# Real Time Monitoring Glycerol Esterification Process with Mid IR Sensors using Support Vector Machine Classification

Iwan Aang Soenandi, Taufik Djatna, Ani Suryani and Irzaman Husein
*Post Graduate Program of Agro-industrial Technology, Bogor Agricultural University*
*E-mail:* iwansoenandi@gmail.com, taufikdjatna@ipb.ac.id, anisuryani.sbrcipb@gmail.com, irzaman@ipb.*ac.id*

***Abstract*—The commercial synthesis of fatty acid esters of glycerol is important as it plays role in other derivative production varieties. This research aims to construct the monitoring system for faster esterification status identification and increase efficiency of energy used for production. The monitoring systems are based on the measurement parameters from two sources LED mid IR 3,4 µm and 5,5 µm and two detectors that connected using the data acquisition system with computer database via USB 2.0 and classifying the esterification status using the Support Vector Machine (SVM) classifier. The purpose of SVM method is to classify the variations of parameter inputs from the LED mid IR sensors in real time monitoring connected with microcontroler. In this research, three esterification status divided for monitoring process in bioreactor. The construction of classification based on SVM deployed in orange system software. In the application of esterification monitoring, the influence of various parameters such as temperature set in the reactor has relation to the process time needed. By monitoring this system status in every minutes, we obtained one of the experiment in optimum process in temperature 210°C is 2 hours.**

**Keywords:** *Glycerol ester, monitoring, esterification, LED IR sensor, SVM.*

## I. INTRODUCTION

The commercial synthesis of fatty acid esters of glycerol is carried out by two different ways: direct esterification of fatty acid with the glycerol and catalysis by a homogenous acid [1]. In Indonesia, synthesis by direct esterification is widely used in the esterification of glycerol because this process is simple and feasible in batch production system [2]. The major factors affecting the conversion efficiency of esterification process are molar ratio of oil, amount of catalyst, reaction temperature, catalyst type and stirring speed according to the reaction duration [3]. However, in this research we focus on monitoring the process by relating it to the temperature and process time needed. Determination of the esterification reaction is highly desirable for glycerol ester product in order to increase the efficiency energy and the cost of production. A current method for determination of esterification is sampling, which needs high cost and time. In this paper, to determine the true esterification status in real time, direct and close online monitoring of the product and critical components is highly desirable. Optical measurement techniques are promising candidates in spectral region i.e. multi-channel NDIR (non-diffractive infrared) absorption or IR spectroscopy. The later method is also used in laboratories, thus allowing better correlation of online data and laboratory results with data acquisition interfaced by database to the computer. SVM method used to find the correlation is used for calibration parameter in online measurement sensors, especially to identify the esterification status. The position of sensors from Sinelli et al. demonstrated a set up with detector array combined with the gradient filter to avoid the need for movable parts [4] was inspirating and applied. Wiesent et al. presented a system with an infrared source, a fluid cell consisting of two sapphire windows and a quadruple infrared detector equipped with different filter windows for analysis of phosphate ester [5]. The objective of this paper is to develop a system for the identification of

This work was supported by Lembaga Dana Pengelola Pendidikan with contract number PRJ1600/LPDP/2014.
I.A.Soenandi is with the Post Graduate Program of Agroindustrial Technology, Bogor Agricultural University, Dramaga, Bogor (iwansoenandi@gmail.com)
T.Djatna is with the Post Graduate Program of Agroindustrial Technology, Bogor Agricultural University, Dramaga,Bogor (taufikdjatna@ipb.ac.id)
A.Suryani is with the Post Graduate Program of Agroindustrial Technology, Bogor Agricultural University, Dramaga,Bogor (anisuryani.sbrcipb@gmail.com)
I. Husein is with Department of Physics, Faculty of Mathematics and Natural Sciences , Bogor Agriculture University , Dramaga Bogor (e-mail: irzaman@ipb.ac.id).

