2019 IEEE 28th Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS) 978-1-7281-4585-3/20/$31.00 ©2020 IEEE 10.1109/EPEPS47316.2019.193199

Comparison of Machine Learning Techniques for Predictive Modeling of High-Speed Links

Hanzhi Ma[1], Er-Ping Li[1], Andreas C. Cangellaris[2], and Xu Chen[2]

[1]College of Information Sci. and Electronic Eng., ZJU-UIUC Institute, Zhejiang University, Hangzhou, China
Email: {mahanzhi, liep}@zju.edu.cn
[2]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL
Email: {cangella, xuchen1}@illinois.edu

***Abstract*—We compare three different machine learning techniques for constructing predictive model for eye opening based on channel length and interconnect cross-sectional geometry. Surrogate model is constructed using sparse grids, support vector regression, and artificial neural networks. Models for training data are generated using quasi-TEM modeling of the interconnect, and eye opening training data is obtained from statistical high-speed link simulation using IBIS-AMI transmitter and receiver models. Numerical results illustrate that all three methods offer reasonable predictions of eye height, eye width and eye width at 10^{-12} bit error rate.**

I. Introduction

In modern computer system development, exhaustive post-physical design board-level analysis is not practical due to CPU time and memory constraints. Presently, board designers tend to make design decisions based on worst-case scenarios, which often lead to over-designing. Ideally, every net should be modeled at the board level, including their frequency responses and effects from nearby aggressors. The models can be used to perform eye diagram simulations, and the electrical performance of the nets are then ranked from best to worst. Thus, creating a machine-learning driven framework for fast, accurate and modular board-level signal integrity analysis to determine eye opening of each net in an electronic system is very desirable. Machine learning techniques have shown promise to quickly assess the performance of high-speed links [1], [2]. This paper studies three different kinds of machine learning techniques, namely sparse grid [3], support vector regression (SVR) [4] and artificial neural networks (ANN) for eye opening prediction of high-speed links using IBIS-AMI behavioral I/O models, including effects of equalization. We examine the performance of the three methods via their application to a differential microstrip channel exhibiting variability in its geometric parameters. While all methods are effective in accurate eye prediction, SVR is found to exhibit best accuracy for the specific study.

II. Problem Statement

A. *Differential Microstrip Line Channel*

For the purposes of this paper, a differential microstrip line channel is considered of cross sectional geometry depicted in Figure 1. Table I provides information about the four geometric parameters of the channel, strip width, w, trace spacing, s, substrate thickness, h, and channel length, l. These parameters are uniformly distributed over the interval.

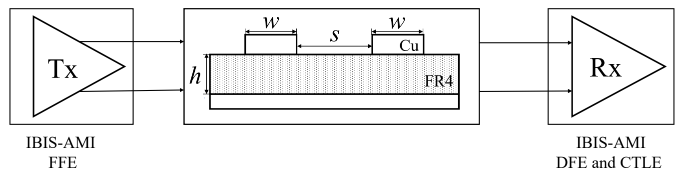

Figure 1. High-speed link simulation topology for eye diagram.

Table I
Design parameter range of values

Parameter	w	s	h	l
Value (mm)	$[0.3, 0.5]$	$[0.5, 0.7]$	$[0.2, 0.4]$	$[50, 100]$

B. *Workflow of Eye Diagram Prediction*

Eye diagram of differential microstrip line channel can be obtained by joint simulation of ANSYS Q3D Extractor and Keysight ADS. ANSYS Q3D Extractor is used to calculate the scattering parameters (S-parameters) of the channel. These, then, are used as input into the Keysight ADS simulator to generate the eye diagram. However, traditional simulation techniques employed for eye analysis for a large amount of parameters which lead to undesirable computation burdens. As suggested by the flow in Figure 2, machine learning methods can be used in place of direct simulation. In other words, once successfully trained, the machine learning method can be used to predict eye diagram directly from the value of uncertainty parameters.

III. Machine Learning Techniques

A. *Sparse Grids*

Sparse Grids [3] is a family of algorithms for constructing multidimensional quadrature and interpolation rules, where the approximation operator is constructed as a linear combination of tensors of multiple one dimensional operators. For predictive modeling we use interpolation, which requires generating training data at a set of prescribed nodes. For

Hanzhi Ma conducted this research while she was a visiting scholar at the University of Illinois at Urbana-Champaign.

