HORA 2019
www.horacongress.com
International Congress on Human-Computer Interaction, Optimization and Robotic Applications
July 5-7, 2019, Urgup, Nevşehir, Turkey
SETSCI Conference Proceedings
www.set-science.com
8: 11 - 14, 2019

Support Vector Machine Classifier of Grain Stored in Silo-bags by Using Wireless Network of Temperature and Moisture Sensors

Shawqi Mohammed Othman Farea[1], Muzaffer Kanaan[2]

[1]*Department of Mechatronics Engineering, Erciyes University, Turkey*
[2]*Department of Mechatronics Engineering, Erciyes University, Turkey*
(shawqifarea91@gmail.com)

Abstract – We propose using support vector machines (SVMs) as a machine learning classifier in a grain-condition-monitoring system. This system is used to proactively monitor the silobag-stored grain condition. The system consists of a wireless sensor network (WSN) and a machine learning classifier. As far as the hardware is concerned, the system relies upon a wireless network of temperature and moisture sensors; and as far as the software is concerned, SVM is used to classify the grain status according to the data coming from the WSN. The grain condition is classified into three classes: safe, risky and dangerous. We use two different multiclass classification algorithms together with SVM. The first algorithm is the one-against-all algorithm and the second is the error-correcting output codes (ECOC). Both algorithms accurately classify the grain status such that the proportion of the misclassified cases from the testing data is zero and, therefore, the classification accuracy is 100%. Nevertheless, the ECOC algorithm is faster than the one-against-all algorithm in terms of the training and testing time. In addition, we compare our results with a previous work, which used artificial neural networks (ANNs) as the classifier, and the current results are obviously better than the previous results.

Keywords –Wireless sensor network, Support vector machine, One-against-all , Error-correcting output codes, Grain status, Silo-bags

I. Introduction

Wireless sensor networks (WSNs) have potential applications in different aspects of modern life, including health care, energy, military surveillance, industry, transportation and agriculture, which is the focus of this paper [1]. By means of WSNs, it is possible to create a smart device which consists of distinct components communicating wirelessly with one another and with their external environment.

In literature, many applications of WSNs have been developed in the area of agriculture and food industry [2]. For continuously monitoring and controlling the condition of stored grain, WSNs are a useful inexpensive method. For example, a WSN can be deployed to monitor the temperature of grain storage, leading to an improvement in the quality control of the commodity [3]. Another application of WSNs in agriculture is a wireless network of moisture and temperature sensors for proactive monitoring and classification of the condition of grain storage [4][5].

In the context of monitoring and controlling the condition of grain storage, many studies have been conducted to address this issue. Some studies concentrate on the overall state of the grain itself to discover whether it is safe or not whereas other studies focus on whether grain storage is infested by insects or not. One of these studies checks the condition of silo-bags filled with soybean and wheat and determines whether storage problems exist according to measurements of the carbon dioxide (CO_2) concentration and moisture content of the storage [6]. Another study classifies cereal grains using four intelligent classification algorithms: statistical classification, fuzzy logic algorithm, hybrid algorithm and ANNs. In that study, ANNs show the best results when compared with the rest where the classification accuracy is 98% [7].

The other group of studies, which detect the insect infestation, use different techniques. Some of the proposed techniques are machine vision, thermal imaging, carbon-dioxide measurements, e-noses and finally acoustic detection, which is the most promising technique among the proposed methods [8]. In fact, some of these methods can be also used to monitor the condition of grain storage and discover when the storage is spoiled and when it is safe. One of the most commonly used method is CO_2 measurement.

Most of the research in the area of grain storage classification relies solely upon the CO_2 measurements. Although CO_2 concentration can be an indicator of the spoilage of the grain storage, it is neither accurate nor sufficient. That is because the temperature and moisture content are assumed to be constant in such systems whereas they are not constant in practice [5]. Yet, the temperature and moisture content of grain storage are two major factors in the process of indicating the condition of the stored grain. In this paper, we use the measurements of the temperature and moisture content of the stored grain. These data come from a wireless network of temperature and moisture sensors and can be used to classify the grain condition into three classes: “safe”, “risky” and “dangerous”. If the condition of the storage is safe, then it is unspoiled and suitable for human and animal.consumption. In contrast, the grain storage that

DOI: 10.36287/setsci.4.5.003

falls under the risky and dangerous classes is spoiled and not appropriate for both human and animal consumption. The difference between the risky and dangerous classes is that the dangerous class is highly spoiled. Thus, in this paper, we propose using support vector machines (together with preprocessing) as an efficient alternative machine learning classifier to artificial neural networks in the application of monitoring the grain status based on a wireless network of temperature and moisture sensors. In addition, we compare our results with a previous work where the artificial neural networks (ANNs) were used as the machine learning classifier.