 ISBN : 978-1-4673-7404-0

esterification status using LED mid IR sources, thermopile detectors and data acquisition system, as well as to find the variable condition related to temperature in the reactor and the duration of the esterification process. In Section 2 of this paper, we briefly explain the materials and methods. In Section 3, we briefly describe the calibration. In Section 4, the review of concept of classifying esterification status with SVM. Finally, in Section 5 discussed about the conclusion.

## II. Materials and methods

### A. Materials

In this research,for the raw material we used pure glycerol with purity level of 85-90% mixed and synthesized it with Oleic acids pro analysis grade specification in 1:1 ratio charged in reactor with heater and stirring blade. For the catalyst, we used Methyl Ester Sulfonic Acid (MESA) of 0.5%.

### B. Equipment

Esterification reactions were carried out in a laboratory-built apparatus. Apparatus consists of 250 ml laboratory conical flask with 30 ml working volume. The esterification reaction was under atmospheric pressure (opened system), and the temperature of the reactor was controlled using hot plate (controlled by internal thermostat) as in Figure 1. All the reactants (oleic acids, pure glycerol and catalyst) were weighted and charged into the reactor. Then, the temperature was increased by adjusting the thermostat. The magnetic stirrer was allowed to operate after 2-3 min (to heat up the mixture). After passing the desired reaction time, the reactor was removed from the hot plate. Samples were withdrawn from the reaction mixture for analysis. The reaction mixture was cooled down to the ambient temperature by immersing it into water bath. The esterification process in closed system was also investigated, where all the reactors were isolated.

### C. Variation Condition

Several reaction and variation condition were tested in Surfactan Bioenergy Research Center (SBRC) IPB Baranangsiang Bogor laboratorium, in order to get various conditions of the temperature and reaction time. The temperature was varied between 150 ºC, 160 ºC and 165ºC. On the other hand, the process time was varied between 180, 240 and 260 minutes.

Fig 1. Apparatus experiment in laboratory

### D. Sensor System

A commercial infrared LED source as original growth of narrow gap semiconductor alloys onto n+-InAs substrate, optical coupling through the use of chalcogenide glasses and Si lenses with antireflection coating (Boston Electronic,2014). Three types Mid IR sensors of 3,4 micron LED-34SR full thread body , 5,5 micron LED-55SR full thread body and 7 micron OPLED 70 full thread body, with the specification in Table 1-3, also the thermopile detectors from Heimann HTIA Dx-Tx used as in Figure 2. For detecting the sample as reach steady state and to record signal amplitude with a good signal to-noise ratio and position layout of the sensor and detector as in Figure 3.

Fig 2. LED MID IR Source and Detector

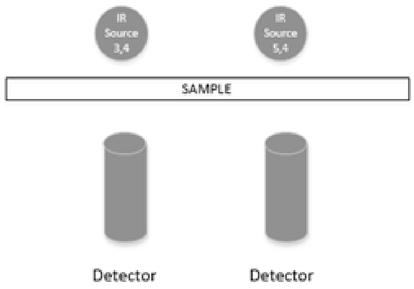



Fig 3. Position LED IR Source and Detector

Table 1.
Specification of LED34Sr

| | | |
|---|---|---|
| Peak wavelength | µm | 3,4±0.05 |
| Pulse Power | mW | 0.25÷0.35 |
| CW Voltage | V | Drive Current 0.2A 0.26÷0.29 |

Table 2.
Specification of LED55Sr

| | | |
|---|---|---|
| Peak wavelength | µm | 5,4÷5.5 |
| Pulse Power | µW | 5÷7 |
| CW Voltage | V | Drive Current 0.2A 1.5÷2.5 |

Table 3.
Specification of OPLED70

| | | |
|---|---|---|
| Peak wavelength | µm | 6,5÷7.0 |
| Pulse Power | µW | 5÷7 |
| CW Voltage | V | Drive Current 0.2A 1.5÷2.5 |

 ISBN : 978-1-4673-7404-0

### E. System Data Acquisition

To ensure the data parameters get collected from the sensors, the system was interfaced with database such as MySQL. In this research, we built the data acquisition system by using microcontroller ATMEGA 8535 with Universal Serial Bus (USB) 2.0 connector to connect the data streaming in real time to the computer (Figure 4). After that, the output was processed using SVM method with personal computer [6].

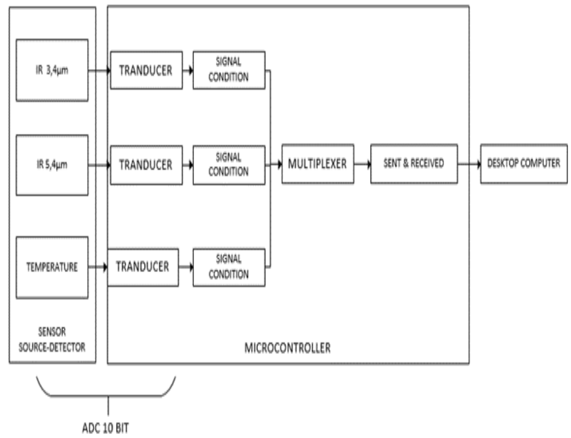



Fig 4. System block diagram for sensors and microcontroller

### F. Relief

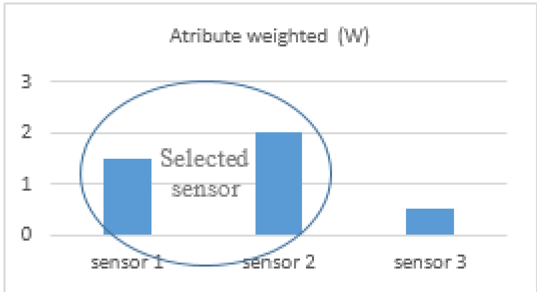
In this work, to analyze the sensors closely related to esterification forming parameter, we used Recursive Elimination of Features (Relief) algorithm, which is generalizable to polynomial classification by decomposing it into a number of binary problems proposed by Kira and Randall [7].The equation of Relief is represented as in (1). To perform Relief in this research, first, defines $i^{th}$ output sensor as $f_i$ where **F** = $\{f_1, f_2,..., f_m\}$, the weight of the $i^{th}$ output sensor as $w_{fi}$ where **W** = $\{w_{f1}, w_{f2},\ldots,w_{fm}\}$, and the result of esterification condition for $m$ dimensional features ($i= 1, 2,\ldots,m$) and $n$ sample size ($t = 1, 2,\ldots, n$) which denotes as $x_{ti}$ where **X** = $\{x_{t1}, x_{t2},\ldots, x_{tm}\}$.

$$w_{f_i} = \frac{\sum_{t=1}^{(n-1)} \sum_{j=(t+1)}^{n} diff\,(x_t, y_j)^2}{(n-1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Traditionally, Relief equation above runs using pairwise between data record ($x_t$ and $y_j$) where $x_t$ indicates an instance of output and $y_j$ indicates an instance neighbor of $x_t$. In equation above, function *diff* $(x_t, y_j)^2$ is calculated as in (3), where $r_{fi}$ is a range unit of $i^{th}$ esterification status to normalize the values of the function into the interval [0, 1].

$$diff\,(x_t, y_j)^2 = \left(\frac{x_t - y_j}{r_{f_i}}\right)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

As in (3), if pairwise between $x_t$ and $y_j$ are in the same class, then the value of function *diff* is negative. Otherwise, function *diff* will have a positive value when $x_t$ and $y_j$ are in a different class. Kononenko [8] said that the relevant features means the features which has weight above 0. The result of weights of design elements of Relief which generated using R language which threshold was set on zero ($\tau = 0$) is represented in Figure 5. As the result of attribute weighted in Relief, the sensors that closely related with esterification identification was 3,4µm and 5,5µm




Note: Sensor 1: 5,5µm, Sensor 2: 3,4 µm and Sensor 3: 7µm

Fig 5. Relief Attribute weighted from 3 sensors for esterification identification

### G. Support Vector Machine (SVM)

Support Vector Machine (SVM) was first heard in 1992, introduced by Boser, Guyon, and Vapnik in COLT-92. Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression [9]. SVM is a useful technique for data classification.

For this type of SVM, training involves the minimization of the error function:

$$\frac{1}{2}\omega^T\omega + C\sum_{i=1}^{N}\zeta_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

subject to the constraints:

$$y_i(\omega^T\phi(x_i)+b) \geq 1-\zeta_i \quad \ldots\ldots\ldots\ldots\ldots(4)$$

$$\text{and } \zeta_i \geq 0, i = 1,\ldots\ldots, N \quad \ldots\ldots\ldots\ldots\ldots(5)$$

where C is the capacity constant, ω is the vector of coefficients, b is a constant, and $\xi_i$ represents parameters for handling non-separable data (inputs). The index i labels the N training cases. Note that $y \in \pm 1$ represents the class labels and xi represents the independent variables. The kernel $\phi$ is used to transform data from the input (independent) to the feature space. It should be noted that the larger the C, the more the error is penalized. Thus, C should be chosen with care to avoid over fitting. A classification task like esterification status process usually involves with training and testing data, which consist of some data instances. Each instance in the training set contains one-target values and several attributes, which tested in laboratory using sample process in each, attribute. The goal of SVM is to produce a model, which predicts target value of data instances in the testing set which are given only the attributes [10]. In this research the target value is esterification status.

## III. CALIBRATION

### A. Calibration

To ensure the precision esterification measurement, this system calibrated by 3 step method: testing with blank sample, full closed sample (using a sheet of paper) and tested glycerol ester FTIR spectrum with parameter 180ºC and 150 minutes process as in Fig 6.

TABLE 4.
BLANK SAMPLE

| Blank Sample | Digital Output | | Transmittance Calibration | |
|---|---|---|---|---|
| Time (Second) | 5,5µm | 3,4µm | 5,5µm | 3,4µm |
| 1 | 205 | 258 | 100 | 100 |
| 2 | 209 | 260 | 100 | 100 |
| 3 | 208 | 258 | 100 | 100 |
| 4 | 210 | 257 | 100 | 100 |
| 5 | 207 | 258 | 100 | 100 |
| 6 | 208 | 256 | 100 | 100 |
| 7 | 209 | 255 | 100 | 100 |
| 8 | 207 | 259 | 100 | 100 |
| 9 | 208 | 257 | 100 | 100 |
| 10 | 206 | 258 | 100 | 100 |
| Average | 207.7 | 257.6 | 100 | 100 |

TABLE 5.
FULL CLOSED SAMPLE

| Blank Sample Time | Digital Output | | Transmittance Calibration | |
|---|---|---|---|---|
| (Second) | 5,5µm | 3,4µm | 5,5µm | 3,4µm |
| 1 | 178 | 203 | 0 | 0 |
| 2 | 178 | 203 | 0 | 0 |
| 3 | 182 | 201 | 0 | 0 |
| 4 | 181 | 202 | 0 | 0 |
| 5 | 182 | 203 | 0 | 0 |
| 6 | 182 | 201 | 0 | 0 |
| 7 | 182 | 201 | 0 | 0 |
| 8 | 181 | 202 | 0 | 0 |
| 9 | 178 | 201 | 0 | 0 |
| 10 | 179 | 201 | 0 | 0 |
| Average | 180.3 | 201.8 | 0 | 0 |

After we get the average value for the two condition (blank sample and closed sample) to described transmittance upper and lower limit value for the detector, we can find the calibration factor for this identification system. The calibration used formula as in Equation 3.

