Data-Driven AI Modelling of Emerging FETs: XGBoost Deep Learning Approach

Authors: Hammad Omar, The American University in Cairo
Lin Zhuofan, Universiti Putra Malaysia

Abstract

This paper tackles the urgent need for advanced modeling techniques for newly developed emerging Field-Effect Transistors (FETs). We aimed to leverage XGBoost for fitting the current-voltage (I-V) characteristics of Carbon Nanotube FETs (CNTFETs), thereby addressing the gap in precise modeling for this novel transistor type through an experimental study. Our methodology encompassed a comprehensive review of FET modeling literature, careful dataset identification, and preprocessing to accommodate CNTFETs' unique characteristics. We applied the XGBoost algorithm, fine-tuning hyperparameters to optimize prediction accuracy. A quantile transformation on the output data (Ids) was employed to normalize the distribution effectively, dealing with small values efficiently. XGBoost, renowned for its effectiveness in the computer vision field, also excels in predicting tabular data for regression problems. We expect the results to demonstrate superior performance, evidenced by low Root Mean Square Errors (RMSEs) and a high R^2 score, even with a constrained dataset. With 90% of the data allocated for training and 10% for validation, we anticipate an R^2 score exceeding 99.8% for the trained data, underscoring the model's precision in predicting I-V characteristics. The significance of our findings lies in offering an efficient and reliable method for predicting I-V characteristics, paving the way for further advancements in the field of emerging transistor types.

Introduction

In response to challenges in accurately modeling Carbon Nanotube Field-Effect Transistors (CNTFETs), this research proposes an approach using XGBoost model to train the emerging FETs datasets effectively and improve the prediction accuracy. The utilization of input parameter sets has proven effective in achieving the desired objectives through multiple iterations, demonstrating a notable level of accuracy that can surpass the precision of expensive instrumentation. While neural networks have already been employed as a viable solution to address this challenge, the integration of deep learning has shown exceptional results. While decision tables are commonly used in standard deep learning techniques, the multilayer perceptron stands out for its precise outcomes, particularly noteworthy when dealing with smaller databases [1]. However, after exploring multiple datasets whereas they are tabular data, in which XGBoost and Random Forest are excelling in predicting those tabular data with a higher accuracies. At first, we have trained the dataset using a three-layer MLP model, in which it gives good accuracies; however, there has been some problems concerning the regression and overfitting.

MLPs can model complex, non-linear relationships in data. This is crucial in our case for capturing intricate patterns, interactions, and dependencies that linear models may struggle to represent. On the other hand, as overfitting and limited data are representing a problem in MLP approach. We used the XGBoost approach and tested the datasets on it. Our XGBoost model improves upon the standard gradient boosting method by adding a regularization term to the loss function to avoid overfitting, optimizing the algorithm for speed and efficiency, and enhancing the model's ability to deal with missing data.

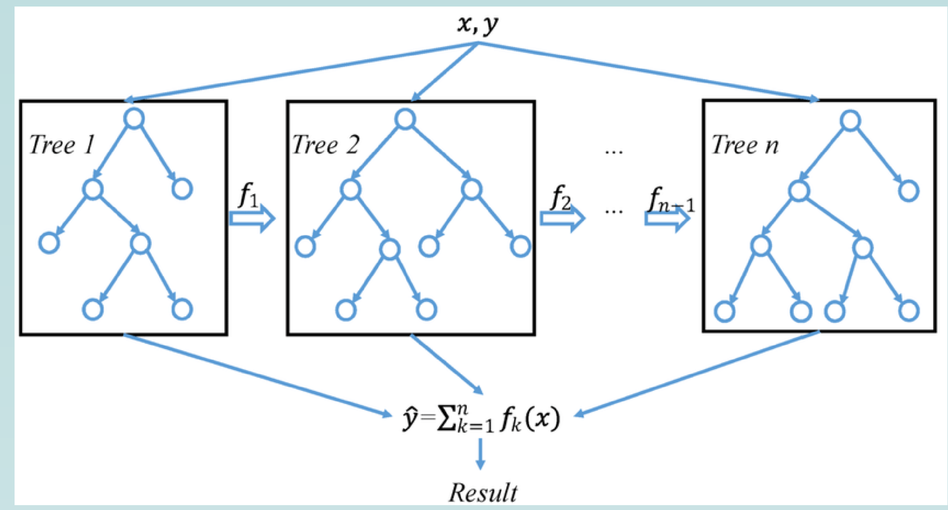

Fig. 1 A general architecture of XGBoost

XGBoost belongs to a family of boosting algorithms which convert weak learners into strong learners in an iterative fashion. It does this by building models sequentially, where each new model attempts to correct the errors of the previous ensemble of models.

