Review on Machine Learning for Analog Circuit Design

Nirali Hemant Patel
Student, Electronics and Communication
Charotar University of Science and Technology, CHARUSAT
Gujarat, India

Abstract— **In this review paper, different methods of Machine Learning are being discussed in this paper. Moreover, major parameters like power efficiency, delay and accuracy is being taken into consideration for verifying the result. The flow of the paper starts with the basic introduction to Analog circuits and Machine learning then different techniques of machine learning, afterwards the usage of machine learning in analog circuits is been discussed using different methods are being discussed and then the conclusion and references.**

Key-words: Analog circuit, Accuracy, delay, Machine learning, Power-Efficiency

I. INTRODUCTION

Machine Learning (ML) is one of the most demanding and popular buzzwords nowadays, but despite the fact that EDA deals with big-data amount of issues it's not made much progress incorporating ML techniques into EDA tools. For several machine learning applications in EDA, algorithm-related parameter selection and training must occur completely within a design customer or industries' computing environment.[1] The perfect interconnect strategy depends upon a really sizable amount of SoC parameters, including floorplan, routing constraints, resources available, connectivity requirements, protocol level dependency, clock characteristics, process characteristics like wire delay, power budgets, bandwidth and latency constraints, etc. the amount of unique dimensions within the design strategies space grows to many hundreds, creating an excessively large design space. Machine learning are generally implemented in EDA. By analyzing billions of knowledge points from previous outputs, we will predict the impact of bugs, design complexity, human resources, licenses, and compute farm throughput on current projects. By identifying bottlenecks in semiconductor designs we will provide forward prediction and identify potential delays. Machine learning can be used for reduction of time for the design and simulation through existing complex algorithms. [5] Machine learning tools are ready to analyze past compilation results to predict optimal synthesis/place-and-route parameters and placement locations out of million number of possible solutions. [3] They infer what tool parameters are best for a design using statistical modeling and machine learning to draw insights from the info to enhance quality of results. Machine Learning has already begun to play a serious role in EDA.

II. MACHINE LEARNING TECHNIQUES

Nowadays many misconceptions are there associated with the words device gaining knowledge of, deep mastering, and artificial intelligence(AI), most of the people suppose all this stuff are the same whenever they hear the phrase AI, they directly relate that phrase to device gaining knowledge of or vice versa. Machine Learning is a technique of parsing statistics, learn from that facts and then apply what they have found out to make the knowledgeable decision. The principle distinction among deep and machine getting to know is, gadget getting to know models become higher progressively but the model still wishes a few guidance. If a system mastering model returns a faulty prediction then the programmer needs to restore that problem explicitly however within the case of deep mastering, the version does it by himself. The automatic vehicle driving system is a great example of deep gaining knowledge of. "AI is a potential of a computer program to feature like a human brain". Machine gaining knowledge of and deep studying is a way of attaining AI, this means that by the use of system getting to know and deep learning we may additionally in a position to obtain AI in destiny however it isn't always AI.

There are various different techniques and plenty of algorithms. Some of techniques are: Regression**,** Classification**,** Clustering**,** Association**,** Dimension Reduction, Anomaly Detection, Sequence mining and Recommendation systems.

TABLE I. DIFFERENCE OF AL, ML AND DL

Artificial Intelligence	ML	Revolution of ML
Computer vision	Classification	Deep learning
Lang processing	Clustering	
creativity	Neutral network	

Along with different techniques, there are different algorithm like supervised and unsupervised.

a. Supervised algorithm:

It is used to observe and to direct the execution of task project or activity. It is done by teaching the model. So load the model by knowledge and we teach is labeled dataset. Moreover, the attributes will indicate ROWS and the features which includes the data indicate COLUMN. Basically, data is in two types: Numerical and characters or categorical. In supervised Learning, the two methods which is included are: classification and regression. Classification is the process of predicting discrete class labels or categories and Regression is the process of predicting continues values.

b. Unsupervised algorithm:

It is not necessary to supervise the model but to let the model work on its onto discover the information which may not be visible to human eye. It is the difficult algorithms. Different Techniques which include: Dimension reduction, Density estimation which is used to find some structure within it, Market basket analysis and Clustering. Whereas, Clustering is

a group of data points or objects that are somehow similar by: Discovering structure, Summarization, Anomaly detection.

