Plant recommendation using Triple Exponential Smoothing and K-Nearest Neighbor based on Internet of Things

M J Habbibie*, Y Yaddarabullah and A Syahputra
Program Studi Teknik Informatika, Universitas Trilogi, Indonesia

*mjurnaliesh@trilogi.ac.id

Abstract. Nutritious Garden Trilogy University is an agricultural land managed by urban farmers as well as students who are also lecturers to practice teaching and learn in the cultivation of types of plant commodities. Cultivation of plant species done by examining the climatic factors of plants on the land. The sustainability reason is one of the obstacles to make sure the results of cultivation in the teaching and learning process. The changing climate makes urban farmers get trouble to determine types of plants to be planted. This study will develop a system to forgive the recommendation of the type of plant according to the change of climate. The climate changes recorded using the internet of things sensors which consist of temperature, humidity, light intensity, and wind speed. The data entered will be processed using the triple exponential smoothing method as a forecast to predict future weather, then classified using the k-nearest neighbor to get the types of plants. The results of forecasting testing from sensors using the mean absolute percentage error obtained values of 9.53% temperature, 16.44% humidity, 3.73% light intensity, 19.42% wind speed.

1. Introduction
The agricultural sector is very vulnerable to crop failure due to climate change and results in a decrease in crop production [1]. Based on research conducted by [2], states that one of the causes of crop failure is a factor of climate change which results in decreased agricultural production. The effects of weather climate as seen from the range of temperature, humidity, humidity, heatwave pressure and drought during the growth of plants can affect yield productivity, causing food prices to go up or down [3,4]. This causes the perpetrators of urban farming in Trilogi University to be difficult in identifying the right type of plants to be planted based on weather climatic factors on the Trilogi University Nutrition Garden land. The problem was found in a solution to facilitate urban farming in identifying crop commodities to be planted using the Internet of Things technology. The use of the Internet of Things technology aims to harmonize between users and devices that can be monitored directly by relying on the internet network. The use of the Internet of Things technology in agriculture is a good step for urban farming in knowing the actual weather conditions. Internet of Things is the application of technology that relies on a set of tools to monitor the state of the object under study, to produce data from several devices in real-time [5]. Weather data that has been obtained from the Internet of Things toolkit will be combined into information data, then processed for forecasting and prediction. Accurate weather forecasting is a solution for providing weather data that takes a prediction of a certain period by using air temperature as a description of the weather conditions in an area [6]. The weather data processing is very necessary for future forecasting as identification of suitable plants to be planted using the Triple Exponential
Smoothing method adjusted based on plant growth requirements data [7]. Plants have many different types that are sometimes difficult to know the name, and therefore the application of classification is needed in terms of determining the type of plants using the K-Nearest Neighbor method. The implementation of the K-Nearest Neighbor method detection system is done by comparing test data with training data that has been stored in the database, serves to find the closest distance between test data and training data [8]. So that the classification process will produce a prediction that is useful for detecting plant commodity types. The purpose of this study is to be able to identify the types of plant commodities using the internet of things tool by applying the triple exponential smoothing method and k-nearest neighbor as a medium for crop recommendations for urban farming during times of uncertain climate and weather.

2. Literature review
The application of the Triple Exponential Smoothing and K-Nearest Neighbor method based on the internet of things in determining the types of plant commodities is the basis of the author to conduct this research. Several studies that have been made are as follows:

The Internet of Things is a very promising technology with solutions to agricultural modernization. Internet of Things data management and analysis can be used in automation processes, situation predictions, real-time activity improvement by creating new applications and services to get more added value. Therefore, the use of the Internet of Things as a real-time weather data collection process is highly recommended so that it can be monitored directly [9]. The Internet of Things can monitor big data and events in real-time by taking data from devices located on sensors for controlled decision making in analyzing, predicting, or detecting event patterns, and making decisions automatically. Because of that, the use of the Internet of Things can help in taking data directly to meet the needs of research [10]. The design of data communication architecture is very important in developing web services that are useful for distributing data. The data to be accessed is by using the application program interfaced (API) which is carried out through the HTTPS protocol [11]. Implementation of the Triple Exponential Smoothing (Winter) method for forecasting passengers of PT. Kereta Api Indonesia next year which relies on time series data in 2006 and 2014. The results of this study show that the data of the train fleet has a pattern of non-stationary fluctuation trends to produce good accuracy above 70%. This research that can be taken is in terms of the method using triple exponential smoothing as a prediction process in the future [12]. The use of Triple Exponential Smoothing and Holt-Winter multiplicative models as a comparison of the forecast number of domestic tourists in the Banyuwangi Regency to get the best model results. The data used in the period 2010 to 2018 to produce forecast error values are MAPE = 42% and MSE = 2773 [13]. Research that develops a system to predict monthly rainfall and the next few months by relying on climatological data from BMKG from 2000 to 2009. The data used consists of seven variables namely temperature, sunlight conditions, air pressure, humidity, wind speed, wind direction, and evaporation. The method used is exponential smoothing to predict future rainfall that relies on climatological data and while the K-Nearest Neighbor method is useful for classifying the determination of rainfall prediction. The results of this study get a predictive accuracy of 58.33% [14]. The use of the K-nearest neighbor method to classify mangrove species based on morphological characteristics seen from the closest distance between test data and training data can be implemented into the website with a success rate of 77.77% [15]. The use of the KNN method in classifying gear fractures based on motor speed and load can be implemented by producing an accuracy of 68% - 100% presentation value [16].

3. Methods
The methodology used in this study consisted of several stages. These stages will result in the determination of types of crop commodities based on the weather climate. The research methodology can be seen in Figure 1 below:
3.1. Data Acquisition

Data acquisition is a process of taking data from a device to get a data value of the object under study so that it can display data information on a computer monitor or Android device [17]. The data acquisition system in this study took data from the Internet of Things devices using temperature sensors, humidity sensors, light intensity, and wind speed sensors. The location point used for data collection is Nutrition Garden at Trilogi University, South Jakarta. Altitude data retrieve data from Google Maps viewed based on longitude and latitude. All data will be collected to be accurate data information. The work process of the data acquisition system on the Internet of Things weather station tool can be seen in Figure 2 below:

In Figure 2 above, it is the work process of a data acquisition system on the Internet of Things weather station tool to get the results of the data as information. Each sensor device takes data consisting of 4 objects namely air temperature, humidity, light intensity, and wind speed. Height data retrieved from the Google Maps website. The data that has been taken will go into Rasperry Pi for sending it to the Cloud Server as a data storage container. The data obtained by the sensor consisting of temperature, humidity, light intensity, and wind speed will be displayed on the Website. So that the data will then enter the forecasting process stage.

3.1.1. System component. The system components used to make the Internet of Things devices are in Table 1 as follows:
Table 1. Device Internet of Things

| No | Component | Function | Picture |
|----|-----------|----------|---------|
| 1. | Raspberry Pi 3+ | As an IoT Gateway that stores data locally from sensors and sends to Cloud Server. | ![Raspberry Pi 3+] |
| 2. | Arduino UNO R3 | As a microcontroller that stores data from the sensor. | ![Arduino UNO R3] |
| 3. | Wemos D1 ESP8266 | As a microcontroller that is connected to various sensors to send data to the Raspberry Pi. | ![Wemos D1 ESP8266] |
| 4. | DHT11 | Get air temperature and humidity data. | ![DHT11] |
| 5. | Light Dependent Resistor (LDR) | Get light intensity data. | ![LDR] |
| 6. | Anemometer | Obtain wind speed data. | ![Anemometer] |