Methodology

I. Data Analysis

In order to achieve better accuracy and develop a robust model, we had to understand the data very well and analyze the behavior of each parameter and its correlation with other parameters in the frame of the dataset. Generalization is one of our aims in this research; therefore, after analyzing the dataset and the behavior of the emerging FETs, we could be able to know the physical and modular behavior of the devices and their responses from the reflected results of the trainings and validations, in addition to, correlation matrices comparisons. To apply data augmentation for the purpose of better performance to the real-world data, data analysis and understanding is a crucial step in this approach.

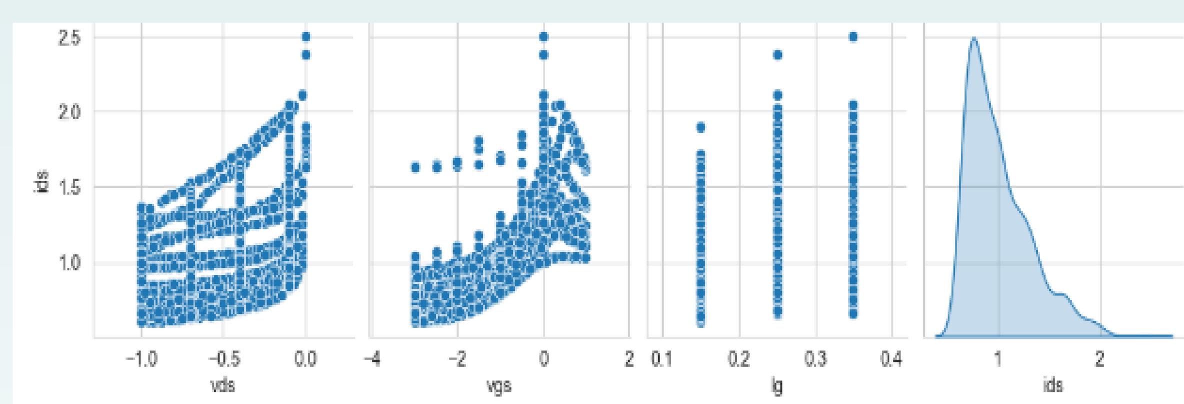

Fig.2 Training Dataset of p-FET before ids normalization

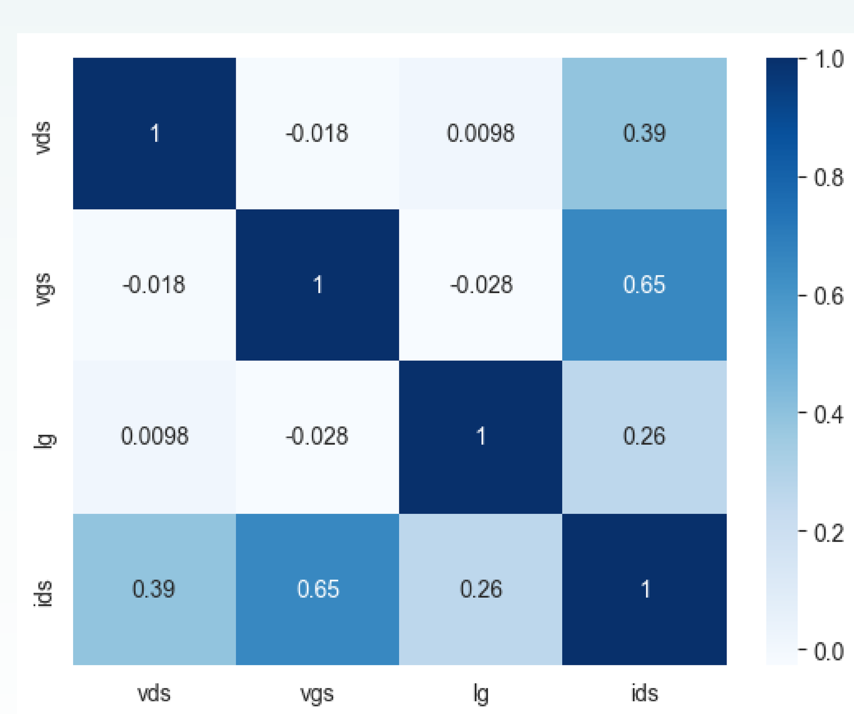

Fig.3 Correlation Matrix of: Vds, Vgs, ld, and ids.

