Support Vector Machine and Fuzzy Logic

József Menyhárt¹, Róbert Szabolcsi²

¹ Corresponding author, Doctoral School of Safety and Security Sciences, Donát Bánki Faculty of Mechanical and Safety Engineering, Óbuda University, Népszínház u. 8, H-1081 Budapest, Hungary, E-mail: jozsef.menyhart@eng.unideb.hu

² Department of Mechatronics, Donát Bánki Faculty of Mechanical and Safety Engineering, Óbuda University, Népszínház u. 8, H-1081 Budapest, Hungary E-mail: szabolcsi.robert@bgk.uni-obuda.hu

Abstract: The article gives a short description of the history of the Support Vector Machine (SVM) method and fuzzy logic and their main parameters. It describes how SVM can be used for classification and regression calculation. In addition, it includes the description of the fuzzy logic and the three main defuzzification methods. Examples will be given to show simultaneous use of the two main Support Vector Machine-methods. It ends with the authors’ description of the new possibilities for the use of the SVM and fuzzy logic in evaluation of the parameter deviation in the batteries of electric vehicles as well as those of robots.

Keywords: SVM; classification; regression; fuzzy

1 Introduction

Nowadays more and more intelligent machines and tools are used in everyday life, which makes life easier for their users. The engineers working on these kinds of machines make efforts to copy human intelligence. Their goal is to create machines with higher efficiency. Machine intelligence is a combination of different types of methodologies, for instance, fuzzy logic, neural networks, neural computing, anytime techniques, probability theory, etc. [1]

The methods mentioned above work well when complementing each other. However, they cannot provide a useful solution if one of them is used by itself. These methods share a feature: they take advantage of the inaccuracy and uncertainty associated with the tolerances and the boundaries to make a better representation model of the real world. [1]

In a previous work [32] lithium polymer battery parameters were analyzed by Fuzzy logic. The result of this paper presents a brand new operational strategy for
the batteries, which are very sensitive to their operational parameters. The authors would like to find an ideal method for Fuzzy logic which enables us to find ideal limits in the batteries operation parameters. Artificial intelligence equipment systems play a key role in the development of modern industrial applications and vehicles. Thanks to these techniques both the efficiency and the availability can be increased. This paper would like to describe the Support Vector Machine (SVM) and the Fuzzy logic methodologies and some applications where they are used simultaneously.

The study gives a short summary and the syntax of the two methods. The last chapter provides the reader with examples in different practical areas.

The authors describe the usage of SVM during energy supply. The empirical results show that batteries can be used with operation parameters different from those prescribed by the producers. But these operation conditions raise a safety technique question as it is essential not to exceed the limit where the batteries still work with reasonable safety. The authors seek the answer to this question.

This question is of crucial importance from a practical perspective. The performance of the machines is strongly influenced by daily unexpected events.

2 Related Works

The importance of battery parameter check is mentioned in the works of G.F. Guo and his co-authors pointing out the problem of SOC for the battery management system. During their work they checked voltage, amperage and temperature parameters and used the PSO-SVM method for optimization. [23]

In [24], which is a study that considers the batteries as complex electro-chemical cells, we can read about their importance. The Relevance Vector Machine and Particle Filters (PF) were used to check the uncertainty properties of the batteries during use.

Another field where the SVM method was used is the energy management of solar systems. Yen Yee Chia and his co-authors checked the energy flow and its optimization in their studies and conducted research work about supercapacitator-battery hybrid energy storage systems which are a part of solar systems. [25]

The studies mentioned above all deal with the optimization of the original energy storage systems. The increase of the availability of the present systems is also of primary importance. The research results prove that the batteries can work with parameter deviation and they can be used for a longer period of time. The SVM method enables us to determine new operation limit(s) for the batteries if we choose the right hyperplane. Since, this method has great importance in practice, maybe it can be seen from a new perspective.
3 Support Vector Machine