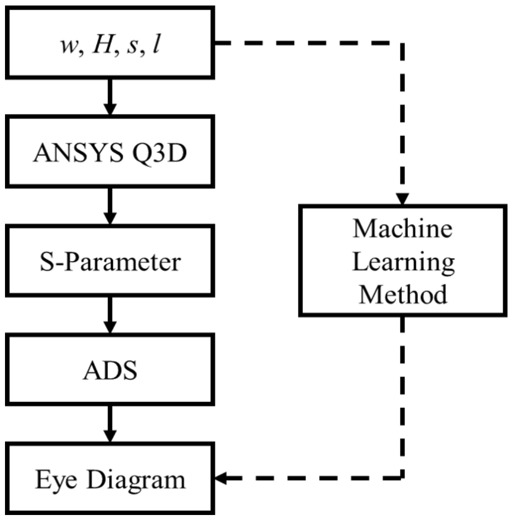

Figure 2. Workflow of eye diagram prediction

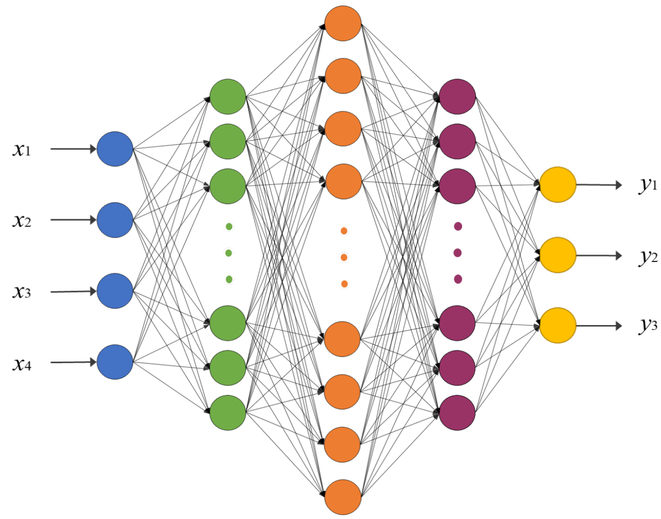

Figure 3. Architecture of ANN

the four-dimensional parameter space problem, 721 nodes are needed for level 7 interpolation, and 1041 nodes for level 8 interpolation. We performed simulation at the nodes in random space, and also re-used the same data set for training SVR and ANN methods.

B. Support Vector Regression

SVR [4], an important application branch of support vector machine, is widely used for regression problems with the characteristics of nonlinear mapping and robustness. For data sample (x, y), instead of using the direct difference between output $f(x)$ of regression model and y to define loss function, SVR allows a margin of tolerance ϵ. In other words, part of the error can be tolerated when the absolute difference between $f(x)$ and y is smaller than ϵ in order to find a suitable hyperplane for minimizing total error. SVR can be expressed as

$$f(x) = \sum_{i=1}^{m} (\hat{a}_i - a_i)\kappa(x, x_i) + b, \tag{1}$$

where $\hat{a}_i \geq 0$ and $a_i \geq 0$ are Lagrangian operators, and b is the displacement of hyperplane. This model is treated as a linear combination of kernel function $\kappa(x, x_i)$. Several different kernel functions can be applied for regression problems. We select Gaussian kernel for this problem, which can be expressed as

$$\kappa(x, x_i) = \exp\left(\frac{-\|x - x_i\|^2}{2\sigma^2}\right), \tag{2}$$

where $\sigma > 0$ is the width of Gaussian kernel.

C. Artificial Neural Network

ANN is an algorithm that mimics the design of human brain neurons, which performs well in regression problems due to the strong abilities of autonomous learning, feature extracting and generalization. Figure 3 illustrates the main architecture of the ANN proposed in this paper, which consists of one input layer, three hidden layers and one output layer. Uncertain design parameters w, H, s, l (labeled as x_1, x_2, x_3 and x_4 in the figure) are fed into this ANN model to predict three characteristics of the eye diagram. There are 10, 100 and 10 neural cells in the first, second and third hidden layer, respectively. Rectified Linear Unit (ReLU) [5] is used as the activation function. The loss function calculates the mean square error between outputs and expected outputs, while stochastic gradient descent (SGD) [6] is adopted as the optimization method with learning rate 0.05.

IV. Results and Comparisons

We consider three predicted eye diagram measurements. "Height" is the distance between the 3σ points of the logic-1 and logic-0 histograms, measured across the eye level boundary. "Width" is the distance between the 3-sigma points of the crossing time histograms. "WidthAtBER" is the maximum width of contour at 10^{-12} bit error rate (BER).

A. Training Data Generation

We use 721 sets of samples of w, H, s, l as training data sets and choose another 320 sets of samples for test. These samples are generated by sparse grids at level 7 and level 8. Eye diagram characteristics (Height, Width and WidthAtBER) of each sample are simulated by ANSYS Q3D Extractor and Keysight ADS.

