SPEKTRA: Jurnal Fisika dan Aplikasinya
Volume 9, Issue 1, April 2024
Homepage: doi.org/10.21009/SPEKTRA

DOI: doi.org/10.21009/SPEKTRA.091.04

Rainfall Classification Analysis Using Naïve Bayes Classifier Based on Air And Wind Temperatures in Serang City

Meilani Nisa Abdilah*, Yayat Ruhiat, Yudi Guntara

Physics Education, Teacher Training and Education, Sultan Ageng Tirtayasa University, Serang, Indonesia

*Corresponding Author Email: nisameilani023@gmail.com

Received: 9 March 2024
Revised: 21 April 2024
Accepted: 25 April 2024
Online: 27 April 2024
Published: 30 April 2024

SPEKTRA: Jurnal Fisika dan Aplikasinya
p-ISSN: 2541-3384
e-ISSN: 2541-3392

ABSTRACT

The city of Serang experiences relatively high annual rainfall, with an average total of more than 100 mm/year. Based on data obtained from BMKG, Serang City, in 2023, there will be a shift in the rainy season due to weather anomalies, affecting the amount of rainfall. Apart from that, Serang City is one of the cities with less rain throughout September, recorded within 15 days. In this case, a determination is needed in the form of rainfall classification. However, an exemplary method is required in order to classify rainfall so that the classification results are accurate. Several studies showed that the naïve Bayes classifier method is the best classification method compared to other analysis methods; namely, it only requires a probability. The parameters used are air temperature and wind speed. So, this research aims to determine the classification of rainfall using the Naïve Bayes classifier based on air temperature and wind in Serang City. The method used is non-experimental quantitative with naïve Bayes classifier analysis. Based on the data analysis using Microsoft Excel software, the results showed that in the Serang city area, February 2024 was classified as a humid month with rainfall of 100 - 200 mm, and March 2024 was classified as a dry month with rainfall <100 mm.

Keywords: rainfall classification, naïve Bayes classifier

INTRODUCTION

Serang City is one of the cities located in Banten Province, Indonesia. Geographically, the city of Serang is situated between 106°07' and 106°25' East Longitude, and South Latitude ranges from 5°99' to 6°22. The area of Serang City is 266.7 km^2. Most of the city of Serang is relatively low, ranging from 0 to 50 meters from sea level. In climate classification, the Serang

City area has an average tropical monitor climate (AM), with the rainy and dry seasons as two different season patterns influenced by monsoon wind movements. But in 2023, there was a shift in season time, which impacted rainfall in Serang. According to the Central Statistics Agency of Serang City (2021), Serang City experienced a relatively high annual rainfall, with an average total of more than 100 mm. Rainfall is the amount of water that falls on a horizontal or horizontal surface within a certain period and is measured in mm units [1]. Rain is mainly produced by rising air and decreased temperature [2]. The size of the rainfall will affect human activity from various fields [3]. Based on information obtained through one of the news media, the occurrence of weather anomalies in 2023, which affected rainfall in the city of Serang, including the lack of rain throughout September, was recorded within 15 days [4]. In this case, a determination is needed in the form of rainfall classification. However, an exemplary method is required in order to classify rainfall so that the classification results are accurate.

Based on the results of [5], the average accuracy in the K-Nearest Neighbor method was 63.68%, whereas in the use of Naïve Bayes Classifier, an accuracy value of 79.71%. In this case, the Naive Bayes Classifier method gets better classification accuracy than the KNN method. This is reinforced in the [6] study, which states the Naïve Bayes Classive method is better than other classification methods. Naive Bayes Classifier (NBC) is used as a classification method. Naïve Bayes Classifier, often called Bayesian Classification, is a classification method that uses the type of statistical and probabilistic methods found by British scientist Thomas Bayes [7]. The advantage of this method is that it only requires relatively small training data to predict the determination of the estimation and parameters needed for classification [8]. Because it is assumed to be an independent variable, the Naïve Bayes classification algorithm method only requires variance in a class to determine the classification of the entire covariance matrix. So that it can make it very easy to use [9]. Using the naïve Bayes Classifier is fundamental to probability factors, so to support this method, a variable is needed as a probability that can affect rainfall.

Based on the results of observations in the Geophysical Climatology Meteorology Agency (BMKG) of the Maritime Station in Serang 1, the most critical climate elements in influencing rainfall are air temperature and wind speed because when the temperature causes low pressure so the occurrence of high air evaporation that will affect against the wind in the formation of rain clouds. In this case, there is no research on the classification of rainfall based on air and wind temperatures in the city of Serang; based on this, a rainfall classification analysis will be carried out using the naïve Bayes Classifier based on air and wind temperatures in Serang City.

