

An Improved Binary Classification Framework for Investment Class Rating

Ajit Kumar Das

International Institute of Information Technology
Bhubaneswar
Odisha

Sudarsan Padhy

Institute of Mathematics and Application
Bhubaneswar
Odisha

ABSTRACT.

This paper presents a binary classification scheme for investment class rating using support vector machine (SVM). The suggested SVM model is trained offline and takes twelve financial ratios as attributes from different standard investment companies as inputs and correctly classify whether it is a good investment grade or bad investment grade company as output. The overall performance of SVM strongly depends on the regularization parameter C and kernel parameter σ . Hence, we propose the PSO based optimization technique using mean square error (MSE) as the fitness function to optimize the value of C and σ . The proposed scheme is implemented using Matlab and Libsvm tool. Comparison is made in terms of different performance measures like classification accuracy, sensitivity, specificity, precision, confusion matrix etc. From experimental results and analysis, it is observed that the proposed scheme has a superior performance as compared to SVM based approach without parameter optimization and neural network based scheme.

General Terms

Support Vector Machine, Binary Classification et al.

Keywords

Support Vector Machine (SVM), Binary Classification, Neural Network, Particle Swarm Optimization (PSO), Mean Squared Error (MSE)

1. INTRODUCTION

The investment class rating provides the investment worthiness of Corporate or Government bonds. Investment class rating is determined by different investment agencies. The investment class rating represents the investment class rating agency's evaluation of qualitative and quantitative information for a Company or Government; based on non-public information obtained by the investment class rating agencies' analysts.

Investment class rating is generally not based on mathematical formula. Instead, investment class rating agencies use their judgment and experience in determining what public and private information should be considered in giving a rating to a particular Company or Government. It is used by individuals and other entities that purchase the bonds issued by Companies and Governments to determine the likelihood that the bond issuer will pay its bond obligations.

A poor investment class rating, which provides an investment class rating agency's opinion based on the agency's analysis of the entity's history and analysis of long term economic

prospects that the Company, Government or the financial agency, has a high risk of defaulting. Investment class ratings are awarded by specialised agencies, the four largest rating agencies being Moody's, S&P, Fitch Ratings, and DBRS. These four ratings agencies are recognized by the European Central Bank for the purposes of determining collateral requirements for banks to borrow from the central bank.

In India, CRISIL is the most reputed global analytical company providing ratings, research, and risk and policy advisory services. They are the India's leading ratings agency and are also the foremost provider of high-end research to the world's largest banks and leading corporations.

Investment class rating is costly to obtain due to large investment of time and human resources by rating agencies to perform investment rating process. Hence, a great deal of effort has been made to simulate the Investment class rating process through statistical and soft-computing methods. The complex relations between financial and other variables are evaluated using soft-computing methods such as Neural Networks (NNs), Fuzzy Systems, Evolutionary Algorithms, Artificial Immune Systems, and Hybrid Systems. The different methods offer different levels of classification accuracy for Investment class rating.

During the last decade, Support Vector Machines (SVMs) developed by Vapnik et al have been used for solving classification problems with a greater degree of classification accuracy. Many variants of SVM have been introduced by different researchers worldwide.

Investment class rating modelling is regarded as a classification problem. The rating process represents a subjective assessment of both quantitative and qualitative factors of a particular object. Very few attempts have been made for classification problem using SVM in investment class rating. So taking this into account in this paper, we have proposed a binary classification approach for investment class rating using SVM. It has been observed that from many instances the overall performance of SVM is completely based on the regularization parameter C and kernel parameter σ . So to obtain an enhanced performance further, we have proposed a PSO based optimization technique to optimize the value of C and σ . The rest of the paper is organized as follows. Section 2 elaborates the related work. Section 3 presents the related background of SVM followed by the proposed binary classification scheme using SVM in Section 4. The overall simulation is divided into a set of experiments and the results obtained are discussed experiment wise in Section 5. Finally, the conclusion is given in Section 6.

2. RELATED WORK

The investment class rating modeling is considered to be a classification problem with the goal of classifying objects. Investment class rating is costly to obtain due to the large amount of time and human resources invested by rating agencies to perform the investment class rating process. There is a great deal of effort made to simulate the investment class rating agencies through statistical [1–3] and soft-computing methods [4, 5]. The difficulty in designing such models lies in the subjectivity of the investment class rating process. This subjectivity is accentuated as the particular modes in which complex relations between financial and other variables are evaluated. Such a complex process makes it difficult to classify rating classes through statistical methods. However, soft-computing methods such as neural networks [4, 5], fuzzy systems [6], evolutionary algorithms [7], artificial immune systems [8], and hybrid systems [9] have been applied for the modeling of such complex relations. As a result, high classification accuracy has been achieved, especially by NNs [8, 10] and support vector machines [9, 11].

