

Performance Comparison of SVM and ANN in Predicting Compressive Strength of Concrete

Kabiru O. Akande¹, Taoreed O. Owolabi², Ssennoga Twaha¹
Sunday O. Olatunji³

¹(Electrical Engineering, King Fahd University of Petroleum and Minerals, Saudi Arabia)

²(Physics, King Fahd University of Petroleum and Minerals, Saudi Arabia)

³(Computer Science, University of Dammam, Saudi Arabia)

Abstract: Concrete compressive strength prediction is very important in structure and building design, particularly in specifying the quality and measuring performance of concrete as well as determination of its mix proportion. The conventional method of determining the strength of concrete is complicated and time consuming hence artificial neural network (ANN) is widely proposed in lieu of this method. However, ANN is an unstable predictor due to the presence of local minima in its optimization objective. Hence, in this paper we have studied the performance of support vector machine (SVM), a stable and robust learning algorithm, in concrete strength prediction and compare the result to that of ANN. It is found that SVM displayed a slightly better performance compared to ANN and is highly stable.

Keywords: Artificial Neural Network, Compressive Strength of Concrete, Mix Proportion, Robust Learning Algorithm, Support Vector Machine

I. Introduction

Concrete is widely used in construction all over the world. It is produced from different components such as water, cement and so on. For quality control and as a measure of performance, 28-day compressive strength test is mandatory for design and construction code requirements [1]. The test is based on standard compression which is very complicated, time consuming and subject to experimental error. More importantly, if the concrete fails the test after such a long wait, there might be no way to effect an adjustment since it is already placed. Therefore, estimation of compressive strength well before 28 day is highly desired and widely practiced since it offers many advantages. It allows advanced scheduling of operations such as prestressing and formwork removal resulting in improved efficiency. It also enhances quality control as structures become more durable due to avoidance of excessive loading at early age. Furthermore, searching for optimal concrete mix proportion is possible using the concrete constituents as prediction data. Hence, many methods exist to predict and estimate concrete strength well before 28 days. Prominent among them is the Abrahams Law which is a linear equation relating concrete strength and its water to cement ratio [2]. In reality, the relationship between concrete strength and its constituents material is highly nonlinear and as such relation as Abrahams law fail to adequately represent this complex relationship and fall short in generalizing to previously unseen data [3].

However, artificial intelligence has proven itself to be a versatile tool in learning complex patterns and is widely used in various fields for classification and regression. It is used in medical field [4], [5], in Oil and gas [6], [7],[8] and financial time-series forecasting [9], design and manufacturing [10], marketing [11], and image recognition [12]. Hence, a widely proposed method is the use of artificial neural network (ANN) in concrete strength prediction and the most widely deployed ANN network is the back propagation network [13]. However, back propagation networks suffered from local minima resulting in instability in the developed model. The convergence to local minima results from the optimization objective of ANN which is multimodal in nature. This instability manifest itself in the fact that different sets of training data produce different models impacting on the ability of the developed model in generalizing to previously unseen data. This limitation of ANN has been widely reported in the literature especially as regards medical prediction and classification [14]. Although this can be overcome by aggregating the outputs of several models developed from the training data a procedure referred to as ‘bagging ,stacking or boosting’ predictors, this result in more computations and a level of uncertainty[15], [16], [17].

However, SVM is a very efficient and stable algorithm which is now being applied in many fields [18], [19]. It exhibit good generalization performance in real-world scenarios where it often outperform many other techniques used in predictions and classifications. Its sound and solid mathematical foundation ensured its adaptation and optimization to varying data analysis problems. In this study, the use of support vector regression (SVR) in concrete strength prediction is investigated and the result compared to that of ANN regression model. It is found that SVM outperform ANN with regards to its accuracy and stability and therefore it has proven to be a viable alternative for concrete strength prediction especially in cases where there is little margin for error as is

often the case in real-life predictions. Errors in construction projects resulting from faulty testing could be fatal leading to building collapses and deaths hence, a stable, robust and accurate method of testing is highly desired. This paper is organized as follows: Section II detailed brief description of SVR. Section III gives the description of the system setup and dataset used. Section IV shows the result of the experiment and discussion. Section V concludes the paper.

II. Support Vector Machine

SVM is a statistical learning algorithm which can be applied to both classification and regression problems and is proposed by Vapnik [20]. In this section, we briefly detailed SVM theory as applied to regression since we aim to predict compressive strength of concrete. This type of SVM is referred to as SV regression. The basic idea behind SVM is that it takes a set of input data \mathbf{x} , transform onto an high-dimensional feature space using a non-linear mapping function, $\varphi(\mathbf{x})$ [21]. Linear regression is then performed in this feature space by constructing a linear model as given below

$$g(\mathbf{x}, \mathbf{w}) = \sum_{k=1}^m w_k \varphi_k + b \dots\dots\dots (1)$$

Where w_k 's are the adjustable parameters of the model and b is a "bias" term which can be eliminated if \mathbf{x} has zero mean. A zero-mean input data can usually be achieved by data preprocessing. φ_k effect non-linear transformation of input data onto the feature space.

A measure of SVM estimation accuracy is the epsilon (ϵ) -insensitive loss function proposed by Vapnik and is given by

$$L_\epsilon(y, g(\mathbf{x}, \mathbf{w})) = \begin{cases} 0, & \text{if } |y - g(\mathbf{x}, \mathbf{w})| \leq \epsilon \\ |y - g(\mathbf{x}, \mathbf{w})| - \epsilon, & \text{otherwise} \end{cases} \dots\dots\dots (2)$$

The constant ϵ is a user-defined parameter stating the margin of error to be tolerated. Theoretical bounds is placed on SVM performance by defining a function called empirical risk based on the loss function described above as

$$E_{emp}(\mathbf{w}) = \frac{1}{m} \sum_{k=1}^m L_\epsilon(y_k, g(\mathbf{x}_k, \mathbf{w})) \dots\dots\dots (3)$$

While ANN seek to control model complexity by limiting the number of features, SVM does this by minimizing the Euclidean norm of the model parameters($\|\mathbf{w}\|$). This is achieved by employing non-negative slack variables. The norm operator finds the length of an object or vector which intuitively means that the minimization objective of SV regression is to reduce the length of the model parameters. In general, the lower the value of the model parameters, the better the model generalization accuracy. This principle is applied in ANN regularization technique where the sum of the model parameters is penalized to address over-fitting problem. Hence, SVM is tailored towards achieving good generalization performance.

SVM generalization performance depends a great deal on selecting optimum combination of parameter C, error-insensitive zone ϵ and kernel specific parameters. The parameter C is a regularization factor used to control the tradeoff between model complexity and SVR errors on training data with smaller C leading to more training errors and larger C leading to hard-margin SVR. The value of ϵ controls the width of the insensitive zone affecting directly the number of support vectors. Bigger value will usually lead to more support vectors. Kernel parameters control the transformation of input data to a high-dimensional features space where linear regression can be carried out. The use of kernel is to enable efficient computation for SVR when constructing non-linear decision surfaces.

As mentioned previously, SVM maintained a number of advantages in comparison with ANN and the following briefly summarize these points. First is the fact that SVM was developed from well-grounded and sound theory before its implementation and application while ANN followed from years of applications and experience known as heuristic approach. While heuristic approach might yield an acceptable solution to a problem it is not necessarily an optimum one.

Similarly, the optimization problem of SVM amount to solving a linearly constrained quadratic programming to which exist an optimal and global solution in contrast to what is obtained in ANN where the optimization problem is multimodal leading to existence of local minima and the resulting non-unique solutions. Furthermore, SVM make use of structural minimization principle which is known to have good generalization accuracy as opposed to empirical risk minimization implemented by ANN which is known to be suboptimal for non-convex functions and is only able to achieve optimality with aggregation procedures. Also, speed and convergence are highly desired for practical reasons and considerations. SVM employ the use of search techniques which eliminate kernel evaluations yielding negligible contribution to the decision function output thus, leading to computational efficiency.

Finally is the case where there is dearth of training data, SVM maintains its good generalization accuracy in this case while ANN tends to over-fit and its generalization performance suffered as a result of that. As was mentioned by [22], increase in generalization performance of ANN is tied to increase in the amount of input training data which may be a limitation in many practical situations. Also, [9] showed that SVM generally has better accuracy than ANN as the data size gets smaller and that it become impossible for ANN to achieve good generalization performance at some small data size where SVM still perform well.