We have noticed during the data analysis the skewness in ids distribution. Also, ids values were too small to work with. Scaling and normalization took part in the pre-processing stage. Here is the training dataset (90% of the original dataset of p-FET) before normalization.

II. Pre-Processing

In the pre-processing stage, we used quantile transformation for the ids values normalization. Quantile transformer has a special methodology to ease the working with the ids values. The methodology involves ranking the values in the original dataset of P-FET. Each value is assigned a rank based on its position in the sorted order of the dataset. After that, the quantiles are computed for each data point based on its rank. Quantiles represent the proportion of data below a certain value. Finally, the transformed data should have a normal (Gaussian) distribution.

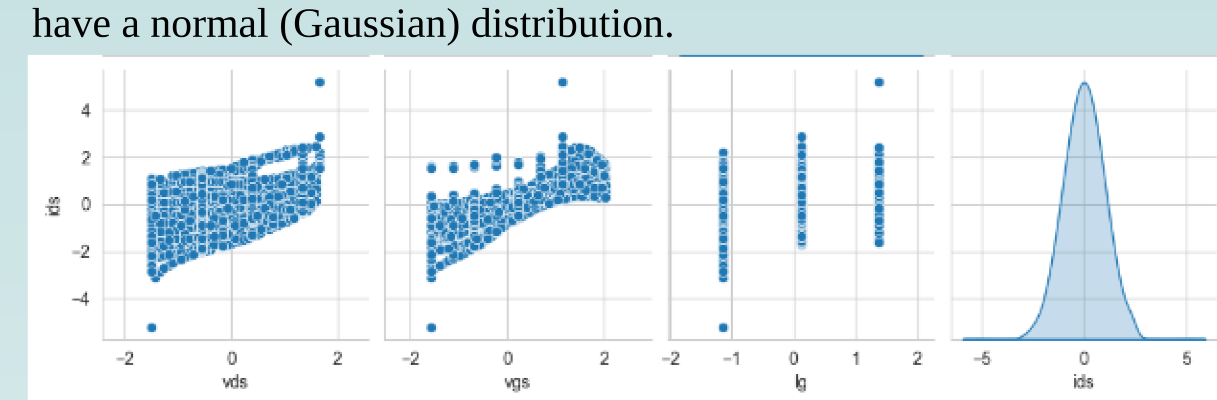

Fig.4 Training Dataset of p-FET after ids normalization

Results

I. XGBoost approach in comparison with K-Fold Validation and MLP structure:

Using XGboost approach showed improvement showed some improvement compared to 5-fold training with MLP structure the same preprocessing.