TABLE II. SUPERVISED V/S UNSUPERVISED

Supervised	Unsupervised

Classification
    Deals with labeled data
	Clustering
	  Finds patterns and grouping from unlabeled data

Regression Predicts trends using previous labeled data	
Has more evaluation methods	Has fewer evaluation methods than supervised learning
Controlled environment	Less controlled environment

III.   REGRESSION TECHNIQUE

a.   Simple linear Regression

It is used to predict continuous value by using other variables and also used to describe relationship between two or more variables. In simple linear regression there are two variables: Dependent and Independent, where dependent value should be continuous and not discrete whereas independent variables can be measured on either categorical or continuous measurement scale.

TOPOLOGY: Simple Linear Regression is used when one independent variable is used to estimate one dependent variable.

Working:

$Y = \Theta_o + \Theta_1 X_1$

Y is the dependent variable

X1 is the independent variable

Θ_o and Θ_1 are the parameters that we must adjust.

Θ_o is known as the intersect and Θ_1 is known as the slope.

Θ_o and Θ_1 are also called the coefficient of linear equation. Linear regression estimates the coefficient of line. So Θ_o and Θ_1 must be calculated to find the best line to fit the data. To adjust the parameters to make the line the best fit for the data, error should be minimizing.

$$MSE = \left(\frac{1}{n}\right)\sum_{i=1}^{n}(y_i - y)^2 \qquad (1)$$

There are two approaches:

1)   Mathematical approach
2)   Optimization approach.

1)   Mathematical approach:

$$\theta 1 = \frac{\sum_{i=-1}^{S}(xi - xb)(yi - yb)}{\sum_{i=1}^{S}(xi - xb)^2} \qquad (2)$$

Where, Xb= mean of independent value and Yb= mean of dependent value and Θ_o = yb - Θ_1 *xb.

Pros of Linear Regression:

1)   Vary Fast
2)   No parameter tuning
3)   Easy to understand and highly interpretable

Model evaluation: The basic approaches are train and test on the same dataset and Train/test Split.

Train and test on the same dataset:

$$Error = \frac{1}{n}\sum_{j=1}^{n}|yj - jy(cap)|. \qquad (3)$$

Where, y is actual value where y(cap) is predicted value. This approach has high **training accuracy** and low **out of sample accuracy.**

Training accuracy is the percentage of correct predictions that the model makes when using the test data set. However, a high training accuracy isn't necessarily good thing. And may result of over-fitting which means the model is overly trained to the dataset, which may capture noise and produce a non-generalized model.

Out of Sample accuracy is the percentage of correct predictions that the model makes on data that the model has not been trained on. Its imp that the model has a high out of sample accuracy so to improve it there are ways: one way is to use another evolution approach called **train/test split.**

Train/Test Split: Involves splitting the data set into training and testing sets respectively which are mutually exclusive. It provides more accurate evaluation on out-of-sample accuracy because the testing dataset is not part of the dataset that has been used to train the data. Issue is the highly dependent on which the datasets the data is trained and tested. Another approach resolves most of the issues know as **K-fold cross-validation.**

K-fold cross-validation: Here k is the number of amount taken to check the data. If k = 4 fold, then it takes first 25% of the data for testing. And rest for training. Then in the next fold the 2nd 25% is taken and continuous for all folds. So the result is average keeping into mind that no data is repeated. k-fold cross-validation in its simplest form performs multiple train/test splits. Using the same data set for each set differently, then the result is average.

Evaluation metrics in regression: Is used to explain the performance of a model. It provides a key role in the development of a model as it provides the insight to areas that require improvement.

ERROR: difference between the data points and trend lines generated by the algorithm

$MAE(\text{Mean absolute error}) = \frac{1}{n}\sum_{j=1}^{n}|yj - jy(cap)|.$

This is the easiest of the metrics to understand. Since it is just an average error.