METHOD

This study uses quantitative methods with research design, namely quantitative non-experimentation. This is because there is no intervention or treatment of data variables. Where the research data source is obtained using the Institutional Survey Method with the type of data in the form of secondary data. The data from the Geophysical Climatology Meteorology Agency (BMKG) of Maritime Class 1 Serang, namely rainfall data and air and wind temperature for ten years from 2014 - 2023. The following is the classification of rainfall according to Oldeman.

TABLE 1. Oldeman Classification

No	Rainfall	Rain classification
1	>200 mm	wet month
2	100 – 200 mm	humid month
3	<100 mm	dry month

The analysis used is naïve bayes classifier. Naïve Bayes classifier is a classification method using probability and statistical methods based on the experience that has occurred. The following is the stage of naïve Bayes classifier.

Average

$$\bar{x} = \frac{\sum xi}{n} \tag{1}$$

Standard deviation

$$\sigma = \sqrt{\frac{\sum (x-\mu)^2}{n}} \tag{2}$$

Class probability

$$p(h \mid x) = \frac{p(x \mid h) \times p(h)}{p(x)} \tag{3}$$

Moving Average

In determining the test data, the most recent data will be used to predict the amount of air temperature and wind speed using the Moving Average calculation. The following is a moving average equation.

$$St = \frac{X_t - X_{t-1} + \cdots + X_{t-n+1}}{n} \tag{4}$$

Rainfall classification

$$p(x_1 \mid x_2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp^{\frac{(x_{12}-\mu)^2}{2\sigma^2}} \tag{5}$$

RESULT AND DISCUSSION

The amount of water that falls at a particular time is called rainfall. Rainfall can also be considered as the amount of rainwater that falls over a specific time, measured in height units. Rainfall patterns in one of the cities in Banten Province, Serang, are prone to monitoring rain. The northwest monsoon wind originating from West Java impacts this rainfall pattern. Every seven months, this precipitation pattern, such as monsoon, creates a cycle of season change. Several factors from other climate elements can cause the size of the rainfall that falls. Based on the results of observations in BMKG, the most dominant climate elements that affect rainfall are air temperature and wind speed. The following are the results of secondary data analysis in the form of monthly rainfall data in Serang City in 2014 - 2023 obtained from BMKG Serang Maritime Meteorology Station in FIGURE 1.

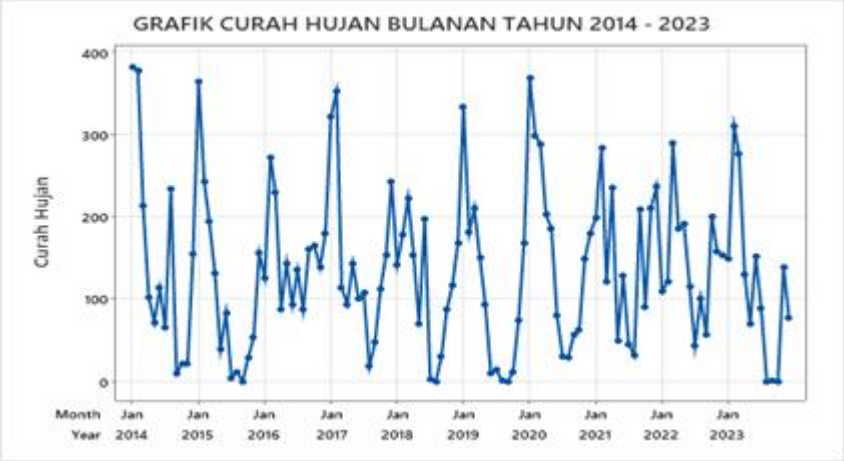

FIGURE 1. Rainfall Graph 2014 - 2023

Based on the graph in FIGURE 1 of the 2014-2023 rainfall in the Serang City area sourced from BMKG Serang Meotorology Station, the average rainfall is 132.7 mm. The maximum rainfall occurred in January 2014 of 376.7 mm. The amount of rainfall that occurs every January is a normal condition because, basically, January - February is the peak of the rainy season, so the amount of rainfall is included in the category of wet months. The minimum rainfall occurred in one of the months in 2018, 2019, and the month of 2023 of 0 mm. Furthermore, the highest rainfall for ten years in the year 2020 was 160.5 mm.