3.1 Theoretical Background of SVM Method

It is undoubtedly true that there is an ever-increasing demand for both mathematical and computer modelling, which runs parallel to the progress in informatics. However, these techniques have to check repetitive and similar schemes, and it is necessary to classify these schemes during their operation. The fastest growing area of artificial intelligence is statistical learning algorithms. In other words, we can call it the theory of machine learning, which enables us to draw conclusions and make generalizations from the learning sets and the sample sets (which means real time observations). The Support Vector (SV) method, which is the most important method, leans on a statistical base. The Support Vector group is made up from different variants of Support Vector Machine (classification) and Support Vector Regression (regression). Their names indicate that not all the elements of the available training set are used for the statistical model. The SVM is a special neural network, or, is a statistical learning theory in some cases. During the last 30 to 40 years we have witnessed formation and development of SVM. V.N. Vapnik has done significant work in this topic. [2] [3] [4] [5] [6] [7] [14] [36]

The main task of the SVM method is to find the optimal solution but at the current level of development it can still be used for approximating functions and classifications as well. The SVM is applied in different areas like character recognition, image processing, bioinformatics, data mining etc. The SVM’s decision-making process is quick and for this reason it can be used in real time applications. The SVM is applicable for real time applications thanks to its ability to make and implement decisions very quickly. Chapter 3 and the subsections are based on [6]. [2] [3] [4] [5] [6] [7] [14] [36]

3.2 SVM used for classification

Suppose the following [6] [14] [36]:

\[(x_1, y_1), \ldots(x_n, y_n) \in \mathbb{R}^d; \quad y_i \in \{-1, 1\}\]

The elements of the samples are derived from two classes. The elements of \(x_i \in \mathbb{R}^d\) from class \(A_1\) then use \(y_i = 1\), if from class \(A_2\) then use \(y_i = -1\). The training set is linearly separable if we know a hyperplane one side of which contains only the elements of class \(A_1\), while the other side contains the elements of class \(A_2\). [6] [14] [36]

So

\[\langle x, \varphi \rangle = c\]
is linearly separable with a hyperplane, if
\[ \langle x_i, \varphi \rangle > c, \text{ if } y_i = 1, \]  
\[ \langle x_i, \varphi \rangle < c, \text{ if } y_i = -1, \]
where \( \varphi \in \mathbb{R}^d \) is an inner product between \( a \) and \( b \) where \( c \in \mathbb{R} \). The SVM always looks for the most optimal hyperplane. Figure 1 shows two non-optimal separator lines and another one, which was made by the SVM method. [6] [14] [36]

The thin lines in Figure 1 in Case b are edges or margins. The thick line in the middle is called the band limit. The points which fit on the margins are the support vectors. The separating hyperplane is as far away as possible from the sample points. The support vectors are the closest to the optimum hyperplane. [6] [7] [14] [36]

### 3.3 SVM Used for Regression

The Support vector based statistical classification and regression methods are parts of the statistical learning algorithms group. The Support Vector Machine (SVM) algorithm was the first method for classification, which appeared in the first form of Support Vector Regression in 1997. The SVM can rank an optional number of vectors into classes in the case of binary classification where the model is based on learning vectors. Figure 2 shows a Support Vector Machine architecture. [5] [6] [14] [15] [36]
The square of the distance is widely used in the classical areas of mathematics. We can find it in the case of L2 and I2 spaces and in the smallest square methodology of numerical mathematics. The variance can be found in probability theory, the classical regression analysis also use this. [5] [6] [14] [15] [36]

Form:

\[ (y - f(x, \alpha))^2, \]  

(5)

Where \( x \) is the input and \( y \) is the output. We use loss functions (like those used in robust mathematics) in the case of SVM, such as the \( \varepsilon \)-insensitive loss function. [5] [6] [14] [36]

The \( \varepsilon \)-insensitive loss function means that the system is not sensitive to deviations which are smaller than \( \varepsilon \). Higher deviations than \( \varepsilon \) are not used quadratically, we use them linearly. [5] [6] [14] [36]

That is

\[ L_\varepsilon(y - f(x, \alpha)) = |y - f(x, \alpha)|_\varepsilon, \]  

(6)

a linear \( \varepsilon \)-insensitive loss function, where

\[ |y - f(x, \alpha)|_\varepsilon = \begin{cases} 0, & \text{if } |f(x, \alpha)| \leq \varepsilon, \\ |f(x, \alpha) - y| - \varepsilon, & \text{if } |f(x, \alpha)| > \varepsilon, \end{cases} \]  

(7)

and it can be seen in Figure 3. [6] [14] [36]
3.4 New Approach of SVM Used for Operating Conditions

In the case of electric and autonomous vehicles the range gets a higher priority. The empirical data show that the batteries can work with parameter deviations.

