**Random Forest Model**

1. Initialize the model
2. Fit the model
3. Evaluate the model using f1 score
4. F1 score = 0.3686013284838245
5. Make predictions on the test data using the trained model
6. Print classification report for the model's predictions on the test data

Precision recall f1-score support

0 0.28 0.08 0.13 1017

1 0.86 0.91 0.88 27522

2 0.36 0.14 0.20 6646

3 0.67 0.85 0.75 14239

4 0.60 0.56 0.58 4644

5 0.15 0.06 0.09 403

6 0.00 0.00 0.00 9

7 0.45 0.34 0.39 41

accuracy 0.75 54521

macro avg 0.42 0.37 0.38 54521

weighted avg 0.71 0.75 0.72 54521

which means that:

the model correctly predicts 75% (accuracy of 0,75) of all instances across all classes, but this can be misleading in imbalanced datasets (where certain classes may dominate)

The macro average considers all classes equally, regardless of their support (number of instances), which can highlight issues in the model's performance on less represented classes. In this case:

1. Precision: 0.42
2. Recall: 0.37
3. F1-score: 0.38

The weighted average takes into account the support of each class, giving more importance to classes with more instances. These metrics indicate that the model performs better overall when accounting for class imbalances. In this case:

1. Precision: 0.71
2. Recall: 0.75
3. F1-score: 0.72

Key words:

Precision: measures the proportion of true positive predictions for a given class out of all the positive predictions made for that class (we want to increase). A high precision value indicates that when the model predicts a class, it is more likely to be correct

Recall: measures the proportion of true positive predictions for a given class out of all the actual instances of that class (we want to increase). A high recall value indicates that the model is good at identifying all instances of a class

F1-score: is the harmonic mean of precision and recall, providing a balanced measure that takes both false positives and false negatives into account (we want to increase)

**Random Forest with GridSearch**

Classification Report:

precision recall f1-score support

0 0.78 0.36 0.50 1017

1 0.84 0.99 0.91 27522

2 0.60 0.04 0.08 6646

3 0.67 0.91 0.77 14239

4 0.76 0.31 0.44 4644

5 0.00 0.00 0.00 403

6 0.00 0.00 0.00 9

7 0.00 0.00 0.00 41

accuracy 0.77 54521

macro avg 0.46 0.33 0.34 54521

weighted avg 0.75 0.77 0.72 54521

Which means:

The accuracy increased 2%

In the macro average :

1. Precision: 0.46 (increased 4%)
2. Recall: 0.33 (decreased 4%)
3. F1-score: 0.34 (decreased 4%)

In the weighted average:

1. Precision: 0.75 (increased 4%)
2. Recall: 0.77 (increased 2%)
3. F1-score: 0.72 (its the same)

Nota: experimentei com o smote e deu ressoltados piores daí ter tirado.