3.2. Forecasting
At this stage, it is a process to predict weather and climate data by relying on short-term forecasting that takes less than three months. The forecasting process is used to predict the data for the next seven days using the time series method with the Triple Exponential Smoothing model. Triple Exponential Smoothing is a smoothing method that is carried out regularly three times by relying on past or present data. The purpose of this method is to minimize irregular seasonal data as a source of making predictive averages from weather climate data. The Triple Exponential Smoothing method uses the following formula equation:

\[ S'_t = \alpha X_t + (1 - \alpha)S'_t \]  
\[ S''_t = \alpha S'_t + (1 - \alpha)S''_t \]  
\[ S'''_t = \alpha S''_t + (1 - \alpha)S'''_t \]  
\[ a_t = 3S'_t - 3S''_t + S'''_t \]  
\[ b_t = \frac{a_t}{2(1-a^2)}[(6-5a)S'_t - (10-8a)S'_t + (4-3a)S'''_t] \]  
\[ c_t = \frac{a_t^2}{(1-a^2)}(S'_t - 2S''_t + S'''_t) \]  
\[ F_{t+m} = a_t + b_t m + \frac{1}{2} c_t m^2 \]
Flowchart forecasting process using the Triple Exponential Smoothing method can be seen in Figure 3, as follows:

\[
\begin{align*}
\text{Start} & \quad \text{input data coming from the internet of things tool} \\
S^t &= \alpha X_t + (1 - \alpha) S^t \\
S''^t &= \alpha S' + (1 - \alpha) S''^t \\
S'''^t &= \alpha S'' + (1 - \alpha) S'''^t \\
at &= 3S^t - 3S''^t + S'''^t \\
b_t &= (\alpha/2(1-\alpha)^2)*[(6-5\alpha)S't-(10-8\alpha)S^t+(4-3\alpha)S'''^t] \\
c_t &= (\alpha^2/(1-\alpha)^2)*(S^t-2S''^t+S'''^t) \\
F_3+m &= at + bt + ct + \frac{1}{2}ctm^2 \\
\text{End}
\end{align*}
\]

\textbf{Figure 3.} Flowchart forecasting work process.

In Figure 3 above, it is a work process of forecasting to determine the weather forecast for the future. Data from the Internet of Things toolkit is taken into test data for the process of determining the weather climate prediction going forward. After the data has been taken, the data will enter the first phase of graduation using the formula equation (1). This process is to determine the prediction for each period so that it will produce the first smoothing value. The second stage of the process is the same as the first stage before using the equation formula (2) which distinguishes the search for the value $S''^t$ by taking the value of the results of the first smoothing. The third stage smoothing process uses the formula equation (3) to find the value of $S'''^t$ by taking the value of the results of the second stage smoothing. All stages of the smoothing process will be obtained values consisting of the values $S^t, S''^t, \text{and} S'''^t$. After that, $a_t, b_t, \text{and} c_t$ values will be searched using the formula equation (4), (5), and (6) to find the value to get the value $F_3+m$ using the formula equation (7). The forecast value is a value for each period to be determined. If all of the data for the period has determined its forecast, it will produce data values and chart patterns for a predetermined period. The results that have been obtained will be tested for accuracy to see forecast errors from forecasting results. Forecast Error is a determination to test the accuracy of forecast errors by using several indicators namely, the average absolute error percentage Mean Absolute Percentage Error (MAPE) and the mean square least mean Square Error (MSE). The formula used to see Mean Absolute Percentage Error (MAPE) and Mean Square Error (MSE) is as follows:
MAPE = \frac{100\%}{N} \cdot \sum_{i=1}^{N} \frac{|x_i - f_i|}{x_i} \quad (8)

MSE = \frac{1}{N} \sum_{i=1}^{N} (x_i - f_i)^2 \quad (9)

### 3.3. Classification

The classification process will use data that has been obtained from the Research and Development of the Ministry of Agriculture to be used as training data in this classification process. Data on conditions for growing plants based on climate to be used can be seen in Table 2 as follows:

| Temperature (Celsius) | Humidity (Percentage) | Height (Mdpl) | Light Intensity (Foot Candles) | Wind Speed (m/s) | Type of Plant |
|----------------------|-----------------------|---------------|--------------------------------|------------------|---------------|
| 26                   | 67,5                  | 500           | 670                            | 3                | Cucumber      |
| 26                   | 40                    | 200           | 775                            | 4                | Watermelon    |
| 22,5                 | 60                    | 1400          | 1050                           | 1,5              | Chili         |
| 19,5                 | 57,5                  | 750           | 700                            | 2,2              | Tomato        |
| 29                   | 55                    | 50,5          | 1275                           | 1,2              | Melon         |
| 26,5                 | 65                    | 250,5         | 925                            | 1,2              | Green Beans   |
| 26                   | 65                    | 300           | 1400                           | 2,5              | Soy           |
| 25                   | 85                    | 1750          | 1750                           | 2,5              | Potato        |
| 27,5                 | 70                    | 1400          | 2750                           | 1,5              | Corn          |

