Medical Informatics in a United and Healthy Europe
K.-P. Adlassnig et al. (Eds.)
IOS Press, 2009

doi:10.3233/978-1-60750-044-5-615

Optimized Detection of Tar Content in the Manufacturing Process Using Adaptive Neuro-Fuzzy Inference Systems

Zikrija AVDAGIC [1], Lejla BEGIC FAZLIC, Samim KONJICIJA
Faculty of Electrical Engineering, University of Sarajevo, Bosnia and Herzegovina

Abstract. The purpose of this study is to model and optimize the detection of tar in cigarettes during the manufacturing process and show that low yield cigarettes contain similar levels of nicotine as compared to high yield cigarettes while B (Benzene), T(toluene) and X (xylene) (BTX) levels increase with increasing tar yields. A neuro-fuzzy system which comprises a fuzzy inference structure is used to model such a system. Given a training set of samples, the Adaptive Neuro-Fuzzy Inference System (ANFIS) classifiers learned how to differentiate a new case in the domain. The ANFIS classifiers were used to detect the tar in smoke condensate when five basic features defining cigarette classes indications were used as inputs. A classical method by High Performance Liquid Chromatography (HPLC) is also introduced to solve this problem. At last the performances of these two methods are compared.

Keywords. ANFIS, fuzzy logic, HPLC, tar detection

1. Introduction

All cigarettes share the same features like nicotine, carbon monoxide (CO), length, filter type and tar concentration with few minor differences. Neuro-fuzzy systems are fuzzy systems which use artificial neural networks (ANNs) theory in order to determine their properties (fuzzy sets and fuzzy rules) by processing data samples. The optimization of a process is to find a setting of the variables so as to obtain the best performances of this process. The cigarette quality process has five basic controllable variables (input variables): nicotine, CO, length, filters type and tar concentration. In the first stage, we identify the relationship between responses and input variables through regression functions. In the second stage, we use optimization techniques to obtain a set of input variables, which give the system the most desirable responses.

In this study, considering the nonlinear responses of a process, a neuro-fuzzy system is employed to model a target system. A specific approach in neuro-fuzzy development is the ANFIS. In ANFIS, the membership function parameters are extracted from a data set that describes the system behavior [1, 2].

[1] Corresponding Author: Prof. Zikrija Avdagic, PhD, Faculty of Electrical Engineering, University of Sarajevo, Zmaja od Bosne bb.; Tel: +387 33 250 737; E-mail: zikrija.avdagic@etf.unsa.ba.

2. Problem Description

The regulatory bodies require that the cigarettes be tested for nicotine and tar content.

The tobacco industry performs these measurements using a standard testing set-up to categorize products as being ultra-low to high tar yield, using extra sophisticated equipment (HPLC) and trained staff. Mainstream (MS) smoke is the smoke which is directly inhaled by the smoker, whereas side stream (SS) or secondhand smoke is the smoke which is released to the environment from the burning cigarette. The tar yield, measured in mg of tar per cigarette, is the amount of tar trapped in a filter during standardized smoking.

The nicotine and the tar content were measured by a machine that smokes each cigarette in exactly the same way. Many cigarette brands have three ventilation slits in or near the filter, which dilute the concentration of tar, CO and other compounds in the MS smoke. There is a similar problem with low-tar cigarettes, which achieve their low tar status by putting small slits in the sides of the filters to dilute the inhaled smoke with air. Smokers can easily cover these vents with their fingers, knowingly or not, resulting in the inhalation of just as much tar as a regular cigarette. After the burning, where the cigarette is artificially 'smoked', the compounds from the cigarette smoke are collected and measured using HPLC and predicted by ANFIS.

Tar is a term used to describe the toxic chemicals found in cigarettes. The concentration of tar in a cigarette determines its rating: high-tar cigarettes contain at least 22 milligrams (mg) of tar; medium-tar cigarettes contain anywhere from 15 mg to 21 mg of tar; low-tar cigarettes contain 5 mg or less of tar [3]. Our goal was to predict the concentration of tar content in a cigarette filter. We used different classes of cigarettes with and without filters. The studies show that when people smoke low-nicotine cigarettes, they tend to inhale more deeply and more frequently than they would with a regular cigarette, compensating for the lower nicotine and tar. Although the cigarette might contain less nicotine and tar, the smoker actually inhales the same amount as in a normal cigarette.

