

## Task 1. Booking challenge.

In 2021, [Booking.com](https://www.booking.com) held a contest aimed at creating a system that could predict the next city in a trip. After the contest, the data set was made available at <https://github.com/bookingcom/ml-dataset-mdt> and problem statement is available in the following article: [https://ceur-ws.org/Vol-2855/challenge\\_short\\_1.pdf](https://ceur-ws.org/Vol-2855/challenge_short_1.pdf).

Your task will be to analyse this data set and implement a solution to this problem basing on the solution from the following article: [https://ceur-ws.org/Vol-2855/challenge\\_short\\_4.pdf](https://ceur-ws.org/Vol-2855/challenge_short_4.pdf). Finally create a http service that provides API to use your model.

1. Analyse the data set and prepare a Jupyter notebook presenting the analysis.
2. Prepare an implementation of the model (either in pytorch or in tensorflow) in the form of python scripts along with instructions for running (please provide it as a git repository).
3. Make at least a few attempts to improve the implemented model or build a better model. Some ideas for improvements can be found here: <https://ceur-ws.org/Vol-2855> but testing your own ideas is very welcome.
4. As one of the attempts to improve the implementation, implement a reranker using, for example, xgboost, which will try to improve the order of candidates selected by the deep learning-based model.
5. Evaluate the implemented solutions and prepare a short report (Jupyter notebook or PDF) showing the obtained results. Make sure to include additionally other metrics than accuracy@4 in model evaluation.
6. Implement a HTTP API service that allows querying the model built by you. Provide it as a git repository (separate from the training script repository).

Your solution should include the following components:

- A Jupyter notebook presenting the dataset analysis
- A Git repository with the implemented models (including training and evaluation scripts) and instructions on how to run the training and evaluation scripts.
- Evaluation results presented in a short report (either in PDF or Jupyter notebook format)
- A Git repository with the implemented web service (ensure it is executable and provide sample requests).

Please work independently and do not share your ideas with colleagues.