Classification is a process method for grouping an object into a certain class based on similarities and differences in the characteristics of the object. This study uses the K-Nearest Neighbor method which is a data classification algorithm to find the closest distance between training data and test data. The identification in this study relies on data values or climate and weather prediction graphs from the forecasting process. The initial approach is by providing a new point from the training data for classification with test data. Then find the K value from the closest point between the training data and test data. The search for K values is determined by the formula equation (10) as follows:

\[ K = \sqrt{n} \quad (10) \]

The formula above is to determine the value of K, with n which is the amount of sample data. The basic K value is an odd number in the determination process to prevent the occurrence of the same distance value. To calculate the closest distance between training data and test data using the Ecludian Distance method. Equation formula (11) Ecludian Distance can be seen as follows:

\[ d_i = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2} = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \ldots + (p_n - q_n)^2} \quad (11) \]

The classification process using the K-Nearest Neighbor method in this study can be seen in Figure 4 as follows:
4. Results and discussion

4.1. Implementation of the Internet of Things

The Internet of Things circuit tool that has been implemented can be seen in Figure 5 as follows:
Figure 5 above is a series of Internet of Things when data was taken at the Nutritious Garden of the University of Trilogy. The picture consists of a DHT11 sensor for temperature and humidity, a light-dependent resistor sensor for light intensity, and an anemometer sensor for wind speed equipped with raspberry pi for its local storage.

4.2. Data sending process

The process of sending data contained on the Internet of Things devices must be by the mechanism of data communication architecture that has been made, namely with the IoT gateway as the recipient of data from the node (sensor) and continue to send to the server. Following Figure 6 is the data communication architecture that exists in the Internet of Things system.

Data transmission system for each device on the Internet of Things consisting of nodes (sensors), gateways, and servers. Each sensor module consists of a microcontroller that is connected to the sensor.

Figure 6. The architecture of data communications internet systems of things.
Then the sensor module will send sensor data to the IoT gateway once every 5 minutes via wireless communication media (WiFi) which is then stored by the IoT gateway into RabbitMQ. The data communication system can be done from the sensor module to the IoT gateway using MQTT, Stomp, Rest and Socket. Data on RabbitMQ will be made into a log into DBMysql. At the IoT gateway, there is a crop commodity identification application that can be accessed locally. Every 5 minutes the IoT gateway will send data to DBMysql to the Cloud Server through the internet network to then be stored by the Cloud Server into DBMysql.

4.3. System implementation

The system implementation phase is a manifestation of the results that have been applied to produce the objectives of the system development that have been achieved in the form of the API address used and the interface on the website.

4.3.1. API access data. Retrieval of data in the form of an API that contains data on each sensor taken and stored in a cloud server. The APIs for each data from the cloud server will be called as follows:

1) Temperature Data
   http://31.220.48.158/projects/iotgateway/raspi02/api/apiGetDataJson_temp.php

2) Humidity Data
   http://31.220.48.158/projects/iotgateway/raspi02/api/apiGetDataJson_hum.php

3) Light Intensity Data
   http://31.220.48.158/projects/iotgateway/raspi02/api/apiGetDataJson_lrd.php

4) Wind Speed Data
   http://31.220.48.158/projects/iotgateway/raspi02/api/apiGetDataJson_ws.php

4.3.2. Website interfaced. The forecasting menu is a menu display to see the prediction results obtained by displaying graphical results from the prediction processing results of each data such as temperature, humidity, light intensity, and wind speed. The graph displays original and predictive data. There are 2 plots in displaying graphs that consist of forecasting graphs from 1-week data that have been taken from sensor devices. The first plot displays the original and predicted data graphs on the 1-week data. While the second plot displays a prediction data graph until the 90th day. The forecasting process also displays the results of the MSE and MAPE values obtained from each forecasting process from each data. The display of the forecasting interface can be seen in Figure 7.