During the research work the authors describe two sets (or classes):

- prescribed data (during normal battery use)
- empirical data (during aperiodic (random) battery use, parameter deviation)

Table 1 and Table 2 contain the data about the two sets.

<table>
<thead>
<tr>
<th>Prescribed data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Voltage</td>
</tr>
<tr>
<td>1. 4,201</td>
</tr>
<tr>
<td>2. 4,204</td>
</tr>
<tr>
<td>3. 4,209</td>
</tr>
<tr>
<td>4. 4,200</td>
</tr>
<tr>
<td>5. 4,205</td>
</tr>
<tr>
<td>6. 4,208</td>
</tr>
<tr>
<td>7. 4,206</td>
</tr>
<tr>
<td>8. 4,200</td>
</tr>
<tr>
<td>9. 4,201</td>
</tr>
<tr>
<td>10. 4,200</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Empirical data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Voltage</td>
</tr>
<tr>
<td>1. 4,392</td>
</tr>
<tr>
<td>2. 4,400</td>
</tr>
<tr>
<td>3. 4,390</td>
</tr>
<tr>
<td>4. 4,410</td>
</tr>
<tr>
<td>5. 4,438</td>
</tr>
<tr>
<td>6. 4,410</td>
</tr>
<tr>
<td>7. 4,378</td>
</tr>
<tr>
<td>8. 4,420</td>
</tr>
<tr>
<td>9. 4,400</td>
</tr>
<tr>
<td>10. 4,397</td>
</tr>
</tbody>
</table>

The prescribed data (normal use) show a linear result, the inaccuracy during the charge is negligible. The empirical results (random use) show a scatter. An optimal hyperplane can be calculated with the SVM method. This new hyperplane can help the users to use the batteries with a satisfactory safety level. The difference between the prescribed and empirical results is shown in Figure 4.
Figure 4
Difference between the measured results

The optimal hyperplane enables the user to choose the right voltage results for a better operational condition. If they choose the new parameters, the range of the vehicles can grow.

The authors address the following questions relating to the topic:

- Where is the ideal place of the optimal hyperplane, between the prescribed and empirical data where it can operate within the proper safety range?
  or

- Where in the empirical data is it demonstrated that are the voltage results higher but the safety level of the vehicle is worse, than in the first case?

The questions are of great importance, in both cases we can find pro and contra arguments. The fuzzy logic may help us to find the answers to these questions.

4 The Theoretical Background of Fuzzy Logic

The first references to fuzzy logic can be found in the works of Lofti Zadeh who achieved outstanding results in this topic in 1965. He examined the blurred boundaries of the truth values of different colloquial concepts from a mathematical approach (calculation with words). He assigned a value between [0;1] (closed interval) to every logical statement during the modelling process. [8] [9] Figure 5 presents the difference between non-fuzzy and fuzzy logic.
We talk about nice weather, high speed, and good restaurants with a blurred meaning in our everyday life. If we were surrounded by exact definitions, verbal communication would become impossible so the lack of precision in the vernacular is necessary in our life. If mathematics is used to describe the complexity of the real world, we have to do it in a numerical form. The Fuzzy is the only way to describe human knowledge, expertise and experience in a mathematical form. [8] [9] [10] [19] [21]

If we happen to read the measuring results incorrectly from the gauges in practice, this mistake can lead to inaccuracy during our work. This is the reason why we can use Fuzzy logic when such problems occur. This inaccuracy can be represented by so-called membership functions. The process in fuzzy systems is illustrated in Figure 5. [8] [9] [10] [21] [22]

The first step is fuzzification, which is in fact (about) giving production values to the system. We have to define the categories and the membership functions for the model. The main factors need to be examined. It is important to select an appropriate number of categories. If the number of the categories is increased, we can get an exact picture of the system but it can lead to unnecessary calculations. Having too much information can be as dangerous as having too little. [8] [10] [21] [22] [33] [34]
Every category has to use a membership function, we know several ways to define the categories. The $\mu(x;a)$ membership function gives the rate of feature $x$ in set $A$. Another important step is scale definition. It is practical to use it in the following scales: 0-10, 1-10, 0-100, 1-100. The target of the scale is to find the simplest possible way to compare and evaluate the examined values. [8] [9] [10]

It is essential that the system uses the right rule in the explanation stage. The rules come from the defined categories. If the rules work well, the system can interpret the values. This is how the rules of a fuzzy model are created. [8] [9] [10]

In the final step, we link the values - which we got from the explanation - which differ form 0 in the light of the features of the controlled process, the fuzzy process. The result of this is a fuzzy set. As this is a preconclusion, this result is not relevant in practical usage. It can be interpreted in the next step called defuzzification. [8] [9] [10]

The defuzzification is the last step of the process. In this step, we have to choose the exact value based on the fuzzy conclusion and this value - depending on the application and the model - is the most representative of the set. The meaning of the defuzzification can be different depending on the usage. We can use different types of defuzzification methods to get the right results. [10] [33] [34]

The most important ones among them are:

- Center of Gravity (COG);
- Center of Area (COA);
- Weighted Mean of Maximum

The Center of Gravity method is one of the most commonly used defuzzification methods. The main advantages of COG are that it is easy to use with triangular and trapezoidal rules and it can show continuous behavior during direct navigating. The COG in general form [10] [33] [34]:

$$Y_{COG} = \frac{\sum_{i=1}^{n} \int_{-\infty}^{+\infty} \mu_i(z) zdz}{\sum_{i=1}^{n} \int_{-\infty}^{+\infty} \mu_i(z) dz}$$  \hspace{1cm} (8)

The Center of Area method is very similar to the Center of Gravity. The difference between the two methods is in the calculation. The COG uses areas which are covered by part conclusion while the COA uses only the cumulative consequences. The form of COA [10] [33] [34]:

$$Y_{COA} = \frac{\sum_{i=1}^{n} \int_{-\infty}^{+\infty} \mu_i(z) dz}{\sum_{i=1}^{n} \int_{-\infty}^{+\infty} \mu_i(z) dz}$$
\[ Y_{COA} = \int_{-\infty}^{\infty} \mu_2(z)dz \]  

The Weighted Mean of Maximum is the most frequently used method for defuzzification. The method’s result shows the biggest membership value. If it reaches the highest result in one interval, then we have to calculate the mean value of this stage. Its form is [10] [33] [34]:

\[ Y_{WMM} = \frac{\sum_{i=1}^{n} \mu_i z_i}{\sum_{i=1}^{n} \mu_i} \]  

Menyhárt and Pokorádi have used Fuzzy logic to examine battery parameter deviations and operation condition monitoring in their previous research. [8]

5 Fuzzy Logic and SVM Applications

The SVM is a machine learning technique based on statistical methods. The SVM can be used for classification or regression analysis and its aim is to find the most optimal hyperplane. Thanks to this feature, it is different from other neural networks. [6]

The variables of Fuzzy systems are based on fuzzy sets. These features make it possible to give a numerical description of empirical and linguistic skills. [10]

They cannot support independent machine learning and adaptation. At first glance (it seems as if) the two machine learning methods cannot be used and combined at the same time. This paper presents how the two methods can work together in different applications in everyday life.

[7] shows a new type of use of Support Vector Machine, which forms a kernel function based on a fuzzy rule base. The proposed network combines the characteristics of SVM and Fuzzy systems. It is high general performance, even if the dimension of the input space is very high, structured and it gives a numerical representation of knowledge and ability using linguistic fuzzy rules, in order to bridge the semantic gap between the low-level descriptors and the high-level semantics of an image. The authors checked different types of images with this special Fuzzy-SVM network putting the main focus on urban and beach pictures. [7]

In order to avoid the scale effect, every feature should contain more or less the same numerical values. In this case the MPEG-7 descriptors are already scaled into integer values of equivalent magnitude. The authors used the Personal Content Services database for their work. It contains 767 high quality color images which are divided into two classes (beach and urban) (Figure 7). 40 images were used from the beach category and 20 from the urban during the system training. The remaining 707 were used for the evaluation. [7]
The Fuzzy SVM network made it possible to extract linguistic fuzzy rules during the classification. An extracted fuzzy rule for the case of the Edge Histogram descriptor [7]:

“IF the number of 0 ± edges on the upper part of the image is low AND the number of 45 ± edges on the upper part of the image is medium AND . . . AND the number of non-directional edges on the lower part of the image is high, THEN the image belongs to class Beach”. [7]

The proposed network was successfully applied to the problem of image classification. The fusion of the two methods was very useful as it can provide a linguistic description of the underlying classification mechanism. The authors refer to their future work: they would like use more MPEG-7 descriptors and more classes. [7]

At the beginning of their work, Yixin Chen and his co-author James Z. Wang give a description of Fuzzy logic: this research topic has been very popular for a long time. The Support Vector Machine method is mentioned, it has a good generalization ability but what is even more important: SVM works well in multidimensional spaces. They describe the connection between fuzzy classification and kernel machines and they look for a link between fuzzy rules and kernel functions and present an algorithm for classification. [12]

Takuya Inoue and his colleague Shigeo Abe at Kobe University in Japan also did a similar research, which was about classification problems. [11]

Boumediene Allaoua and Abdellad Laoufi - unlike the previous examples – used the methods mentioned above in the vehicle industry. They present a new sliding mode fuzzy control scheme for torque control of induction motors. The control principle which they developed is based on sliding mode fuzzy control combined with SVM. The sliding mode fuzzy control contributes to the robustness of induction motor wheel drives and the space vector modulation improves many other properties of the electric system. They used the Lyapunov direct method with fuzzy logic. Figure 8 shows their system. [13]
Compared with the classical PI-control method, the conventional SMC method and SMFC with SVM technique, this new scheme has low torque ripple, low current distortion, and high performance dynamic characteristics of the developed system. This new control scheme can achieve high accuracy in torque tracking to various reference torque signals and shows robustness to external load disturbances. In summary, their system is simple, accurate and it has high reliability. [13]

**Conclusion**

These examples perfectly represent the combination of Fuzzy logic and SVM methods in different areas. There is an infinite number of problems where we can use these methods but the most important thing is in [13]. In vehicle industry and other industrial applications we would like to use machines with high reliability and outstanding operating performance. Current knowledge shows that Fuzzy logic and SVM can be used to create more than sufficient regulatory systems.

Intelligent power supply systems and power stations are very important during the development of alternative fuel vehicles and autonomous vehicles. [8] presents a Fuzzy rule base for the monitoring of an electric car’s battery condition. These points are the base for future work. This work would like to present an energy regulation system with Fuzzy logic and SVM.

This study contains a short description of the history of SVM. The reader can get information about the SVM method for regression analysis and classification. The main focus of this method is to find the most optimal hyperplane.
The article contains another machine learning technique, which is the Fuzzy logic. It provides the reader with details of the story of Fuzzy logic and the three most widely used defuzzification methods.

The study presents the advantages of the simultaneous combination of Fuzzy and SVM. It is possible to find solutions to different types of classification problems with these methods. But we can find the most substantial opportunity in the vehicle industry. The articles contains two examples, one of classification and one of vehicle industry.

The focus of our future work will be the determination of the optimal hyperplane. We also have to distinguish a milk run robot from a military robot. In the case of military different missions and events. The next research work focuses on the simultaneous usage of the Support Vector Machine and Fuzzy logic.

The system has to check the telemetry data in real time and it is supposed to be able to define which hyperplane is the best in practice, in different real situations.

References