In general, any kernel-based algorithm is composed of two modules, a general purpose learning machine and a problem-specific kernel function [12]. Support vector machines represent an essential kernel-based method. Many variants of SVMs have been developed since SVMs were proposed, e.g. least squares SVMs and robust SVMs [12]. Concerning the problem of classification, other methods (NNs, fuzzy systems, etc.) have also been extended to incorporate maximizing margins and feature space mapping [1]. SVM are affiliated to NNs and create the category of so-called kernel machines [12 - 14]. The main principle of SVMs consists in the creation of the decision hyper plane between classes, so that the margin between positive and negative patterns is maximized. More precisely, SVMs represent the approximate implementation of a structural risk minimization method [13]. This principle is based on the fact that testing error is limited, with the sum of training error and the expression depending on the Vapnik–Chervonenkis dimension. Several literatures exist in computational finance area for classification and regression. In a nutshell, we have reviewed some of the relevant reported literature and presented in sequel.

In 2010, Jan-Henning et al. [15] have proposed the relative performance between least-squares support vector machines and logistic regression models for default classification and default probability estimation. The financial ratios from a data set of more than 78,000 financial statements from 2000 to 2006 are used as default indicators. The main focus of this paper was on the influence of small training samples and high variance of the financial input data and the classification performance measured by the area under the receiver operating characteristic. The resolution and the reliability of the predicted default probabilities were evaluated by decompositions of the Brier score. It was shown that support vector machines significantly outperform logistic regression models, particularly under the condition of small training samples and high variance of the input data. Further, in 2010 Peter Hajek and Vladimir Olej [16] have presented the modeling possibilities of kernel-based approaches to a complex real-world problem, i.e. corporate and municipal credit rating classification. Based on a model design that includes data pre-processing, the labeling of individual parameter vectors using expert knowledge, the design of various support vector machines with supervised learning as well as kernel-based approaches with semi-supervised learning, this modeling was undertaken in order to classify

objects into rating classes. The results showed that the rating classes assigned to bond issuers can be classified with high classification accuracy using a limited subset of input variables. This holds true for kernel-based approaches with both supervised and semi-supervised learning. Tony Gestel et al. [17] have proposed a support vector machine approach for credit scoring. In this reported literature they have proposed least square support vector machine to obtain a better credit scoring. Das & Padhy [18] have presented SVM based method for prediction of future prices in Indian stock market.

It has been observed from the literature that SVM is a potential tool to apply in the field of investment class rating, which motivated us to propose a binary classification scheme for Indian Companies using SVM for investment class rating with PSO based parameter optimization technique.

3. RELATED BACKGROUND OF SVM

Support Vector Machine (SVM) methods are supervised learning algorithms first introduced by Vapnik [19]. Consider, a training set S consisting of labeled input vectors (x_i, y_i) ,

$i = 1, \dots, m$ where $x_i \in R^n$ and $y_i \in \{\pm 1\}$. A classification

rule or decision function, f can be specified by

$$f(x) = \langle w, x \rangle + b, \quad (1)$$

where, $w \in R^n$ and $b \in R$, $\langle w, x \rangle$ is the inner product or scalar product between two vectors. The vector w is known as the weight vector, and the scalar b is called the bias that translates the hyper plane with respect to origin. The decision boundary is the hyper plane $\{x \in R^n : \langle w, x \rangle + b = 0\}$,

where w is a normal vector to the hyper plane. The data for a two-class learning problem consists of objects labeled with one of two labels: positive examples with label (+1) and negative examples with label (-1).

Thus the decision boundary of the classifier f is the boundary between the positive and negative regions and margin of a linear classifier is defined as the distance of the closest example to the decision boundary. A natural question is to choose a hyper plane from many choices of such separating hyper planes so that it will ensure correct classification not only for existing dataset, but also for any other examples unseen by the classifier at the training time. Hyper plane classifiers work better if the hyper plane separates the examples correctly with a large margin. Hard margin SVM is the classifier with maximum margin among all classifiers that correctly classify all the input examples. It can be shown

that the margin is $\frac{1}{\|w\|}$, where $\|w\| = \sqrt{\langle w, w \rangle}$. Thus to

compute w and b corresponding to the maximum margin hyper plane, one has to solve the following optimization problem:

$$\underset{(w,b)}{\text{Minimize}} \frac{1}{2} \|w\|^2, \text{ subject to } y_i(\langle w, x_i \rangle + b) \geq 1; \quad i = 1, \dots, n, \quad (2)$$