![Figure 7. Forecasting menu.](image_url)
The classification menu is a display of the results of the classification of crop recommendations suitable for planting in the area. The classification must enter the longitude and latitude to adjust the coordinates as a reference to get the height of the area. The results of the classification will be sorted so that the earliest plants are highly recommended crops to be planted in the area. The display in the classification process will display each euclidian distance obtained. The recommended crop that is suitable to be applied is the top 3 plants as the best recommendation. The appearance of the classification interface can be seen in Figure 8.

![Classification menu](image)

**Figure 8.** Classification menu.

### 4.4. MAPE and MSE test result

The test results of Mean Absolute Percentage Error (MAPE) and Mean Square Error (MSE) were carried out to determine the error value of forecasting. Based on the recording of the error value of each data predicted or predicted consists of temperature, humidity, light intensity, rainfall, and soil moisture. Then the forecasting results obtained from each of which are seen in Table 3 below:

| No | Data          | MAPE  | MSE  |
|----|---------------|-------|------|
| 1  | Temperature   | 9.53% | 3.36 |
| 2  | Humidity      | 16.44%| 6.90 |
| 3  | Light Intensity | 3.73% | 27.82 |
| 4  | Wind Speed    | 19.42%| 0.59 |

### 4.5. Ecludian distance test result

The results of ecludian distance testing are conducted to determine the distance value of each type of plants such as watermelons, green beans, soybeans, chilies, corn, potatoes, tomatoes, and cucumbers measured from the value of training data that has been obtained from the Ministry of Agriculture. The results of the ecludian distance value of each type of plant can be seen in Table 4 as follows:
Table 4. Ecludian distance test result.

| No | Type of Plant | Ecludian Distance |
|----|---------------|-------------------|
| 1  | Green Beans   | 256.03            |
| 2  | Watermelon    | 269.57            |
| 3  | Soybeans      | 554.55            |
| 4  | Cucumber      | 634.79            |
| 5  | Chili         | 786.21            |
| 6  | Tomato        | 943.84            |
| 7  | Melon         | 1276.14           |
| 8  | Potato        | 1900.06           |
| 9  | Corn          | 2242.33           |

5. Conclusion
Based on the discussion that has been explained in the previous explanation, that identification of crop commodity types based on climatic weather using triple exponential smoothing and K-nearest neighbors can be concluded that Triple Exponential Smoothing and K-Nearest Neighbor based on the Internet of Things can be implemented on a website system to recommend types plants based on climate weather factors. The system can retrieve data online that is accessed using the API and display forecasting graphically from the original data and prediction data and present the results of recommendations for plants that are suitable for planting in the Nutritious Garden Location of the Trilogi University, South Jakarta. Internet of Things devices can produce data derived from each sensor such as DHT11, light-dependent resistor, and anemometer. The process of collecting data with the Internet of Things device takes 1 week. Triple Exponential Smoothing applied to the system gets the MAPE value from each data that is temperature = 9.53%, humidity = 16.44%, light intensity = 3.73%, wind speed = 19.42%. Meanwhile, the MSE values obtained were temperature = 3.36, humidity = 6.90, light intensity = 27.82, and wind speed = 0.59. the resulting accuracy is relatively small because if the value of MAPE and MSE gets smaller the percentage of accuracy results in no error gets bigger and can be said to be good. If the MAPE value is 9.53%, then 90.47% is declared no error. K-Nearest Neighbor successfully displays 9 types of plants by getting the value of the ecludian distance from the value of the closest distance to the farthest by the value obtained from the ecludian distance. The top 3 plants can be classified as recommended in crop matching so that the lowest value obtained will be made a recommendation for the type of plant suitable for planting.

