Principles of AI Engineering

Exercise 3

Prof. Dr. Steffen Herbold

ENGINEERING
UNIVERSITY OF PASSAU
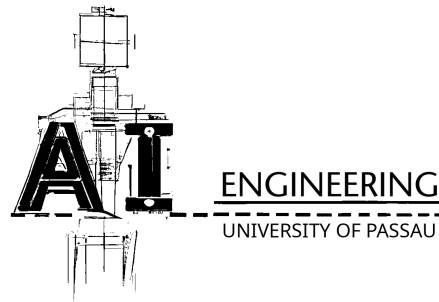
Author: Lukas Schulte

Project task

Build and train a random forest model **pipeline** with sklearn. Re-use a version of the text pre-processing from exercise 02.

Hint: In a later exercise, you will need to integrate the model into a FLASK backend. This means you will need to modularize your pipeline and the pre-processing, such as both can be re-used in the application.

Questions

1. Describe what the components of your model pipeline do.
2. Describe how you will modularize your model to be integrated into the project's application (outside the Jupyter notebook).
3. Describe the drawback of a Random Forest model in the application considering concept drift. Introduce a different model which you could use in the application that would mitigate concept drift.