Project Instructions

Pick a textual dataset from https://www.kaggle.com/datasets

Make sure you pick a classification problem (not regression). You should split the data to train and test.

You should write a classifier using KNN and a second classifier using Naïve Bayes. You should compare the results you obtained from both methods.

Your code should use Pandas. Your code should report accuracy, precision, recall, and f-measure. Report your accuracy also on the train and also on the test. (Which would you expect to be higher?)

You will be presenting your work during our last lesson on 6/3 (create a PowerPoint presentation). You will be asked questions to ensure your understanding of the material. You must come with the final report printed and hand it in.

Final report: the final report must be at least 2 pages long. You should have a section about KNN and a section about Naïve Bayes. You should also compare the results you obtained in both methods. The report should include some screenshots of the results you obtained.

Do not include any code in your report.

Good luck!