Classification of Ammonia in water for Oil and Gas Industry using Case Based Reasoning (CBR)

Muhamad Faruqi Zahari[1], Muhammad Sharfi Najib[1], Kamarul Hawari Ghazali[1], Mohd Razali Daud[1], Fathimah Abdul Halim[1], Abd. Aziz Mohd Azoddein[2]
[1] Faculty of Electrical and Electronic Engineering, University Malaysia Pahang, Pekan, 26600, Pahang
[2] Faculty of Chemical and Natural Resources, University Malaysia Pahang, Gambang, 26300, Pahang
E-mail: sharfi@ump.edu.my[1]

***Abstract*-- Toxic gasses are exists in environment such as benzene, ammonia and others. Ammonia highly dissolves in water which is sources of human and other species. If the ammonia have high concentration, the effect of human health will be dangerous. Then, using proper monitoring and wastewater management the hazard can be prevented. This paper proposed the intelligence classification technique using an Electronic Nose (E-nose) measurement. The sensor array in the E - nose are used for the inputs of the Case Based Reasoning (CBR) for intelligent classification. The experimental result shows that the technique accomplished to classify with high accuracy which is 100% of accuracy.**

***Index Terms*--CBR, e-nose, ammonia, classification, water.**

I. INTRODUCTION

Nowadays, there are common toxic gasses exists in an environment such as benzene, ammonia, and others [1]. There are several industries that contribute to this toxic gasses in an environment such as medical , food packaging, agriculture, oil and gas and others [2]–[4]. The industries have own regulations to protect human or earth from hazards such as wastewater regulations, hazard managements and others [5]–[8]. Furthermore, water environmental issues have continued since long years ago, which source of the clean water being contaminated due to poor management and monitoring by various unpermitted activities [9]–[12].

One of the famous chemicals in water is Ammonia which caused lung edema, failure of the nervous system, acidosis and kidney damage [13], [14]. In addition, ammonia have highly dissolve in water, which have colorless fluid and pungent smell [15]–[17]. Thus, monitoring and management of wastewater that consist of toxic substance are compulsory to prevent environmental pollution [10], [17].

There are several methods for monitoring such as thermal image processing, Light Detection and Ranging (LIDAR), Electronic Nose (e-nose) and others [18]. E-nose using a concept sensor array for odor classification device [13]. The widely used sensor for chemical detection was metal-oxide gas sensors which used for varied application [19]. Metal-oxide sensor have advantages of low cost, short time response and high sensitivity [20], [21].

The E - nose was a device that functions as human olfactory system which detect the odor and the human brain will classify the odor based on the knowledge [22]. Thus, e-nose will take over the human nose to detect toxic gasses and using intelligence classification in order to define the o (1) toxic gasses [23].

The e-nose detection data will analyze using normalization techniques which enables the comparison between ammonia concentration [24], [25]. In addition, the normalization technique was corrected systematic error which often present and can be removed by normalization method [26], [27]. Furthermore, normalization has several methods such as Range scale, relative scale, baseline subtraction, global method and local method [24], [28]. While, the data analyze using range scale which is [0, 1] as a fixed range for all samples [29], [30].

The simple boxplot displays the several categories of a discrete variable by separating the continuous variable of five statistic which is minimum, first quartile, median value, third quartile and maximum value [31]. The boxplot can differentiate the features of each sensor for each sample by statistical features as mentioned[32], [33].

Regression one of the methods to determine the validity of data point through a set of data point, regression analysis can be used in combination with statistical techniques [34], [35].

There are several classification methods such as an Artificial Neural Network (ANN), k-Nearest Neighbor (KNN), Case Based Reasoning (CBR) and others [36]. CBR was created as a four step process of human or computer reasoning which called CBR cycled [37]. CBR cycled contain of retrieve, reuse, revise and retain which the crucial step of CBR was the retrieval case in order to classify similar cases that stored in the library[36], [38]. CBR used the retrieval method using similarity percentage which based on nearest distance between new cases and stored cases was illustrated in Eq. 1 below [37], [39]–[41]:

$$\text{Similarity}(N_i, S_i) = \frac{\sum w_i \times f(N_i, S_i)}{\sum w_i} \quad (1)$$