III. Empirical Study

3.1 Data Set Description

The data sets used in this paper is that retrieved from [23]. We have separated those results corresponding to 28 day testing and used them in our analysis. The reason for this is that building and design code requirements for concrete are specified in terms of 28-day compressive strength and hence this is what we trained our model to predict. The total data sets is 425 (m = 425). The components making up this data set and their statistical analysis obtained from excel spreadsheet are given in table 3.1.

3.2 Description of the experiment

All the programming tasks relating to this work were carried out using MATLAB computing environment. The data sets corresponding to 28-day compressive strength were first separated from the original data retrieved from [23]. The resulting data is then normalized in order to prevent components with large attributes from biasing the resulting model. A random shuffling of normalized data is further carried out to separate the data into training and testing sets in ratio 8:2 respectively. Data randomization was done to ensure that all parts of the data were fairly represented in the training/testing partition. We proceed to training ANN and SVM regression model using the training set and validate the trained model with the testing data set. The use of testing set for validation ensures that we are able to accurately measure the generalization performance of our trained model.

3.3 System setup

ANN is set up using feedforward network and has been trained with back propagation algorithms. It is well known that a feedforward network with one hidden layer and enough neurons in the hidden layer is capable of fitting any input-output data [3]. Hence, in this network we have used one hidden layer and vary the number of neurons to obtain the optimum performance. SV regression was set up by varying the major parameters in order to obtain optimum outputs. Both (Gaussian) radial basis function and polynomial function were used in combination with varying values of C, epsilon and kernel options during the parameter search procedures.

Table 3.1. Statistics analysis of dataset

Variable	Mean	Maximum	Minimum	Standard Deviation
Cement (kg/m3)	86.29	359.40	0.00	87.72
Blast Furnace Slag (kg/m3)	62.79	200.10	0.00	66.15
Fly Ash (kg/m3)	183.06	247.00	121.75	19.31
Water (kg/m3)	6.99	32.20	0.00	5.39
Superplasticizer (kg/m3)	956.06	1145.00	801.00	83.70
Coarse Aggregate (kg/m3)	764.38	992.60	594.00	73.03
Fine Aggregate (kg/m3)	36.75	81.75	8.54	14.69
28-day Concrete compressive strength (MPa)	265.44	540.00	102.00	104.55

3.4 Performance Measure

The generalization performance of the trained model was accessed using three different criteria. The first is coefficient of correlation (cc) defined as

$$cc = 1 - \left[\frac{\sum_1^n \left\{ \frac{S_{exp} - S_{pred}}{S_{exp}^2} \right\}^2}{n} \right] \dots \dots \dots (4)$$

For very good performance, the value of CC should be approaching 1.

Second is the root mean square error (rmse) criteria which is widely used and defined as

$$rmse = \sqrt{\frac{1}{n} \left[\sum_1^n \left\{ \frac{S_{exp} - S_{pred}}{S_{exp}^2} \right\}^2 \right]} \dots \dots \dots (5)$$

The lower the value of rmse, the better the model generalization performance. And finally we take the absolute error given as

$$Ea = \sum_1^n \left\{ \frac{S_{exp} - S_{pred}}{n} \right\}^2 \dots\dots\dots (6)$$

Also, lower value of absolute error is indicative of good performance. In all of the above equations, S_{exp} is the measured value of the concrete strength while S_{pred} is the predicted value by our trained model. n is the size of the dataset.

IV. Result and Discussion

Optimum parameters were obtained for both ANN and SVM regression through optimum search approach. Although there are some proposed methods of parameter selection in the literature [24], trial and error approach is still widely practiced and guarantee good result. The optimum parameters for both SVM and ANN are shown in Table 4.1 based on the parameter search procedure implemented in this work.

These optimum parameters correspond to the best generalization performance for the trained model as measured by the performance criteria previously defined. Kernel option parameter had the most profound effect on SVM performance and is shown in Fig 1. It is seen that the generalization performance of SVM model increased steadily in proportion to increase in kernel option till a value of 85 which correspond to the best CC. Beyond the value of 85, the performance then become constant indicating that increase in value of kernel option beyond this point no longer has significant effect on SVM performance.

