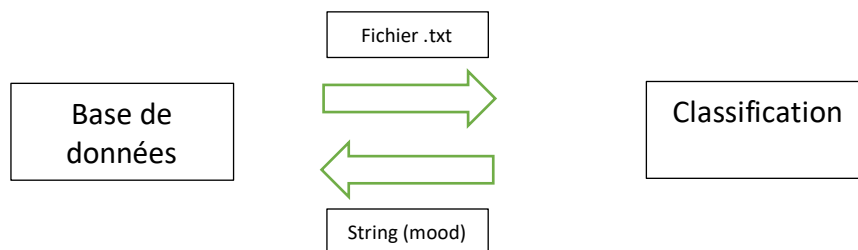


Module Classification :

Le module Base de données utilise la fonction de classification développée dans le module Classification. Autrement dit, la base de données donne en entrée un fichier .txt à traiter, issu du découpage préalable du livre téléchargé par l'utilisateur, au classifieur. Celui-ci renvoie en sortie un string qui désigne l'une des 48 classes (ou ambiances) sur lesquelles il a été entraîné.

On peut par ailleurs noter que le découpage préalable en fichiers .txt se fait selon les paragraphes des pages HTML du fichier epub associé, afin de faciliter l'affichage de ces fichiers epub sur l'interface graphique développée par le module Android.

Enfin, pour que la base de données puisse utiliser la fonction, il a fallu sauvegarder et lire les données d'entraînement du classifieur, ainsi que toutes les fonctions permettant le traitement du fichier .txt dans des fichiers binaires.



Nous avons effectué une batterie de tests concernant la qualité du classifieur développé par le module Classification, dont vous pourrez retrouver les fonctions sur GitLab à l'adresse [Pact35/modules/TestIntegration/src/PythonTest/TestMoodFinal.py](https://gitlab.com/Pact35/modules/TestIntegration/src/PythonTest/TestMoodFinal.py).

Pour cela, nous avons testé la fonction sur près de 4000 fichiers .txt dont nous connaissons leur « mood » associé et avons incrémenté notre score d'une unité à chaque fois que le classifieur renvoyait un string correspondant exactement à son « mood » préalablement défini. Nous avons également enregistré les temps d'exécution du classifieur ainsi que le nombre de textes testés, et ce pour chaque classe. En voici les résultats. En moyenne, sur toutes les classes, nous avons une classification qui avoisine les 35,505% d'accuracy (avec un critère de réussite très strict à savoir l'ambiance retournée correspond-elle oui ou non à celle pensée au préalable).

