Original Article

Nanomaterials Characterization and Risk Assessment Using Fuzzy Support Vector Machines

Odedele, T.O*, Obasi Seb. C and Ibrahim H.D

Raw Materials Research and Development Council (RMRDC) (Federal Ministry of Science and Technology) Abuja, Nigeria

Address for Correspondence

Raw Materials Research and Development Council (RMRDC) (Federal Ministry of Science and Technology) Abuja, Nigeria.
E-mail: odedelermrdc@yahoo.com

ABSTRACT

Nanotechnology is the process that develops novel materials at size of 100 nm or less and has become one of the most promising areas of human endeavor. In this paper, Fuzzy Support Vector Machines (Fuzzy SVM) model is developed to predict/assess the toxicity of nanomaterials. Because of their novel and unique properties, nano-particles are commonly applied in medicine, Engineering, environmental and agricultural industries However, several toxicological research results have indicated evident toxicity of some nanoparticles to living organisms (toxicity), and their potentially negative impact on environmental ecosystems (ecotoxicity) for which relatively simple testing procedures are available for their characterization. However, because of the large number of nanoparticles and the variety of their characteristics particularly sizes and coatings it is only rational to develop an approach that avoids testing every single nanoparticle produced. Therefore, the main motivation of this study is to assist the users of nanomaterials in classifying nanomaterials and assessing the risk of toxicity. The hybrid Fuzzy Support Vector Machine (Fuzzy SVM) model will be developed to predict the toxicity of nanomaterials based on the trained datasets. The proposed method uses the dataset information to expose the nanomaterials exhibiting toxicity.

Keywords: Characterization, Support vector machines, Fuzzy support vector machines, Nanomaterials, Nanotechnology, Membership function, Linguistic variables, Fuzzy rules, Risk assessment, Toxicity.

INTRODUCTION

The term "nanotechnology" covers processes associated with the creation and utilization of structures in the 1 nanometer (nm) to 100 nm range.

The unique properties of these [nanotechnology] materials are a double edged sword because they can be tailored for beneficial properties and at the same time may also have unknown new toxicological and environmental impacts.

On the Environmental, Health, Safety (EHS) issues, nanomaterials could play some harmful roles in their distribution through environment, ecosystem and human body. Their novel biological activities/or unique properties have made it easy to gain access into human body system through the skin, lungs, gastrointestinal tract. Several toxicological research works have reported that nanomaterials can be cytotoxic, neurotoxic, genotoxic and ecotoxic[1]. These apprehensions of the potential EHS effects of nanomaterials constitute serious barrier to nanotechnology transfer towards business perspectives. There is the need to develop screening protocol to assess, address, and manage the potential risks. To accomplish this, it is imperative to develop sensitive analytical methodologies, tools and an acceptable protocol for screening, characterization and monitoring the application of nanomaterials. Therefore, considering the EHS issues there is serious need to develop and design predictive models for nanomaterials toxicity using computational intelligent systems.

OBJECTIVE

Recent advances in classifiers have provided attractive alternatives for constructing interpretation models of complex nanomaterials. Here, fuzzy Support Vector Machines (SVM), a class of a hybrid classifier has been explored to determine its capabilities for determining the relationship between physicochemical properties and human health.

The objective is to develop computational/predictive model used to establish knowledge base, risk modeling and nanoinformatics capabilities to reliably assist decision making

Therefore, in order to accomplish this, the following are necessary:

- Development of computational intelligent predictive models for nanomaterials toxicity.
- Development of standardized methods, risk evaluation, risk assessment and management protocol.
- Information sharing, common database for research that uses standard protocols to generate knowledge

This paper will therefore focus on the capability of fuzzy SVM to model physicochemical properties and toxic effect of nanomaterials in view of the imprecision and uncertainty surrounding the prediction of nanomaterials toxicity.

Section I gives a brief introduction. Section II highlights the barriers to Nano-technology transfer towards business perspectives. Section III highlights the physicochemical characteristics for nanomaterial characterization. Section IV describes the proposed SVM technique. Section V describes the process of fuzzy logic modeling, control and decision making. Section VI discusses methodology of nanomaterials characterization risk assessment system. In Section VII, detailed numerical data for training and testing the model. Section VIII discusses the results of the study. Section IX highlights the conclusion of the study.