$MSE(\text{Mean square error}) = \frac{1}{n}\sum_{j=1}^{n}(yj - jy(cap)^2).$

RMSE (Root Mean Square Error) =

$$\sqrt{\frac{1}{n}\sum_{j=1}^{n}(yj - jy(cap)^2)}. \qquad (4)$$

This is one of the most popular.

$$RAE(\text{Relative absolute error}) = \frac{\sum_{i=-1}^{S}|yj - yj(cap)|}{\sum_{i=1}^{S}|yj - yb|}. \qquad (5)$$

$$RSE(\text{Relative square error}) = \frac{\sum_{i=-1}^{S}(yj - yj(cap))^2}{\sum_{i=1}^{S}(yj - yb)^2}. \qquad (6)$$

It is used to calculated R^2 = 1- RSE.

The higher the R^2 the better the model fits the data.

b.   Multiple linear regression

It is the extension of simple linear regression model. The applications are:

1)   Independent variable effectiveness on prediction: If we would like to identify the strength of the effect that the independent variables have on the dependent variables.
2)   Predicting impacts of changes: It can be used to predict the impact of changes that is to understand that how the dependent variable changes when we change independent variable.

It is the method of predicting continuous values. It is very useful for examine which variables are significant predictors of the outcome variable.

$Y(cap) = \Theta_0 + \Theta_1 x1 + \Theta_2 x2 + \ldots + \Theta_n x_n.$

So, $y(cap) = \Theta^T.X$

Θ is the parameters and x is the feature set and independent value and y is dependent. To achieve this goal, we must minimize the error.to find the optimized parameters we should understand the optimized parameters and the way to do it. Optimized parameters are the once which lead to a model with the fewest error. To estimate Θ, there are approaches known as:

1. Ordinary Least square it tries to estimate the value of the coefficient by minimizing thee MSE. It uses Linear algebra operation but it takes long time for large datasets (10k+ rows).
2. Optimization algorithm: to find the best parameters that is you can use a process of optimizing the value of the coefficient by iteratively minimizing the error of the model on your training data.

Like gradient descent: it is a proper approach if you have large dataset.

c. Non-linear regression

It is the method to model non-linear relationship between the dependent variable and a set of independent variables. For a model to be consider non-linear y(cap) must be non-linear function of the parameters Θ, not necessarily the features x. It can be shapes of exponential, logarithmic and logistic or many other types. That is in non-linear regression the model is non-linear by parameters. In contract to linear regression method, we cannot use the ordinary least square methods to fit the data in non-linear regression. As the estimation of parameters is not easy.

d. Introduction to classification

Is a supervised learning approach which means that it is categorized some unknown items into a discrete set of categories or classes, classifications attempts to learn a relationship between a set of featured variables and a target variable of interest. The target attribute in a classification ia s categorical variable discrete values. Classification determines the class label for an unlabeled test case. Classification Applications: Email filtering, speech recognition, biometric identification, document classification and much more. Classification algorithms in ML

1) Decision trees (ID3, C4.5, C5.0)
2) Naïve Bayes
3) Linear discriminant analysis
4) K-nearest neighbor
5) Logistic regression
6) Neutral networks
7) Support vector machines(SVM).

IV.   MACHINE LEARNING IN ANALOG DESIGN

1. Machine Learning for fast and Accurate Mixed-Signal Design optimization [1]: A machine learning model called Artificial Neural Network Meta-model (ANNM) have been explored to capture the highly non-linear nature of analog block. OP_AMP and PLL is being constructed over here where firefly optimization algorithm is applied to an OP_AMP design in the iVAMs 2.0 framework. Whereas, ABC (Artificial Bee Colony) based algorithm performs optimization over the ANN meta-model of PLL. It is observed that ANN achieve more accurate results than

polynomial meta-models with shorter optimization time. The step which are taken into considerations are:

   a. Concept of iVAMs (Intelligent meta-model integrated Verilog-AMS).
   b. Schematic design of OP_AMP
   c. Intelligent meta-model generation with neural networks.
   d. Firefly algorithm using iVAMs
   e. Construction of a neural network in Verilog-AMS
   f. ABC algorithm & ANN based met-modeling for fast & accurate Nano-CMOS mixed-signal design exploration.
   g. A non-polynomial meta-model based design optimization flow of analog/ mixed-signal circuit is presented.
   h. Practical use of non-polynomial ANN & PD of 180nm CMOS PLL is undertaken.
   i. ABC algorithm is used for optimization of PLL physical design that uses the meta-model instead of actual circuit.
   j. Demonstration of PLL is faster and accurate compared to the polynomial meta-model.

