Homework Challenge

Hello students,

congratulations for having completed Part 7 - Natural Language Processing.

If you are up for some practical activities, here is a little challenge:

1. Run the other classification models we made in Part 3 - Classification, other than the one we used in the last tutorial.

2. Evaluate the performance of each of these models. Try to beat the Accuracy obtained in the tutorial. But remember, Accuracy is not enough, so you should also look at other performance metrics like Precision (measuring exactness), Recall (measuring completeness) and the F1 Score (compromise between Precision and Recall). Please find below these metrics formulas (TP = # True Positives, TN = # True Negatives, FP = # False Positives, FN = # False Negatives):

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Precision = TP / (TP + FP)

Recall = TP / (TP + FN)

F1 Score = 2 * Precision * Recall / (Precision + Recall)

3. Try even other classification models that we haven't covered in Part 3 - Classification. Good ones for NLP include:

- CART
- C5.0
- Maximum Entropy

Submit your results in the Q&A for this Lecture or by pm and justify in few words why you think it's the most appropriate model.

Enjoy Machine Learning!

Best to all,

Hadelin