Table 4.1. Optimum parameters for the trained models

SVM		ANN	
Parameter	Value	Parameter	Value
C	300	Hidden Layer	1
Lamda	1.00E-07	Hidden Neuron	15
Epsilon	0.2	Training function	trainlm
Kernel Option	85		
Kernel	Gaussian		

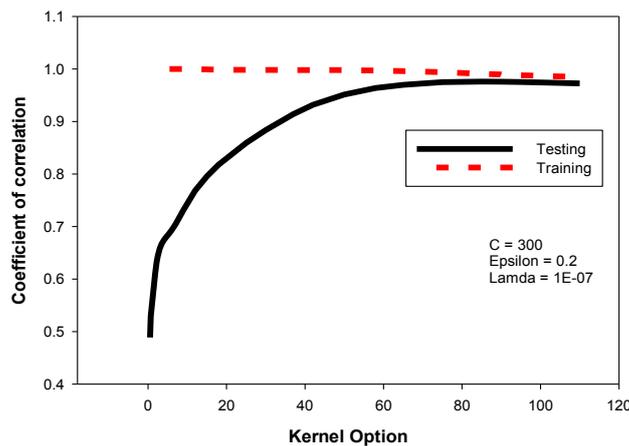


Figure 1: Effect of kernel option on SVM model performance

The results obtained using the optimum parameters are shown in Table 4.2

Table 4.2. Best results for the trained models

Model	Training			Testing		
	CC	RMSE	EA	CC	RMSE	EA
SVM	0.9920	14.31	2.19	0.9773	23.14	4.89
ANN	0.9796	23.13	6.6	0.9671	27.15	8.03

It is seen that the SVM outperform ANN with CC of 99.20% and 97.73% for training and testing respectively while ANN produce corresponding values of 97.96% and 96.71%. This shows that SVM has good generalization performance in predicting the compressive strength of concrete and is slightly better compared to ANN. Furthermore, SVM is 38% better than ANN in terms of root mean square error for training data and 15% better for testing.

Fig 2 shows the correlation between ANN predicted output and the actual measured value of the concrete strength. It is seen that it achieve good performance and as we have already shown above, the best result is CC of 96.71%.

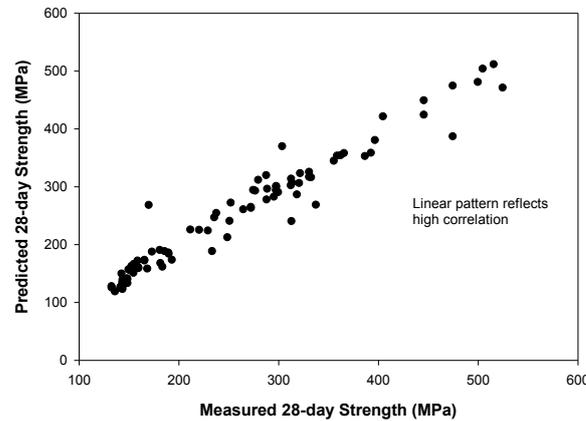


Figure 2. Correlation of ANN predicted concrete strength to measured concrete strength for testing data

Fig 3 shows the graphical representation of the generalization accuracy of trained SVM model in concrete strength prediction. Specifically it shows the correlation between SVM predicted strength and the measured values. It is seen that the values lies closer together than those of ANN which is indicative of better correlation and better predictive ability. We have already shown above that SVM generalize better than ANN with a CC of 97.73% as against that of ANN that stood at 96.71%.

The generalization performance of both ANN and SVM trained model is further shown in Fig. 4 and 5 respectively. The figures reflected the ability of the models to adequately follow the measured concrete strength which is the target of our prediction. It is seen that SVM is able to match ANN performance and showed good tracking ability depicting its potential to predict correctly and accurately.

Finally, the optimum parameters obtained above, as discussed earlier, were used to carry out 10 different experiments using the same dataset partitioning and size. It was observed that while SVM maintained its stability producing same result for all the experiments, ANN had different result for each of the experiments which is testament to its instability. The statistics of the results from the experiments are given in Table 4.3. It is seen that SVM has standard deviation of 0.000 for both training and testing data while ANN has 0.0033 and 0.014 for training and testing respectively. Thus, SVM distinguished itself as a very stable and viable concrete strength prediction tool.