2. Regression & ANN models for Electronic Circuit Design [2]: In this particular method, the circuit contain four Resistors whose value are kept on changing for better accuracy. Nominal values of Resistors are decided of the amplifier circuit. Values of resistor are changed from LOW to NOMINAL and then HIGH. After the variation of R1, R2, R3, R4; Gain $\Delta(V) \, and \, \left(\frac{I_c}{I_B}\right)$ are calculated. The dataset is collected form the output and split into training and test set and the graph is Estimated ANN response surface for $\left(\frac{I_c}{I_B}\right)$.

Use Regression polynomial for these two outputs. Different datasets and graphs are created from Interaction effect plot showing Gain $\Delta(V) \, and \, \left(\frac{I_c}{I_B}\right)$ obtained with the ANN. Higher accuracy was obtained with the ANN than with conventional Regression technique.
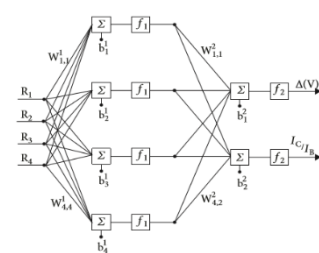

Fig. 3 ANN structure [2]

3. MOM-SVM [3]: Multi-output modelling with Support Vector Machine. Basically, VCO (Voltage Controlled Oscillator) is design in cadence Virtuoso as a test circuit then in SPICE simulator for obtaining the dataset and then k-cross validation, method is being used for tuning. And MATLAB is been used because of LSSVM for modelling is been used.
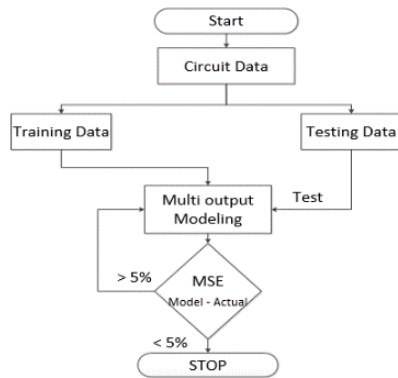
Fig.1 Flowchart for the MOM_SVM analog circuit [3]

Fig. 2 Flowchart for the method of analog circuit using Sparse Regression [4]

4. Regression Methodologies on Analog Circuit Design [5]: The main purpose of this is to save simulation time with the minimum amount of accuracy loss. The basic features if regression method are: Training Error, Build time, Testing Error, testing time, Variable space and Training data set, among which Build time, Testing Error and Testing Time are the important measures. Gaussian Processes are used as regression methodologies because they are easy to train and use. The process includes:

    a. GP Scikit: it will smooth the prediction when noisy training dataset is being used. It is implemented in Python.

    b. GP Rasmussen: It is implemented in MATLAB.

    c. GPLP Park: It is implemented in MATLAB.

    d. SPGP: It will help to choose pseudo-inputs among training datasets. It is implemented in Python.

    e. KRG: It is implemented in Python on Scikit-learn.

    f. KRC: It is used to sub-divide input space during training. It is implemented in Python.

    g. Py XGPR: It is implemented in Python.

    h. FFX: It is used to predict the data and implemented in Python.

    i. RIWD: It works with large dataset and also for improving prediction error. It is implemented in Python.

5. Analog Circuit Optimization using Sparse Regression and Error Margining [4]: Run Cadence Virtuoso for sample points. In which first, build linear models using Sparse regression. Along with that, derive the equation for worst case model using:

$$g_{WC}^{l,u} = \sum_{i=1}^{n} \delta_i x_i \pm c - 3 * \sigma \geq G^{l,u} \qquad (7)$$

Draw the schematic of circuit (OP_AMP), then find the parameters like power consumption, Gain Frequency, Gain, Phase Margin, Slew Rate, Input Offset, Output Swing.