Since minimizing $\|w\|^2$ is equivalent to maximizing the margin. When the data are not linearly separable, and even if they are, a larger margin can be achieved by allowing the classifier to misclassify some points. The resulting margin known as soft margin generally provides better performance than the hard margin SVM. To incorporate larger margins, the

equation (2) can be modified as $y_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i$, $i = 1, 2, \dots, n$, where $\xi_i \geq 0$, are slack variables that allow an example to be within the margin or to be misclassified. This leads to the soft margin SVM formulation [20] of (2) as

$$\begin{aligned} & \underset{(w,b)}{\text{Minimize}} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i, \text{ subject to} \\ & y_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i, \quad i = 1, 2, \dots, n \end{aligned} \quad (3)$$

The constant $C > 0$ sets the relative importance of maximizing the margin and minimizing the amount of slack. The dual formulation of the primal problem specified in equation (3) is obtained by using the Lagrange multipliers α_i and is expressed as:

$$\max_{\alpha_i} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j \langle x_i, x_j \rangle \quad (4)$$

subject to (i) $\sum_{i=1}^n y_i \alpha_i = 0$, and (ii) $0 \leq \alpha_i \leq C$

Then, the weight vector W can be expressed in terms of the examples and the solution α_i^* of the above optimization problem as,

$$w = \sum_{i=1}^n y_i \alpha_i^* x_i \quad (5)$$

Choose an index i for which $\alpha_i^* > 0$. Then b can be obtained as:

$$b = y_i - \langle w, x_i \rangle = y_i - \sum_{j=1}^n y_j \alpha_j^* \langle x_j, x_i \rangle \quad (6)$$

Then the classification rule f specified in equation (1) reduces to

$$f(x) = \sum_{i=1}^n y_i \alpha_i^* \langle x_i, x_p \rangle + b \quad (7)$$

The x_i for which $\alpha_i > 0$ are called support vectors i.e. the points with positive Lagrange multiplier α_i are the support vectors. It is observed from equation (5) that W is a linear combination of the support vectors alone. Thus the set of support vectors contains all the information about the dataset used by the SVM to learn a discrimination function. It is observed from equations (4), (6), (7) that learning a linear classifier and predicting the class of a new point only involves the points in the training set through their dot products. When the data are not linearly separable we cannot find a hyperplane which separate the two classes. However, the use of kernels make it possible to map the data to some inner product space known as feature space where the dot products of images of the input data can be computed and train a linear machine in such a space. A kernel is a function

$K: X \times X \longrightarrow R$ such that for all $x, x' \in X$ (input space) satisfies:

- I. $K(x, x') = \langle \Phi(x), \Phi(x') \rangle$, where Φ is a mapping from the input space X to an inner product space (feature space) F .
- II. K is symmetric and positive definite, that is, $K(x, x') = K(x', x)$ for any two objects $x, x' \in X$,

and $\sum_{i=1}^n \sum_{j=1}^n c_i c_j K(x_i, x_j) \geq 0$, for any choice of $n > 0$ objects $x_1, \dots, x_n \in X$, and any choice of real numbers $c_1, \dots, c_n \in R$.

Examples of some frequently used kernels in practice are Polynomial, Gaussian etc. Thus, using some kernel, the SVM method in the feature space can be stated in the form of the following algorithm.

Algorithm for General SVM

Step 1: Choose a problem specific kernel K and find

$\alpha = (\alpha_1, \dots, \alpha_n)$ that maximizes

$$\sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j K(x_i, x_j), \quad \text{Subject to}$$

constraints:

(i) $\sum_{i=0}^n y_i \alpha_i = 0$ and (ii) $0 \leq \alpha_i \leq C$ for $i = 1$ to n

Step 2: Find an index i with $0 < \alpha_i < C$, and set b as

$$b = y_i - \sum_{j=1}^n y_j \alpha_j K(x_j, x_i).$$

Step 3: The classification of a new object $x \in X$ is then based on the sign of the function

$$f(x) = \sum_{i=1}^n y_i \alpha_i K(x_i, x) + b.$$

4. PROPOSED SVM BASED INVESTMENT CLASS RATING

This section presents a SVM based approach to classify the investment class rating for Indian companies. The SVM model is trained offline using twelve attributes like profitability, liquidity etc. (described later) as input and correctly classify whether it is a good or bad investment grade entity.

Performance measures are made in terms of classification accuracy, sensitivity, specificity, precision, confusion matrix, and F1 score. It is observed that the proposed scheme has a superior performance as compared to SVM based approach without parameter optimization technique and neural network based scheme.

