DetoxiPy

LSTM Pipeline

Text Processing stage:

- **Homog selected proper nouns**
    - **Homog contractions**
- **Remove selected punctuation**
- **Remove identity words**

Pre-stage NN to train embeddding:

- **concat(pre-trained embedd, own embedd)**
- **("sequence")-level embedding as output**

Initial word embedding matrix
[v[wi]] i=1,…,|V|

Xt

(Validation: tune hyperparams)

Stage t:

(forget)

$$g_t = \sigma(W_{igt} x_{igt})$$

$$*. tanh$$

$$f_t = \sigma(W_{ft} x_t) \,|\, i_t = \sigma(W_{it} x_t)$$

$$o_t = \sigma(W_{gt} x_{gt})$$

ht-1

Overall perf

Bias

Predictive perf

OUTt = ht