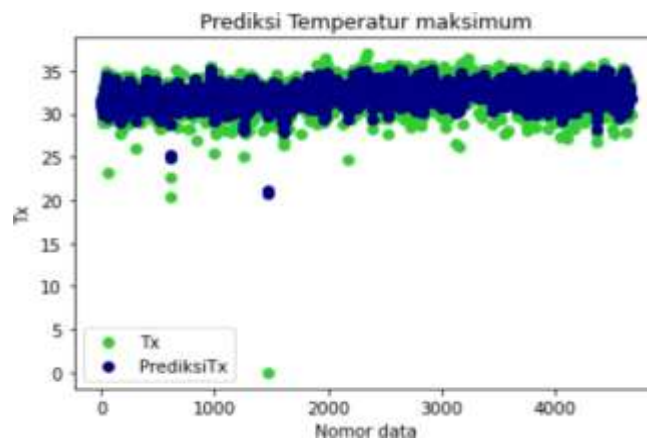
Data sebenarnya dan data prediksi jika ditunjukkan dalam bentuk *scatterplot* akan terlihat seperti pada Gambar 5, Gambar 6, dan Gambar 7 berikut.



Gambar 5. *Scatterplot* Temperatur Minimum



Gambar 6. *Scatterplot* Temperatur Rata-rata



Gambar 7. Scatterplot Temperatur Maksimum

Hasil prediksi dievaluasi menggunakan *Mean Squared Error* (MRE) dan *Root Mean Squared Error* (RMSE). Hasil evaluasi dapat dilihat pada Gambar 8 berikut.

```
MSE dan RMSE Tn,  
MSE : 0.7472358072358073  
RMSE : 0.8644280231666528  
  
MSE dan RMSE Tavg,  
MSE : 1.0001506268172935  
RMSE : 1.0000753105728055  
  
MSE dan RMSE Tx,  
MSE : 0.3698972782306116  
RMSE : 0.6081918103942304
```

Gambar 8. Hasil Evaluasi

Hasil Mean Square Error dan Root Mean Square Error terkecil ada pada atribut temperatur maksimum sementara hasil Mean Square Error dan Root Mean Square Error terbesar ada pada temperatur rata-rata. Terdapat perbedaan MSE dan RMSE antar temperatur. Pada temperatur maksimum, perbedaannya sebesar 0.238, pada temperatur minimum, perbedaannya sebesar 0.117, dan pada temperatur rata-rata, perbedaannya sebesar 0.0001. Metode *moving average* yang dijalankan memiliki hasil akurasi terbilang cukup baik. Namun, nilai evaluasi hasil prediksi tidak dapat memprediksi data temperatur yang perbedaan datanya berada pada rentang yang jauh pada suatu hari tetapi perbedaan hari sebelum dan sesudahnya cenderung dekat, sehingga menyebabkan peningkatan drastis angka MSE dan RMSE.

5. KESIMPULAN

Setelah melakukan *forecasting* dengan metode *moving average* dapat disimpulkan bahwa untuk menganalisis data cuaca *time series* di Kota Bandar Lampung memerlukan perbaikan karena hasil yang tergolong sudah cukup baik namun masih memiliki masalah. Permasalahan dari metode ini adalah jika ada sedikit data yang perbedaannya terlalu jauh, maka data tersebut tidak dapat diprediksi. *Moving average* memberikan hasil yang baik pada data yang konsisten. Saran untuk penelitian selanjutnya dengan tema yang sama yaitu *forecasting* data *time series* pada Kota Bandar Lampung ini dapat mencoba bereksperimen dengan menggunakan metode-metode lain yang lebih modern dan menambahkan lebih banyak data agar mencapai hasil yang lebih variatif namun juga tetap akurat.

DAFTAR PUSAKA

- [1] L. Susanti, P. Hasanah dan W. , “Peramalan Suhu Udara dan Dampaknya Terhadap Konsumsi Energi Listrik di Kalimantan Timur,” *Barekeng: Jurnal Ilmu Matematika dan Terapan*, vol. 3, no. 14, pp. 397-410, 2020.
- [2] S. dan A. Rohmah, “Penerapan Jaringan Syaraf Tiruan Metode Hopfield Untuk Prediksicuaca Di Kota Palangka Raya,” *Jurnal Sains Komputer dan Teknologi Informasi*, vol. 1, no. 3, pp. 52-59, 2020.
- [3] A. D. Achmad, “Metode Moving Average dan Metode Support Vector Machine Untuk Prediksi Variabel Meteorologi,” *JTRISTE*, vol. 1, no. 4, pp. 45-50, 2017.
- [4] T. Triwahyuni, I. Husna dan M. Andesti, “Hubungan Curah Hujan dengan Kasus Demam Berdarah Dengue di Bandar Lampung 2016-2018,” *ARTERI: Jurnal Ilmu Kesehatan*, vol. 3, no. 1, pp. 184-189, 2020.
- [5] Z. Ivanovski, A. Milenkovski dan Z. Narasanov, “Time Series Forecasting Using A Moving Average Model for Extrapolation of Number of Tourist,” *UTMS Journal of Economics*, vol. 2, no. 9, pp. 121-132, 2018.
- [6] F. N. Adnan dan A. Nahrul, “Analisis Penentuan Data Latih pada PeramalanCurah Hujan Menggunakan Metode Simple Moving Average,” *Journal of Information System*, vol. 1, no. 3, pp. 49-58, 2018.