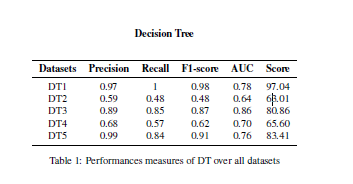
DT.

Le tableau 1 montre les mesures de performance (précision, rappel, mesure F et précision) des résultats de notre classificateur DT après expérimentation sur tous nos jeux de données. L’observation montre que le classifier DT présente les meilleurs score sur les jeux de données DT1, DT3 et DT5 avec des scores de 97.04%, 80.86% et 83.41% respectivement. On retrouve également les plus grandes valeurs des AUC (Area Under the Curve) pour ces mêmes jeux de données avec 0.78, 0.86 et 0.76 respectivement. Cependant on note que les valeurs de la sensibilité sont plus élevées que les valeurs de la spécificité sur les jeux de données DT1 et DT5 et qu’ils sont sensiblement identiques pour les jeux de données DT2, DT3 et DT4. Ceci signifie que la DT est plus enclin à prédire aussi bien si patient donné est atteint de paludisme ou s’il ne l’est pas sur les jeux de donnés DT2, DT3 et DT4 contrairement sur les jeux de donnés DT1 et DT5 ou notre classifieur est performant uniquement pour prédire si un patient donné est atteint de paludisme. Cette même tendance est observée sur les F-scores des valeurs variant de 0.91 à 0.98 sur les jeux de données DT1, DT3 et DT5.

Table 1 shows the performance measures (precision, recall, F measure and precision) of the results of our Decision Tree classifier after experimentation on all our datasets. The observation shows that the best scores of our classifier are achieved on the datasets DT1, DT3 and DT5 which are 7.04%, 80.86% and 83.41% respectively. Noticing that sensitivity values are higher than specificity values for Also AUC (Area Under the Curve) values are higher ​​for these same datasets which are 0.78, 0.86 and 0.76 respectively. However, we note that the sensitivity values ​​are higher than the specificity values ​​on the datasets DT1 and DT5 so that they are substantially identical for the datasets DT2, DT3 and DT4. This means that DT is more inclined to predict as well whether a given patient has malaria or he doesn’t, on the datasets DT2, DT3 and DT4, while our classifier on the datasets DT1 and DT5 our classifier is only efficient in predicting whether a given patient has malaria. This same trend is observed on the F-scores which higher values varying between 0.91 and 0.98 on the datasets DT1, DT3 and DT5.



**Discussion**

The research has shown that there is not necessarily a single best classification tool, but instead the best performing algorithm will depend on the features of the dataset to be analysed, with particular emphasis on health care data, which are discussed in the paper.

**Conclusion**

This research has indicated that in practice there is no single best classification tool, but instead the best technique will depend on the features of the dataset to be analysed and any preferences of end-users. The research has made a start in investigating what these features are with particular emphasis on health care data. A summary of the main findings are as follows: