Applied K-Nearest Neighbors (KKN) on Dust Suppression Prototype

Agung Kurniawan Nurhasan, Hendro Agus Widodo, R. Y. Adhitya*, Rachmad Andri Atmoko

Automation Engineering, Politeknik Perkapalan Negeri Surabaya 60111, Indonesia

ABSTRACT

Steam power plants that use coal as fuel have serious problems during operation. Before heading to the combustion process, coal is stored in an open field area. However, this results in fine particles of coal dust being exposed to wind and polluting the surrounding environment. The purpose of this study is to minimize the impact of pollution from coal dust by using the dust suppression tool. The tools that have been run manually or conventionally can be operated automatically to facilitate the operator in controlling dust suppression without the need to go to the field. This research proposes a prototype dust suppression equipped with dust and temperature sensors, the sensor data is a representation of the condition of the coal storage area which is processed using the K-Nearest Neighbors method to classify whether the condition of the storage area is normal or dusty. When conditions are dusty, the pump activates and directs bursts of water at the coal to minimize dust. In the application of the K-Nearest Neighbors method, center point 1 is obtained for normal conditions, with a dust density of 0.4353 mg/m³ and a temperature of 27.5818 °C. Whereas center point 2 for dusty conditions has a dust density of 2.374 mg/m³ and a temperature of 28.2667 °C. From 40 testing data in real-time, a success rate of 87.5% was obtained.

Keywords: Coal dust, pump control, K-Nearest neighbors

Introduction

Steam power plants that utilize coal as the main fuel are the types of power plants with the largest capacity and quantity that Indonesia has. The abundance and low cost of using coal have made it the primary fuel choice for building power plants in Indonesia. But the issue of the environment becomes a matter that must be considered, especially fine particles of coal dust is a major cause of acute respiratory disease [1], mercury is nerve damage to the development of children under five and fetuses in the womb of pregnant women.

SNI Regulation 19-0232-2005 concerning the threshold value of chemical substances in the workplace air also mentions the threshold value for coal dust is 2 mg/m³. Many industries in Indonesia also refer to the dust threshold value in general listed in Permenakertrans No. PER.13 / MEN / X / 2011. It also concerns the threshold value of physical and chemical factors in the workplace, which are 2 mg/m³ for boiler areas that use coal. This makes it possible that the level of exposure to coal dust in various workplaces in Indonesia could be higher than 2 mg/m³ so that the chances of respiratory problems for workers are higher.

The impact of environmental pollution resulting from fine particles of coal dust is overcome by using a dust suppression tool as shown in Figure 1 to wet the coal surface. Coal that has been moistened with water will not be easily carried by the wind, thereby reducing the environmental pollution in the area around the industry. Dust suppression has been controlled conventionally so far with the assistance of operators. The valve used to channel water to dust suppression is also still activated manually. The distance from the operator’s room to the coal storage area is quite far resulting in the...
operator having to walk several meters to get to the coal storage area. This is certainly less efficient because it takes a long time just to turn on / turn off dust suppression.

These problems can be solved by using a system optimization to help the operator operate a dust suppression device automatically. Therefore, the authors took the initiative to create innovation in the form of Dust Suppression System Design with the K-Nearest Neighbors (K-NN) Method to simplify the operation and increase the efficiency of dust suppression tools.

**Methods**

The following is the process of the prototype in this research:

1. **Silo**
2. **Servo Motor**
3. **GP2Y1010AU0F Dust Sensor**
4. **DHT11 sensor**
5. **Spray**
6. **Pump Motor**
7. **Storage**
8. **Proximity**
9. **Conveyors**
10. **Combustion Process Site**

The basic concept of this prototype is to decide when the pump is activated to minimize coal dust in the storage area. GP2Y1010AU0F [4] dust sensor and DHT11 temperature sensor [5] will send information to the Arduino Mega 2560 microcontroller. The ADC facility contained in Mega 2560 will convert the information in the form of analog voltage signals from the two sensors into 10-bit numbers (0-1024) [6] and adjusted according to the unit of
K-Nearest neighbors

K-Nearest Neighbors (K-NN) algorithm is an algorithm used to classify an object, based on k training data which is the closest distance to the object [2]. The condition for k values cannot be greater than the amount of training data, and k values must be odd [3]. There are several ways to measure the proximity between new data and old data (training data), one of them is Euclidean distance. Euclidean distance is a method of finding the proximity of two values of distance variables. In this study, the Euclidean distance search method is used, because there are 2 variables with the same type of data that are used to determine the value of the distance between the two. The two variables are new data or test data and old data or training data. Euclidean distance equation for measuring the distance of two objects is as follows:

\[ d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \]  

Information:
\[ x_1, y_1 = \text{new data} \]
\[ x_2, y_2 = \text{old data} \]

After knowing the distance value from new data, it can be seen the classification of new data based on the distance closest to existing data.

Results and Discussion

After designing the prototype and understanding the concept of KNN, the next step is the application of KNN to the microcontroller. Before applying the KNN algorithm to the microcontroller, first, we must determine the center point value for both dusty and normal conditions. Center Point (CP) is a reference point to the test point so that when a new test point appears, the KNN algorithm can be applied to find the difference between the test point with CP_D (dusty center point) and the test point with CP_N (normal center point). Through the search for the difference, two different values will appear which will be chosen the smallest value as the basis for determining whether the condition is classified as dusty or normal. To get the CP value in this study, dust and temperature data collection was taken under normal conditions and in dusty conditions, the following is a data collection table for CP_N and CP_D:

Table 1. Data Collection

<table>
<thead>
<tr>
<th>CP</th>
<th>Dust Intensity (mg/m³)</th>
<th>Temp(°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP_N</td>
<td>0.4353</td>
<td>27.5818</td>
</tr>
<tr>
<td>CP_D</td>
<td>2.374</td>
<td>28.2667</td>
</tr>
</tbody>
</table>

