

## Input Sentence

$n$  words,  $K$  features

feature 1 (text)

**the cat sat on the mat**

feature 2

$s1(1) s1(2) s1(3) s1(4) s1(5) s1(6)$

...

feature  $K$

$sK(1) sK(2) sK(3) sK(4) sK(5) sK(6)$

## Lookup Tables

$(d1+d2+...dK)*n$

$LT_W^1$



$\vdots$

$LT_W^K$



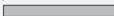
## Convolution Layer

$\#hidden\ units * (n-2)$



## Max Over Time

$\#hidden\ units$



## Optional Classical NN Layer(s)

## Softmax



$\#classes$