A Neuro-Fuzzy Application for Clustering Greenhouse Gases Produced by Activities in University Dormitor

Narissara Eiamkanitchat
Department of Computer Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand
Email: narisara@eng.cmu.ac.th

Abstract—This research applies a Neuro-fuzzy method for clustering greenhouse gases produced by student activities. The partitional clustering algorithm is combined with Neuro-fuzzy. A standard dataset including iris and breast cancer is used to test the ability of the clustering algorithm. All activities who live in university dormitories are used to calculate the coefficient greenhouse gas emissions. These gases are further clustered by using Neuro-fuzzy techniques. The analysis results portray students’ behavior and can be used to promote the reduction of their greenhouse gas emissions.

Index Terms—neuro-fuzzy, clustering, greenhouse gas

I. INTRODUCTION

Computational intelligence is a consistently evolving technology. Many researchers proposed various high accuracy approaches for clustering and classification. A Neuro-fuzzy approach proposed in many researches can present high accuracy classification results. [1], [2] present an efficient Neuro-fuzzy clustering technique. This technique involves three classification layers, i.e. input layer, hidden layer and output layer. The hidden layer is a combination of neuron network structures with Gaussian membership function. Three values, i.e. small, medium or large, are possible results from membership computation. After that, the values are further classified and determined whether they are important or not. Important information is classified and assigned as (1) , In contrast, other unimportant remainings are classified as (-1). Data from the second layer is passed as input to the next layer, using Sigmoid as an activation function. Due to the very good results with several data sets [1], [2], this study is the application of the Neuro-fuzzy technique above for clustering data.

Either intentionally or unintentionally, human activities often produce greenhouse gases. Greenhouse gases considering in this research include Carbon dioxide (CO2), Methane (CH4) and Nitrous oxide (N2O). Each of them has a different Global Warming Potential (GWP) which designates their ability to absorb heat. Carbon dioxide is generally used as a standard of comparison of GWP [3]. For example, Methane and Nitrous oxide have a GWP of 21 and 310 more times than Carbon dioxide [4]. This study involves the entire life cycle of student activities which produce greenhouse gases. The calculation of greenhouse gas emissions is based on the standards set by related organizations [3]-[8]. This research hypothesizes that student dormitory can be clustered based on greenhouse gas emissions based on their descriptive similarities. The results of this study can be used for a campaign for reduction of activities which produce high emission of greenhouse gases.

II. DATA COLLECTION AND PREPARATION

This study uses data from student dormitories of Chiang Mai University in 2010 [9]. All 18 dormitory halls are thoroughly investigated. These facilities involve 6 male dormitories, 11 female dormitories and 1 unisex dormitory. These dormitories can accommodate up to 8,346 students. Nevertheless, in 2010, they are occupied by only 7,951 students.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Type of collection</th>
<th>Unit/Collection method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electricity consumption</td>
<td>Amount</td>
<td>KWh/data are collected continuously</td>
</tr>
<tr>
<td>Fuel used by students to travel between dormitories and classroom buildings.</td>
<td>Details to calculate the distance traveled.</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>Water consumption.</td>
<td>Amount</td>
<td>Cubic meters/ data are collected continuously</td>
</tr>
<tr>
<td>Waste management. Transporting waste to the disposal site.</td>
<td>Amount</td>
<td>Kg per year/ Data collection and evaluation of a standard method.</td>
</tr>
</tbody>
</table>

Each facility provides different greenhouse gas emissions, both directly and indirectly. In order to simplify the evaluation and management of the emissions or absorption of the organization, as specified in ISO 14064-1 and the GHG protocol, the emissions are divided into three scopes [3]. The first scope involves direct emission and absorption of greenhouse gases. The second scope includes indirect emission and absorption of energy...
consumption, which is mainly electricity in this study. The final scope involves other indirect emission and absorption. This study implements a variety of methods to collect major activity data as display in Table I.

The data will be used as an emission coefficient [3] to calculate the amount of greenhouse gas emissions.