Both conditions can be achieved by conditioning the coal material in the storage area. After determining the CP value of the sensor readings, if there are new test data then through equation (1) we can calculate the difference value or the value of d (distance) of new data against CP_N and the value of new data d against CP_D. The following is an example of the calculation of distance for 6 new test data:

Table 2. Distance calculation results

<table>
<thead>
<tr>
<th>Id</th>
<th>Dust (mg/m³)</th>
<th>Temp (°C)</th>
<th>dc1</th>
<th>dc2</th>
<th>c1</th>
<th>c2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.25</td>
<td>27</td>
<td>1.00</td>
<td>2.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.58</td>
<td>28</td>
<td>0.19</td>
<td>1.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.33</td>
<td>29</td>
<td>1.00</td>
<td>1.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.01</td>
<td>28</td>
<td>1.62</td>
<td>1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2.37</td>
<td>30</td>
<td>2.81</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2.52</td>
<td>29</td>
<td>2.35</td>
<td>0.22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Where:
\[ d_{c1} = \text{distance value to CP}_N \]
\[ d_{c2} = \text{distance value to CP}_D \]
\[ c_1 = \text{classification results included as normal conditions} \]
\[ c_2 = \text{classification results included as dusty conditions} \]

After obtaining the distance values, the next step is to determine the classification results. To determine the classification results simply we can use the minimize function to find the smallest difference and logical operations using if-then-else statements, here is the process of determining classification results:
1. \( d_{cm} = \text{MIN}(d_{c1}; d_{c2}) \), find the smallest value of the distance value between CP\(_N\) and CP\(_D\), the smallest difference results will be stored in the \( d_{cm} \) variable.
2. if \( d_{cm} == d_{c1} \) then result: = "normal" and pump: = LOW. If the value of \( d_{cm} \) is equal to \( d_{c1} \), it is concluded in the normal class and pump deactivated.
3. if \( d_{cm} == d_{c2} \) then result: = "dusty" and pump: = HIGH. If the value of \( d_{cm} \) is equal to \( d_{c2} \), it is classified as dusty and the pump is activated.

To test the reliability of the KNN method applied to the prototype, this study conducted 40 times each of 20 samples for testing in normal conditions and 20 samples for testing in dusty conditions. Here are the results of the test recapitulation under normal conditions:

Table 3. Table of system responses for normal conditions

<table>
<thead>
<tr>
<th>Id</th>
<th>Dust</th>
<th>Temp</th>
<th>Output</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.49</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>1.41</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>1.02</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>0.43</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>5</td>
<td>0.45</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>6</td>
<td>0.11</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>7</td>
<td>1.31</td>
<td>26</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>0.38</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>9</td>
<td>0.23</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>10</td>
<td>0.82</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>11</td>
<td>0.43</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>12</td>
<td>0.03</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>13</td>
<td>0.09</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>14</td>
<td>0.28</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>15</td>
<td>0.75</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>16</td>
<td>0.05</td>
<td>26</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>17</td>
<td>0.68</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>18</td>
<td>0.99</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>19</td>
<td>0.62</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
<tr>
<td>20</td>
<td>0.78</td>
<td>25</td>
<td>N</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 3 shows the KNN algorithm is able to classify normal conditions with a 100% success rate for 20 sampling data. And here are the results of the test recapitulation for dusty conditions:
Table 4. System response table for dusty conditions

<table>
<thead>
<tr>
<th>Id</th>
<th>Dust</th>
<th>Temp</th>
<th>Output</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.49</td>
<td>29</td>
<td>N</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>0.26</td>
<td>29</td>
<td>N</td>
<td>x</td>
</tr>
<tr>
<td>3</td>
<td>1.72</td>
<td>28</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.14</td>
<td>29</td>
<td>N</td>
<td>x</td>
</tr>
<tr>
<td>5</td>
<td>2.20</td>
<td>29</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.81</td>
<td>29</td>
<td>N</td>
<td>x</td>
</tr>
<tr>
<td>7</td>
<td>2.15</td>
<td>29</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2.24</td>
<td>29</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.48</td>
<td>29</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.36</td>
<td>29</td>
<td>N</td>
<td>x</td>
</tr>
<tr>
<td>11</td>
<td>1.88</td>
<td>28</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2.16</td>
<td>29</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>2.13</td>
<td>30</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1.41</td>
<td>29</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1.28</td>
<td>29</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>2.24</td>
<td>29</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1.62</td>
<td>29</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.31</td>
<td>30</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>2.43</td>
<td>28</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2.15</td>
<td>29</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 shows that the KNN algorithm can classify normal conditions with a 75% success rate for 20 sampling data. KNN can correctly predict dust conditions 15 times.

**Conclusion and Recommendation**

The use of the KNN method is highly recommended for systems with two input parameters and classification problems that can be solved by a stochastic-based algorithm. In the application of dust suppression prototype, the center point of the normal conditions has dust density of 0.4353 mg/m³ and temperature of 27.5818 °C. Whereas the center point of dusty condition has dust density of 2,374 mg/m³ with a temperature of 28.2667 °C. In testing the system by applying the K-Nearest Neighbors method for 40 test data with the same proportions for normal and dusty conditions obtained a success rate of 87.5%. The application of metaheuristic methods such as fuzzy subtractive clustering can be used as a comparison for this study.

**Acknowledgment**

The authors say thanks to Automation Engineering Politeknik Perkapalan Negeri Surabaya for supporting this research.

**References**

AK Nurhasan et al., 2020 / Applied K-Nearest neighbors (KKN) on dust suppression prototype

doi: https://doi.org/10.23960/elc.v9n2.168