3. ANFIS Architecture

From Sugeno Fuzzy Model [1], ANFIS was proposed by Roger Jang in 1992 [2, 4]. The architecture of a two-input two-rule ANFIS is shown in Figure 1.

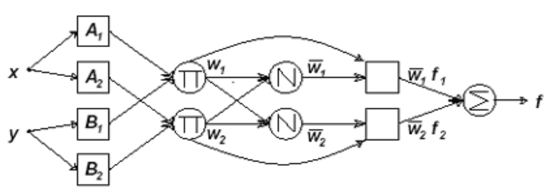

Figure 1. ANFIS architecture

The ANFIS has five layers, in which node functions of the same layer have the same function type. Layer 1**:** Every node *i* in this layer is an adaptive node with node function, see Eq. (1) or Eq. (2),

$$O_{j,i} = \mu_{A_i}(x) = \frac{1}{1+\left(\frac{x-c_i}{a_i}\right)^{2b_i}} \tag{1}$$

$$O_{j,i} = \mu_{B_i}(y) = \frac{1}{1+\left(\frac{y-c_i}{a_i}\right)^{2b_i}} \tag{2}$$

where $\{a_i, b_i, c_i\}$ are premise parameters updated through Back Propagation Learning Algorithm.

Layer 2: Every node i in this layer is a fixed node labeled Π, whose output is the product of all the incoming signals, see Eq. (3).

$$O_{2,i} = \varpi_i = \mu_{A_i}(x) \times \mu_{B_i}(x) \tag{3}$$

Layer 3: Every node i in this layer is a fixed node labeled N. The *ith* node calculates the ratio of the *ith* rule's firing strength to the sum of all rules' strengths, see Eq. (4).

$$O_{3,i} = \varpi_i = \frac{\varpi_i}{\varpi_1 + \varpi_2} \tag{4}$$

Layer 4: Every node i in this layer is adaptive node with node function refer to equation in the text as Eq. (5),

$$O_{4,j} = \varpi_i f_i = \varpi_n (p_i(x) + q_i(y) + r_i) \tag{5}$$

where $\{p_i, q_i, r_i\}$, are consequent parameters [2, 4], updated through Recursive Least-Squares Estimation.

Layer 5: The single node in this layer is a fixed node labeled Σ, which computes the overall output as the summation of all the incoming signals, see Eq. (6).

$$O_5 = \sum_i \varpi_i f_i = \frac{\sum_i \varpi_i f_i}{\sum_i \varpi_i} = \frac{\varpi_1 f_1 + \varpi_2 f_2}{\varpi_1 + \varpi_2} \tag{6}$$

Functionally speaking, the ANFIS architecture is completely equivalent to a Sugeno fuzzy inference system.

4. Experimental

The prediction of tar concentration in a filtered cigarette is a typical nonlinear regression problem, in which several attributes of the cigarette's profile information are used to predict another continuous attribute, that is, the tar concentration. We used data collected from cigarettes of various manufactures.

The five input attributes are: type of filter, length of a cigarette/mm, CO content of cigarette/g and nicotine/mg; the output variable to be predicted is the tar concentration. Data base investigated in this study consisted of 586 analyzed data cases – a part of the data is shown in Table 1 [5].

Table 1. Used data set

Nic	Tar	CO	Brand Name	Type
.1	<.5	<0.5	Carlton	100 Filter Hard Pack Ultra Light
.1	<.5	<0.5	Now	King Filter Hard Pack
.1	1	1	Bristol	King Filter Soft Pack Lowest
.5	4	5	True	King Filter Soft Pack
.5	5	4	Capri	100 Non Filter Hard Pack Ultra Light

We partitioned the data set into a training set and a checking set and attempted to find the input attributes that have a better prediction power for ANFIS modeling. The feature 'Nicotine' had a bigger impact on the tar in a cigarette than the type of filter. The training and checking errors are distinguished, indicating the outset of overfitting. The least squares method can be used to identify the optimal values of these parameters. The output of the ANFIS is calculated by employing the consequent parameters found in the forward pass. The output error is used to adapt the premise parameters by means of a standard backpropagation algorithm [4]. We trained the sixth ANFIS classifier to combine the predictions of the five ANFIS classifiers. The outputs of the five ANFIS classifiers were used as the inputs of the sixth ANFIS classifier. The training and the checking data are shown in Figure 2.