The rest of this paper follows the following order: Section II summarizes the methods and the machine learning algorithms (SVMs) used in this paper, Section III presents the results of using SVMs for classifying the collected data set and discuss the results, and Section IV concludes this paper.

II. Materials and Method

In this section, we describe the grain-status-monitoring system that was constructed to collect the data set. Then, we illustrate the data set and how it is partitioned into different parts. Lastly, we explain the SVM algorithm and the preprocessing stages.

A. Grain-status-monitoring System

The grain-status-monitoring system consists mainly of two distinct parts: hardware and software. The hardware part is the wireless network of temperature and moisture sensors, which is used to collect the data set that can be used later for both training and testing the machine learning classifier, namely SVMs. The detailed illustration of the WSN used as the hardware part of this system is found in [5]. The software part is the machine learning algorithms (SVMs) used to classify the condition of grain storage. These algorithms are executed on MATLAB, which runs on the server of the existing WSN. Therefore, the server performs the classification process according to the data coming from the sensors and shows the status of the grain storage to the user.

B. Data set Structure

SVMs, like other machine learning algorithms, need training data to build a robust model that can be used later to classify the testing data or any new example. The data set, that is used in this paper, contains 300,000 readings of the temperature and moisture of the stored grain. This data set has 100,000 items of the "safe" class, 100,000 items of the "risky" class and 100,000 items of the "dangerous" class.

After the original raw data have been shuffled and preprocessed as illustrated in the next section, 80% of the data set is used as the training set and 20% of the data set is used as the testing set. The purpose of the training set is to try to find the optimal decision hyperplane which separates the positive class from the negative class (in our case, the optimal hyperplane is a two-dimensional curve since the feature space is a two-dimensional plane with two features). After the model (decision hyperplane) has been constructed, this model can be used to classify the testing data.

C. Machine Learning Classifiers

We have classified the stored grain status into three separate classes via a machine learning algorithm: SVMs. In fact, SVM and ANN algorithms are very powerful and common among machine learning experts. ANNs are the core of the deep learning, a new and very powerful trend in machine learning. Like ANNs, the SVM is one of the supervised learning techniques that can build a model by means of labeled training data. Furthermore, SVMs produce better results than ANNs in some cases and specific applications [9]. Unlike ANNs, SVMs still perform well even with sparse data and the computation that is required is less as compared with their counterpart: ANNs. Thus, we try to classify the stored grain using the SVM algorithms and compare the results with a previous classifier trained using ANNs, as mentioned in [4][5]. Figure 1 shows the different stages of constructing the final SVM classifier, which can be used later to classify the testing data. The raw data are a 300,000-by-3 matrix with 300,000 rows (items) and 3 columns (the first two columns are temperature and moisture measurements, respectively, and the third column is the class labels such that the safe class is labeled as 1, the risky class is labeled as 2 and the dangerous class is labeled as 3).

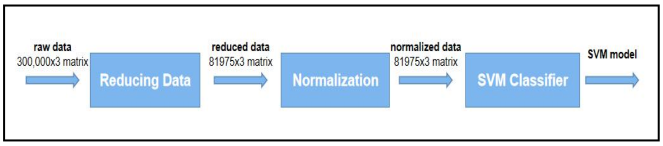

Fig. 1 Stages of building the SVM model

C.1 Preprocessing Data

Prior to feeding the data into the SVM algorithm, we need to do some preprocessing to the data to improve the training process and reduce the required time for training. The preprocessing consists of two significant stages : reducing the data and normalization.

The data are a collection of temperature and moisture measurements. Therefore, one particular sample can have the same temperature and moisture readings as some other samples. That is, the data are most likely to contain redundancy and we have to get rid of redundant data and keep only the unique data. Accordingly, we get rid of any redundant data and reduce the data in the first stage of the preprocessing. We compare the rows of the dataset matrix and eliminate any repetition of any row such that every row appears only once in the prepreocessed dataset. In fact, the data dropped significantly from 300,000 items to below 82,000 items, and that is a decrease by nearly 73%.

As any machine learning algorithm, normalization is an important preprocessing step that has to be taken before training the SVM algorithm. Normalization maps all the features (probably with different ranges) into the same range [0,1], which improves the training process and reduces the training time. The normalization step is done according to Eq.1:

$$\frac{x_i - \min(\mathbf{x})}{\max(\mathbf{x}) - \min(\mathbf{x})} \tag{1}$$

Where x_i is one element of the vector-valued **x**. **x** is the vector of all the values of the first feature (temperature). **y** is the vector of all the values of the second feature (moisture). We need also to normalize the second feature **y** according to Eq.1 by replacing x_i and **x** by y_i and **y** respectively.