4.1 Data Collection and Preprocessing

We have collected investment and business related data as well as company related data of about 900 companies in the financial year ending 2013. Credit ratings and financial parameters have been collected from the CRISIL database and CMIE database respectively.

After data collection, it is found that certain data were incomplete i.e., for a particular company, certain attributes were not provided. That is why data preprocessing is required to prepare quality data. During data preprocessing/data preparation, outliers have been identified and removed or replaced wherever necessary. And also missing values are filled in manually with the most probable value or attribute mean. This is called data cleaning. Another preprocessing used is data scaling. All the data points are scaled to fall within a small, specified range. The main advantage of scaling

is to avoid attributes in greater numeric ranges dominating those in smaller numeric ranges. Another advantage is to avoid numerical difficulties during the calculation. Because kernel values usually depend on the inner products of feature vectors, e.g. the linear kernel and the polynomial kernel, large attribute values might cause numerical problems.

For this experiment the training and testing data points are scaled into the range of [-1; 1] as the data points include both positive and negative values.

In this experiment, initially we have taken 600 training samples and after that varied the training samples as 500, 450, 300, 250 etc. to study the performance measure. Similarly we have taken 281 testing samples and varied the testing samples to study the performance measure.

Table 1 shows the test data of the proposed approach.

Table 1: A tabular representation of test data

A	B	C	D	E	F	G	H	I	J	K	L	M
1	0.3818	0.0874	0.2143	0.0132	1.1497	6.095709	0.184	0.148	0.121	0.131	0.153	0.166
1	0.2972	0.0388	0.1083	0.0172	0.8230	3.2005988	0.131	0.069	0.102	0.052	0.065	0.166
1	0.2477	0.0332	0.1525	0.0485	1.6889	2.6504684	0.09	0.035	0.051	0.039	0.049	0.167
1	0.3955	0.0895	0.2282	0.0397	0.7483	4.8636079	0.294	0.24	0.227	0.141	0.167	0.167
1	0.1562	0.0253	0.1311	0.0106	3.6868	1.4803463	0.035	0.016	0.019	0.039	0.067	0.168
-1	-0.0428	0.0445	0.1253	0.1309	1.0065	0.7083333	0.123	0.052	0.065	0.045	0.053	0.168
1	0.2201	0.0248	0.1580	0.0395	2.1907	1.5873494	0.072	0.013	0.031	0.025	0.044	0.168
1	0.1371	0.1002	0.2985	0.0158	1.0731	2.0748954	0.269	0.187	0.174	0.141	0.168	0.168
1	-0.0160	0.0099	0.0926	0.0013	0.0000	0.4744101	0.965	0.178	0.13	0.014	0.015	0.168
1	0.2509	0.0878	0.1861	0.0217	1.7176	2.9159946	0.107	0.069	0.075	0.111	0.14	0.169

4.2 Proposed SVM based Investment Class Rating Approach

The proposed architecture of SVM based binary classification is shown in **Figure 1**.

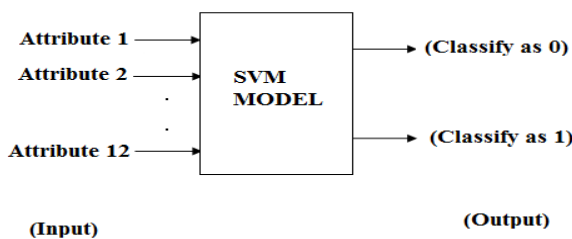


Figure 1: Architecture of Proposed SVM based Classification Approach

The following attributes chosen by considering the diversity of companies for credit rating classification are fed as the input to the SVM model:

- x_1 = (current assets-current liabilities) / Total assets,
- x_2 = Retained profits / Total Assets,
- x_3 = EBDIT / Total Assets,
- x_4 = Total Capital / Total liabilities,
- x_5 = Sales / Total assets,
- x_6 = current assets / Current liability,
- x_7 = PBDITA / Total income,

$$x_8 = \text{PBT} / \text{Total income},$$

$$x_9 = \text{Cash profit} / \text{Total income},$$

$$x_{10} = \text{PAT} / \text{Total assets},$$

$$x_{11} = \text{PAT} / \text{Capital employed, and}$$

$$x_{12} = \text{PAT} / \text{Net worth}$$

The proposed scheme for investment class rating using Support Vector classification technique is implemented using LIBSVM tool (supports SVC i.e., support vector classification).

After data preprocessing, both the training and testing data set which are in xlsx format need to be converted to LIBSVM specific format. For this, MATLAB programming is done which converts the xlsx files to respective tool specific file.

Once the training and test files are created, LIBSVM tool is used to create a model which is trained and tested with the training file and test file.