Physicochemical characteristics dependent toxicity

Considering the harmful effects of fibrous particles (such as asbestos), the most important factors that determines the adverse health effects of nanoparticles are dose, dimension, and durability (the three D's)[3]. However, recent studies show different correlations between various physicochemical properties of nanoparticles

and the associated health effects, raising some uncertainties as to which are the most important parameters in deciding their toxicity, or all together. In the following section we will discuss what are considered to be the most important nanoparticle characteristics associated with their toxicity.

Dose-dependent toxicity

Dose is defined as the *amount* or *quantity* of substance that will reach a biological system. The dose is the product of exposure or the concentration of substance in the relevant medium (air, food, water) and the duration of contact. Generally, the negative health effects of nanoparticles are not determined by nanoparticle mass dose[4]. For instance, TiO_2 nanoparticles with different sizes, it is noted that the low dose (20 mg/m^3) exposure to 20 nm diameter particles resulted in a greater lung tumor incidence than the high dose (250 mg/m^3) exposure of 200nm diameter particles[5]. The measure that correlates with the effects is the surface area and not the mass dose[4,6].

Size-dependent toxicity

From various toxicological research works, it has been shown that small nanoparticles (<100 nm) cause adverse respiratory health problems, than larger particles made from the same material[4,7,8]. For instance, Rat inhalation of titanium oxide particles with two sizes, 20nm and 250nm diameter, having the same crystalline structure show that smaller particles exhibited a more pronounced inflammatory reaction in the lungs compared to larger size particles.

Surface area-dependent toxicity

A greater toxicity was observed from nanoparticles than from their larger counterparts or the same mass of particles with the same chemical composition and crystalline structure. It can be concluded that the inflammatory effect may be dependent on the surface area of nanoparticles. Actually, smaller nanoparticles have higher surface area and particle number per unit mass compared to larger particles. Larger surface area leads to increased reactivity[10] and is an increased source of reactive oxygen species, as shown *in vitro* experiments[3].

Concentration-dependent toxicity

It has been shown that a high concentration of nanoparticles would promote particle aggregation[9,10], and therefore reduce toxic effects compared to lower concentrations[9]. Most aggregates are observed to be larger than 100 nm, a size that seems to be a threshold for many of the adverse health effects of small particles.

Aspect ratio dependent toxicity

It was found that the higher the aspect ratio, the more toxic the particle is[15]. More exactly, lung cancer was associated with the presence of asbestos fibers longer than 10 microns in the lungs, mesothelioma with fibers longer than 5 microns, asbestos is with fibers longer than 2 microns[15].

Overview of support vector machines

Vapnik[16] proposed the support vector machines (SVMs) which was based on statistical learning theory. The governing principles of support vector machines is to map the original data x into a high dimension feature space through a non-linear mapping function and construct hyper plane in new space. The problem of classification can be represented as follows. Given a set of input-output pairs Z = {(x1, y1), (x2, y2),..., (xℓ, yℓ)}, construct a classifier function f that maps the input vectors x € X onto labels y € Y . In binary classification the set of labels is simply Y = {−1, 1}. The goal is to find a classifier f ∈ F which will correctly classify new samples.

There are two main cases to consider when we use a separating hyper-plane:
1. A linearly separable case
2. The data might not be linearly separable.

SVMs tackle the first problem by finding the hyper-plane that realizes the maximum margin of separation between the classes[12]. A representation of the hyper-plane solution used to classify a new sample xi is:

$$Y=f(x)= w_i\phi(x)+b \quad (1)$$

Where wi, ϕ(x) is the dot-product of the weight vector w and the input sample, and b is a bias value. The value of each element of w can be viewed as a measure of the relative importance of each of the sample attributes for the classification of a sample. Various research studies have shown that the optimal hyperplane can be uniquely constructed through the solution of the following constrained quadratic optimization problem[2].

$$\text{Minimise } 1/2\|w\|+C\sum_{i=1}^{l}\xi_1 \quad (2a)$$

$$\text{Subject to } \_\ y_i\ (\|w\|+ b) \geq 1 - \xi i,\ i = 1,\ldots,\ \ell$$

$$\xi i \geq 0, i=1,\ldots,\ell \quad (2b)$$

In linearly separable problem, the solution minimizes the norm of the vector w which increases the flatness (or reduces the complexity) of the resulting model and hence the generalization ability is improved. With non-linearly separable hard-margin optimization, the goal is simply to find the minimum ||w|| such that the hyperplane f(x) successfully separates all ℓ samples of the training dataset. The slack variables ξi are introduced to allow for finding a hyperplane that misclassifies some of the samples (soft-margin optimisation) because many datasets are not linearly separable. The complexity constant C>0 determines the trade-off between the flatness and the amount by which misclassified samples are tolerated. A higher value of C means that more importance is attached to minimising the slack variables than to minimising ||w||. Instead of solving this problem in its primal form of (2a) and (2b), it can be more easily solved in its dual formulation by introducing Lagrangian multiplier α [13]:

Maximize

$$W(\alpha)=\sum_{i=1}^{l}\alpha i_+ \tfrac{1}{2}\sum_{i,j=1}^{l}\alpha i\alpha j y i y j\langle xi, xj\rangle \quad (3a)$$

$$\text{Subject to } C\geq\alpha i\geq 0, \sum_{i=1}^{l}\alpha i y i=0 \quad (3b)$$

In this solution, instead of finding w and b the goal now is find the vector α and bias value b, where each αi represents the relative importance of a training sample I in the classification of a new sample. To classify a new sample, the quantity f(x) is calculated as:

$$f(x)=\sum_{i=1}^{sv}\alpha i y i K\langle xi, xj\rangle +b \quad (4)$$

Where b is chosen so that yif(x) = 1 for any i with C > αi > 0. Then, a new sample xs is classed as negative if f(xs) is less than zero and positive if f(xs) is greater than or equal to zero. Samples xi for which the corresponding αi are non-zero are called as *support vectors* since they lie closest to the separating hyperplane. Samples that are not support vectors have no influence on the decision function.

Training an SVM entails solving the quadratic programming problem of (3a) and (3b). There are many standard methods that are be applied to SVMs, these include the Newton method, conjugate gradient and primal-dual interior-point methods[15]. But this study used the Sequential Minimal Optimization[13].

In SVMs, kernel functions are used to map the training data into a higher dimensional feature space via some mapping φ(x) and construct a separating hyperplane with maximum margin. This yields a non-linear decision boundary in the original input space. Typical types of kernels are:

- Linear Kernel: $K(x, z) = \langle x, z\rangle$
- Polynomial Kernel: $K(x, z) = (\langle x, z\rangle)^d$
- RBF Kernel: $K(x, z)= \exp(-\|x-z\|^2/2\sigma 2\ )$
- Sigmoid Kernel: $K(x, z) = \tanh(\gamma*\langle x, z\rangle - \theta)$

This condition ensures that the solution of (3a) and (3b) produces a global optimum. The functions that satisfy Mercer's conditions can be as kernel functions.

As promising as SVM is compared with ANN as regards generalization performance on unseen data, the major disadvantage is its black box nature. The knowledge learnt by SVM is represented as a set numerical parameters value making it difficult to understand what SVM is actually computing.

Fuzzy logic overview

Fuzzy Logic which was introduced by Lotfi A. Zadeh was based on fuzzy sets in 1965[16]. The basic concept of fuzzy logic is to consider the intermediate values between [0, 1] as degrees of truth in addition to the values 1 and 0. The following sections will briefly discuss the general principles of fuzzy logic, membership functions, linguistic variables, fuzzy IF-THEN rules, combining fuzzy sets and fuzzy inference systems (FISs).

Fuzzy inference system

Fuzzy inference systems (FISs) are otherwise known as fuzzy-rule-based systems or fuzzy controllers when used as controllers. A fuzzy inference system (FIS) is made up of five functional components. The functions of the five components are as follows:

1. A *fuzzification is an* interface which maps the crisp inputs into degrees of compatibility with linguistic variables.
2. A *rule base is an interface* containing a number of fuzzy if-then rules.
3. A *database* defines the membership functions (MFs) of the fuzzy sets used in the fuzzy rules.
4. A *decision-making* component which performs the inference operation on the rules.
5. A *defuzzification* interface which transforms the fuzzy results of the inference into a crisp output.