C.2 Support Vector Machine

We applied the SVM classifier to classify the status of the stored grain. Originally, the SVM algorithm is used for binary classification (two-class classification) such that the algorithm tries to find the optimal decision boundary (hyperplane) with the maximum margin between the positive class and the negative class. In order to extend the capability of the SVM algorithm, two main algorithms can be used for multiclass classification. The first one is called one-against-all algorithm whereas the other one is the error-correcting output codes (ECOC) algorithm.

C.2.1 One-against-all Algorithm

In this algorithm, the number of SVM models, to be constructed, is the same as the number of classes. In our application, we have three classes (safe, risky and dangerous); therefore, we need to construct three SVM models. The first SVM model will classify the safe class against the other two; in other words, the safe class is the positive class of the model while the risky and dangerous classes together are the negative class. The second SVM model will classify the risky class against the the other two. Finally, the third model will classify the the dangerous class against the other two. Table 1 shows the three binary SVM models and the positive and negative classes of each model. The +1 and -1 in the table denote the positive class and the negative class respectively. Then any test example will be entered as input into the three models simultaneously and classified into the positive class whose model has the highest score.

Table 1. One-against-all Algorithm

Class	Model 1	Model 2	Model 3
Safe Class	+1	-1	-1
Risky Class	-1	+1	-1
Dangerous class	-1	-1	+1

C.2.2 ECOC Algorithm

This algorithm is more accurate than the previous method since it is more robust against error. Moreover, the ECOC algorithm is faster than the one-against-all algorithm, particularly for classification problems with a few classes. In general, it divides into two steps: coding and decoding. The coding step assigns a codeword to each class such that the codeword is a string of bits or symbols where each bit determines the membership of the class for the corresponding binary classifier. Therefore, the coding step build a K-by-N coding matrix with the K classes as rows and the N binary classifiers as columns. The rows of the code matrix should be different from one another and so should the columns. In addition, the negative of any particular column have to be different form the other columns as well. The coding step has two different methods: the binary and ternary coding. In the binary coding, the codeword of each class contains only 1 and -1 whereas the codeword contains 1, 0 and -1 in the ternary coding. The standard binary coding schemes are the one-against-all method (thus the one-against-all algorithm can be considered as a special case of the ECOC algorithm) and the dense random method. The standard ternary coding schemes are the one-against-one method (it is the one used in this paper) and the sparse random method. In the decoding step, the closest class codeword is determined, given a specific test sample. Some of the famous decoding techniques are the Hamming distance and Euclidean distance [10].

In the one-against-one approach, the number of binary classifiers or models is k(k-1)/2 where k is the number of classes. Thus, we need 3 binary SVM models in our problem. Table 2 shows these three SVM models with their positive and negative classes such that 0 denotes that the data corresponding to that class are not included while training that SVM model. For example, Model 1 is to separate the safe class from the risky class. For testing, the new data point will be classified into the class that has the majority of the votes of the three model. For instance, if one data point is classified to belong to the negative class, the negative class and the positive class by Model 1, Model 2 and Model 3 respectively, then it will be classified to belong to the risky class.

Table 2. ECOC (One-against-one) Algorithm

Class	Model 1	Model 2	Model 3
Safe Class	+1	+1	0
Risky Class	-1	0	+1
Dangerous class	0	-1	-1

III. Results and Discussion

In this section, we will present the results of using SVMs as the classification algorithm and draw a comparison between the SVM classifier and ANN classifier.

We used two different algorithms of SVM multiclass classification algorithms: one-against-all and ECOC algorithms. In both algorithms, the number of misclassified examples is zero and, consequently, the classification accuracy, in terms of how many examples are correctly classified, is 100% for both algorithms in this grain-condition-monitoring application. The confusion matrices are used to evaluate the performance of a machine learning classifier. The confusion matrices show how much of the data set is correctly classified and how much is misclassified. Thus, the confusion matrices, for classification problems, are more important than the other performance measures, such as the mean sum of squares or cross-entropy, and indicate the real performance of any classifier. The confusion matrix in Fig. 2 shows the proportions of the testing data that are correctly classified (in blue and on diagonal) and that are misclassified (in yellow or red and off diagonal) as well, together with the predicted class and true class. As shown in the confusion figure, all the testing data are correctly classified, with zero misclassified data.