$$\textit{Transmitance (\%)} = \frac{x_n - \frac{\sum x_{iclose}}{n_{iclose}}}{\frac{\sum x_{iblanko}}{n_{iblanko}} - \frac{\sum x_{iclose}}{n_{iclose}}} \quad ....... \ (6)$$

Where:

Xn= bit number digital output;

$$\frac{\Sigma X_{i\ Close}}{n_{i\ close}} = average\ bit\ number(closed);$$

$$\frac{\Sigma x_{i\ blanko}}{n_{i\ blanko}} = average\ bit\ number\ (blanko)$$

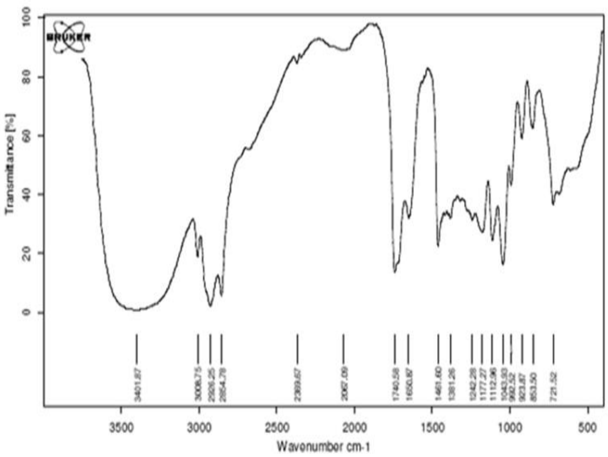
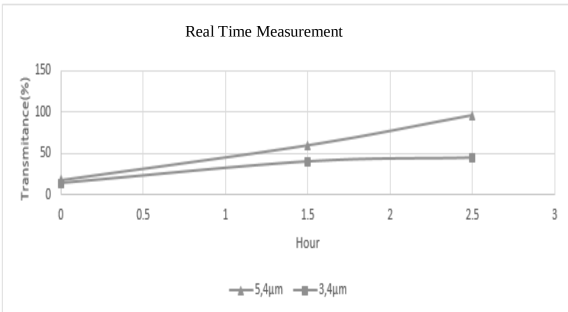


Fig 6. FTIR spectrum of Glycerol Ester 180ºC-150 minutes

## IV. RESULT AND DISCUSSION

### A. Online Measurement

In the laboratory, we doing experimental process using this sensor and data acquisition with temperature from 160ºC until 210ºC with various process time. The graphical data was describe in Fig 7. With data acquisition system, we can get the real time data plotted in interval of 3 seconds and average it in each hours to compare with SVM method




In our research, we used SVM method to classify the input parameter identified by sensor as a transmittance parameter. The performance of SVM was tested by using ROC analysis.

### B. *Esterification Status with SVM*

For the data acquisition in this research, we used online data measurement interfaced with SQL database and identified the parameter inputs by using optical sensors MID IR that attached in bioreactor . The input database was collected and clustered into esterification status by using SVM method. Using computer software application Orange ver 2.6.1, the SVM method was trained and tested by processing the data file of 250 examples, 2 attributes and Classification Discrete Class with 3 values as in Figure 8 and the knowledge flow Orange software schema as in Figure 9 with C-SVM using C=1,00 and Kernel type:Sigmoid.