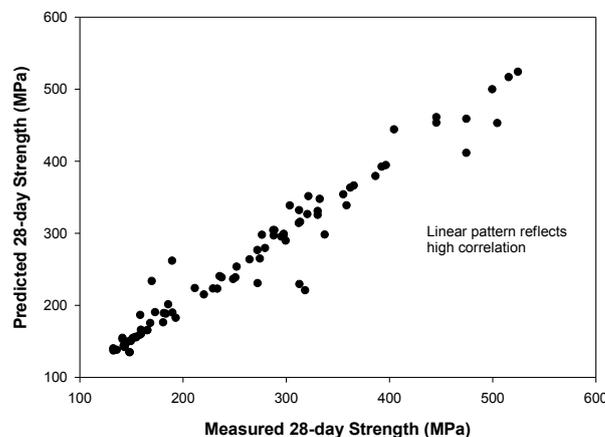


Figure 3. Correlation of SVM predicted concrete strength to measured concrete strength for testing data

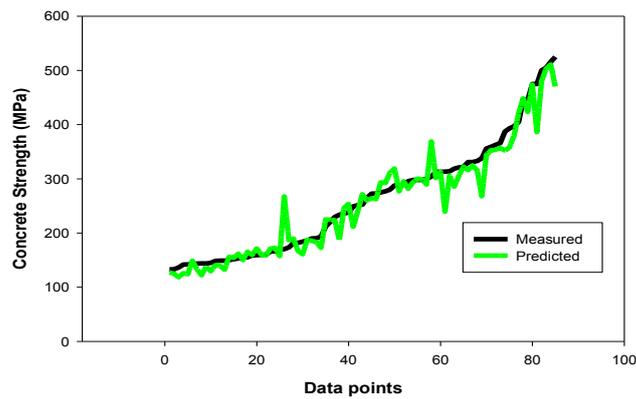


Figure 4. Tracking of measured concrete strength by ANN (Testing data)

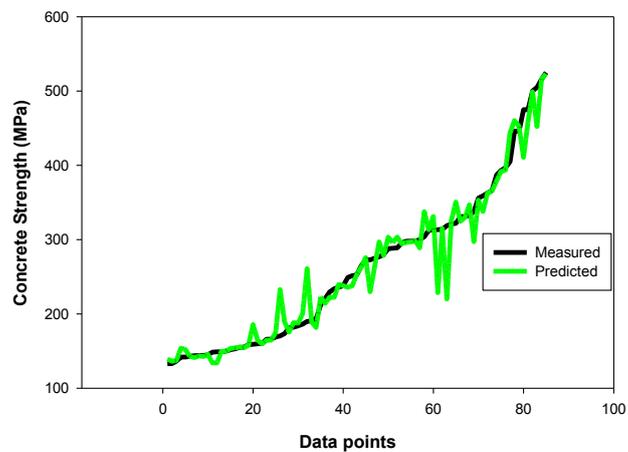


Figure 5. Tracking of measured concrete strength by SVM (Testing data)

Table 4.3. Statistics of CC for trained models (10 iterations)

	Training		Testing	
	SVM	ANN	SVM	ANN
Minimum	0.9920	0.9749	0.9773	0.9254
Maximum	0.9920	0.9854	0.9773	0.9770
Average	0.9920	0.9806	0.9773	0.9675
Standard deviation	0.0000	0.0033	0.0000	0.0144

V. Conclusion and recommendation

Compressive strength of concrete has a highly non-linear dependence on the constituents of the concrete which make its modelling a very difficult task. Hence, ANN is widely used in predicting 28-day concrete strength from its constituents in order to indicate its overall quality. Support vector machine, a highly promising alternative, has been studied in this work for predicting concrete strength and it achieved a better performance in addition to its stability as compared to ANN in concrete strength prediction. It is further shown that ANN is an unstable predictor producing differing results at different iterations for the same optimum parameters, dataset partitions and data size while SVM maintain its stability. It can therefore be concluded, based on the findings of this work, that SVM is a viable alternative to ANN in concrete strength prediction due to its stability and good performance. SVM is therefore recommended for predicting the compressive strength of concrete so as to improve the safety and design of buildings and constructions. Future research could be in the areas of using SVM in the prediction of other properties of materials.