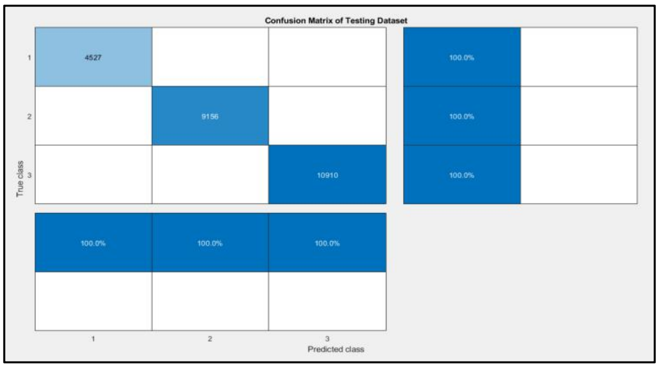

Fig. 2 Confusion matrices of testing data

Despite this, the time taken by each algorithm is a significant factor and we compare them according to this major factor. Table 3 shows how long each algorithm takes to train and test its model. As far as the platform used for training and testing is concerned, the SVM models were trained and tested using MATLAB R2018b on an ASUS laptop with Intel Core i5-5200U processor and 4GB RAM. According to the table, the one-against-all algorithm takes about 109 seconds while the ECOC algorithm takes roughly 6 seconds. Consequently, the latter algorithm is approximately 17 times faster than the former one.

Table 3. Speed of SVM Algorithms

SVM Algorithm	**Elapsed time (seconds) (training and testing)**
One-against-all algorithm	109.08
ECOC algorithm	6.23

Furthermore, we can compare the results we have got with the results shown in a previous work [5]. In the previous work, the classification accuracy was below 99% using ANNs with three different training algorithms as the machine learning classifier. In this work, the classification accuracy is 100% using both SVM classification algorithms. Table 4 summarizes the comparison.

Table 4. Comparison Between SVMs and ANNs

Classification Algorithm		**Classification Accuracy**
ANNs	Gradient descent with momentum backpropagation	98.9924%
	Scaled conjugate gradient backpropagation	98.9912%
	Levenberg-Marquardt backpropagation	98.9855%
SVMs	One-against-all algorithm	100%
	ECOC algorithm	100%

IV. CONCLUSION

In this work, we tried to use the SVMs as the machine learning classifier instead of the ANNs in the application of a wireless network of temperature and moisture sensors for classifying silobag-stored grain into three classes: ‘safe’, ‘risky’ and ‘dangerous’. We used two different algorithms (namely one-against-all and ECOC algorithms) to use the binary SVM classifiers as a multiclass classifier. Although the classification accuracy of the two algorithms are 100% such that all the testing data are correctly classified, the ECOC algorithm takes shorter time. When compared with ANNS, SVM algorithm as the classification algorithm outperforms ANNs. Therefore, a WSN with a SVM classifier can be used to efficiently monitor the condition of grain storage.

ACKNOWLEDGMENT

The authors would like to thank Erciyes University Scientific Research Projects Support Office and the Department of Mechatronics Engineering for their support of the work reported in this paper.

REFERENCES

[1] J. A. Stankovic, "Wireless Sensor Networks," *in Computer*, vol. 41, no. 10, pp. 92-95, Oct. 2008.

[2] N. Wang, N. Zhang and M. Wang, “Wireless sensors in agriculture and food industry—Recent development and future perspective,” *Computers and Electronics in Agriculture,* vol. 50, Issue 1, 2006, pp.1-14.

[3] P. Armstrong, 'Wireless data transmission of networked sensors in grain storage,” 2003 ASAE Annual Meeting, Las Vegas, USA, 27-30 July 2003.

[4] C.K., Baykara, “Wireless Network Based Grain Control System Design for Grain Bagging System (Silobag)”, Erciyes University, Master’s Thesis (in Turkish), pp. 35-69, 2018.

[5] M. Kanaan and C. K. Bavkara, "Proactive Monitoring and Classification of Stored Grain Condition via Wireless Sensor Networks and Machine Learning Techniques," 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, 2018, pp. 1-4.

[6] R. Bartosik, L. Cardoso, and J. Rodríguez, "Early detection of spoiled grain stored in hermetic plastic bags (silo-bags) using CO2 monitoring," In Proceeding of 8th International Conference on Controlled Atmosphere and Fumigation in Stored Products, pp. 550-554. 2008.

[7] A. Douik and M. Abdellaoui, “Cereal Grain Classification by Optimal Features and Intelligent Classifiers,” *International Journal of Computers, Communications & Control (IJCCC)*, vol. 5, 2010, pp. 506-516.

[8] K. S. Banga, N. Kotwaliwale, D. Mohapatra and S. K. Giri, “Techniques for insect detection in stored food grains: An overview,” *Food Control,* vol. 94, 2018, pp. 167-176.

[9] L. Wang (Ed.), *Support vector machines: theory and applications,* vol. 177. Springer Science & Business Media, 2005, pp. 1-47.

[10] S. Escalera, O. Pujol, P. Radeva, “On the decoding process in ternary error-correcting output codes,” *IEEE transactions on pattern analysis and machine intelligence,* vol. 32,2010, pp.120-134.