In fuzzy logic, the rule base and the database in a FIS are both referred to as the "knowledge base". The steps of fuzzy reasoning are:

1. Input variables are compared with the MFs on the premise part to obtain the membership values (or degree of match) of each linguistic label. This first step is also known as "fuzzification".
2. The membership values on the premise part are combined through fuzzy set operations such as: min, max or multiplication to get firing strength (weight) of each rule.
3. The qualified consequent (either fuzzy or crisp) of each rule is obtained depending on the firing strength.
4. The qualified consequents are combined to produce crisp output according to the defined methods such as: centroid of area, bisector of area, mean of maximum, smallest of maximum and largest of maximum etc. This final step is also known as "defuzzification"[16-18]. The major disadvantage of standard fuzzy logic is the curse of dimensionality nature for high dimensional input space. For instance, if each input variable is allocated m fuzzy sets, a fuzzy system with n inputs and one output needs on the order of m^n rules.

METHODOLOGY

In this section, we will first give an insight into how to extract fuzzy rules from Support Vector Machine (SVM), and then explain the process of optimizing the fuzzy rules system and highlight an algorithm that will convert SVM into interpretable fuzzy rules. This method has both good generalization performance and ability to work in high dimensional spaces of support vector machine algorithm with high interpretability of fuzzy rules based models.

Extracting fuzzy rules from support vector machine

As mentioned earlier, Support vector machine (SVM) is a useful method of classifying dataset. This is a new machine learning method based on the Statistical Learning.

Suppose a set of training dataset denotes the input space patterns. Their main concept is to construct a hyperplane that acts as a decision space such that the margin of separation between positive and negative samples is maximized. This is generally referred as the Optimal Hyperplane". This property is achieved as the support vector machines are an approximate implementation of the method of structural risk minimization[14]. Despite the fact that a support vector machine does not provide domain-specific knowledge, it provides good generalization ability, a unique property among the different types of machine learning techniques.

Instead of solving this problem in its primal form of (2a) and (2b), it can be more easily solved in its dual formulation by introducing Lagrangian multiplier α [13]: as highlighted in section II.

The crucial step in fuzzy SVM is to build a reliable model on training samples which can correctly predict class label and extract fuzzy rules from SVM.

On the other hand, fuzzy rule-base which consists of set of IF-THEN rules constitutes the core of the fuzzy inference[3,6]. Suppose there are m fuzzy rules, it can be expressed as following forms:

Rulej: If x_1 is Aj^1 AND x_2 is Aj^2 and ……. x_n is. Aj^n THEN bj (5)

Where x_k is the input variables; b_j is the output variable of the fuzzy system; and A^k are linguistic terms characterized by fuzzy membership functions a_j^k. If we choose product as the fuzzy conjunction operator, addition for fuzzy rule aggregation, and height defuzzification, then the overall fuzzy inference function is,

$$F(x) = \frac{\sum_{j=1}^{m} bj \prod_{k=1}^{n} a_j^k(xk)}{\sum_{j=1}^{m} \prod_{k=1}^{n} a_j^k(xk)} \quad (6)$$

Where F(x) is the output value when the membership function achieves its maximum value.

If on the other hand, the input space is not wholly covered by fuzzy rules, equation (5) may not be defined. To avoid this situation, Rule0 can be added to the rule base,

Rule0: If $A0^1$ AND $A0^2$ AND ……. $A0^n$ THEN b0

$$F(x) = \frac{b0 + \sum_{j=1}^{m} bj \prod_{k=1}^{n} a_j^k(xk)}{1 + \sum_{j=1}^{m} \prod_{k=1}^{n} a_j^k(xk)} \quad (7)$$

In a binary classification, sign (F(x)) shows the class label of each input x and since the denominator is always positive, class label of each input is computable by,

$$Label(x) = sign\left(b0 + \sum_{j=1}^{m} bj \prod_{k=1}^{n} a_j^k(xk)\right) \quad (8)$$

In order to let equation (4) and (8) are equivalent, at first we have to let the kernel functions in (4) and the membership functions in (8) are equal. The Gaussian membership functions can be chosen as the kernel functions to satisfy the Mercer condition (1). Besides, the bias term of the expression (4) should be zero. If the Gaussian function is chosen as the kernel function and membership functions, and the number of rules equals the number of support vectors. Then (4) and (8) becomes equal and then output of fuzzy system (8) is equal to the output of SVM (4).

A schematic of fuzzy SVM nanomaterials characterization and risk assessment system is shown in figure 1. The system is designed to assess the risk of using nanomaterials. The size, surface area and concentration produce the symptoms of increased toxicity. The toxicity signature is extracted on measuring the above

parameters. The fuzzy model was simulated using Fuzzy controller software. The toxicity assessment is carried out by analyzing the fault signature through the fuzzy rules derived from expert's knowledge and experimental data. The simulation procedure is explained in the section V.