Based on these data, rainfall in 2020 is still in the normal category. The rainy and dry seasons occur according to the condition of the Indonesian atmosphere in general. At the same time, the lowest Average of the lowest annual rainfall was in 2019 of 103.9 mm. In 2019, until the end of the month of rains, it was still shallow; it happened because of the shift in the degree of the earth when orbiting the sun, which caused a reduction in clouds because of the dry season, making the solar radiation higher in Indonesia in the tropics. Bright weather conditions are still quite dominating, so the anomaly of cold temperatures in Indonesian waters causes rain clouds to be challenging to form. Wind patterns are still blowing from the east, which causes the rainy season to be insignificant so that in September 2019, the city of Serang has not entered the rainy season and Still in the dry month category. The size of the rainfall is

influenced by other climate elements, namely air temperature. The following are the graph results showing the relationship between air temperature and precipitation in 2014 - 2023.

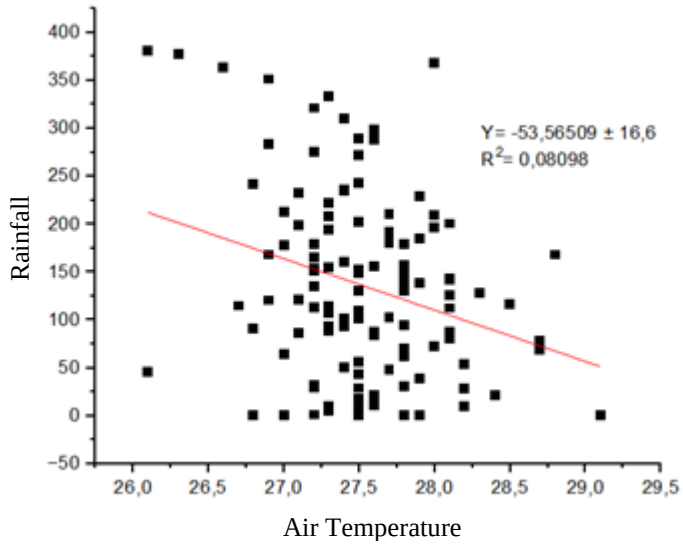

FIGURE 2. Graph of Relationship of Air Temperature and Rainfall in 2014 – 2023

Based on FIGURE 2, the relationship between air temperature and rainfall is obtained by R Squared 0.08098. The results of the R Squared show how many variations a variable is interrelated, with the value received, which means the air temperature can affect rainfall by 8.1%. This means that 91.9% of other factors, including air humidity, affect rainfall. This is by previous research [10], which shows that the relationship between rainfall and air humidity is 57%. In addition to the air temperature, the climate element used in this study is wind speed. Following are the results of the Windrose graph in 2023 Serang City Region.

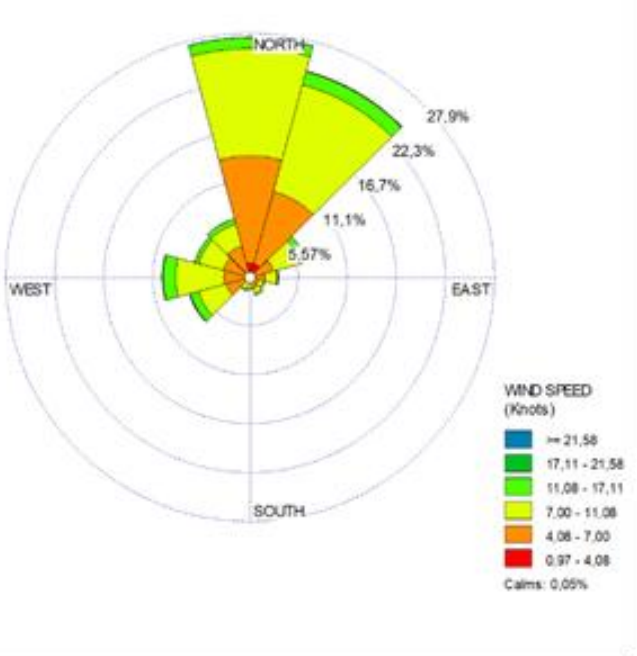

FIGURE 3. Windrose Direction and Wind Speed in 2014 - 2023

Based on FIGURE 3, Windrose Graphics Direction and Speed of Wind in 2023 in Serang City show that the most dominant wind direction is in the north. Low wind conditions will also impact low rainfall in 2023 so that 2023 experienced a shift in the rainy season which was sufficient to affect the weather of higher air temperatures.