**TABLE II. THE POTENTIAL TO CAUSE GLOBAL WARMING OF THE GAS.**

<table>
<thead>
<tr>
<th>Type of gas</th>
<th>GWP value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO₂</td>
<td>1</td>
</tr>
<tr>
<td>CH₄</td>
<td>25</td>
</tr>
<tr>
<td>N₂O</td>
<td>298</td>
</tr>
<tr>
<td>HFC – 32</td>
<td>675</td>
</tr>
<tr>
<td>HFC – 134a</td>
<td>1430</td>
</tr>
<tr>
<td>HFC – 22</td>
<td>1810</td>
</tr>
<tr>
<td>HFC - 125</td>
<td>3500</td>
</tr>
<tr>
<td>SF₆</td>
<td>22800</td>
</tr>
</tbody>
</table>

The details of greenhouse gas emissions in each activity are calculated based on their scopes. The first scope, annual direct emission from fuel combustion from vehicles (Emission\(_{GHG,Fuel}\)), can be calculated from

\[
E_{GHG,Fuel} = \sum \left( FC \times EF_i \right).
\]

where

- \( FC \) = Consumption of fuel or other energy per year. (Units / year).
- \( EF_i \) = Coefficient of greenhouse gas emission type I per fuel or other energy consumption 1 unit (kg/unit)

The second scope, indirect emission, involves release and absorption of greenhouse gases from energy purchased or imported from external agencies. In this research, the major activity of this scope is electricity consumption. The coefficient of greenhouse gas emissions for production of electricity evaluated by Thailand Greenhouse Gas Management Organization is equal to 0.0264 kgCO₂-eq/kWh [4].

**TABLE III. THE COEFFICIENT OF GREENHOUSE GAS EMISSIONS FROM FUEL COMBUSTION IN VEHICLES [3]**

<table>
<thead>
<tr>
<th>Fuel</th>
<th>Vehicle</th>
<th>FCU* (km/L)</th>
<th>CO₂</th>
<th>CH₄</th>
<th>N₂O</th>
<th>GWP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquid Petroleum</td>
<td>Small</td>
<td>17.770</td>
<td>2.1816</td>
<td>0.003</td>
<td>0.0535</td>
<td>2.238</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>15.238</td>
<td>2.1816</td>
<td>0.003</td>
<td>0.0535</td>
<td>2.238</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>12.248</td>
<td>2.1816</td>
<td>0.003</td>
<td>0.0535</td>
<td>2.238</td>
</tr>
<tr>
<td>Compressed Natural Gas (LPG)</td>
<td>8.929</td>
<td>1.6797</td>
<td>0.0413</td>
<td>0.0016</td>
<td>1.7226</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gas (CNG)</td>
<td>11.905</td>
<td>2.1262</td>
<td>0.0872</td>
<td>0.0339</td>
<td>2.2472</td>
</tr>
<tr>
<td>Diesel</td>
<td>Pickup Trucks</td>
<td>11.111</td>
<td>2.6987</td>
<td>0.0036</td>
<td>0.0423</td>
<td>2.7446</td>
</tr>
</tbody>
</table>

*FCU: Fuel consumption unit

The final scope depicts other indirect emissions including fuel combustion from vehicles and water consumption. This also includes waste management and disposal. The coefficient of greenhouse gas emissions from fuel combustion in vehicles from IPCC Guidelines for National Greenhouse Gas Inventories (2006) is shown in Table III [3].

The coefficient of greenhouse gas emissions evaluates by Thailand Greenhouse Gas Management Organization is equal to 0.0264 kgCO₂-eq/kWh. And is equal to 0.47 kgCO₂-eq/kgWaste and 0.0874 kgCO₂-eq/tkm for waste management and transport the waste to the disposal site, respectively [4].

**III. RESEARCH METHODOLOGY**

Data collected from previous steps is converted to the coefficient of greenhouse gas data. The Algorithm for clustering this dataset is as follows.

1. Obtained the input dataset (\( g \)), number of clusters (\( C_n \)).
2. Create prototype cluster.
3. Train prototype cluster in Neuro-fuzzy
4. Repeat :
   - Calculate Mean (\( M_c \)) and Error (\( e_c \)) of each cluster
   - Use \( M_c \) as desired output of each cluster
   - Restart Neuro-fuzzy process
   - Calculate distance between output and \( M_c \) of each cluster
   - Assign object to cluster with the minimum distance
5. Until : sample not change cluster

Since the objective of this experiment is to identify clusters of high and low greenhouse gas emissions, the amount of clusters is set to 2. Dissimilarity matrix generated from the Euclidean distance is displayed in equation (2).

\[
d(i,j) = \sqrt{\sum_{k=1}^{m} (x_{ik} - x_{jk})^2}
\]

The mean of each cluster is used as the cluster prototype calculate from

\[
M_c = \frac{1}{n} \sum_{i=1}^{n} x_i
\]

where \( x_i \) is the member in each cluster, the error of member in each cluster is calculated from

\[
e^2 = \sum_{i=1}^{n} (x_i - M_c)^2
\]

This error is the sum of the squared distance between clusters mean and each data point. The square error of all members in the cluster is calculated from

\[
E^2 = \sum_{c=1}^{N} e^2
\]

As aforementioned, the high performance Neuro-Fuzzy classification is applied in this study [1], [2]. The structure of clustering model in this experiment is illustrated in Fig. 1. The class label from clustering are defined to each datum and are attached to the original dataset.