References
[1] Hidayati I N and Suryanto 2015 Pengaruh Perubahan Iklim Terhadap Produksi Pertanian dan Strategi Adaptasi pada Lahan Rawan Kekeringan J. Ekon. dan Stud. Pembang. 16, April
[2] Ruminta 2016 Analisis penurunan produksi tanaman padi akibat perubahan iklim di Kabupaten Bandung Jawa Barat Analysis of decreasing production of paddy due to climate change in Bandung district West Java Pendahuluan J. Kultiv. 15, 1 p. 37–45
[3] Siebert S Webber H and Rezaei E E 2017 Weather impacts on crop yields - searching for simple answers to a complex problem August p. 10–13
[4] Chatzopoulos T Pérez I Zampieri M and Toreti A 2019 Climate extremes and agricultural commodity markets : a global economic analysis of regionally simulated events Weather Clim. Extrem. April 2018 p. 100193
[5] Susanti E and Triyono J 2017 Prototype Alat IoT ( Internet of Things ) Untuk Pengendali dan Pemantau Kendaraan Secara Real Time December 2016
[6] Anwar S 2018 Peramalan Suhu Udara Jangka Pendek di Kota Banda Aceh dengan Metode Peramalan Suhu Udara Jangka Pendek di Kota Banda Aceh dengan Metode Autoregressive Integrated Moving Average ( ARIMA ) MJ Mech. Sci. Technol. 5, June 2017
[7] Affandi L Syulistyo A R and Putra F R 2018 Pengembangan Aplikasi Mobile Peramalan Cuaca untuk Penentuan Tanaman Pangan Menggunakan Metode Triple Exponential Smoothing J.
Inform. Polinema 4, 2 p. 117–124

[8] Syahid D Jumadi and Nursantika D 2016 Syahid, D., Jumadi, & Nursantika, D. (2016). Daun Philodenron Menggunakan Metode K-nearest Neighbor ( KNN ) Berdasarkan Nilai HUE , Saturation , Value ( HSV ), I(1), 20–23.Daun Philodenron Menggunakan Metode K-nearest Neighbor ( KNN ) Berdasarkan Nilai HU JOIN, 1 p. 20–23

[9] Tzounis A Katsoulas N and Bartzanas T 2017 Internet of Things in agriculture , recent advances and future challenges Biosyst. Eng. 164 p. 31–48

[10] Mazon-olivo B Hernández-rojas D Maza-salinas J and Pan A, 2018 Rules engine and complex event processor in the context of internet of things for precision agriculture Comput. Electron. Agric. 154, September p. 347–360

[11] Yaddarabullah Muttaqin M and Rafiansyah M 2019 Service-Oriented Architecture for E-Marketplace Model Based on Multi-Platform Distributed System IOP Conf. Ser. Mater. Sci. Eng. 662 p. 42028

[12] Setiawan W Juniati E and Farida I 2016 The Use of Triple Exponential Smoothing Method ( Winter ) in Forecasting Passenger of PT Kereta Api Indonesia with Optimization Alpha , Beta , and Gamma Parameters ICSITech p. 198–202

[13] Feby I Y and Dwi H A 2018 Metod Triple Exponential Smoothing Holt-Winters untuk Peramalan Jumlah Wisatawan Nusantara Di Kabupaten Banyuwangi J. UJMC 4 p. 1–6

[14] Prihastuti H Mulyana I and Ofik Mulyana A 2012 Prediksi Curah Hujan Bulanan Menggunakan Time Series (Single Exponential Smoothing) Dan KNN (Studi Kasus : Kabupaten Padang Pariaman) Proc. KNASTIK 2012

[15] Wardani S H Rismawan T and Bahri S 2016 Aplikasi Klasifikasi Jenis Tumbuhan Mangrove Berdasarkan Karakteristik Morfologi Menggunakan Metode K-Nearest Neighbor ( KNN ) Berbasis Web J. Coding, Sist. Komput. Utan 4, 3 p. 9–21

[16] Wang D 2015 K -nearest neighbors based methods for identification of different gear crack levels under different motor speeds and loads : Revisited Mech. Syst. Signal Process. p. 1–8.

[17] Suryatini F and Fauzandi F I 2018 Sistem Akuisisi Data Suhu Dan Kelembaban Tanah Pada Irigasi Tetes Otomatis Berbasis Internet Of Things SEMNASTEK p. 1–6