If train file is train1 and test file is test1, following steps are followed to use LIBSVM tool. The command svm-train reads the training data file 'train1' and output the model file as 'train1.model'. Then svm-predict uses the model file 'train1.model' to predict the test file 'test1'. It produces output as 'op' and outputs the 'classification accuracy' and 'number of correct classification'. The command svm-scale scales the input data file 'train1' and produces 'train1.scale' file. It then scales the test data file 'test1' and produces 'test1.scale' file as output. In the next command svm-train is used to read the scaled training data file and it produces the model file 'train1.scale.model' as output.

Then svm-predict uses this scaled model to predict the scaled test file, 'test1.scale' and stores output in 'scaledop' file.

For complete analysis of the Binary Classification for Investment Class Rating using the proposed scheme, Algorithm 1 lists the steps of the proposed scheme.

Algorithm 1: Proposed SVM based Binary Classification for Investment Class Rating

Input: A training dataset containing twelve attributes related to ratios and parameters for the corresponding investment companies

Output: Investment class rating classification accuracy

Training Phase:

1. Select a training dataset consisting of attributes (related to the corresponding investment company).
2. Convert it into the format of the respective classification tool
3. Train the tool using the training dataset.

Testing Phase:

1. Select a testing dataset consisting of same attributes as that of training
2. Convert it into the format of the respective classification tool
3. Classify the information of the testing dataset using the tool.

4.3 Proposed Parameter Optimization Using PSO

In the present work, the SVM approach towards investment class rating is improved by optimizing the parameters of SVM. A genetic algorithm (GA) based approach has been proposed [21] to optimize the SVM parameters. Authors in [22] have also suggested a GA based approach for optimizing the SVM parameters. The PSO is utilized in this work for faster convergence and easier implementation. PSO has also fewer adjusting parameters than GA. PSO is discussed in a nutshell followed by detailed procedure of our proposed method in sequel.

PSO is a stochastic optimization technique initially developed by Eberhart and Kennedy [23] and subsequently modified to a more generalized form [24, 25]. It is an evolutionary computation technique based on intelligent behavior of swarm. A swarm consists of particles, which represent the solution. They fly in a multidimensional search space. Each particle changes its position according to its own experience, and experience collected from the neighboring particles. In this way, the particles move towards the best solution. The performance of each particle is measured using a fitness function which is application dependent. In the present work, MSE is used as the fitness function.

The present particle position (*presentx*) and velocity (*V*) of each particle are updated using the following two update equations defined as,

$$V = V + c1 * rand() * (pbest - presentx) + c2 * rand() * (gbest - presentx) \quad (8)$$

and

$$presentx = presentx + V \quad (9)$$

Where, *rand* is a random number between 0 and 1 and *c1*, *c2* are two weighting constants or accelerating constants. The local best solution of the particle is defined as the *pbest* and *gbest* is the global best solution of all the particles achieved so far. The SVM is trained with a random initial set of parameters. The parameters of the SVM i.e. *C*, σ are updated till the MSE is below a threshold level (*T*). In this work, a threshold value of 0.01 is used in the experiments. The flow chart of the proposed SVM-PSO hybrid technique is shown in Figure 2.

5. RESULTS AND DISCUSSION

In order to evaluate the efficacy of the proposed scheme, different experiments have been carried out on financial datasets obtained for Indian Companies. The proposed scheme has been implemented using Matlab and LIBSVM tool. To evaluate the overall performance of the proposed SVM based model with parameter optimization, it is compared with the neural network scheme and SVM based approach without parameter optimization scheme. The overall experimental study is divided into three experiments and are discussed in sequel

Experiment 1: Analysis of classification accuracy in (%) with different number of training samples

The proposed SVM model with parameter optimization scheme, Neural Network and SVM based model without parameter optimization scheme are simulated using common dataset obtained by us. Table 2 shows the accuracy in (%) with different number of training samples. During the experiment we have varied the number of training samples and computed the correct classification accuracy. Figure 3 shows the accuracy in (%) vs. number of training samples. It is observed that in most of the cases the proposed scheme outperforms in terms of classification accuracy, as number of samples increase, in contrast with the Neural Network scheme and SVM based scheme without parameter optimization. In particular, the proposed scheme shows superior performance in terms of classification accuracy as compared to the earlier schemes.

