



MaChAmp

#068



#### Attacks:

seq	classification
seq-bio	dependency
seq-multi	mlm
string2string	seq2seq

MaChAmp is a multi-task NLP toolkit, it can seemingly effortlessly handle multiple NLP tasks simultaneously. It has functionality for joint training, continuous training, dataset smoothing, loss weights and dataset embeddings.

More information on:  
[machamp-nlp.github.io](https://machamp-nlp.github.io)

## Notes

How to use MaChAmp.

This is what the dataset configuration file looks like:

```
{
  "UD": {
    "train_data_path": "data/ewt.train",
    "validation_data_path": "data/ewt.dev",
    "word_idx": 1,
    "tasks": {
      "lemma": {
        "task_type": "string2string",
        "column_idx": 2,
      }
    }
  }
}
```

Then I can train with the following command:

```
python3 train.py -dataset_config ewt.json
```

And predict with:

```
python3 predict.py logs/ewt/model.tar.gz
data/ewt.dev preds/ewt.dev.out
```