Dear editor and reviewer ,

We really appreciate your helpful comments and suggestions, and we have carefully revised the manuscript（No.:88203）according to each suggestion.Taking reviewers' comments into consideration, we have rephrased the manuscript. All the changes we made in the manuscript have been highlighted in colorful text.

The following are the details.

editors' comments:

We have found a professional language polishing agency to revise the article.

Reviewers' comments:

Reviewer #1:

1.The introduction is skillfully written and provides valuable background information. However, it is worth noting that the authors could have strengthened their work by including additional literature that elucidates the critical role of machine learning in facilitating the identification of lymph node metastasis in gastric cancer. Additionally, the authors should consider a more comprehensive discussion of gastric cancer lymph node metastasis in the background section.

Re:The working mechanism of machine learning and gastric cancer lymph node metastasis related literature have been added to the introduction

2. Logistic regression is not the best model, both in the textual description of the results section and in the table presentation, but the authors used logistic regression as an important model for presenting the results. Can the author's team give a reasonable explanation?

Re: 1. Although logistic regression is not the best model, logistic regression, as a traditional model building method, is the most widely used in machine learning. 2. In

this study, the establishment of logistic regression can also be used as a comparison with other models, which can reflect the accuracy and reliability of other models. 3. On the basis of logistic regression, nomogram can be developed as a visualization of the model to make the internal structure of the model more intuitive, which is not capable of other models. Therefore, the characteristics of logistic regression are described extensively in this paper.

3. The conclusion part of the article describes that GBM has the best performance, the highest predictive value and accuracy. Through this study, machine learning can tap into the ability of clinical data to reflect disease, which can help clinicians assess patients' conditions and make better treatment decisions. However, we think the author team just illustrated that GBM works better compared to other machine learning models, and did not prove how the algorithm helps clinicians assess patients' conditions and make decisions accordingly. I hope the author team can give a reasonable explanation.

Re: Thank you very much for raising this question. To solve this problem, we established a web risk calculator based on the GBM model (figure5). The probability of lymph node metastasis in patients with gastric cancer can be obtained by inputting clinical characteristics of patients. This calculator is simple to operate and can directly help clinicians make diagnosis and treatment judgments. You can visit our web site visit: https://gastric.shinyapps.io/gbm4lymph/.

4. We found that Table 1 of the article uses non-English descriptions, which is not in line with international journal publication standards. And Table 1 does not

describe the abbreviations accordingly.

Re: Has been changed as you requested

5. We wish the author team could have accurately depicted the exclusion and inclusion criteria for the studies in Figure 1 to make the picture more clear and concise.

Re: Figure 1 has been changed as you requested

6. Regarding the ethical aspects of the study, the article describes that the study was approved by the Ethics Committee of Xuzhou Medical University Hospital. However, this study used relevant data from two regional hospitals, and the ethics of the other hospital was not described accordingly.

Re: Our group has contacted the Ethics Committee of Jining First People's Hospital and has been authorized, and the paper ethical documents are being approved.

Reviewer #2:

I am really grateful to review this manuscript. In my opinion, this manuscript can be published once some revision is done successfully. I made one suggestion and I would like to ask your kind understanding. This study used numeric data from 369 patients, applied seven machine learning models and achieved the areas under the curves of 92% with the random forest and boosting for the prediction of lymph node metastasis in gastric cancer. This study presented variable importance results as well. I would argue that this is a good achievement. However, it can be noted that the Shapley Additive Explanations (SHAP) summary plot is very effective to identify the direction of association between lymph node metastasis in gastric cancer and its major

predictor derived from variable importance. In this context, I would like to ask the authors to derive the SHAP summary plot.

Re：We have created a SHAP summary plot and described it.

Final：Thank you for your valuable comments. We have revised and explained your comments.