Class 3 - Discrimination & Fairness: Pre-Class Task

In the second class, we will dive into the challenge of fairness of machine learning models.

In this pre-class task, you will develop a simple classication model that takes a short textual biography of a person and returns its occupation.

The model doesn't need to be fancy, you can aim for a "baseline" model with basic preprocessing that achieves reasonable accuracy of at least

80% on the test dataset. For example, Linear Regression on Bag of Words should be sucient, but feel free to explore more powerful model

families. We recommend keeping it simple and use the sklearn package. Finally, you will store your model so that you will be able to load it into

the notebook in class.

Please go through the whole notebook before you start coding. You could plan your work better if you have an overview of the task.

As usual, if you have any questions, please post them in the #class-3 channel in Discord or join the office hours.

Let's start!