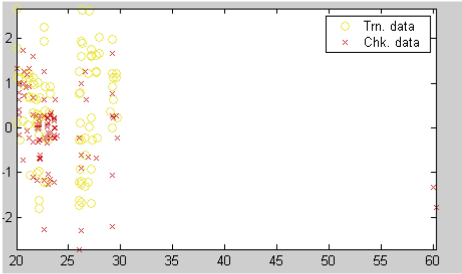

Figure 2. Training and checking data

The input-output surface of the best two-input ANFIS model for tar prediction is shown in Figure 3**.** This is a surface in which the predicted tar increases with the increase in 'Nicotine' and decrease in 'Filter type'.

The training of the Root Mean Squared Error (RMSE) is 0.0338; the checking of RMSE is 0.0321% [5]. Prediction results of the amount of tar in manufacturing process require a number of additional chemical analyses. The results of nicotine and tar measured via HPLC are detectable with 0.19% error: from 1.25 µg/ml, and the standards are linear up to 4,000 µg/ml. The preliminary data shows that BTX levels (in nicotine) increase with increasing tar yields. ANFIS access enables faster prediction of tar amount with a permitted deviation which is less than 0.03387%.

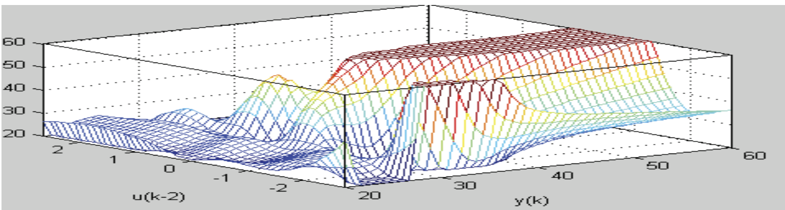

Figure 3. Input-output surface

5. Conclusion

This paper presented a new application of the ANFIS model for the detection of tar in different types of cigarettes. The five ANFIS classifiers were used to detect tar in cigarettes. The predictions of the ANFIS classifiers were combined with the sixth ANFIS classifier. The total classification accuracy of the ANFIS model was with error less than 0.03387%. For comparison, results of nicotine and tar measured via HPLC is detectable with 0.19% error and simple linear regression [6] using all input candidates, results in a training RMSE of 1.32, and a checking RMSE of 1.02.

Therefore, we conclude that the proposed ANFIS model can be used in detecting the tar content in cigarettes. Also, low yield cigarettes contain similar levels of nicotine when compared to high yield cigarettes while BTX levels increase with increasing tar yields. ANFIS access is to be successfully used in biomedical engineering and genomics. By defining and training ANFIS classifiers, as well as by combining results to a new classifier, we can achieve good results especially in diagnosis of a disease with similar symptoms.

References

[1] Dubois, D., Prade, H. (1998) An introduction to fuzzy systems. *Clinica Chimica Acta* 270:3–29.

[2] Jang, J.-S.R. (1993) ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics* 23(3):665–685.

[3] Wahdan, A., Lesouëf, P.N., Devadason, S.G. (2005) Quantification of nicotine and toxic components from cigarette output by HPLC. *Respirology* 10(Suppl 1):A13.

[4] Jang, J.-S.R. (1992) Self-learning fuzzy controllers based on temporal backpropagation. *IEEE Transactions on Neural Networks* 3:714–723.

[5] Avdagic, Z., Begic Fazlic, L., Fatkic, H. (2007) Automatic detection of CO content in smoke condensate using adaptive neuro fuzzy inference systems. In *Proceedings of the International Symposium on Innovations in Intelligent Systems and Applications, Inista 2007*, 86–91.

[6] McIntyre, L. (1994) Using cigarette data for an introduction to multiple regression. *Journal of Statistics Education* 2(1).