Classification of rainfall using Naïve Bayes Classifier

In this study, the analysis of the naïve Bayes Classifier will look for the results of the classification of rainfall in February 2024 based on past data using the amount of 120 data; the data includes rainfall data, air temperature, and wind speed in the last ten years which consists from 2014 - 2023. The Naïve Bayes Classifier was calculated through several stages, including finding an average value, standard deviation, and class probability, determining test data through moving averages, and calculating the results of the classification of rainfall, which will occur in February 2024. Following are the results of the classification of Oldeman rainfall from 2014 – 2023.

TABLE 2. Rainfall Classification 2014 - 2023

	Month											
Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
2014	B	B	L	K	L	K	L	K	K	K	L	L
2015	B	B	L	L	K	K	K	K	K	K	K	L
2016	L	B	B	K	L	K	L	K	L	L	L	L
2017	B	B	L	K	L	L	L	K	K	L	L	B
2018	L	L	B	L	K	L	K	K	K	K	L	L
2019	B	L	B	L	K	K	K	K	K	K	K	L
2020	B	B	B	B	L	K	K	K	K	K	L	B
2021	L	B	L	B	K	L	K	K	B	K	B	B
2022	L	L	B	L	L	K	L	K	B	L	L	L
2023	L	B	B	L	K	L	K	K	K	K	L	K

Oldeman classification results based on rainfall data in 2014 - 2023 are found in TABLE 2, obtained from the Serang City Maritime Meteorology Station with each wet month classification (B), moist (L), and dry month (K). The table is the result of the classification of rainfall that has been grouped to find the data included in the dry, humid, and wet months, which each average value, standard deviation, and class probability will seek. In addition, in TABLE 1, it also appears that the average wet month occurred in January - February, humid months occurred in November, and the dry month occurred in August. Seasonal factors influence this. Based on these data, it can be seen that 2019 has the lowest rainfall level between other years. This happens because of the shift in the degree of the earth when orbiting the sun, which causes reduced clouds due to drought, thus making solar radiation higher in Indonesia in the tropics. Sunny weather conditions are still quite dominating, and the anomaly of cold temperatures in Indonesian waters causes rain clouds to be difficult to form, so the city of Serang has not entered the rainy season as usual. In addition, the shift in the rainy season also occurred in 2023. ideally, the rainy season began in October and November, but in 2023, in December, the rainfall only reached 77 mm, so it entered the category of dry months; this

happened because of climate changes and skills of the phenomenon La Nina to El Nino. After rainfall for ten years, it is categorized based on rainfall, according to Oldeman. Then, the average and standard deviation of the air temperature and wind speed are determined, which is the parameter used to classify rainfall in Serang City.

TABLE 3. Average and Standard Air Temperature Deviation

Classification	Average	Standard deviation
Wet	27.244444	0.4540626
Moist	27.506522	0.408817
Dry	27.742553	0.5262249

TABLE 4. Average and Standard Deviation of Wind Speed (Knot)

Classification	Average	Standard deviation
Wet	1.962963	0.400137
Moist	1.747826	0.43625
Dry	1.7	0.354905

The average results and standard deviations of the air temperature data can be seen in TABLE 3. The average result in the wet month is 27.244°C with a standard deviation of 0.4541, the average humid month is 27.5065°C with a standard deviation of 0 and 4088 Average Dry Moon is 27.743°C with a standard deviation of 0.5262. Furthermore, the average results and standard deviations of the wind speed are in TABLE 4; the average result in the wet month is 1.963 knots with a standard deviation of 0.4001, and the average in a humid month is 1.7478 knots with a standard deviation of 0.4363 The Average in the dry month was 1.7 knots with a standard deviation of 0.354905. Based on the average air temperature results, the most significant average value is found in the dry month, and the lowest air temperature is in the wet month, so it can be said to be a negative relationship between air temperature and rainfall. The highest average wind speed is found in the wet month, and the lowest wind speed is in the dry month, so it can be said to be a positive relationship between wind speed and rainfall.

TABLE 5. Class Probability

Class Probability	Mark
Wet	0.225
Moist	0.3781513
Dry	0.3983051

After obtaining the average results and standard deviations, the probability value is used as an opportunity to determine the possibilities that will occur based on past data. These results can be seen in TABLE 5 the wet month obtained a probability value of 0.225 in a humid month and 0.3781513 in the dry month of 0.354905. To find out whether or not the probability results in each class can be proven by adding up all the probability values of each class obtained, if the results are 1 then the probability value is correct, but if the results are not 1 then the probability value is wrong, meaning there is a calculation error. After the probability value has been confirmed to be correct, the next stage is moving the average in advance to get test data from air temperature and wind speed in February 2024.

