

## Aufgabe 8

### Sentiment-Analyse mit LSTMs

Laden Sie die Sentimentanalyse-Daten an der Adresse <http://www.cis.uni-muenchen.de/~schmid/lehre/Experimente/data/sentiment-data.zip> herunter und trainieren Sie damit einen Klassifizierer auf Basis von LSTMs. Die Zahl vor dem Tabulator gibt jeweils die Bewertung des nachfolgenden Satzes an, die das System lernen soll.

Das neuronale Netzwerk besteht aus

- einer Embedding-Ebene
- einem bidirektionalen LSTM mit zwei Ebenen (`num_layers=2`)
- einer Max-Operation über die Länge des Eingabesatzes
- einer linearen Projektion auf einen Ausgabevektor der Länge 5

Für das Einlesen der Daten und die Generierung der Batches verwenden Sie das `Torchtext`-Modul. Sie können den Code aus dem `Torchtext`-Tutorial als Vorlage nehmen und anpassen. Halten Sie sich aber an die obigen Vorgaben für die Implementierung des neuronalen Netzes. Das Vokabular soll alle Wörter umfassen, die mindestens zweimal auftraten. Verwenden Sie einen *BucketIterator*.

Im Training minimieren Sie die negative Loglikelihood (mit dem `CrossEntropyLoss` von `PyTorch`). Nach jeder Epoche evaluieren Sie das System auf den Entwicklungsdaten und geben die erzielte Genauigkeit aus.

Schicken Sie mir Ihr Programm und die Ausgabe mit den erzielten Genauigkeiten.