References

- [1] A. M. Neville, *Properties of Concrete*. 2012, p. 872.
- [2] D. A. Abrams, "Water-Cement Ratio as a Basis of Concrete Quality," *ACI*, vol. 23, no. 2, pp. 452–457.
- [3] "COMPARISON OF CONCRETE STRENGTH PREDICTION TECHNIQUES WITH ARTIFICIAL NEURAL NETWORK," vol. 56, no. 1, pp. 23–36, 2008.
- [4] S. Olatunji and H. Arif, "IDENTIFICATION OF ERYTHEMATO-SQUAMOUS SKIN DISEASES USING EXTREME LEARNING MACHINE AND ARTIFICIAL NEURAL NETWORK," *ictactjournals.in*, vol. 6956, no. October, 2013.
- [5] G. D. Magoulas, A. Prentza, N. Technical, and G.- Athens, "MACHINE LEARNING IN MEDICAL APPLICATIONS George D. Magoulas 1 and Andriana Prentza 2 1," vol. 1999, 1999.
- [6] S. O. Olatunji, A. Selamat, and A. Abdulraheem, "A hybrid model through the fusion of type-2 fuzzy logic systems and extreme learning machines for modelling permeability prediction," *Inf. Fusion*, vol. 16, pp. 29–45, Mar. 2014.
- [7] S. O. Olatunji, A. Selamat, and A. A. Abdul Raheem, "Improved sensitivity based linear learning method for permeability prediction of carbonate reservoir using interval type-2 fuzzy logic system," *Appl. Soft Comput.*, vol. 14, pp. 144–155, Jan. 2014.
- [8] S. O. Olatunji, A. Selamat, and A. A. A. Raheem, "Predicting correlations properties of crude oil systems using type-2 fuzzy logic systems," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 10911–10922, Sep. 2011.
- [9] K.-S. Shin, T. S. Lee, and H. Kim, "An application of support vector machines in bankruptcy prediction model," *Expert Syst. Appl.*, vol. 28, no. 1, pp. 127–135, Jan. 2005.
- [10] D. Stoneking, "Improving the manufacturability of electronic designs," *IEEE Spectr.*, vol. 36, no. 6, pp. 70–76, Jun. 1999.
- [11] S. Ben-David and M. Lindenbaum, "Learning Distributions by Their Density Levels: A Paradigm for Learning without a Teacher," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 171–182, Aug. 1997.
- [12] E. Osuna, R. Freund, and F. Girosit, "Training support vector machines: an application to face detection," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 130–136.
- [13] S. Lai and M. Serra, "Concrete strength prediction by means of neural network," *Constr. Build. Mater.*, vol. 11, no. 2, pp. 93–98, 1997.
- [14] A. Cunningham, J. Carney, and S. Jacob, "Stability problems with artificial neural networks and the ensemble solution," vol. 20, pp. 217–225, 2000.
- [15] Y. Freund, R. E. Schapire, and M. Hill, "Experiments with a New Boosting Algorithm," in *International Conference on Machine Learning*, 1996, pp. 148–156.
- [16] D. H. Wolpert, "Stacked generalization," *Neural Networks*, vol. 5, no. 2, pp. 241–259, Jan. 1992.
- [17] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.
- [18] J. M. Moguerza and A. Muñoz, "Support Vector Machines with Applications," *Stat. Sci.*, vol. 21, no. 3, pp. 322–336, Aug. 2006.
- [19] S. M. Gupta, "Support Vector Machines based Modelling of Concrete Strength," pp. 305–311, 2007.
- [20] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 297, pp. 273–297, 1995.
- [21] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory - COLT '92*, 1992, pp. 144–152.
- [22] S. Haykin, "Neural Networks: A Comprehensive Foundation," Oct. 1994.
- [23] I.-C. Yeh, "Modeling of strength of high-performance concrete using artificial neural networks," *Cem. Concr. Res.*, vol. 28, no. 12, pp. 1797–1808, Dec. 1998.
- [24] Y. M. Vladimir Cherkassky, "Practical selection of svm parameters and noise estimation for svm regression," *Neural Networks*, vol. 17, pp. 113–126, 2004.