B. Prediction Results

As mentioned above, sparse grids, SVR and ANN are applied to predict the performance of the channel model. Figures 4, 5, and 6 offer a comparison of the three methods on the basis of their error in predicting Height, Width, and WidthAtBER, respectively. The prediction error is defined as follows:

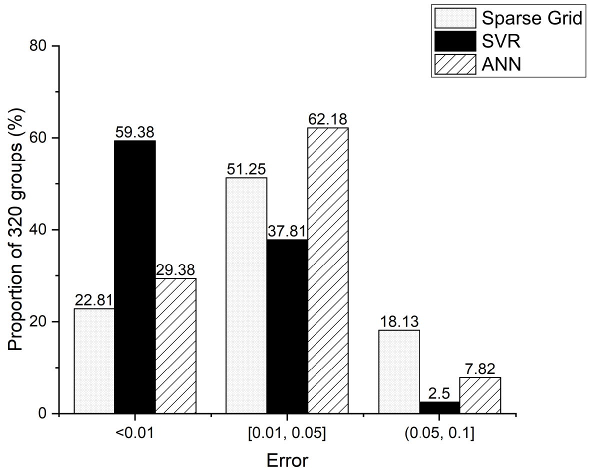

Figure 4. Prediction Error of Height

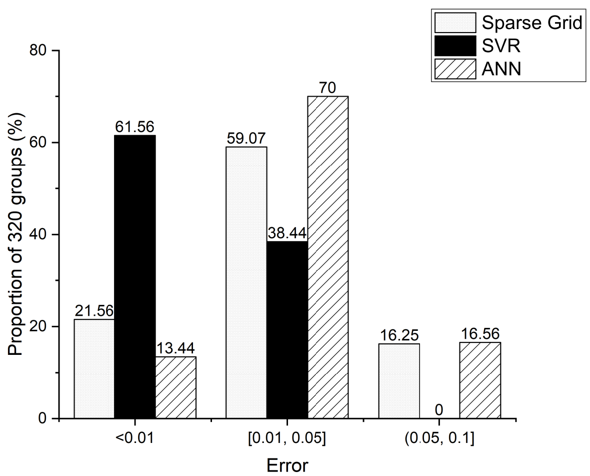

Figure 5. Prediction Error of Width

$$Error = \frac{|y_{prediction} - y_{simulation}|}{y_{simulation}}. \quad (3)$$

The numerical results in these figures show that sparse grid, SVR and ANN can be used for eye diagram prediction with good performance. All these three methods can almost keep every prediction error less than 10%. For this example, SVR has the best performance in terms of prediction error. It is worth noting that ANN model can be further improved for better prediction results after adding more hidden layers or doing other topology optimization.

V. Conclusion and Future Work

In this paper, we examined the performance of three different machine learning methods, sparse grid, SVR and ANN, in predicting the eye diagram of differential microstrip line channel including equalization. All three methods are shown

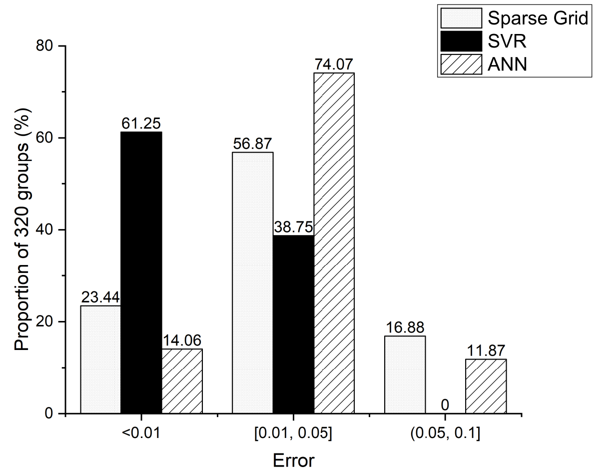

Figure 6. Prediction Error of WidthAtBER

to exhibit acceptable accuracy in predicting eye height, eye width, and eye width at 10^{-12} BER in the presence of variability in strip width, strip spacing, substrate height and channel length. SVR demonstrates its strong learning and prediction ability for the specific study and can be treated as a promising method for more complex channel model prediction. Sparse grid and SVR require all data available during training, while ANN is more flexible since it can stream data as they become available. An important feature of any machine learning method we choose will be its scalability with respect to the number of variable design parameters. It will be important to investigate how each of these methods work with problems with larger number of design parameters exhibiting variability.

Acknowledgment

This work was supported by the National Science Foundation under Grant No. CNS 16-24811 - Center for Advanced Electronics through Machine Learning (CAEML) and its industry members.

References

[1] R. Trinchero, M. Larbi, H. M. Torun, F. G. Canavero, and M. Swaminathan, "Machine learning and uncertainty quantification for surrogate models of integrated devices with a large number of parameters," *IEEE Access*, vol. 7, pp. 4056–4066, 2019.

[2] T. Lu, J. Sun, K. Wu, and Z. Yang, "High-speed channel modeling with machine learning methods for signal integrity analysis," *IEEE Transactions on Electromagnetic Compatibility*, vol. 60, pp. 1957–1964, Dec 2018.

[3] M. Stoyanov, "User manual: Tasmanian sparse grids," Tech. Rep. ORNL/TM-2015/596, Oak Ridge National Laboratory, One Bethel Valley Road, Oak Ridge, TN, 2015.

[4] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, pp. 199–222, Aug. 2004.

[5] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, (USA), pp. 807–814, Omnipress, 2010.

[6] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.