TABLE 6. Moving Average February 2024

Moving Average
Air temperature $(^{\circ}C)$
Wind velocity (Knot)

Moving Average is one of the methods of time series used to determine forecasting when the data does not yet exist, to get the results of rainfall classification. The parameters used in classifying rainfall are air temperature and wind speed with vulnerable time per 11 months. Air temperature and wind speed data to be used in this calculation comes from the latest data in 2023. The results found in TABLE 6 show that the air temperature in February amounted to 27.8°C with a wind speed of 1.1 knots. The data that has been obtained will be used as test data to classify rainfall in March and April 2024.

TABLE 7. Results of rainfall classification for February 2024.

Classification	**Air temperature**	**Wind velocity**	**Results**
Test data	27.9	1	Moist
Wet	0.2088497	0.034856	0.001637924
Moist	0.392733	0.13902	0.020646278
Dry	0.5260115	0.095767	0.020064509
			0.020646278

The final results of the rainfall classification using the Naïve Bayes Classifier, calculated using Microsoft Exel in February, are in TABLE 7. Based on the table the results of rainfall classification in Serang City using Naïve Bayes Classifier in February with an air temperature of 27.8°C and Wind speed 1.1 knots produce a wet moon rainfall classification of 0.003885894

moist months with 0.036587466 and a dry month of 0.034945913. So the maximum value created in a humid month means that in February 2024, rainfall is predicted to be included in the sweltering month classification, ie, rainfall will be 100 - 200 mm. This is by the Rainfall Data in February obtained from the Serang City BMKG station that the amount of rainfall that occurred was 197.3 mm which was included in the category of wet month classification, so that Naive Bayes Classifier could classify rainfall well and accurately.

CONCLUSION

The results of the rainfall classification using Naïve Bayes Classifier in Serang City in February 2024 with an air temperature of 27.8°C and the wind speed of 1.1 knots received a maximum value in a humid month of 0.028721149, thus showing rainfall in February of 100 - 200 mm. The result of the graph of the relationship between air temperature and precipitation is an effect of 8.1% which shows a reasonably small influence, so there is another climate element that can affect the more considerable rainfall one of which is the humidity.

REFERENCES

[1] H. Wibowo, "Desain Prototipe Alat Pengukur Curah Hujan Jarak Jauh Dengan Pengendali Komputer," 2014.
[2] H. J. Fowler et al., "Anthropogenic intensification of short-duration rainfall extremes," Nature Reviews Earth & Environment, vol. 2, no. 2, pp. 107-122, 2021.
[3] N. Sunarmi et al., " Analisis faktor unsur cuaca terhadap perubahan iklim di Kabupaten Pasuruan pada tahun 2021 dengan metode Principal Component Analysis," Newton-Maxwell Journal of Physics, vol. 3, no. 2, pp. 56-64, 2022.
[4] BMKG, "Daerah Kurang Hujan Terpanjang Hingga Tengah November 2023," www.instagram.com. https://www.instagram.com/infobmkg/p/CzlIJWkBUVo/?img_index=1 (accessed Apr. 28, 2024).
[5] F. Tempola et al., "Perbandingan Klasifikasi Antara KNN dan Naive Bayes pada Penentuan Status Gunung Berapi dengan K-Fold Cross Validation," Jurnal Teknologi Informasi dan Ilmu Komputer, vol. 5, no. 5, p. 577, Oct. 2018, doi: https://doi.org/10.25126/jtiik.201855983.
[6] D. Xhemali et al., "Naïve Bayes vs. Decision Trees vs. Neural Networks in the classification of training web pages," Loughborough University, 2009. Accessed: Jan. 10, 2024. [Online]. Available: https://hdl.handle.net/2134/5394.
[7] T. Rosandy, "Perbandingan Metode Naive Bayes Classifier Dengan Metode Decision Tree (C4. 5) Untuk Menganalisa Kelancaran Pembiayaan (Study Kasus: KSPPS/BMT Al-Fadhila)," Jurnal Teknologi Informasi Magister, 2(01), 52-62, 2016.
[8] A. Almahjoub and D. Qiu, "Random Access in IoT Using Naïve Bayes Classification," WSEAS Transactions on Circuits and Systems, vol. 20, May 2021, doi: https://doi.org/10.37394/23201.2021.20.10.
[9] M. J. Sánchez-Franco et al., "A naive Bayes strategy for classifying customer satisfaction: A study based on online reviews of hospitality services. Journal of Business Research," vol. 101, pp. 499-506, 2019.
[10] M. I. Jumarang, "Analisis Hubungan Kelembaban Udara dan Suhu Udara Terhadap Parameter Tebal Hujan di Kota Pontianak," Prisma Fisika, vol. 4, no. 3, pp. 80–83, 2016.