Table- 2: Accuracy in % obtained by different models

No. of Samples	Neural Network	SVM without parameter optimization	SVM with parameter optimization
600	75.33	81	84.19
500	68.20	76	82
450	64.44	71.33	79.33
325	49	49.6	66.4

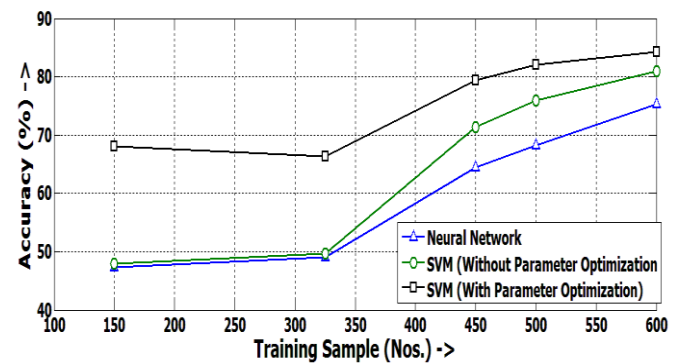


Figure 2: Training sample number vs. Accuracy in %

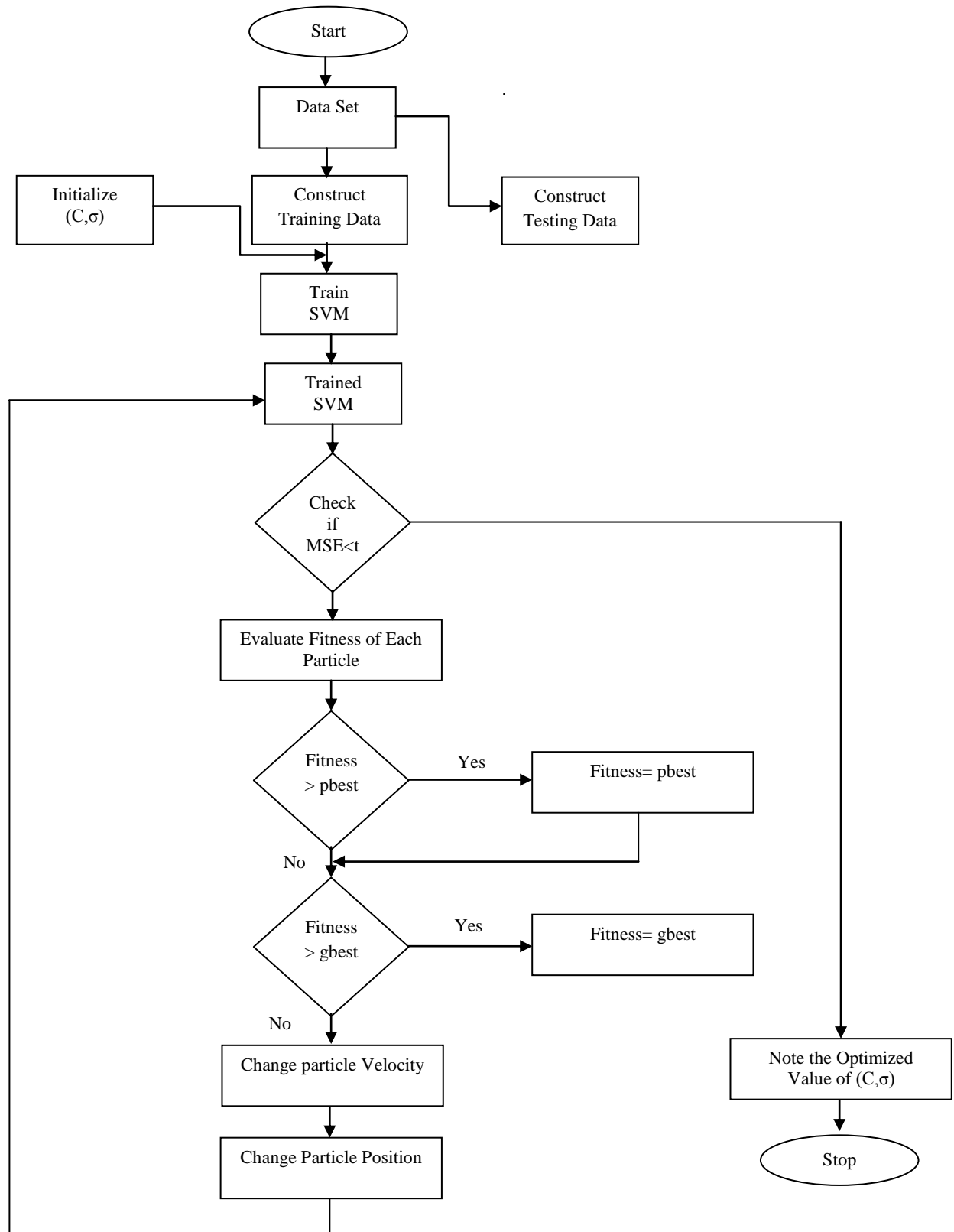


Figure 3: Flowchart of Proposed PSO based Optimization

Experiment 2: Analysis of the parameters like sensitivity, specificity, precision, and F₁-score

Experiment 2 computes the different classification performance parameters like sensitivity, specificity, precision, and F₁-score. In the proposed scheme we compute these parameters along with the existing neural network and SVM based scheme without parameter optimization. It is observed that the proposed scheme shows improved performance in all respects as compared to its competent schemes. Tables 3(a)-3(c) list the value of sensitivity, specificity, precision, and F₁-score parameters for the three schemes.

