

BASIC REPORTING

This study focuses on an important issue in health. The study is particularly significant due to the use of genetic algorithms to optimize the performance of deep neural networks. However, the authors did not provide a justification for using CNN and BiLSTM, especially when research has shown that models based on transformer architectures perform much better than CNN and BiLSTM models.

There is need to correct line 44, *“Therefore, it is essential to develop more efficient non-invasive e predictive systems that can”*

On line 73 the authors stated *“Current studies on the prediction of VDD levels using ML and standard features have limitations [1,3].”* **But they didn’t state what are the current limitations in the status quo.**

The research gap is not clearly defined or adequately justified. While the authors highlighted in the literature review that there is existing work on predicting vitamin D, they did not sufficiently explain why further research on this subject is necessary.

Additionally, the authors did not provide information on the F1 score, Recall, Precision, and Accuracy of the existing work, which would help clarify what they aim to improve with their approach. It would be highly beneficial for the authors to justify the need for using CNN, BiLSTM, and Genetic Algorithms (GA) in the context of the existing F1 score, Recall, Precision, and Accuracy metrics.

EXPERIMENTAL DESIGN

The authors did an excellent job of collecting data from scratch. This is a significant undertaking and makes the work highly contextual to their specific circumstances. The experimental design is generally sound and clear enough to be understood and reproduced. It is commendable that the authors employed an oversampling technique to balance their training data.

It is also recommended that the authors explore other encoding techniques (such as one-hot encoding or leave-one-out encoding) and evaluate their impact on the performance of the ML models. The use of convolutional layers to extract features built by the embedding layer is an interesting approach. However, it would be beneficial for the authors to provide readers with a diagrammatic representation of the model’s architecture for better understanding.

The statement in line 321 is somewhat confusing: “The Sigmoid technique was employed to determine the probability of each category: 'VDD1,' 'VDD2,' 'VDD3,' and 'VDD4.’” This appears inconsistent with the earlier claim that the activation function for the output layer is SoftMax.

Additionally, line 352 requires restructuring, as the content is not clearly visible.

It is recommended that the authors run an experiment comparing the use of Genetic Algorithms (GA) to optimize the DNN with conventional methods, such as grid search, to validate whether GA is indeed the most effective optimization method for this task.

There is further confusion in line 365. The authors mention using the "Mean Squared Error (MSE) metric for predicting vitamin D level"; however, given that the activation function for the output layer is SoftMax, this implies they were predicting a class, not a continuous value. In the dataset, there are two variables: "Vitamin D Level (ng/mL)" and "VDD Deficiency Level." It is unclear which of these two variables was actually predicted. Clarification on this point would significantly improve the manuscript.

VALIDITY OF THE FINDINGS

Table 3 does not clearly show how the iterations of the CNN+LSTM configurations resulted in the 10 different metrics. Additionally, it is unclear from Tables 2 and 4 how the results presented in Table 3 were derived. As previously mentioned, the metrics used in Table 3 indicate that this task was a classification task. However, the authors used MSE to optimize the CNN+LSTM model, which could lead to confusion. The authors are encouraged to maintain consistency throughout the paper to avoid misunderstandings.

It is noteworthy that the authors successfully ran five other models on the same dataset and compared the results with their model. It is particularly interesting that their method performed well compared to the existing models.

However, the general complexity of the algorithm is not clearly defined. Based on the provided function, it is unclear whether the complexity falls within the linear or polynomial range.