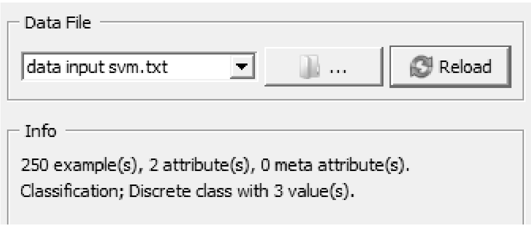



Fig 8. Data input from database (Orange Software)

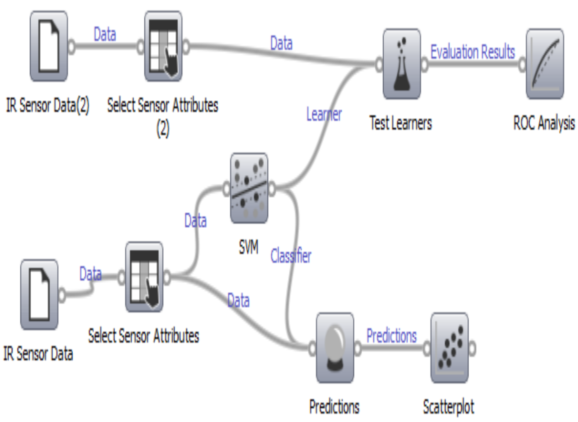



Fig 9. Knowledge flow schema in Orange software for data mining

### C. *ROC Analysis*

To measure the performance of SVM classifier, we used test learners with sampling Cross validation 5 number folds. The evaluation result was as in Table 6 with Classification Accuracy 95,58%, sensitivity 97.81% and specifity 95.54%. And another test method this classification was tested by ROC curve is a graphical plot that illustrates the performance of a classifier system as its discrimination threshold is varied. The curve generated by Orange Software data mining was shown by plotting the true positive rate against the false positive rate at various threshold settings. From the three predicted classes in Figure 10, we found that the classifying performance of SVM was excellent. The interpretation of ROC curve is similar to a single point in the ROC space, the closer the point on the ROC curve to the ideal coordinate, the more accurate the test is. The closer the points on the ROC curve to the diagonal, the less accurate the test is.

TABLE 6.
TEST LEARNER EVALUATION RESULT

| CA | Sens | Spec |
|---|---|---|
| 0.9558 | 0.9781 | 0.9554 |

### D. *Effect of temperature and time*

Both temperature and time of the process had major effects on the conversion of the esterification process. Accordingly, they were studied together. The obtained results showed that by increasing the reaction temperature, the reaction conversion also increased rapidly. Table 5. Shows that after 2 hours, the esterification reached a well-formed esterification status.

Therefore, another esterification reaction was carried out within a temperature range of 120°C-210°C (maximum heater temperature) as shown in Table 7. The results revealed that by increasing the esterification time, the esterification yields also increased up to a maximum conversion. To determine the time needed of esterification process, besides considering the maximum yields of esterification, it was also necessary to take the time required to reach the reaction temperature into account (Table 8). The length for heating time to reach the reaction temperature was certainly longer and the energy consumption was surely greater for higher reaction temperature. Consequently, a faster reaction at a set temperature was desirable.

TABLE 7.
EXAMPLE IDENTIFICATION ESTERIFICATION WITH SVM

| Time (minutes) | Temperature(ºC) | Esterification Status With SVM |
|---|---|---|
| 40 | 120 | Initialize |
| 78 | 180 | On Process |
| 120 | 210 | Finish |

TABLE 8.
EXAMPLE DATA TESTING WITH SVM

| TEMPERATURE (CELCIUS) | TIME (MINUTES) | ESTERIFICATION STATUS |
|---|---|---|
| 170 | 180 | Finish |
| 190 | 140 | Finish |
| 210 | 120 | Finish |
| 210 | 135 | On Process |

