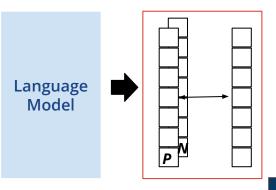


Class: Positive sentiment (*P*) Description: ["like good

recommend"]

**Class descriptions** 

Encode descriptions & keywords using a neural Language Model



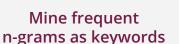
Assign each keyword its "closest" class

 $\begin{array}{l} \lambda_1 \colon \textit{LF ("like"} \to \textit{P)} \\ \lambda_2 \colon \textit{LF ("hate"} \to \textit{N)} \\ \lambda_3 \colon \textit{LF ("funny"} \to \textit{P)} \end{array}$ 

Generate probabilistic labels using DP i.e. estimate  $P(Y | \lambda_1, \lambda_1, ..., \lambda_m)$ 

Y = [(0.65,0.35), ...]

Proba. labels





Labeling functions

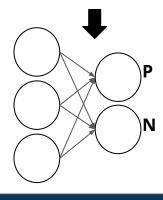


X = ["I liked this film myself. Loved that one scene at...", ... ]



["like", "hate", "funny" ... l

Keywords



**Unlabeled Text Data** 

Downstream classifier