Table 3(a): Neural Network based scheme

Number of Samples	600	500	450	325	125
Sensitivity	0.82	0.8339	0.7217	0.67	0.58
Specificity	0.68	0.6180	0.6776	0.65	0.52
Precision	0.82	0.7726	0.73	0.651	0.54
F1-Score	0.8214	0.802	0.729	0.66	0.56

Table 3(b): SVM without parameter optimization based scheme

Number of Samples	600	500	450	325	125
Sensitivity	0.91	0.86	0.8077	0.4	0.51
Specificity	0.25	0.28	0.1	0.6	0.2
Precision	0.85	0.843	0.8537	0.87	0.85
F1-Score	0.89	0.82	0.83	0.61	0.63

Table 3(c): SVM with parameter optimization based scheme

Number of Samples	600	500	450	325	125
Sensitivity	0.9514	0.9123	0.896	0.72	0.69
Specificity	0.2273	0.2759	0.3	0.27	0.28
Precision	0.8736	0.8814	0.8898	0.85	0.83
F1-Score	0.9109	0.8966	0.87	0.78	0.76

Experiment 3: Analysis of confusion matrix for different schemes

Confusion matrix, also known as a contingency table or an error matrix, is a specific table layout that allows visualization of the performance of an algorithm, typically a supervised learning one (in unsupervised learning it is usually called a matching matrix). Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. The name stems from the fact that it makes it easy to see if the system is confusing two classes. It is an important performance measure in classification problem. In the experiment we computed confusion matrix for the proposed SVM based scheme with parameter optimization. Similar computations are also made

for the neural network based scheme, and SVM based scheme without parameter optimization. Figures 4(a)-4(c) show the different confusion matrices for training sample of size 600.

6. CONCLUSION

This paper presents a binary classification scheme using SVM with PSO based parameter optimization for investment class rating of Indian companies. The model network is initially trained with twelve financial ratios as attributes related to investment companies' data as input and classify whether it is a good investment or bad investment as output.

Further, a PSO based optimization technique is applied to optimize the value of C and σ for enhanced performance. From different simulation results it is observed that the proposed scheme outperforms in terms of classification accuracy, sensitivity, specificity, precision, confusion matrix etc. as compared to traditional SVM without parameter optimization and neural network scheme. There exist further scope of research in terms of introducing multiple kernels in SVM and dimensionally reduction method in SVM hopefully to obtain better result and extend the same for multiple classification.

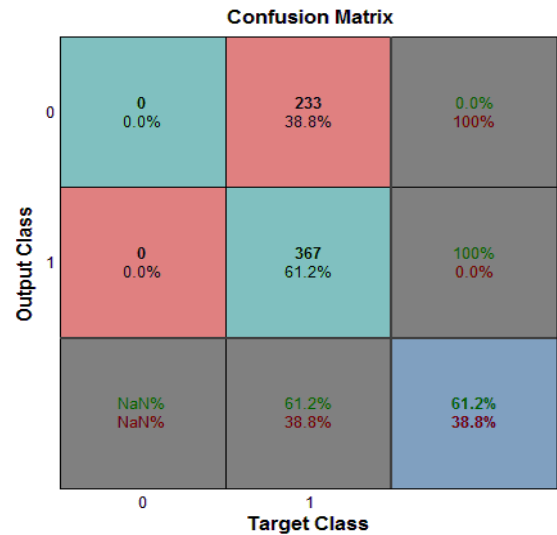


Figure 4(a): Confusion matrix for neural network based scheme

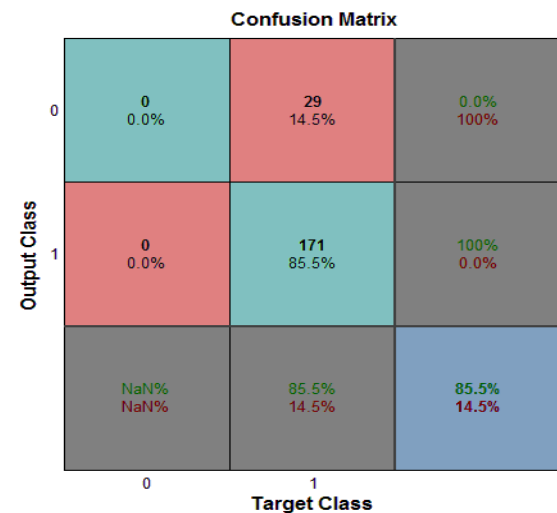


Figure 4(b): Confusion matrix for SVM without parameter optimization based scheme