	Average	
	R2 Score	MAE
K fold / MLP # 5	0.9953	1.66E-05
XGBoost	0.9992	4.324E-6

Table 1 Results Comparison between MLP and XGBoost Approaches

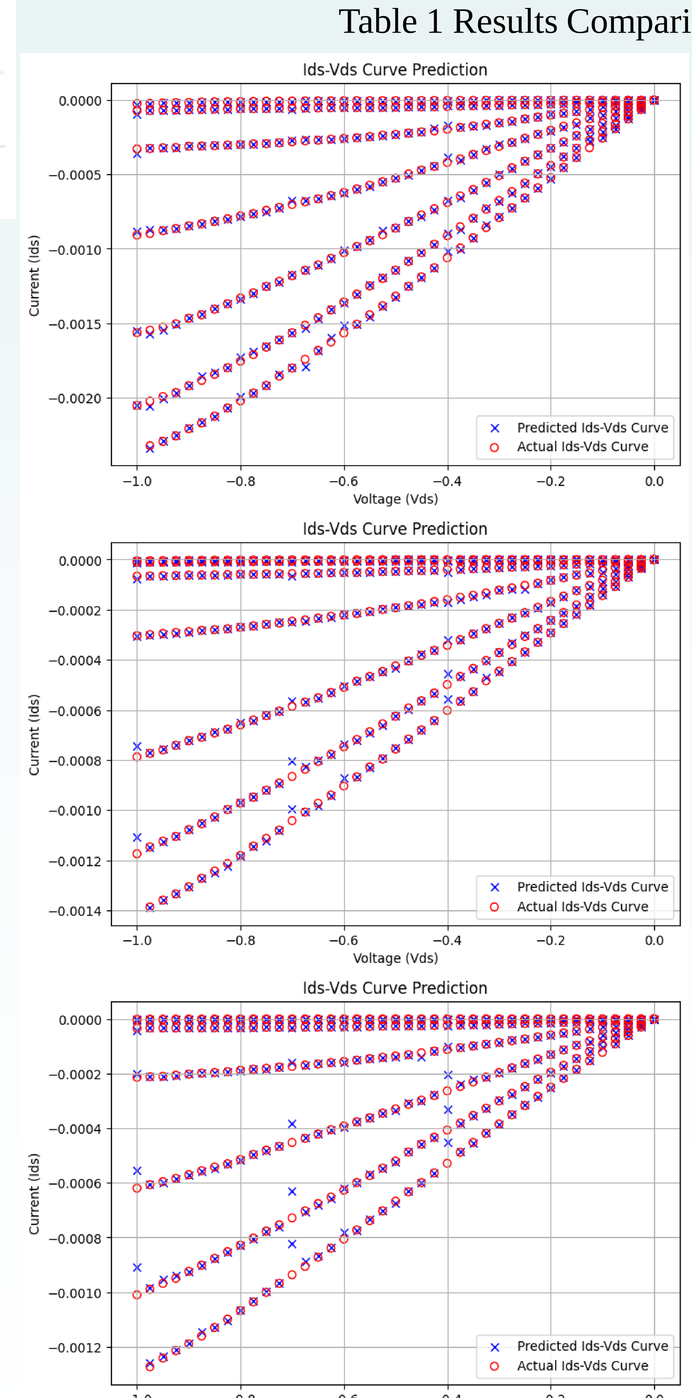

Fig.5 Ids vs Vds curve for PFET

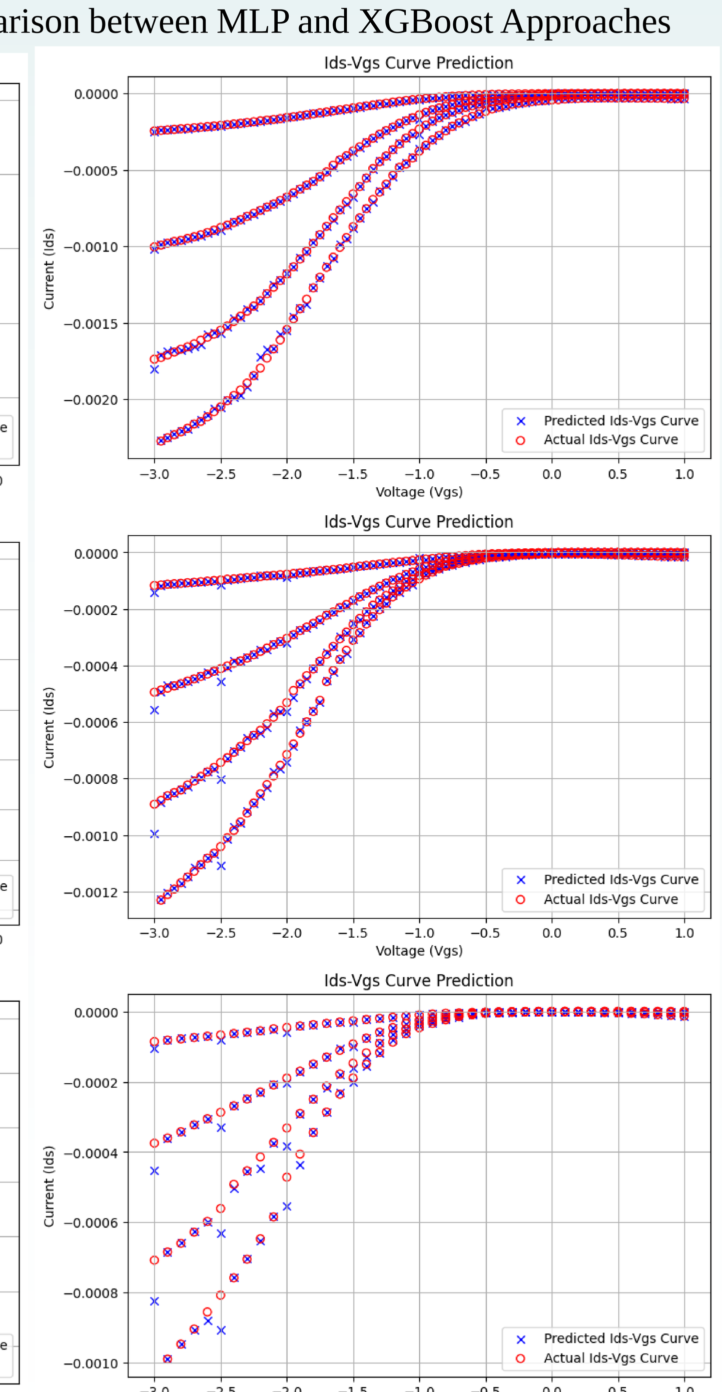

Fig.6 Ids vs Vgs curve for PFET

Discussion

Looking at the plots, the predicted Ids-Vgs and Ids-Vds curves are closely aligned with the actual curves, indicating that the model can effectively capture the behavior of the FETs under various voltage conditions. This tight alignment suggests that the model has learned the underlying physical phenomena governing the FETs' behavior with a high degree of precision.

For XGBoost in comparison with K-Fold Validation and Multilayer Perceptron (MLP) structure, the improvement you've noted likely stems from XGBoost's efficient handling of non-linear relationships and its capacity for handling various types of data. XGBoost is also less likely to overfit compared to a deep learning approach like MLP, especially when the dataset is not exceedingly large or when the signal-to-noise ratio is high, which seems to be indicated by the high R2 score and low MAE

