

The goal of the assignment is to build a machine learning model for Automatic cyberbullying detection, Task 6-2: Type of harmfulness <http://2019.poleval.pl/index.php/tasks/task6> and deploy it as a simple but well designed service.

ML part:

You need to build and validate the model and comment on its fitness. You don't need to create the best model available (focus on building prototype of end-to-end solution), however you can compare the outcome with competition results: <http://2019.poleval.pl/index.php/results/>. If you want you can borrow the ideas and try to reproduce the results. You can also use models mentioned here <https://klejbenchmark.com/leaderboard/> (CBD task is based on PoEval dataset). We want to see how you approach the problem and how you understand model building tasks. Please include all the assumptions that you made.

Deployment part:

After training and saving the model, propose how to deploy it for scalable usage in production. Please provide text discussion and deployment validation, along with deployment files e.g. Dockerfile, JSON/YAML deployment definition etc. with the example usage and results; printout of the results are sufficient.

You will get additional points for:

- Docker image smaller than 4GB
- Solution runs on CPU faster than 1 second per observation (attached performance test)
- CD4ML/MLOps good practices included
- Deployment with Kubernetes
- Clean, self-explanatory code with comments were necessary
- Assumptions made and argumentation behind them included
- Appropriate model validation
- Creative modelling approach
- Balance between simplicity and performance

We suggest coding in Python and using a Python analytical stack (e.g. scikit-learn/Keras/Tensorflow/PyTorch) for model implementation (or libraries that are built on top of these DL engines). Submit your result as a Jupyter notebook, along with necessary dependencies like python source files, requirements and deployment files

Show us your outstanding skills and prove your experience within the provided solution.