Review of: "A Graphical User Interface Based on Logistic Regression Approach for Malarial Detection"

Kim Gaik Tay[1]

1 Universiti Tun Hussein Onn Malaysia

Potential competing interests: No potential competing interests to declare.

The authors compared the performance of Random Forest (RF), Support Vector Machine (SVM), and Logistic Regression (LR) in classifying malaria using k-fold (k = 2, 4, 5, 10) cross-validation. They found that LR outperformed the other two classifiers, leading them to develop an LR-based GUI for malaria detection. The paper is well-written; however, some concerns need to be addressed to improve its quality as follows:

1. Add a paragraph on the research gap and contribution, respectively.
2. Include the source of the malaria dataset. List the initial features and their correlation matrix. Also, list the reduced features using correlation coefficients.
3. Write out the full names of PTB and CA on their first mention.
4. For abbreviations, use the full name only on the first mention and abbreviations thereafter. For example, original dataset (D).
5. List down hyperparameters used for RF, SVM, and LR.
6. Clarify how k=10 was determined to be the best fold, as k=4 yielded the highest accuracy and precision for LR.
7. Add a verification / comparative of the current results with existing studies.
8. Show and describe the initial / main GUI. Do users need to input all feature values? Specify what 0 and 1 represent.