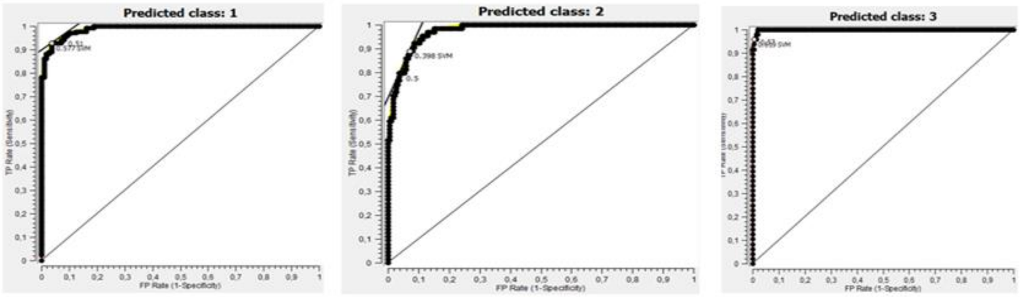



Fig 10. ROC analysis curve

## V. CONCLUSION

The system of real time monitoring glycerol esterification process with mid IR sensors classyfying with SVM, was contributed to support the identification esterification status in every minutes and to get information for the time needed for the esterification process. This esterification status achieved a good performance when classify into 3 status: Initialize, on process and finish. This classification was trained and tested in Orange Software for data mining using SVM method whereas the performance of the classifier was tested using ROC analysis. In applying for esterification optimization, the influence of various parameters, such as the temperature set in the reactor, had a relation to the process time needed. By using this monitoring system based on the measurement and classification of esterification forming using SVM from two inputs of LED mid IR 3,4 and 5,5 µm sensors, we obtained the optimum process condition was 210ºC and the time needed for the process was 120 minutes.

## ACKNOWLEDGMENT

The authors would like to acknowledge the financial support of Lembaga Pengelola Dana Pendidikan (LPDP) Kementerian Keuangan RI number: PRJ-1600/LPDP/2014

## REFERENCES

[1] D.Isabel,M. Federico, M., Joaquın, P. and Enrique, S. Synthesis of MCM-41 materials dialkyl silane groups and their catalytic activity in functionalised with the esterification of glycerol with fatty acids. Applied Catalysis A: General, 242, 161-169. doi:10.1016/S0926-860X(02)00501-X .2003

[2] Pardi.2005. Optimasi proses produksi gliserol monooleat dari gliserol hasil samping pembuatan biodiesel [tesis].Medan(ID):Sekolah Pascasarjana Universitas Sumatera Utara.

[3] N.A Mostafa, A. Maher,"Production of Mono-,Di,-, and Triglycerides from waste fatty acid through esterification with glycerol". *Advanced in Bioscience and Biotechnology* Vol ,4 pp.900-907.2013

[4] N.Sinelli, de Dionigi, S, M.A Pagani, M. Riva, P.Belloni,"Application of NIR spectroscopy in on-line monitoring of doughproofing and bread staling". *Tecnica Molitoria* Vol 55,pp. 1075–1092.2004

[5] B.R Wiesent, D. Dorigo, and A.Koch: Suitability of tunable Fabry-Perot spectrometers for condition monitoring purposes of gear oils in offshore wind turbines, Proceedings IRS2 2011 – 12th International Conference on Infrared Sensors and Systems, Nuremberg, 82–87, doi:10.5162/irs11/i4.2, 2011.

[6] J.Kwok,"Moderating the outputs of support vector machine classifier". *IEEE Trans. Neural Networks* vol 10,pp. 1018–1031.July 1999.

[7] K. Kira ,L.A Rendell . The feature selection problem : traditional methods and a new algorithm. Proceedings of Tenth National Conference on Artificial Intelligence. di San Jose, California. Publication by The AAAI Press,Menlo Park, California:pp.129-134. 12-16 July, 1992

[8] I.Kononenko. Estimating attributes: Analysis and extensions of relief. In ECML-94: Proceedings of the European Conference on Machine Learning, Secaucus, NJ, USA. New York: Springer. 171–182.1994

[9] V. Vapnik, S. Golowich, and A. Smola. Support vector method for function approximation, regression estimation, and signal processing. In M. Mozer, M. Jordan, and T. Petsche, editors, Advances in Neural Information Processing Systems 9, pages 281– 287, Cambridge, MA, 1997. MIT Press.

[10] N. Cristianini and J.S.Taylor, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods", Cambridge University Press, 2000.