Draw the schematic of different circuit (Low Noise Amplifier), then find the parameters like Noise figure, Forward Reflection, Forward Transmission, Reverse Reflection.
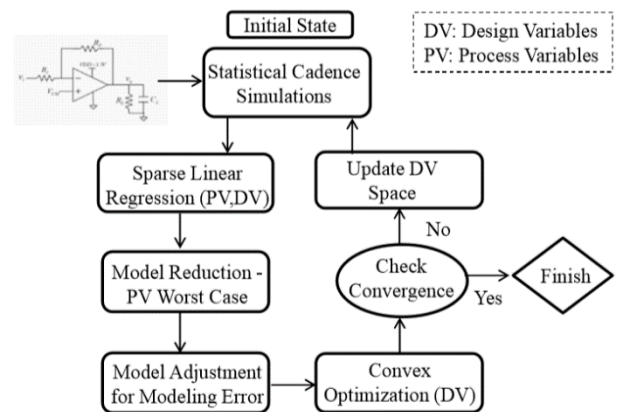
CONCLUSION

Analog means having various parameters and various data, if data changes (minor) i.e., if temperature changes the output may get affected, as there are various data; it is manually not possible to get the proper output. If the data change (major) i.e., if the input voltage gets fluctuated from 0.8V to 1.2V the data will change automatically which will change the output and every time it is not possible to change & update the output manually. So, Machine learning is used over here. Basically in ML we are teaching the machine with different languages like python or R to learn something. Here, we are having a dataset from the output which may be given to different topologies of ML like in Regression we have simple linear regression, multiple regression, Decision making tree, polynomial and many more. Now, we have different parameters to check like power consumption, delay, accuracy which is one of the important parameter when Machine learning is taken into consideration, i.e. in regression model there are various sub-models; and also apart from that there are different models of ANN and CNN which is chosen according to the accuracy of output. So, as close loop we can do is to save simulation time, to provide better accuracy and we can do multi-output system.

REFERENCES

[1] Saraju P. Mohanty Elias Kougianos, "iVAMS2.0: Machine-Learning-Metamodel-Integrated Intelligent Verilog-AMS for Fast and Accurate Mixed-Signal Design Optimization", arXiv:1907.01526v1 [eess.SP] 11 Jun 2019.

[2] M. I. Dieste-Velasco, M. Diez-Mediavilla and C. Alonso-Tristán, "Regression and ANN Models for Electronic Circuit Design", Hindawi Complexity Volume 2018, Article ID 7379512, doi.org/10.1155/2018/7379512.

[3] B.Shivalal Patro, Sushanta K. Mandal, "A Multi Output Formulation for Analog Circuits Using MOM-SVM", Indonesian Journal of Electrical Engineering and Computer Science Vol. 7, No. 1, July 2017

[4] Mohamed Baker Alawieh, Fa Wang, Rouwaida Kanj, Xin Li and Rajiv Joshi, "Efficient Analog Circuit Optimization Using Sparse Regression and Error Margining", 17th Int'l Symposium on Quality Electronic Design, 2016 IEEE.

[5] Ivick Guerra-G´omez, Trent McConaghy and E. Tlelo-Cuautle, "Study of Regression Methodologies on Analog Circuit Design", Esteban Tlelo-Cuautle, DOI: 10.1109/LATW.2015.7102504, 2015.

[6] Anirban Sengupta, Saraju Mohanty, "VLSI Circuits and Systems Letter", IEEE, Volume 5, Issue 1, 2019.

[7] Ramin M. Hasani, Dieter Haerle, Christian F. Baumgartner, Alessio R. Lomuscio and Radu Grosu, *"Compositional Neural-Network Modeling of Complex Analog Circuits"*, Internatinal joint conference on Neural Networks, 2017.

[8] Hao Cai, Jean-François Naviner, Hervé Petit, "Statistical Methods Applied to CMOS Reliability Analysis -A Survey", HAL Id: hal-01570282, 2017.

[9] Mriganka Chakraborty, "Artificial Neural Network for Performance Modeling and Optimization of CMOS Analog Circuits", International Journal of Computer Applications (0975 – 8887) Volume 58– No.18, November 2012

[10] Li Yu, "Efficient IC Statistical Modeling and Extraction Using a Bayesian Inference Framework", MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 2009.

[11] João Pedro da Silva Rosa, "Using Artificial Neural Networks to Size Analog Integrated Circuits", Electrical and Computer Engineering, 2018.