Forecast and Simulation **Summer Term 2023**
`http://ml3.leuphana.de/lectures/summer23/FS/`
Machine Learning Group, Leuphana University of Lüneburg

Ulf Brefeld (`brefeld@leuphana.de`)
Yannick Rudolph (`yannick.rudolph@leuphana.de`)

Exercise 2

Due: Thursday, April 27, 2023

Task 1: HMM Training

This is a continuation of Task 1 from Exercise 1.

In the last weeks, Alice, the unorthodoxy weather scientist, has been noticing that her model is making a lot of mistakes. She heard that one can learn the parameters of an HMM from data. Since she has access to several weeks of both weather conditions and Bob's activities, she thought of hiring someone to learn better parameters for the HMM she designed. Since you now know how to train an HMM model in a supervised fashion,[1] you are in charge of implementing the training.

You can find the data in a numpy *.npz* file (*exercise-02-data.npz*). There are 1000 weeks of 5 weekdays each recorded. 'x' holds the activity data encoded as *walk* = 0, *shop* = 1, *clean* = 2. 'y' holds the weather data encoded as *sunny* = 0, *rainy* = 1. Your task is to estimate the parameters of the model.

Bonus task: can you generate data alike the training data (*exercise-02-data.npz*) from the newly trained model?

Task 2: HMM for Named Entity Recognition

Named Entity Recognition (NER) is a Natural Language Processing (NLP) task that consists of identifying and classifying names of entities (people, location, companies, among others) in a given sentence. Build upon the provided code in *exercise-02-NER.ipynb* where we explain the setup, data and evaluation and have implemented a simple baseline.

i) As in Task 2, you are asked to implement the algorithm for supervised training of an HMM. You should then be able to evaluate predictions for latent states made by the Viterbi algorithm. Evaluate the model on the evaluation (dev) set alike the evaluation of the baseline, report the obtained results and compare them to the baseline.

ii) In [Rabiner, 1989, Section III.B] there is another method discussed for predicting latent states. Implement this approach and compare the evaluation against the results with the baseline and with the Viterbi algorithm. Prepare to discuss your findings in class.

[1] At least it is part of the slides provided in the lecture. Tip: it is a "rather simple" maximum likelihood approach involving counting transitions and emissions.

References

[Rabiner, 1989] Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286.