Numerical experiments

Prediction performance of the resulting models depends on the size and quality of the training data. Each data record consists of input and output data. Input data are derived from physicochemical properties of the materials as shown in Table 1.

Purpose of study

The objective of this study is to classify and assess the risks associated with the use of nanomaterials based on size, surface area, exposure time, aspect ratio, concentration and relative toxicity index. The flow chart for nanomaterial characterization/classification is as shown in Fig. 1.

Step1: Data preprocess and variable selection

In this study, the measured attribute are size, surface area, dose, exposure time, aspect ratio, concentration and relative toxicity index (Nanomaterial Class:-1(non-toxic material), 1(toxic material)).

In Computational Intelligent Nanomaterials Toxicity (CINT) software (developed by the author), classification of toxic nanomaterials is performed.

The result confirmed that the classification precision of the SVM with radial function (RBF) kernel function was high as 100% when γ and C where 0.55 and 0.1. Then the best parameter of C and γ was selected to train the whole training set, we have 20 support vector index sets.

The outputs from NCIS software are;
Accuracy=100%.
MSE=0.0
Squared correlation coefficient=1.

RESULTS AND DISCUSSION

The sample data used for testing are as shown in Table 3. There are two types of errors namely Type I and Type II errors. Type I refers to a situation when toxic material was classified as non- toxic material. Type II refers to non- toxic material being classified as toxic material. The predicted result is as listed in Table 3. The results of testing (external validation check were summarized in Table 3. We observed form these results that the hybrid Fuzzy-support vector machines modeling scheme performed satisfactorily for predictive correlations. The model showed a high accuracy in predicting toxicity class with a stable performance, and achieved the lowest absolute percent relative error type I and type II errors, lowest root mean square error, and the highest correlation coefficient among other correlations for the used two distinct data sets. A plot of the experimental and predicted data versus the input data is as shown in Fig. 2.

CONCLUSION

This study developed a novel fuzzy SVM to characterize and assess nanomaterials toxicity. The classification of nanomaterials (non-toxic, Low risk and high risk) is a work that is aimed at with an in-depth study and extraction of rules from support vectors. The study and understanding of the fuzzy rule based support vector machines and its roles in classification tasks were done. This technique was then implemented in the Microsoft C# programming language to perform data classification task for the nanomaterial toxicity data set. This approach compensated for the shortcomings of Fuzzy logic and standard SVM.

REFERENCES

1. Agnieszka Gajewicz a, Bakhtiyor Rasulev b, Tandabany C. Dinadayalane b, Piotr Urbaszek a, Tomasz Puzyn a,Danuta Leszczynska c, Jerzy Leszczynski (2012) Advancing risk assessment of engineered nanomaterials: Application of computational approaches.
2. John F. Sargent Jr. "Nanotechnology and Environmental, Health and Safety: Issues for Consideration" (2011).
3. Oberdörster G 2002 Toxicokinetics and effects of fibrous and nonfibrous particles *Inhalation Toxicol.* 14 29-56 and references therein.
4. Oberdörster G, Oberdörster, E, Oberdörster J 2005 Nanotoxicology: an emerging discipline evolving from studies of ultrafine particles *Environ. Health. Perspect.* 113 823-839 and supplemental material found at http:// www.ehponline.org/members/2005/7339/7339.html.
5. Hoet P H M, Bruske-Hohlfeld I, Salata O V 2004 Nanoparticles - known and unkown health risks *J. Nanobiotechnol.* 2 12-27 and references therein.
6. Oberdörster G, Ferin J, Lehnert B E 1994 Correlation between particle size, *in vivo* particle persistence, and lung injury *Environ. Health Persp.* 102 Suppl 5 173-179.
7. Ferin J, Oberdörster G, Penney D P 1992 Pulmonary retention of ultrafine and fone particles in rats *Am.J. Respir. Cell Mol. Biol.* 6 535-552.
8. Stoeger T, Reinhard C, Takenaka S, Schroeppel A, Karg E, Ritter B, Heyder J, Schultz H 2006 Instillation of six different ultrafine carbon particles indicates a surface area threshold dose for acute lung inflammation in mice *Environ. Health. Perspect.* 114 328-333.
9. Takenaka S, Karg E, Roth C, Schulz H, Ziesenis A, Heinzmann U, Schramel P, Heyder J 2001 Pulmonary and systemic distribution of inhaled ultrafine silver particles in rats *Environ. Health Persp.* 109 (Suppl. 4) 547-551 and references therein.
10. Roduner E 2006 Size matters: why nanomaterials are different *Chem. Soc. Rev.* 35 583-592.
11. Lippmann M 1990 Effects of fiber characteristics on lung deposition, retention, and disease *Environ.Health Perspec.* 88 311-317.
12. Sch¨olkopf, S. Mika, C. Burges, P. Knirsch, K.-R. Muller, G. R¨atsch, and A. Smola "Input space vs. feature space in kernel-based methods". *IEEE Transactions on Neural Networks*, 10(5):1000–1017 1999.
13. Sch¨olkopf, A. Smola, R. C. Williamson, and P. L. Bartlett "New support vector algorithms". *Neural Computation*, 12:1207–1245 2000.
14. V. Cherkassky and Y. Ma "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks*, vol. 17, pp. 113–126 2004.
15. Z. Y. Luo, P. Wang, Y. G. Li, W. F. Zhang, W. Tang and M. Xiang, "Quantum-inspired evolutionary tuning of SVM parameters," *Progress in Natural Science*, vol. 18, pp. 475–480, 2008.
16. Zadeh, L. A., 1965. Fuzzy sets. *Information and Control*, vol, 8, pp, 338–353.
17. Zadeh, L. A., 1973. Outline of a new approach to analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 28-44.
18. L.A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," *Information Sciences*, no. 8, pp. 199–249, 301-357, 1975.