A dataset with class label is fed to the Neuro-fuzzy. At the first layer, the membership value will be calculated by using the following equation:
\[
\mu_{ij} = e^{-\frac{(y^i_j - c_j)^2}{2\sigma^2_j}}.
\]  
(6)

The Sigmoid function is then used as activation function the output for next layer. Equation (7) and (8) are used in this calculation.

\[
y^o_k = \frac{1}{1 + e^{-z_k}}
\]  
(7)

\[
z_k = \sum_j w_{jk} y^i_j.
\]  
(8)

The error backpropagation algorithm is used in the backward pass of the proposed model. Considering the \(k^{th}\) neuron of the output layer at iteration \(n\), the error signal is defined by the following equation:

\[
e_k(n) = d_k(n) - y^o_k(n)
\]  
(9)

where \(e(n)\) and \(d(n)\) represent the error signal and desired output, respectively. The output signal of neural \(k\) in output layer is represented by \(y^o_k(n)\).

The clustering iteration will stop if the total error of each cluster less than the specified threshold.

IV. EXPERIMENTAL RESULT

Use An iris and breast cancer dataset [10] are used to verify the accuracy of clustering algorithms. High accuracy are acquire from the experimental results. A total of 10 experiments with different initial parameters is completed on each dataset. The result of clustering the iris has an accuracy of 96.20% while the breast cancer dataset results in 95.24% accuracy rate. A comparison between this proposed algorithm and other well-known algorithms are displayed in Table IV. This obviously proves the high efficiency of the proposed algorithms.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>BC</th>
<th>Iris</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed algorithm</td>
<td>95.24</td>
<td>96.20</td>
</tr>
<tr>
<td>k-means[12]</td>
<td>89.30</td>
<td>96.10</td>
</tr>
<tr>
<td>FCM[13]</td>
<td>89.30</td>
<td>95.60</td>
</tr>
</tbody>
</table>

The dormitory dataset has 18 data points represent each building. Each data point involves 33 features representing coefficient greenhouse gas calculated from the student activities. An identification number of 1 to 18 is used instead of dormitory name. A total of 10 experiment has been conducted in this dataset with different initial parameter. The experimental result shows that the first cluster involving building 1, 2, 3, 4, 5, and 17 display a higher ratio if greenhouse gas emission while building 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 and 18 which are classified in the second cluster produce less greenhouse gases. Interestingly, almost all members of the first cluster are men's dormitories except building 17 which is the only unisex facility. In contrast, all women's dormitory is
unexpectedly classified in the second cluster. This could imply that students of different gender have different habits which produce a significantly different amount of greenhouse gases.

V. CONCLUSION

This study attempts to apply a Neuro-Fuzzy technique proposed in [1], [2] to classify emission rates of greenhouse gases. Data from 18 student dormitories is used as the main dataset. All facilities are geographically distributed in random within the campus. Most of them are nearby while a few are rather isolated. After implementing the algorithm, it is found that greenhouse gas production of the samples can be clearly classified into two groups. Interestingly, despite the fact that gender is not utilized as a variable in this study, it is found that male and unisex dormitories produce significantly higher greenhouse gases than all-female facilities. The result of this study clearly suggests a focal and starting point for a strategic campaign on greenhouse gas reduction.

The result further strengthens the validity of the proposed algorithm. Additionally, it can be undoubtedly implemented in other similar research which need high efficiency clustering techniques.

ACKNOWLEDGMENT

The preliminary dormitory and activities information favored from “Report of carbon foot print account of Ching Mai University Dormitory” [9].

REFERENCES


Narissara Eiamkanitchat is working as a lecturer at Department of Computer Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand. She received her Bachelor of Computer Engineering from Chiang Mai University. In 1997 she worked as a production engineer at the Thai Asahi Electronic Device. From 1998 until 2000 she received a scholarship from the Heiwa Nakajima Foundation to be and research student at Mie University, Japan. During that time her research interests are studying on the soft fruit harvesting robot to recognition of leaf’s hidden fruit by using image processing. From 2000 to 2002 she received the Royal Thai Government Scholarship to pursue Master degree in Electrical Engineering from Chiang Mai University. From 2008 until 2010 she received the Thai Government Science and Technology scholarship to pursue Ph.D. degree in Electrical Engineering from Chiang Mai University, Thailand. Her main research areas involve Neuro-fuzzy, computational intelligence for data analysis and data mining for business.