## **DetoxiPy**LSTM Pipeline

ref: https://medium.com/@andre.holzner/lstm-cells-in-pytorch-fab924a78b1c

Xt

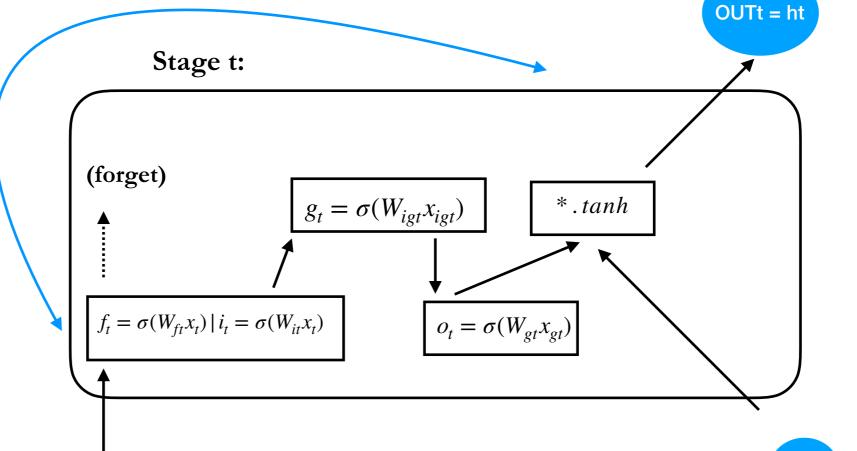
## **Text Cleaning stage:**

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

## Pre-stage NN to train embeddding:

- Use pre-trained embeddings?
- Sentence
  ("sequence")-level
  embedding as output

## (Validation: best hyperparams)



Overall perf

ht-1

Bias

Predictive perf

Initial word embedding vector [v[wi]] i=1,...,|V|