Table 1. Nanomaterials training samples

Listed species of a nanomaterial	Nanomaterial Size (nm)	Surface area (cm^2)	Exposure Time (mins)	Aspect Ratio	Concentration $\mu g/m^3$	Nanomaterial Class
1	123.0	2.0	5.0	20.0	20.0	-1
2	118.0	8.0	3.0	10.0	21.0	-1
3	116.0	8.0	5.0	30.0	17.0	-1
4	119.0	9.0	4.0	40.0	18.0	-1
5	120.0	2.2	6.0	20.0	19.0	-1
6	124.0	2.0	6.0	20.0	20.0	-1
7	123.0	14.0	5.0	22.0	20.0	-1
8	20.0	18.0	20.0	100.0	7.0	1
9	18.0	19.0	30.0	200.0	8.0	1
10	19.0	12.0	35.0	200.0	8.0	1
11	122.0	3.0	10.0	35.0	15.0	-1
12	16.0	20.0	40.0	400.0	7.0	1
13	19.0	19.0	45.0	250.0	7.0	1
14	115.0	8.0	4.0	26.0	20.0	-1
15	118.0	4.0	3.0	42.0	18.0	-1
16	112.0	7.0	6.0	32.0	13.0	-1
17	121.0	8.0	5.0	28.0	12.0	-1
18	119.0	8.0	6.0	38.0	19.0	-1
19	122.0	9.0	7.0	31.0	14.0	-1
20	120.0	5.0	4.0	30.0	19.0	-1

Table 2. Nanomaterials test sample

Listed species of a nanomaterial	Nanomaterial Size (nm)	Surface area (cm^2)	Exposure Time (mins)	Aspect Ratio	Concentration $\mu g/m^3$	Nanomaterial Class
1	123.0	2.0	5.0	20.0	20.0	-1
2	118.0	8.0	3.0	10.0	21.0	-1
3	116.0	8.0	5.0	30.0	17.0	-1
4	119.0	9.0	4.0	40.0	18.0	-1
5	120.0	2.2	6.0	20.0	19.0	-1
6	124.0	2.0	6.0	20.0	20.0	-1
7	123.0	14.0	5.0	22.0	20.0	-1
8	20.0	18.0	20.0	100.0	7.0	1
9	18.0	19.0	30.0	200.0	8.0	1
10	19.0	12.0	35.0	200.0	8.0	1
11	122.0	3.0	10.0	35.0	15.0	-1
12	16.0	20.0	40.0	400.0	7.0	1
13	19.0	19.0	45.0	250.0	7.0	1
14	115.0	8.0	4.0	26.0	20.0	-1
15	118.0	4.0	3.0	42.0	18.0	-1
16	112.0	7.0	6.0	32.0	13.0	-1
17	121.0	8.0	5.0	28.0	12.0	-1
18	119.0	8.0	6.0	38.0	19.0	-1
19	122.0	9.0	7.0	31.0	14.0	-1
20	120.0	5.0	4.0	30.0	19.0	-1
21	118.0	7.0	3.0	29.0	18.0	?
22	118.0	15.0	20.0	115.0	7.0	?
23	115.0	19.0	30.0	125.0	8.0	?
24	120.0	6.0	4.0	30.0	19.0	?
25	125.0	5.0	6.0	35.0	17.0	?

