The results of classification with DT over embeddings of 2 most populated regulatory rules in the whole DPA dataset:(with legalBERT)

The exact result depends on the random seed as well, i.e., how may positive sample per label exist in the training and test set?

```
Binary classification for R7
Metrics for R7:
Accuracy: 0.8752436647173489
Precision: 0.8977272727272727
Recall: 0.7745098039215687
F1 Score: 0.831578947368421
              precision    recall  f1-score   support

           0       0.86      0.94      0.90       309
           1       0.90      0.77      0.83       204

    accuracy                           0.88       513
   macro avg       0.88      0.86      0.87       513
weighted avg       0.88      0.88      0.87       513
```

```
Binary classification for R16
Metrics for R16:
Accuracy: 0.7953216374269005
Precision: 0.7142857142857143
Recall: 0.046296296296296294
F1 Score: 0.08695652173913043
              precision    recall  f1-score   support

           0       0.80      1.00      0.88       405
           1       0.71      0.05      0.09       108

    accuracy                           0.80       513
   macro avg       0.76      0.52      0.49       513
weighted avg       0.78      0.80      0.72       513
```