Key features of XGBoost Approach:

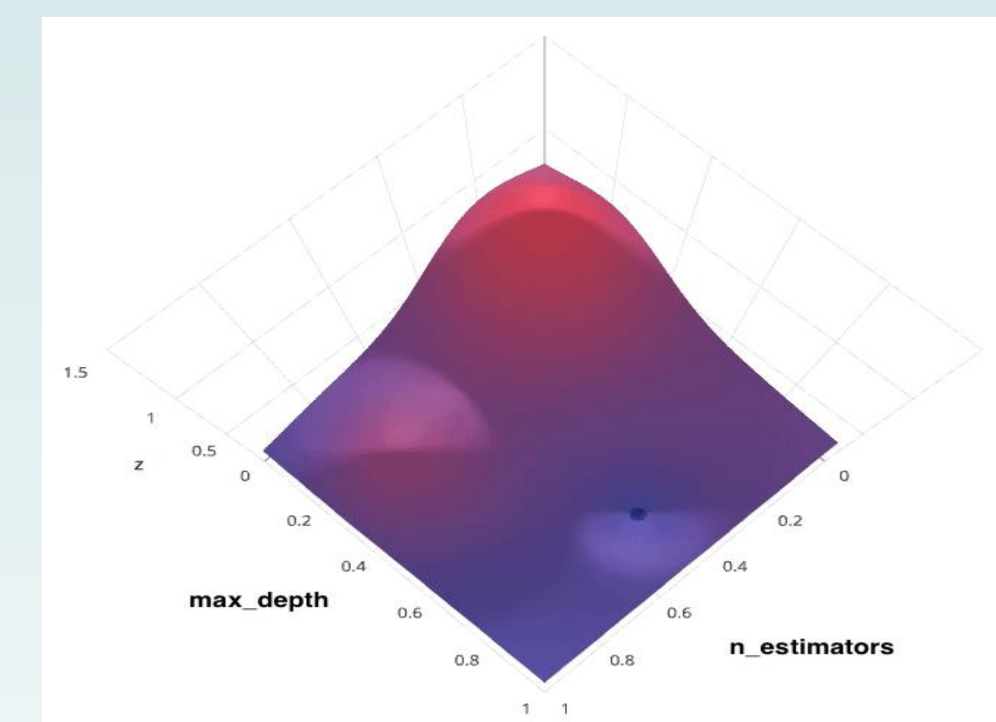

Fig.7 Hyperparameter Search Space

The hyperparameter space is the set of all possible combinations of hyperparameters that can be used to train a machine learning model. It is a multidimensional space, with each dimension representing a different hyperparameter.

Acknowledgements

We greatly acknowledge Prof. Yanan Sun, Shanghai Jiao Tong University and Prof. Yongfu Li, Shanghai Jiao Tong University for their great mentoring and supporting during this research time period.

References

[1] Deyasi, A., Bhattacharjee, A. K., Mukherjee, S., & Sarkar, A. (2021). Multi-layer perceptron based comparative analysis between CNTFET and quantum wire FET for optimum design performance. Solid State Electronics Letters, 3, 42-52.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Ghosh, J., Lim, S. Y., & Thean, A. V. Y. (2021, September). Bridge-Defect Prediction in SRAM Circuits Using Random Forest, XGBoost, and LightGBM Learners. In 2021 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) (pp. 259-262). IEEE.