Table 3. Toxicity prediction results

Method	Number of samples	Type I error	Type II error	Error	Accuracy
Fuzzy SVM	20	0%	0%	0	100%

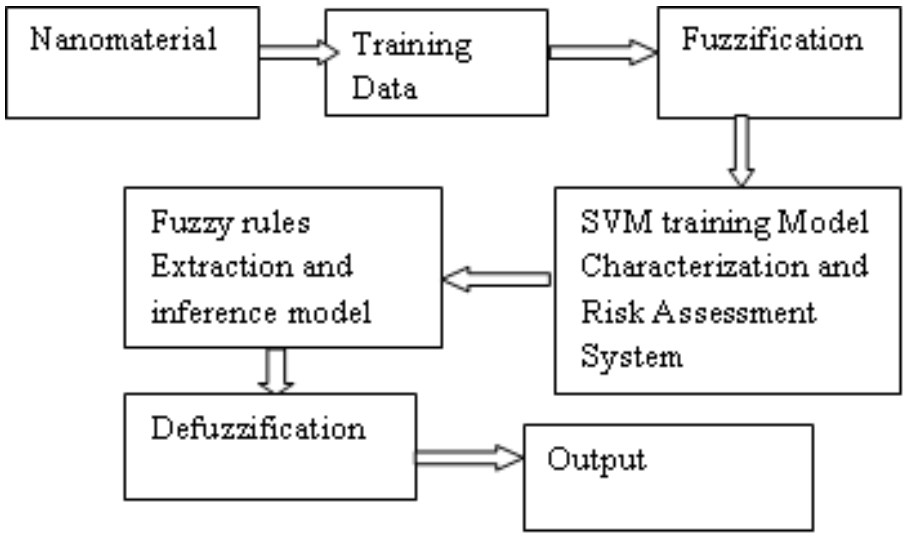

Figure 1. Fuzzy SVM nanomaterials characterization and risk assessment system

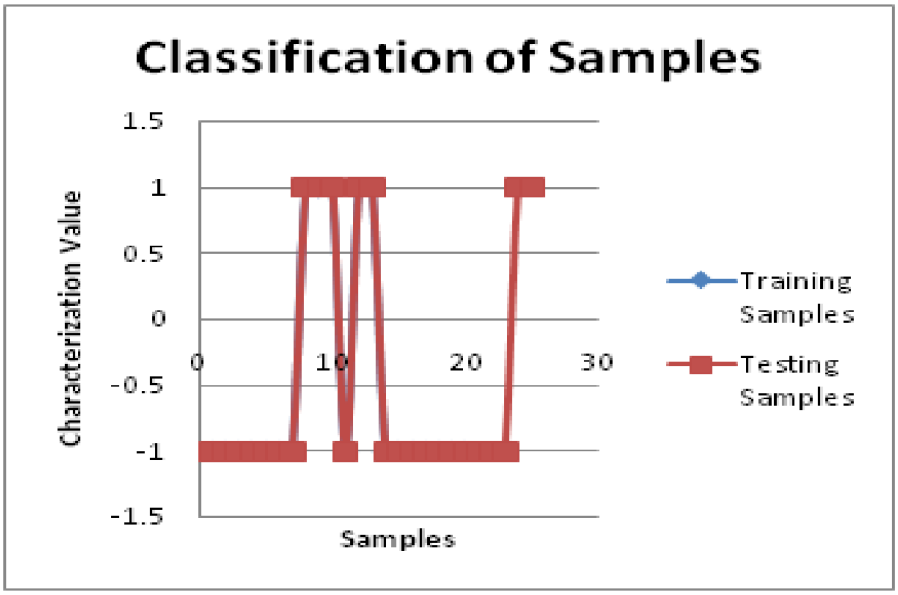

Figure 2. Classification of sample