Confusion Matrix

Output Class	0	1	2
	0	1	2
	0	1	2
0	0 0.0%	27 9.3%	0.0% 100%
1	0 0.0%	264 90.7%	100% 0.0%
	NaN% NaN%	90.7% 9.3%	90.7% 9.3%
	0	1	
	Target Class		

Figure 4 (c): Confusion matrix for SVM with parameter optimization based scheme

7. REFERENCES

- [1] Jackson JD, Boyd JW (1988) A statistical approach to modelling the behaviour of bond raters. *J Behav Econ* 17:173–193
- [2] Kamstra M, Kennedy P, Suan TK (2001) Combining bond rating forecasts using logic and Finance Rev 37:75–96
- [3] Kaplan RS, Urwitz G (1979) Statistical models of bond ratings: a methodological inquiry. *J Bus* 52:231–261
- [4] Brennan D, Brabazon A (2004) corporate bond rating using neural networks. In: *Proc of the conf on artificial intelligence*, Las Vegas, pp 161–167
- [5] Huang Z, Chen H (2004) Credit rating analysis with support vector machines and neural networks: A market comparative study. *Decision Support System* 37:543–558
- [6] Ammar S, Duncombe W, Hou Y (2001) Using fuzzy rule-based systems to evaluate overall financial performance of governments: An enhancement to the bond rating process. *Public Budget Finance* 21:91–110
- [7] Brabazon A, O'Neill M (2006) Credit classification using grammatical evolution. *Inform* 30:325–335
- [8] Delahunty A, OCallaghan D (2004) artificial immune systems for the prediction of corporate failure and classification of corporate bond ratings. University College Dublin, Dublin
- [9] Kim KS, Han I (2001) The cluster-indexing method for case based reasoning using self-organizing maps and learning vector quantization for bond rating cases. *Exp System Application* 21:147–156
- [10] Garavaglia S (1991) An application of a counter-propagation neural networks: Simulating the Standard & Poor's corporate bond rating system. In: *Proc. of the 1st int conf on artificial intelligence on Wall Street*, pp 278–287
- [11] Lee YCh (2007) Application of support vector machines to corporate credit rating prediction. *Exp Syst Appl* 33:67–74
- [12] Abe S (2005) Support vector machines for pattern classification, Springer-Verlag, London
- [13] Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, Cambridge
- [14] Haykin S (1999) Neural networks: a comprehensive foundation. Prentice-Hall, New Jersey
- [15] Jan Henning Trustorff, Paul Markus Konrad, Jens Leker (2010), Credit risk prediction using support vector machines. *Rev Quant Finance Acc* 36:565-581
- [16] Peter Hajek, Vladimir Olej (2011), Credit rating modeling by kernel-based approaches with supervised and semi-supervised learning. *Neural Computing & Application* 20:761-773
- [17] Tony Van Gestel , Bart Baesens , Dr. Ir , Joao Garcia , Peter Van Dijke (2003), A support vector machine approach to credit scoring, *journal of machine learning*, Springer, 4(3)
- [18] Shom Prasad Das, Sudarsan Padhy (2012), Support vector machines for prediction of futures prices in Indian stock market, *International journal of computer application* 41(3)
- [19] Vapnik, V.N., "The Nature of Statistical Learning Theory", Springer, 2nd edition, 1999
- [20] Shawe-Taylor, J., and Cristianini, N. "Kernel Methods for Pattern Analysis" , Cambridge UP, 2004.
- [21] Hung Wua, Gwo-Hshiung Tzeng, and Rong-Ho Lin. A novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression. *Expert Systems with Applications*, 36:4725-4735, 2009.
- [22] A GA based feature selection and parameters optimization for SVM, *Expert system with applications*, 31:231-240, 2006.
- [23] J. Kennedy and R. C. Eberhart. Particle Swarm Optimisation. In *Proc. IEEE International conf. on neural networks*, volume IV, pages 265 – 270, November
- [24] Kennedy and R. C. Eberhart. A discrete binary version of the particle swarm algorithm. In *IEEE International Conference on Computational Cybernetics and simulation* volume 5, pages 4104 –4108, October 1997
- [25] Huiyuan Fan. A modification to particle swarm optimisation algorithm *Engineering Computations*, 19(8):970–989, 2002